HHS Public Access

Author manuscript

Circ Res. Author manuscript; available in PMC 2020 August 02.

Published in final edited form as:

Circ Res. 2019 August 02; 125(4): 379–397. doi:10.1161/CIRCRESAHA.118.314578.

Transcriptomic Profiling of the Developing Cardiac Conduction System at Single-Cell Resolution

William R. Goodyer[1,2], **Benjamin M. Beyersdorf**[1,3], **David T. Paik**[1], **Lei Tian**[1], **Guang Li**[4], **Jan W. Buikema**[1,5], **Orlando Chirikian**[1,6], **Shannon Choi**[1], **Sneha Venkatraman**[1], **Eliza L. Adams**[7], **Marc Tessier-Lavigne**[7], **Joseph C. Wu**[1,8], **Sean M. Wu**[1,2,8]

[1]Cardiovascular Institute, Stanford University School of Medicine, Stanford, CA 94305 USA [2]Pediatrics, Stanford University, Stanford, CA 94305 USA [3]Cardiovascular Surgery, German Heart Center Munich, Technische Universität München, Lazarettstraße 36, 80636 Munich, Germany [4]Developmental Biology, University of Pittsburgh School of Medicine, Pittsburgh, PA 15201 USA [5]Cardiology, Utrecht Regenerative Medicine Center, University Medical Center Utrecht, Utrecht University, 3508 GA Utrecht, The Netherlands [6]Molecular, Cellular, and Developmental Biology, UC Santa Barbara, Santa Barbara, CA 93106 USA [7]Biology, Stanford University, Stanford, CA 94305 USA [8]Cardiovascular Medicine, Department of Medicine, Stanford University School of Medicine, Stanford, CA 94305 USA.

Abstract

Rationale—The cardiac conduction system (CCS) consists of distinct components including the sinoatrial node (SAN), atrioventricular node (AVN), His bundle, bundle branches (BB) and Purkinje fibers (PF). Despite an essential role for the CCS in heart development and function, the CCS has remained challenging to interrogate due to inherent obstacles including small cell numbers, large cell type heterogeneity, complex anatomy and difficulty in isolation. Single-cell RNA-sequencing (scRNA-seq) allows for genome-wide analysis of gene expression at single-cell resolution.

Objective—Assess the transcriptional landscape of the entire CCS at single-cell resolution by scRNA-seq within the developing mouse heart.

Methods and Results—Wild-type, embryonic day 16.5 mouse hearts (n=6 per zone) were harvested and three zones of microdissection were isolated, including: Zone I – SAN region; Zone

Address correspondence to: Dr. Sean M. Wu, Division of Cardiovascular Medicine, Stanford Cardiovascular Institute, Institute of Stem Cell Biology and Regenerative Medicine, Stanford University School of Medicine, Room G1120A Lokey Stem Cell Research Building, 265 Campus Drive, Stanford, CA 94305, Tel: (650)724-4498, smwu@stanford.edu.

AUTHOR CONTRIBUTIONS

W.R.G and S.M.W. conceived the study. G.L, J.B. and O.C. provided significant input on experimental design. W.R.G performed all described wet experiments in collaboration with B.B. (immunofluorescence, RNAscope, iDISCO+, image acquisition/analysis), O.C. (immunofluorescence, image analysis), J.B. (immunofluorescence, image analysis), and S.C. (immunofluorescence, *in situ* hybridization). Embryo microdissection performed by W.R.G. scRNA-seq analysis was performed by W.R.G., D.T.P., L. T., G.L. and S.C. W.R.G. and S.V. performed mouse husbandry and maintenance. W.R.G., B.B. and E.A., with support by M.T.-L., performed iDISCO+, accompanying image acquisition and analysis. B.B., J.C.W. and M.T.-L. made critical revisions of the manuscript. S.M.W. provided financial support for the studies. W.R.G. and S.M.W. wrote the manuscript.

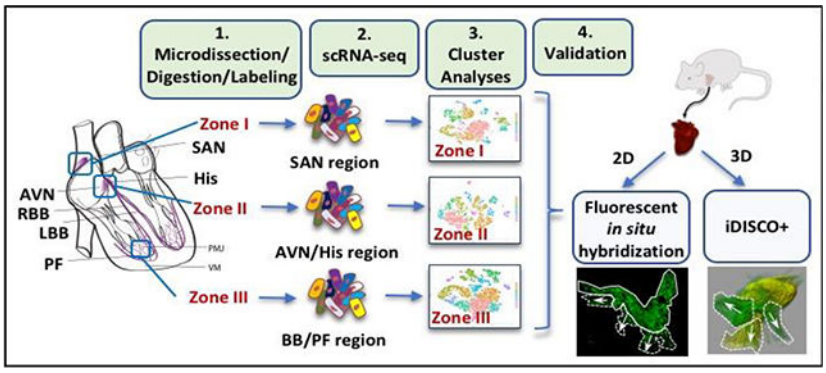
DISCLOSURES

There are no conflicts of interest to declare by any of the authors.

II – AVN/His region; and Zone III – BB/PF region. Tissue was digested into single cell suspensions, isolated, reverse transcribed and barcoded prior to high-throughput sequencing and bioinformatics analyses. scRNA-seq was performed on over 22,000 cells and all major cell types of the murine heart were successfully captured including bona fide clusters of cells consistent with each major component of the CCS. Unsupervised weighted gene co-expression network analysis led to the discovery of a host of novel CCS genes, a subset of which were validated using fluorescent *in situ* hybridization as well as whole mount immunolabelling with volume imaging (iDISCO+) in three-dimensions on intact mouse hearts. Further, subcluster analysis unveiled isolation of distinct CCS cell subtypes, including the clinically-relevant but poorly characterized "transitional cells" that bridge the CCS and surrounding myocardium.

Conclusions—Our study represents the first comprehensive assessment of the transcriptional profiles from the entire CCS at single-cell resolution and provides a gene atlas for facilitating future efforts in conduction cell identification, isolation and characterization in the context of development and disease.

Graphical Abstract



Keywords

Single-cell RNA-sequencing; cardiac conduction system; development; bioinformatics; cardiac development; cardiac electrophysiology; conduction; gene transcription

Subject Terms

Cell Biology/Structural Biology; Developmental Biology; Electrophysiology; Gene Expression and Regulation

INTRODUCTION

The cardiac conduction system (CCS) is made up of specialized heart cells that establish the rhythmic beating of the heart through coordinated contraction of its chambers.[1] This intricate system is comprised of distinct components including the sinoatrial node (SAN), atrioventricular node (AVN), His bundle (His), bundle branches (BB) and Purkinje fibers (PF).[2] The CCS is essential for the formation and normal function of the heart and

disturbance to the CCS can result in severe clinical manifestations including arrhythmias, decreased cardiac output and even sudden death.[3]

Each component of the CCS consists of unique cardiac cell types with their own physiologic and electrochemical properties.[4] Further, the CCS components each have significant intracomponent cell type heterogeneity.[5] The most clinically relevant example of this heterogeneity has been the discovery of “transitional cells”, that exist in each CCS component and provide a cellular bridge to the surrounding working myocardium.[6–10] They are hypothesized to play a role in facilitating the spread of depolarization, providing a high resistance barrier as well as amplifying the current before passing it on to the surrounding myocardium.[9] Transitional cells have been implicated in several important clinical disorders including sinus node dysfunction, heart block and even ventricular fibrillation. However, they remain poorly understood due to challenges in their identification, isolation and, ultimately, molecular characterization.[6,8,9,11]

Currently, the limited number of distinct molecular markers that are known for the different CCS cell types present a number of challenges for additional investigation into their specification, patterning and function. While individual factors have been examined thus far, these studies have been performed at tissue but not at single-cell resolution.[12–14] Specific hurdles to better understanding the molecular signature of the CCS cells have included: 1) low total number of conduction cells in the heart; 2) complex three-dimensional anatomy of the CCS; 3) inability to isolate these cells from the surrounding working myocardium; and 4) the aforementioned significant inter- and intra-component cell type heterogeneity.[1] While prior studies have assessed gene expression within individual components of the CCS using sophisticated techniques such as microdissection, laser capture or even fluorescence-activated cell sorting[12–21], these analyses are unable to discern cell type heterogeneity due to bulk tissue analysis and/or a reliance on transgenic fluorescent reporter models.

Single-cell RNA sequencing (scRNA-seq) allows for global gene expression analysis at single-cell resolution and thereby circumvents many of the aforementioned hurdles including small cell numbers, complex and variable anatomy as well as distortion of unique transcriptional profiles by cell type heterogeneity or contamination from non-conduction cell types. Here, we performed scRNA-seq on over 22,000 cells from wild-type, developing mouse hearts, successfully capturing all components of the CCS including previously unattainable conduction cell subtypes including transitional cells. Using unsupervised weighted gene co-expression network analysis[22,23], we discovered sets of uniquely expressed genes within the entire conduction system, individual CCS components as well as conduction cell subtypes. Validation of both established and novel markers derived from our scRNA-seq data was performed at single-cell resolution using fluorescence RNA *in situ* hybridization (RNAscope). Additionally, whole-mount immunostaining and volume imaging using iDISCO+ (immunolabeling-enabled three-dimensional imaging of solvent-cleared organs) and light sheet microscopy was utilized to visualize, three-dimensionally, the entire conduction system in intact whole hearts. Overall, our studies represent the first step in the deconvolution of the molecular and cellular identity of the cardiac conduction system at single-cell resolution, resulting in the discovery and validation of a host of new conduction-specific genes and an unprecedented profiling of previously elusive conduction cell

subtypes. Our molecular analysis of individual cells in the CCS provides a new foundation for future efforts to understand the functional role of this anatomically complex cellular network and to improve our ability to diagnose and treat diseases of the conduction system *in utero* and during adulthood.

METHODS

All data and materials have been made publicly available at the Geo Repository and can be accessed at https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE132658

Accession numbers

All scRNA-seq raw data have been deposited into the NCBI/GEO database under accession number GEO: GSE132658.

Mice

Wild-type, timed pregnant CD1 mice were acquired from Jackson Laboratory (Sacramento, CA). Embryonic pups or postnatal mice at indicated ages were used in accordance with the Institutional Animal Care and Use Committee of Stanford University. Both female and male mice were used for all experiment types described at an 1:1 ratio.

Tissue isolation and single-cell sequencing using the 10x Genomics© platform

Single cells were isolated and analyzed using the droplet-based platform by 10x Genomics, Inc per recommended company guidelines. Single cells were prepared following the protocol from 10x Genomics, Inc (Pleasanton, CA). Briefly, embryonic day 16.5 (E16.5), wild-type CD1 mouse hearts were harvested and three zones of microdissection were isolated based on anatomical landmarks and entailed: **Zone I** - Sinoatrial node (SAN) region (SVC/right atrial junction), **Zone II** - Atrioventricular node (AVN)/His region (crux of heart) and **Zone III** - Bundle branch (BB)/Purkinje fiber (PF) region (luminal side of ventricles). Specifically, Zone II was dissected as a large area at the crux of the heart from the base of the interatrial septum (including the triangle of Koch) to below the plane of the mitral annulus, from the posterior-most aspect of the heart to the anterior-most. Tissues from a total of six different embryos were pooled for each zone of dissection. Use of E16.5 hearts for analysis was influenced by: 1. size restriction associated with cell capture within microfluidic channels of the Chromium ™ platform from 10x Genomics; and 2. the opportunity to investigate the genetic program required not only for the function but also for the embryonic development of the CCS. The libraries were sequenced using Illumina HiSeq 4000. Additional details provided in Supplemental Methods.

Bioinformatics analysis

The Droplet platform data was de-multiplexed and mapped to mouse genome MM10 using CellRanger from 10x Genomics with default parameters. Cell filter, data normalization, and unsupervised analysis were carried out in Seurat version 2 per their recommended steps (Butler et al., 2018; Macosko et al., 2015). Significance is presented as an “adjusted p-value”, which is based on the Bonferroni correction using all features in the dataset. Briefly, the cells were filtered by their gene number and UMI number. The threshold we used for

gene number is 500 to 60,000, and UMI number is 1,000 to 5 million. Next we used the LogNormalize function to normalize gene expression in each cell. Specifically, we calculated the expression value of genes by following this formula: log {(each gene expression level / total gene expression value)*10,000}. The Y-axis of all ViolinPlots indicates this normalized gene expression. Average log fold change (avg_log FC) described in all data provided (Online Tables II–VII) represents the log fold-change of the average expression between the two groups. To remove the unwanted sources of variations, we scaled the data with the "vars.to.regress" parameter based on the number of UMIs, percentage of ribosome genes, and Rn45s expression value. Furthermore, we found all the variable genes and used them to perform principal component analysis (PCA). Within all the PCs, we used the top 10 PCs to do clustering and tSNE analysis. In the tSNE analysis, we set the seed.use as 10 and perplexity as 30. Finally, we used the FindAllMarker function to identify the genes differentially expressing in the cell clusters. To be detected, the genes have to express in at least 25% of cells in one of the two comparing clusters and the differential expression level also should be higher than 25%. Finally, Gene Ontology (GO) Expression Analysis was performed using the DAVID Bioinformatics Resources 6.8, NIAID/NIH (https://david.ncifcrf.gov/summary.jsp). GO/KEGG term enrichment analyses were performed using geneAnswers R package with hypergeometric test. Statistically significant cardiac-related terms with at least 2 genes were used for each gene list.

Immunofluorescence

Immunofluorescence staining was carried out by following a previous protocol with minor modifications[24]. Specifc details can be found in Supplemental Methods. Primary antibodies used included: Anti-mouse Igfbp5 Goat Polyclonal Antibody (R&D systems/Fisher Scientific: AF578) at 1:100 dilution; Anti-mouse Connexin 40 Rabbit Polyclonal antibody (Alpha Diagnostics, Cx40-A) at 1:100 dilution; Anti-mouse Hcn4 Rat Monoclonal [SHG 1E5] antibody (Abcam, ab32675) at 1:75 dilution. The following secondaries were used at a 1:500 dilution: Donkey anti-goat IgG Alexa Fluor 555 (Invitrogen, A-21432), Chicken anti-Rabbit IgG Alexa Fluor 488 (Invitrogen, A-21441), Donkey anti-Rabbit IgG Alexa Fluor 647 (Invitrogen, A-31573) and Chicken anti-Rat IgG AlexaFluor 488 (Invitrogen A-21470). All images were taken with Axioimager microscope at Neuroscience Microscope Service (NMS) facility at Stanford University. Negative controls for immunostaining included the use of primaries or secondary antibodies alone. A minimum of 4 biological (different hearts) and 4 technical (different slides/heart) replicates were used for each antibody staining.

iDISCO+

For detailed protocol, please see https://idisco.info/idisco-protocol/. Protocol was followed strictly with primary and secondary incubation periods lasting 2 days each. The following primary antibodies were used: Anti-mouse Igfbp5 Goat Polyclonal Antibody (R&D systems/ Fisher Scientific: AF578) at 1:200 dilution (total 1.4 mL/embryo); Anti-mouse Cx40 Rabbit Polyclonal antibody (Alpha Diagnostics, Cx40-A) at 1:200 dilution (total 1.4 mL/embryo); Anti-mouse Hcn4 Rat Monoclonal [SHG 1E5] antibody (Abcam, ab32675) at 1:200 dilution (total 1.4mL/embryo). The following secondaries were used at a 1:700 dilution: Donkey anti-goat IgG Alexa Fluor 555 (Invitrogen, A-21432), Donkey anti-Rabbit IgG Alexa Fluor 647 (Invitrogen, A-31573) and Chicken anti-Rat IgG AlexaFluor 647 (Invitrogen A-21472).

At least one day after clearing, iDISCO+ samples were imaged on a light sheet microscope (Ultramicroscope II, LaVision Biotec) equipped with a sCMOS camera (Andor Neo) and a 2×/0.5 NA objective lens (MVPLAPO 2x) equipped with a 6 mm working distance dipping cap. Version v285 of the Imspector Microscope controller software was used. We imaged using 488-nm, 561-nm, and 640-nm lasers. The samples were scanned with a step-size of 3 μm using the continuous light-sheet scanning method. A minimum of 8 biological (different hearts) replicates were used for each immunolabelling within optically cleared hearts.

RNAscope in situ hybridization

RNAscope® Multiplex Fluorescent v2 (Cat. #323100) was used per manufacturer suggested protocol. The following murine probes were used: Mm-Hcn4-C2 – 421271-C2, Mm-Smoc2-C1 - Cat No. 318541, Mm-Rgs6-C1 - Cat No. 521211, Mm-Rgs6-C1 - Cat No. 521211, Mm-Cpne5-C3 - Cat No. 496711-C3 and Mm-Ntm-C1 - Cat No. 489111. All images were taken with Axioimager microscope at Neuroscience Microscope Service (NMS) facility at Stanford University. A minimum of 3 biological (different hearts) and 4 technical (different slides/heart) replicates were used for each *in situ* hybridization.

RESULTS

Single-cell isolation and expression profiling of cardiac conduction cells

In order to obtain the transcriptional profiles of individual cardiac conduction cells, wild-type, embryonic day 16.5 (E16.5) mouse hearts (n=6 per zone) were harvested and three zones of microdissection were isolated based on anatomical landmarks (Figure 1A). These three zones included the SAN region (Zone I), the AVN/His region (Zone II) and the BB/PF region (Zone III). Our designed workflow then entailed tissue digestion and single-cell capture on a microfluidic panel, automated reverse transcription, barcoding, library generation, and high-throughput sequencing and bioinformatics analysis (Figure 1A). We acquired high quality samples from each zone, collecting 5,919 cells from Zone I, 5,625 cells from Zone II and 10,918 from Zone III (Figure 1B). Unsupervised dimensionality reduction of the scRNA-seq data by t-Distributed Stochastic Neighbor Embedding (t-SNE)[25] from each zone, or all zones combined, demonstrated expected clusters of the major cell types contained within the E16.5 heart including fibroblasts, endocardial, epicardial, endothelial and smooth muscle cells in addition to cardiomyocytes (Figure 1B, Online Figure I, Online Table I). Interestingly, when we analyzed genes that were previously reported to be differentially expressed in the SAN using the most stringent bulk RNA sequencing approaches to date[20] (e.g., transgenic reporter plus laser capture microdissection), we found a significant portion (25%) of these genes to be enriched in other cell types such as endothelial cells and fibroblasts rather than SAN cells (Online Figure II), demonstrating the power and specificity of the single cell transcriptomic approach. Similarly, more recent efforts[21] using cell sorting and knock-in *Tbx3* reporter mice to perform transcriptome analysis of the SAN showed similar issues of contamination (Online Figure III). When compared to our dataset, 12 of the top 16 (75%) reported SAN-enriched genes from E17.5 mice appear to be enriched within non-conduction cell clusters including fibroblasts (Clusters 2, 5, 6, 11; eg. *Tnxb*), endothelial (Cluster 13; eg. *Tgfb1*), endocardial

(Clusters 4 and 7; eg. *Sox18*) and/or neuronal (Cluster 14; eg. *Ngfr*) rather than the *Hcn4+/Shox2+/Tbx3+* SAN cluster (Cluster 9).

Identification of a Bona Fide Sinoatrial Node (SAN) Cell Cluster within Zone I

Zone I cells underwent unsupervised clustering by t-SNE (Figure 2A), and gene enrichment analysis (Figure 2B) revealed six cardiomyocyte clusters (*Tnnt2*$^{+}$/*Actn2*$^{+}$ double positive) (Clusters 0, 1, 3, 9, 12, 15) (Figure 2A–B). To identify the cell cluster(s) containing putative SAN cells, we assessed for the expression of established SAN markers including *Hcn4, Isl1, Shox2,* and *Tbx3*[5,26,27] in addition to a cardiomyocyte signature. Only one (Cluster 9) of all sixteen clusters showed significant enrichment of nodal markers (Figure 2C–D). Further, gene ontology (GO) functional cluster analysis of all enriched genes for Cluster 9 revealed multiple statistically significant GO terms associated with not only cardiac development but also heart rate regulation and SAN development and function (Figure 2E). Finally, in addition to the enrichment of established nodal genes, our analysis revealed a host of significant novel genes not previously involved in SAN development or function (Online Table II), including *Igfbp5* (Insulin growth factor binding protein 5) which showed the greatest amount of enrichment (2.12 avg log FC, adjusted p value = 2.63×10^{-126}). Immunostaining of murine heart sections confirmed expression of Igfbp5 protein within the compact SA node (cSAN) but not in the surrounding atrial working myocardium (Figure 2F) nor additional cardiac cell types such as the SAN artery (Online Figure IV). *Igfbp5* gene expression was also enriched within Clusters 4 (endocardial) and 10 (epicardial) (Figure 2B); however, protein expression could not be detected within these additional cell populations by immunofluorescence (data not shown). Other novel genes found to be significantly enriched within Cluster 9 included *Smoc2* (SPARC-related modular calcium-binding protein 2; 1.69 avg log FC, adjusted p=5.2×10^{-164}), *Ntm* (Neurotrimin; 0.75 avg log FC, adjusted p value = 4.97×10^{-81}), and *Cpne5* (Copine 5; 0.57 avg log FC, adjusted p value = 2.90×10^{-37}) as well as the previously identified *Rgs6* (regulator of G-protein signaling type 6; 0.6 avg log FC; adjusted p value = 2.92×10^{-61}) (Figure 2D).[28] Interestingly, *Smoc2* was also recently reported by the Christoffels group as a new SAN marker.[21] We subsequently validated these genes by either immunostaining or FISH, showing their enrichment within the SAN as compared to the surrounding atrial myocardium (Online Figure V).

Analysis of Cluster 9 in Zone I reveals unique compact SAN and transitional cell subtypes

To further evaluate SAN cell type heterogeneity, Cluster 9 was isolated as an independent Seurat object and unbiased subcluster analysis was then performed (Figure 3A).[29] Two distinct subclusters were identified within Cluster 9, both showing high expression of *Actn2* consistent with cardiomyocyte identities (Figure 3B–C). One cluster (named “cSAN” for compact SAN) showed high levels of established SAN markers including *Hcn4, Hcn1, Gjc1, Isl1, Shox2, Tbx3* and *Tbx18*, with low or no expression of known atrial cardiomyocyte gene markers *Gja5 (Cx40), Scn5a* and *Nkx2–5*, consistent with isolation of compact SAN cells (Figure 3C–D). Conversely, while the second cluster (named “Tz” for transitional cells) expressed these same nodal markers, they were expressed at lower levels. Additionally, the Tz cluster was reciprocally enriched for atrial cardiomyocyte markers consistent with the isolation of a hybrid or transitional SAN cell population (Figure 3D).[2,7,8,30,31] Differential

gene expression was assessed between the cSAN and Tz cell clusters (Online Table III). Notably, *Smoc2* was found to be present in both the cSAN and Tz subclusters (Figure 3E). Consistently, high resolution fluorescent *in situ* hybridization (RNAscope) and co-immunostaining analyses of wild-type murine heart sections of the SAN region further validated these results. Clear subpopulations of cells entailing: 1) cSAN cells (*Hcn4*$^{+}$/*Smoc2*$^{+}$/*Cx40*^{neg}); 2) Tz cells (*Hcn4*$^{low/neg}$/*Smoc2*$^{+}$/*Cx40*^{low}); and 3) atrial cardiomyocytes (*Hcn4*^{neg}/*Smoc2*^{neg}/*Cx40*$^{+}$) (Figure 3F–G). Additional novel SAN markers, including *Igfbp5, Cpne5, Rgs6 and Ntm*, were also found to be expressed in both the cSAN and transitional cell subpopulations (Online Figure VI) and validated by immunostaining or FISH (Online Figure V).

In order to further investigate possible subdomains within the cSAN, cardiomyocytes (CMs) expressing the nodal gene *Hcn4* (*Hcn4*$^{+}$/*Actn2*$^{+}$ double positive) were next isolated *in silico*. Notably, two distinct clusters emerged consistent with previously recognized “Head” and “Tail” subdomains within the cSAN (Online Figure VIIA).[5,8] While both clusters were expectedly enriched for the nodal markers *Hcn4* and *Shox2*, the “Head” cluster showed increased *Tbx18* and decreased *Nkx2–5* expression as per prior reports (Online Figure VIIB).[2,31] Conversely, the “Tail” cluster reciprocally showed downregulation of *Tbx18* and upregulation of *Nkx2–5*.[2,5,31] These subclusters represent highly distinct cell populations consistent with the previously recognized functional subdomains of the SAN. Comparison of these two subpopulations have revealed a host of significant, differentially expressed genes not previously reported (Online Table IV).

Identification of an Atrioventricular Node (AVN)/His Cluster in Zone II

Unbiased clustering of Zone II single cells by t-SNE revealed 14 total clusters of which five (Clusters 0, 3, 4, 5, 7) represented cardiomyocyte clusters by gene expression analysis (Figure 4A). Only one (Cluster 4) showed enrichment of established AVN/His markers including *Hcn4, Kcne1*, *Cacna2d2* and *Cacna1g*[6,26,32–35] consistent with isolation of a legitimate AVN/His cluster (Figure 4C–D). GO functional clustering analysis of all significantly enriched genes in Cluster 4 also revealed several significant GO terms associated with cardiac development and regulation of heart rate in addition to Bundle of His development (Figure 4E). Finally, Cluster 4 showed highly significant enrichment of several novel genes not previously reported in AVN/His cells (Online Table V), including *Cpne5* (0.73 avg log FC, adjusted p value < 2.22×10^{-308}) but not the SAN-enriched gene, *Smoc2* (Figure 4D, 4F). Consistent with this scRNA-seq data, immunostaining within murine heart sections confirmed robust and specific signal within the AVN for *Cpne5* but not *Smoc2* (Figure 4G). Of note, while our scRNA-seq data showed that *Smoc2* was mildly enriched within Clusters 2, 8, 11 (Fibroblasts) and Cluster 13 (Epicardial) within Zone II (Figure 4F), *Smoc2* could not be detected by immunofluorescence within these cell types (data not shown). Single-cell FISH analysis of *Cpne5* expression additionally confirmed robust and specific signal within the His bundle but not the surrounding ventricular myocardium (Figure 4H). Additionally, most of the other novel genes found to be enriched within the SAN cluster were also present within the AVN/His cluster including *Igfbp5* (1.19 avg log FC, p = 1.49×10^{-120}), *Rgs6* (0.39 avg log FC; p = 9.75×10^{-153}) and *Ntm* (0.55 avg log FC,

$p = 1.20\times10^{-101}$). Consistently, expression of *Igfbp5, Rgs6* and *Ntm* were all detected within the murine AV node (Online Figure VIII).

Analysis of Cluster 4 in Zone II Unveiled Distinct AVN, His Bundle and Transitional Cell Subtypes

Prior studies have described at least four distinct cell subtypes within the AVN/His region through detailed immunofluorescence and electrophysiologic analyses, including the compact AVN (cAVN), nodal AV Ring (NAVR) and His bundle[32,36] as well as transitional cell types such as the transitional AV ring (TAVR).[6,33,37] In order to evaluate for transcriptional heterogeneity within the AVN/His region and their relationship to these previously reported cell subtypes, we focused on Cluster 4 from Zone II and performed a subcluster analysis (Figure 5A–B). Six distinct subclusters were identified and these show consistency with previously reported gene profiles[6] including: 1) cAVN ($Hcn4^{hi}/Tbx3^{hi}/Kcne1^{hi}/Gjc1^{+}/Gja1^{neg}/Gja5^{neg}/Scn5a^{neg}/Cacna2d2^{+}$); 2) NAVR ($Tbx3^{+}/Kcne1^{+}/Gjc1^{+}/Gja1^{neg}/Gja5^{neg}/Scn5a^{neg}/Cacna2d2^{+}$); 3) His bundle ($Hcn4^{+}/Tbx3^{+}/Kcne^{hi}/Gjc1^{+}/Gja1^{neg}/Gja5^{+}/Etv1^{+}/Scn5a^{+}/Cacna2d2^{+}$); and 4) TAVR ($Hcn4^{neg}/Tbx3^{low/neg}/Kcne1^{+}/Gjc1^{+}/Gja1^{+}/Gja5^{neg}/Scn5a^{+}/Cacna2d2^{low}$). Two additional, distinct transitional cell types were isolated which we have termed: 5) Atrial Transitional Zone (ATZ) ($Myh6^{hi}/Tbx3^{low/neg}/Kcne1^{low}/Gjc1^{+}/Gja5^{hi}/Cacna2d2^{+}$); and 6) Ventricular Transitional Zone (VTZ) ($Myl2^{hi}/Myh7^{hi}/Tbx3^{low/neg}/Kcne1^{hi}/Gjc1^{+}/Gja1^{+}/Gja5^{neg}/Cacna2d2^{low}$) based on their atrial/nodal and ventricular/nodal hybrid expression signatures, respectively (Figure 5C–E). Differential gene expression between each cell cluster revealed a number of genes that are significantly different in their expression level, further refining their unique signatures (Online Table VI). When assessing *Cpne5* gene expression within the AVN cell subtypes, it was found to be enriched in all of the *Hcn4* positive cell populations, including the cAVN and His bundle cells (Online Figure IX). Consistently, *in situ* hybridization revealed perfect co-localization of *Cpne5* mRNA expression with *Hcn4* expression (Figure 4H).

Identification of an immature Purkinje Fiber (PF) Cell Cluster in Zone III

Unsupervised clustering of Zone III cells by t-SNE revealed 15 distinct clusters, nine of which (Clusters 0, 1, 2, 3, 6, 7, 8, 10, 13) were ventricular cardiomyocytes based on their gene signatures (Figure 6A–B). Of all clusters, only one (Cluster 13) demonstrated significant enrichment of the ventricular conduction system genes *Gja5 (Cx40), Scn5a, Irx3* and *Cacna2d2* among others (Figure 6C).[17,26,38–40] Given the isolation of putative Purkinje fiber cells within Cluster 13 of Zone III occurs at embryonic day 16.5 these cells are considered to be “immature” Purkinje fiber cells. The concordant expression of these known PF genes, however, strongly supports the identification of Cluster 13 as a bona fide PF cluster. Further, GO term analysis of all significantly enriched genes within Cluster 13 revealed functional categories such as heart development, regulation of heart rate and cardiac conduction (Figure 6D). Enrichment of several novel genes were detected within this PF cluster (Online Table VII) including *Igfbp5* (0.96 avg log FC, adjusted p value = 1.11×10^{-100}), *Cpne5* (0.68 avg log FC, adjusted p value < 2.22×10^{-308}) and *Ntm* (0.61 avg log FC, adjusted p value = 6.32×10^{-110}) (Figure 6E, 6F). FISH and immunostaining of murine heart sections confirmed highly specific expression of these genes within the ventricular conduction system including the bundle branches and Purkinje fiber network as

compared to the surrounding ventricular working myocardium (Figure 6G **and** Online Figure X).

Analysis of Cluster 13 in Zone III identified distinct PF and transitional PF Cell subtypes

To further evaluate subpopulations of cells within Cluster 13, subcluster analysis was performed. Two distinct cell clusters were detected, representing standard Purkinje fibers (PF) as well as transitional PF cells (Tz), both known to exist at the Purkinje-myocyte junction (Figure 7A–B).[9,10,41,42] Consistent with these prior reports, the standard PFs were enriched for *Gja5 (Cx40), Etv1, Cacna2d2* and *Sema3a* while the transitional PF cells expressed these distal conduction genes at lower levels as well as the conduction marker *Gjc1*. Conversely, Tz PF cells were enriched for the ventricular myocardial gap junction gene *Gja1 (Cx43)* (Figure 7D–E). Several novel genes were noted to be significantly enriched within the Tz PF cells that were largely found to be expressed within the ventricular myocardium as well, consistent with a transitional or hybrid phenotype (Online Table VIII).

Overall, Cluster 13 showed significant upregulation of many genes not previously associated with ventricular conduction cells including *Ntm*.[43] Interestingly, upon subcluster analysis, *Ntm* was found to be expressed in both standard and transitional PF cell types, with a trend towards enrichment within the standard PF subcluster (0.53 avg log FC, adjusted p = 1.22×10^{-01}) (Figure 7F). Consistently, FISH of wild-type murine heart sections of the ventricular conduction region reinforced these *in silico* results, showing cell subpopulations including 1) standard immature Purkinje fiber cells ($Hcn4^{+}/Ntm^{hi}$); 2) transitional PF cells ($Hcn4^{neg}/Ntm^{low}$); and 3) ventricular cardiomyocytes ($Hcn4^{neg}/Ntm^{neg}$) (Figure 7G). *Cpne5* expression was similarly detected within transitional PF cells in addition to standard PF cells (Online Figure X).

Optical clearing and 3D volumetric analyses delineate the architecture of the entire CCS and SA Nodal substructure within intact murine hearts

Comparative gene analysis of all conduction cardiomyocytes throughout the CCS (SAN = Zone I Cluster 9; AVN/His = Zone II Cluster 4; BB/PF = Zone III Cluster 13) as compared to all other cell types revealed multiple enriched genes including *Igfbp5, Cpne5, Rgs6* and *Ntm* (Online Tables II, V and VII). For example, the expression of *Igfbp5* was dramatically increased in all conduction clusters including the SAN (Cluster 9, Zone I = 2.12 avg log FC, adjusted p value = 2.63×10^{-126}), AVN/His (Cluster 4, Zone II = 1.19 avg log FC, adjusted p value = 1.49×10^{-120}) and BB/PF (Cluster 13, Zone III = 0.96 avg log FC, adjusted p value = 1.11×10^{-100}) (Online Figure XIA). Given the challenges of visualizing the complex three-dimensional anatomy of the CCS, whole-mount immunostaining and 3D imaging was undertaken using tissue clearing (iDISCO+)[44] and light sheet microscopy with volume rendering on intact wild-type mouse hearts (Figure 8A). Consistent with the enrichment of *Igfbp5* expression within the conduction system by scRNA-seq, anti-Igfbp5 protein signal was detected within the entire CCS, including SAN, AVN, His, bundle branches and Purkinje fiber network (Figure 8B). *Igfbp5* was also detected within connecting tracts between these major CCS components including the internodal tracts and right and left atrioventricular ring bundles (Figure 8B, Online Figure XI, Online Movie I). Similarly,

Cpne5, another novel marker, was shown to mark the entirety of the CCS in 3D (Online Movie II).

In order to further evaluate the transitional and compact SAN subpopulations uncovered in our transcriptome analyses, iDISCO+ cleared wild-type hearts were co-immunolabeled for Hcn4 and Rgs6, one of the novel CCS markers we uncovered in this study (Figure 8C–D). At least 2 distinct transitional sinoatrial conduction pathways (SACPs) ($Hcn4^{neg}/Rgs6^{+}$) exiting the SA node were visualized including one from SAN body directed rightward towards the crista terminalis and another from the SAN head directed inferiorly and leftward (Figure 8C–D, Online Movie III). Additionally, a third $Hcn4^{+}/Rgs6^{+}$ pathway emerged from the tail to give rise to the internodal tracts (Figure 8D). These transitional SACPs were faithful across multiple timepoints examined (E16.5 through P12) and, consistent with the aforementioned 2D data (Figure 3, Online Figure V), also labelled by other novel markers including Smoc2, Igfbp5 and Cpne5 protein (Online Figure XII).

DISCUSSION

While the cardiac conduction system remains a crucial element to heart development and function, progress in our understanding of its intricate cellular and molecular landscape remains incomplete.[5] Significant limitations within the field have included: 1) low conduction cell numbers within the heart; 2) intra- and intercomponent cell type heterogeneity; 3) challenges in conduction cell isolation; and 4) complexity of the three-dimensional anatomy of the CCS. As a result, our molecular understanding of the conduction system of the heart has largely originated from painstaking evaluation of the overlapping expression of a handful of known markers within each component, including ion channels (*Hcn4, Cacna2d2*), transcription factors (*Isl1, Tbx3, Tbx5, Tbx18*) and gap junction genes (*Gjc1/Cx45, Gja5/Cx40*), with only a limited number of unique and validated conduction-specific markers (*Hcn4, Contactin 2*).[18,45] Further, bulk RNA sequencing approaches are plagued by non-CCS cell contamination (Online Figure II) even when coupled with cell-sorting strategies using historical SAN reporter genes such as *Tbx3* (Online Figure III). Here, we have employed microdissection coupled with single-cell isolation and RNA sequencing in order to overcome the aforementioned technical hurdles. Specifically, we have successfully isolated cell types representing the entirety of the cardiac conduction system including rare CCS subtypes for deep sampling in order to assess differential genome-wide expression. Our analyses have uncovered a host of novel conduction markers as well as unique molecular signatures of the various CCS cell subtypes not previously attainable to provide the foundation for a molecular blueprint of the conduction system (Figure 9). Further, high-resolution *in situ* hybridization (RNAscope) and volume-rendering whole-mount immunofluorescence with iDISCO+ tissue clearing have provided platforms for gene expression validation both at single-cell resolution and in complex three-dimensional space within the intact murine heart.

The SAN, the primary pacemaker of the heart, has long been recognized as a multicomponent structure containing functional subdomains, including “Head” and “Tail” regions.[5] The Head represents the leading pacemaker region; however, this can shift (“wandering pacemaker”) to the tail in response to various stimuli both physiologic and

extrinsic.[30,46] These distinct but partially redundant subdomains have even been suggested to provide important “fail-safes” to prevent from rhythm failure[47] further arguing for additional insight into these poorly defined structures.[48] Additionally, in between the compact SAN and surrounding working atrial myocardium, the existence of transitional cells have also been recognized by electron microscopy and patch clamp experiments.[8,49] They are suspected to facilitate the spread of depolarization and may play a crucial clinical role (e.g., sinus node exit block); however, the molecular and cellular identity of these cell types remains virtually unknown.[5] By combining microdissection with scRNA-seq, we have demonstrated the isolation of not only a bona fide SAN cell cluster (Cluster 9 from Zone I) (Figure 2) but unique SAN cell subtypes including Head and Tail cells within the compact node (Online Figure V) and transitional SAN cells (Figure 3), thereby allowing differential gene expression analysis at unprecedented resolution (Online Tables III and IV). As a result, a host of novel SAN genes were uncovered including *Igfbp5, Cpne5, Rgs6, Ntm* and *Smoc2*. Interestingly, *Smoc2*, a gene previously associated with endothelial cell proliferation and migration and angiogenesis in non-cardiac tissues[50], was enriched in the compact SAN, transitional SAN and internodal cell populations but excluded from the rest of the CCS. Finally, enrichment of *Rgs6* within the SAN is consistent with prior reports demonstrating *Rgs6* as a necessary modulator of parasympathetic innervation in the heart[28] and implicated in resting heart rate variability in humans by GWAS.[51] Homozygous *Rgs6* knockout mice demonstrated exaggerated bradycardia in response to carbachol and significantly enhanced the effect of carbachol on inhibition of spontaneous SAN cell action potential firing.[28]

Interestingly, all of these novel SAN markers were also present within the transitional SAN cell population, providing some of the first unique markers for this previously elusive cell type. Further, high-resolution, 3D studies using iDISCO+ have provided insight into SA nodal architecture, detailing the exact locations of these transitional cells ($Hcn4^{low/neg}$/$Igfbp5^{+}$/$Cpne5^{+}$/$Rgs6^{+}$/$Smoc2^{+}$) in at least 2 clear transitional SAN conduction pathways (SACPs) (Figure 8, Online Figure XII). These findings are remarkably consistent with prior activation mapping research of the SAN and surrounding atria in canines and humans showing at least 2 SACPs exiting the SAN in similar locations.[8,14,52,53] These electrophysiological studies demonstrated that SACPs not only allow for direct activation of the surrounding atrial tissue but also play an important role in protecting the SAN against overdrive activation during atrial arrhythmias.[53] Further studies are now needed to understand a possible role for these other novel genes uncovered within the conduction system, in particular the transitional SAN cell population.

The AVN remains critical to the CCS in transmitting and delaying the electrical impulses generated by the SAN, thereby allowing the ventricles to be filled before their contraction is initiated.[33] Unfortunately, due to its complex anatomy, buried within the crux of the heart, and significant cell type heterogeneity, the AVN has remained a challenge to define molecularly. Our study has not only demonstrated successful isolation of AVN cells but also uncovered a host of novel AVN-specific genes and represents the first genome-wide analysis of the AVN region at single-cell resolution (Figure 4A–D). Notably, one of the most significantly enriched genes within the original AVN cluster (Cluster 4, Zone II) was *Cpne5* (Figure 4D), a member of the copine calcium-dependent, phospholipid-binding family of proteins thought to possibly play a role in membrane trafficking.[54] While its true function

remains unknown, *Cpne5*, along with its paralog *Cpne8*, have been associated with heart rate variability in humans by GWAS.[55] Unsupervised analysis of the AVN cluster additionally unveiled six subclusters representing previously suspected AVN subtypes.[6,33,37] With our ability to identify these AVN subtypes we are well-positioned to resolve the unique molecular signature of each cell type (Figure 5, Online Table VI) and begin to understand the factors intrinsic in establishing their function and, possibly, development.

The distal portion of the conduction system, including the bundle branches and Purkinje fibers, allow for the rapid transmission of electrical impulses throughout the ventricular myocardium thereby allowing for their coordinated contraction.[4] Within Zone III ventricular cells, a legitimate immature Purkinje fiber cluster was detected (Cluster 13) (Figure 6A–E). Differential gene expression analysis of this cluster revealed a host of novel genes including *Igfbp5, Cpne5* and *Ntm*. Neurotrimin (*Ntm*), a member of the IgLON immunoglobulin domain-containing cell adhesion molecules[43] was found to be significantly enriched in ventricular conduction cells by scRNA-seq data (Figure 7F) and validated by FISH (Figure 7G). *Ntm*, not previously associated with the CCS, was found to be expressed within all components of the CCS (Figure 7, Online Figures V and VIII).

To date, while a unified description of Purkinje fiber subtypes is lacking within the literature, several groups have concretely showed at least two distinct cell types in the distal portions of the CCS by electron microscopy and functional patch clamp analysis: 1) standard PF cells[56] and 2) transitional PF cells - broader cells that serve as an intermediate layer between the PF and the working cardiomyocytes.[9,41] Poised at the Purkinje-myocyte junction, these transitional PF cells are hypothesized to play a role in facilitating the spread of depolarization, providing a high resistance barrier, shielding the Purkinje system from electrotonic loading, as well as amplifying the current before passing it on to the surrounding myocardium.[9] Clinically, Purkinje fiber and transitional PF cells are believed to be particularly vulnerable for the development of serious and even life-threatening tachyarrhythmias including bundle branch or fascicular re-entry and polymorphic ventricular tachycardia/fibrillation (VT/VF).[11] Upon subcluster analysis of Cluster 13, two distinct populations were isolated that suggest their identities as standard PFs and transitional PFs (Figure 7A–E), identified by at least two novel markers including *Ntm* (Figure 7G) and *Cpne5* (Online Figure X). To our knowledge, these finding represent the first transcriptomic characterization of these cell subtypes and provides a unique opportunity for uncovering new molecular markers and key regulators for standard and transitional PFs.

Finally, in addition to providing insight into the rare conduction cell subtypes of each CCS component, our analyses have also allowed for the systematic discovery of a host of genes enriched throughout the entire conduction system including *Igfbp5, Cpne5* and *Ntm*. Our dataset has also allowed for the discovery of conduction genes enriched within specific CCS components such as *Smoc2* that was found to be specifically enriched within the SAN and internodal tracts but notably absent from the AVN. Further studies are currently underway to validate additional conduction-specific genes as well as employing CRISPR-Cas9 technology in order to generate systemic knockdowns of top candidates for the assessment of their possible functional roles in CCS development and/or function. Altogether, our study represents the first comprehensive assessment of transcriptional profiles from the entire CCS

at single-cell resolution (Figure 9) and provides bioinformatics tools to facilitate future efforts in conduction cell identification, isolation, characterization in the context of development and disease and to improve our ability to diagnose and treat diseases of the conduction system *in utero* and during adulthood.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

The authors thank John Coller and Dhananjay Wagh at Stanford Functional Genomics Facility (SFGF) for their help with the 10X experiments and data alignment. The authors would additionally like to thank Andrew Olson and the Stanford Neuroscience Microscopy Service, supported by NIH grants NS069375 and 1S10OD025091, for image acquirement.

SOURCES OF FUNDING

This work was supported by the NIH Office of Director's Pioneer Award LM012179-03, the American Heart Association Established Investigator Award 17EIA33410923, the Department of Pediatrics and Division of Pediatric Cardiology at Lucille Packard Children's Hospital, the Stanford Cardiovascular Institute, the Stanford Division of Cardiovascular Medicine, Department of Medicine, the Institute for Stem Cell Biology and Regenerative Medicine, and an endowed faculty scholar award from the Stanford Child Health Research Institute/ Lucile Packard Foundation for Children (S.M.W). This work was also supported by NIH R01 HL141371, NIH R01 HL145676, NIH R01 130020 (J.C.W.), Multi-Disciplinary Training Program in Cardiovascular Imaging (EB009035) (D.T.P.), Stanford BioX Bowes Fellowship (E.L.A.) and the Training Grant entitled Research Training in Myocardial Biology at Stanford (NIH 2 T32 HL094274) (W.R.G.).

Nonstandard Abbreviations and Acronyms

scRNA-seq	Single-cell RNA-sequencing
CCS	Cardiac conduction System
SAN	Sinoatrial node
AVN	Atrioventricular node
His	His bundle
BB	Bundle branches
PF	Purkinje fibers

REFERENCES

1. Park DS, Fishman GI. The Cardiac Conduction System. Circulation. 2011;123:904–915. [PubMed: 21357845]
2. Christoffels VM, Smits GJ, Kispert A, Moorman AFM. Development of the pacemaker tissues of the heart. Circ Res. 2010;106:240–254. [PubMed: 20133910]
3. Vedantham V New Approaches to Biological Pacemakers: Links to Sinoatrial Node Development. Trends Mol Med. 2015;21:749–761. [PubMed: 26611337]
4. van Eif VWW, Devalla HD, Boink GJJ, Christoffels VM. Transcriptional regulation of the cardiac conduction system. Nature Reviews Cardiology. 2018;15:617–630. [PubMed: 29875439]
5. Liang X, Evans SM, Sun Y. Development of the cardiac pacemaker. Cellular and Molecular Life Sciences. 2017;74:1247–1259. [PubMed: 27770149]

6. Aanhaanen WTJ, Mommersteeg MTM, Norden J, Wakker V, de Gier-de Vries C, Anderson RH, Kispert A, Moorman AFM, Christoffels VM. Developmental Origin, Growth, and Three-Dimensional Architecture of the Atrioventricular Conduction Axis of the Mouse Heart. Circulation Research. 2010;107:728–736. [PubMed: 20671237]
7. Chandler NJ, Greener ID, Tellez JO, Inada S, Musa H, Molenaar P, DiFrancesco D, Baruscotti M, Longhi R, Anderson RH, Billeter R, Sharma V, Sigg DC, Boyett MR, Dobrzynski H. Molecular Architecture of the Human Sinus Node: Insights Into the Function of the Cardiac Pacemaker. Circulation. 2009;119:1562–1575. [PubMed: 19289639]
8. Csepe TA, Zhao J, Hansen BJ, Li N, Sul LV, Lim P, Wang Y, Simonetti OP, Kilic A, Mohler PJ, Janssen PML, Fedorov VV. Human sinoatrial node structure: 3D microanatomy of sinoatrial conduction pathways. Progress in Biophysics and Molecular Biology. 2016;120:164–178. [PubMed: 26743207]
9. Tranum-Jensen J, Wilde AA, Vermeulen JT, Janse MJ. Morphology of electrophysiologically identified junctions between Purkinje fibers and ventricular muscle in rabbit and pig hearts. Circulation Research. 1991;69:429–437. [PubMed: 1860183]
10. Vigmond EJ, Stuyvers BD. Modeling our understanding of the His-Purkinje system. Prog Biophys Mol Biol. 2016;120:179–188. [PubMed: 26740015]
11. Haissaguerre M, Vigmond E, Stuyvers B, Hocini M, Bernus O. Ventricular arrhythmias and the His-Purkinje system. Nat Rev Cardiol. 2016;13:155–166. [PubMed: 26727298]
12. Kim EE, Shekhar A, Lu J, Lin X, Liu F-Y, Zhang J, Delmar M, Fishman GI. PCP4 regulates Purkinje cell excitability and cardiac rhythmicity. J Clin Invest. 2014;124:5027–5036. [PubMed: 25295538]
13. Marionneau C, Couette B, Liu J, Li H, Mangoni ME, Nargeot J, Lei M, Escande D, Demolombe S. Specific pattern of ionic channel gene expression associated with pacemaker activity in the mouse heart: Ion channel expression in the murine heart. The Journal of Physiology. 2005;562:223–234. [PubMed: 15498808]
14. Tellez JO, Dobrzynski H, Greener ID, Graham GM, Laing E, Honjo H, Hubbard SJ, Boyett MR, Billeter R. Differential Expression of Ion Channel Transcripts in Atrial Muscle and Sinoatrial Node in Rabbit. Circulation Research. 2006;99:1384–1393. [PubMed: 17082478]
15. Gaborit N, Le Bouter S, Szuts V, Varro A, Escande D, Nattel S, Demolombe S. Regional and tissue specific transcript signatures of ion channel genes in the non-diseased human heart: Regional ion channel subunit gene expression in the human heart. The Journal of Physiology. 2007;582:675–693. [PubMed: 17478540]
16. Horsthuis T, Buermans HPJ, Brons JF, Verkerk AO, Bakker ML, Wakker V, Clout DEW, Moorman AFM, 't Hoen PAC, Christoffels VM. Gene Expression Profiling of the Forming Atrioventricular Node Using a Novel *Tbx3* -Based Node-Specific Transgenic Reporter. Circulation Research. 2009;105:61–69. [PubMed: 19498200]
17. Kim K-H, Rosen A, Hussein SMI, Puviindran V, Korogyi AS, Chiarello C, Nagy A, Hui C-C, Backx PH. Irx3 is required for postnatal maturation of the mouse ventricular conduction system. Sci Rep. 2016;6:19197. [PubMed: 26786475]
18. Pallante BA, Giovannone S, Fang-Yu L, Zhang J, Liu N, Kang G, Dun W, Boyden PA, Fishman GI. Contactin-2 expression in the cardiac Purkinje fiber network. Circ Arrhythm Electrophysiol. 2010;3:186–194. [PubMed: 20110552]
19. Shekhar A, Lin X, Lin B, Liu F-Y, Zhang J, Khodadadi-Jamayran A, Tsirigos A, Bu L, Fishman GI, Park DS. ETV1 activates a rapid conduction transcriptional program in rodent and human cardiomyocytes. Sci Rep. 2018;8:9944. [PubMed: 29967479]
20. Vedantham V, Galang G, Evangelista M, Deo RC, Srivastava D. RNA Sequencing of Mouse Sinoatrial Node Reveals an Upstream Regulatory Role for Islet-1 in Cardiac Pacemaker Cells. Circulation Research. 2015;116:797–803. [PubMed: 25623957]
21. van Eif VWW, Stefanovic S, van Duijvenboden K, Bakker M, Wakker V, de Gier-de Vries C, Zaffran S, Verkerk AO, Boukens BJ, Christoffels VM. Transcriptome analysis of mouse and human sinoatrial node cells reveals a conserved genetic program. Development. 2019;146.
22. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics. 2008;9:559. [PubMed: 19114008]

23. Li G, Plonowska K, Kuppusamy R, Sturzu A, Wu SM. Identification of cardiovascular lineage descendants at single-cell resolution. Development. 2015;142:846–857. [PubMed: 25633351]
24. Goodyer WR, Gu X, Liu Y, Bottino R, Crabtree GR, Kim SK. Neonatal β cell development in mice and humans is regulated by calcineurin/NFAT. Dev Cell. 2012;23:21–34. [PubMed: 22814600]
25. van der Maaten L, Hinton G. Visualizing Data using t-SNE. Journal of Machine Learning Research. 2008;9:2579–2605.
26. Bartos DC, Grandi E, Ripplinger CM. Ion Channels in the Heart. Compr Physiol. 2015;5:1423–1464. [PubMed: 26140724]
27. Choudhury M, Institute of Cardiovascular Sciences, University of Manchester, Manchester, UK, Boyett MR, Institute of Cardiovascular Sciences, University of Manchester, Manchester, UK, Morris GM, Institute of Cardiovascular Sciences, University of Manchester, 46 Grafton Street, Manchester, M13 9NT, UK. E: Gwilym.Morris@manchester.ac.uk. Biology of the Sinus Node and its Disease. Arrhythmia & Electrophysiology Review. 2015;4:28. [PubMed: 26835096]
28. Yang J, Huang J, Maity B, Gao Z, Lorca RA, Gudmundsson H, Li J, Stewart A, Swaminathan PD, Ibeawuchi S-R, Shepherd A, Chen C-K, Kutschke W, Mohler PJ, Mohapatra DP, Anderson ME, Fisher RA. RGS6, a modulator of parasympathetic activation in heart. Circ Res. 2010;107:1345–1349. [PubMed: 20864673]
29. Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nat Biotechnol. 2018;36:411–420. [PubMed: 29608179]
30. Monfredi O, Dobrzynski H, Mondal T, Boyett MR, Morris GM. The anatomy and physiology of the sinoatrial node--a contemporary review. Pacing Clin Electrophysiol. 2010;33:1392–1406. [PubMed: 20946278]
31. Wiese C, Grieskamp T, Airik R, Mommersteeg MTM, Gardiwal A, de Gier-de Vries C, Schuster-Gossler K, Moorman AFM, Kispert A, Christoffels VM. Formation of the Sinus Node Head and Differentiation of Sinus Node Myocardium Are Independently Regulated by Tbx18 and Tbx3. Circulation Research. 2009;104:388–397. [PubMed: 19096026]
32. Aanhaanen WTJ, Brons JF, Domínguez JN, Rana MS, Norden J, Airik R, Wakker V, de Gier-de Vries C, Brown NA, Kispert A, Moorman AFM, Christoffels VM. The *Tbx2*$^{+}$ Primary Myocardium of the Atrioventricular Canal Forms the Atrioventricular Node and the Base of the Left Ventricle. Circulation Research. 2009;104:1267–1274. [PubMed: 19423846]
33. Bakker ML, Moorman AFM, Christoffels VM. The Atrioventricular Node: Origin, Development, and Genetic Program. Trends in Cardiovascular Medicine. 2010;20:164–171. [PubMed: 21742272]
34. Greener ID, Monfredi O, Inada S, Chandler NJ, Tellez JO, Atkinson A, Taube M-A, Billeter R, Anderson RH, Efimov IR. Molecular architecture of the human specialised atrioventricular conduction axis. Journal of Molecular and Cellular Cardiology. 2011;50:642–651. [PubMed: 21256850]
35. Moskowitz IPG, Kim JB, Moore ML, Wolf CM, Peterson MA, Shendure J, Nobrega MA, Yokota Y, Berul C, Izumo S, Seidman JG, Seidman CE. A molecular pathway including Id2, Tbx5, and Nkx2–5 required for cardiac conduction system development. Cell. 2007;129:1365–1376. [PubMed: 17604724]
36. Bakker ML, Boukens BJ, Mommersteeg MTM, Brons JF, Wakker V, Moorman AFM, Christoffels VM. Transcription Factor Tbx3 Is Required for the Specification of the Atrioventricular Conduction System. Circulation Research. 2008;102:1340–1349. [PubMed: 18467625]
37. Li J, Greener ID, Inada S, Nikolski VP, Yamamoto M, Hancox JC, Zhang H, Billeter R, Efimov IR, Dobrzynski H, Boyett MR. Computer three-dimensional reconstruction of the atrioventricular node. Circ Res. 2008;102:975–985. [PubMed: 18309098]
38. Goodyer WR, Wu SM. Fates Aligned: Origins and Mechanisms of Ventricular Conduction System and Ventricular Wall Development. Pediatr Cardiol. 2018;39:1090–1098. [PubMed: 29594502]
39. Li Y, Tian X, Zhao H, He L, Zhang S, Huang X, Zhang H, Miquerol L, Zhou B. Genetic targeting of Purkinje fibres by Sema3a-CreERT2. Scientific Reports [Internet]. 2018 [cited 2018 10 23];8 Available from: http://www.nature.com/articles/s41598-018-20829-9 [PubMed: 29311689]

40. Shekhar A, Lin X, Liu F-Y, Zhang J, Mo H, Bastarache L, Denny JC, Cox NJ, Delmar M, Roden DM, Fishman GI, Park DS. Transcription factor ETV1 is essential for rapid conduction in the heart. J Clin Invest. 2016;126:4444–4459. [PubMed: 27775552]
41. Garcia-Bustos V, Sebastian R, Izquierdo M, Molina P, Chorro FJ, Ruiz-Sauri A. A quantitative structural and morphometric analysis of the Purkinje network and the Purkinje-myocardial junctions in pig hearts. Journal of Anatomy. 2017;230:664–678. [PubMed: 28256093]
42. Severs NJ, Bruce AF, Dupont E, Rothery S. Remodelling of gap junctions and connexin expression in diseased myocardium. Cardiovasc Res. 2008;80:9–19. [PubMed: 18519446]
43. Struyk AF, Canoll PD, Wolfgang MJ, Rosen CL, D'Eustachio P, Salzer JL. Cloning of neurotrimin defines a new subfamily of differentially expressed neural cell adhesion molecules. J Neurosci. 1995;15:2141–2156. [PubMed: 7891157]
44. Renier N, Adams EL, Kirst C, Wu Z, Azevedo R, Kohl J, Autry AE, Kadiri L, Umadevi Venkataraju K, Zhou Y, Wang VX, Tang CY, Olsen O, Dulac C, Osten P, Tessier-Lavigne M. Mapping of Brain Activity by Automated Volume Analysis of Immediate Early Genes. Cell. 2016;165:1789–1802. [PubMed: 27238021]
45. Liang X, Wang G, Lin L, Lowe J, Zhang Q, Bu L, Chen Y, Chen J, Sun Y, Evans SM. HCN4 dynamically marks the first heart field and conduction system precursors. Circ Res. 2013;113:399–407. [PubMed: 23743334]
46. Verheijck EE, van Kempen MJ, Veereschild M, Lurvink J, Jongsma HJ, Bouman LN. Electrophysiological features of the mouse sinoatrial node in relation to connexin distribution. Cardiovasc Res. 2001;52:40–50. [PubMed: 11557232]
47. Li N, Hansen BJ, Csepe TA, Zhao J, Ignozzi AJ, Sul LV, Zakharkin SO, Kalyanasundaram A, Davis JP, Biesiadecki BJ, Kilic A, Janssen PML, Mohler PJ, Weiss R, Hummel JD, Fedorov VV. Redundant and diverse intranodal pacemakers and conduction pathways protect the human sinoatrial node from failure. Sci Transl Med. 2017;9.
48. Froese A, Breher SS, Waldeyer C, Schindler RFR, Nikolaev VO, Rinné S, Wischmeyer E, Schlueter J, Becher J, Simrick S, Vauti F, Kuhtz J, Meister P, Kreissl S, Torlopp A, Liebig SK, Laakmann S, Müller TD, Neumann J, Stieber J, Ludwig A, Maier SK, Decher N, Arnold H-H, Kirchhof P, Fabritz L, Brand T. Popeye domain containing proteins are essential for stress-mediated modulation of cardiac pacemaking in mice. J Clin Invest. 2012;122:1119–1130. [PubMed: 22354168]
49. Boyett MR, Honjo H, Kodama I. The sinoatrial node, a heterogeneous pacemaker structure. Cardiovasc Res. 2000;47:658–687. [PubMed: 10974216]
50. Rocnik EF, Liu P, Sato K, Walsh K, Vaziri C. The novel SPARC family member SMOC-2 potentiates angiogenic growth factor activity. J Biol Chem. 2006;281:22855–22864. [PubMed: 16774925]
51. Nolte IM, Munoz ML, Tragante V, Amare AT, Jansen R, Vaez A, von der Heyde B, Avery CL, Bis JC, Dierckx B, van Dongen J, Gogarten SM, Goyette P, Hernesniemi J, Huikari V, Hwang S-J, Jaju D, Kerr KF, Kluttig A, Krijthe BP, Kumar J, van der Laan SW, Lyytikäinen L-P, Maihofer AX, Minassian A, van der Most PJ, Müller-Nurasyid M, Nivard M, Salvi E, Stewart JD, Thayer JF, Verweij N, Wong A, Zabaneh D, Zafarmand MH, Abdellaoui A, Albarwani S, Albert C, Alonso A, Ashar F, Auvinen J, Axelsson T, Baker DG, de Bakker PIW, Barcella M, Bayoumi R, Bieringa RJ, Boomsma D, Boucher G, Britton AR, Christophersen I, Dietrich A, Ehret GB, Ellinor PT, Eskola M, Felix JF, Floras JS, Franco OH, Friberg P, Gademan MGJ, Geyer MA, Giedraitis V, Hartman CA, Hemerich D, Hofman A, Hottenga J-J, Huikuri H, Hutri-Kähönen N, Jouven X, Junttila J, Juonala M, Kiviniemi AM, Kors JA, Kumari M, Kuznetsova T, Laurie CC, Lefrandt JD, Li Y, Li Y, Liao D, Limacher MC, Lin HJ, Lindgren CM, Lubitz SA, Mahajan A, McKnight B, zu Schwabedissen HM, Milaneschi Y, Mononen N, Morris AP, Nalls MA, Navis G, Neijts M, Nikus K, North KE, O'Connor DT, Ormel J, Perz S, et al. Genetic loci associated with heart rate variability and their effects on cardiac disease risk. Nature Communications. 2017;8:15805.
52. Fedorov VV, Schuessler RB, Hemphill M, Ambrosi CM, Chang R, Voloshina AS, Brown K, Hucker WJ, Efimov IR. Structural and functional evidence for discrete exit pathways that connect the canine sinoatrial node and atria. Circ Res. 2009;104:915–923. [PubMed: 19246679]

53. Fedorov VV, Glukhov AV, Chang R. Conduction barriers and pathways of the sinoatrial pacemaker complex: their role in normal rhythm and atrial arrhythmias. Am J Physiol Heart Circ Physiol. 2012;302:H1773–1783. [PubMed: 22268110]

54. Creutz CE, Tomsig JL, Snyder SL, Gautier MC, Skouri F, Beisson J, Cohen J. The copines, a novel class of C2 domain-containing, calcium-dependent, phospholipid-binding proteins conserved from Paramecium to humans. J Biol Chem. 1998;273:1393–1402. [PubMed: 9430674]

55. den Hoed M, Eijgelsheim M, Esko T, Brundel BJJM, Peal DS, Evans DM, Nolte IM, Segrè AV, Holm H, Handsaker RE, Westra H-J, Johnson T, Isaacs A, Yang J, Lundby A, Zhao JH, Kim YJ, Go MJ, Almgren P, Bochud M, Boucher G, Cornelis MC, Gudbjartsson D, Hadley D, van der Harst P, Hayward C, den Heijer M, Igl W, Jackson AU, Kutalik Z, Luan J, Kemp JP, Kristiansson K, Ladenvall C, Lorentzon M, Montasser ME, Njajou OT, O'Reilly PF, Padmanabhan S, St Pourcain B, Rankinen T, Salo P, Tanaka T, Timpson NJ, Vitart V, Waite L, Wheeler W, Zhang W, Draisma HHM, Feitosa MF, Kerr KF, Lind PA, Mihailov E, Onland-Moret NC, Song C, Weedon MN, Xie W, Yengo L, Absher D, Albert CM, Alonso A, Arking DE, de Bakker PIW, Balkau B, Barlassina C, Benaglio P, Bis JC, Bouatia-Naji N, Brage S, Chanock SJ, Chines PS, Chung M, Darbar D, Dina C, Dörr M, Elliott P, Felix SB, Fischer K, Fuchsberger C, de Geus EJC, Goyette P, Gudnason V, Harris TB, Hartikainen A-L, Havulinna AS, Heckbert SR, Hicks AA, Hofman A, Holewijn S, Hoogstra-Berends F, Hottenga J-J, Jensen MK, Johansson A, Junttila J, Kääb S, Kanon B, Ketkar S, Khaw K-T, et al. Identification of heart rate-associated loci and their effects on cardiac conduction and rhythm disorders. Nat Genet. 2013;45:621–631. [PubMed: 23583979]

56. Challice CE, Virágh S. The architectural development of the early mammalian heart. Tissue Cell. 1974;6:447–462. [PubMed: 4432234]

NOVELTY AND SIGNIFICANCE

What Is Known?

- Normal function and formation of the heart relies on specialized cardiac cells known as the cardiac conduction system (CCS).
- A limited number of conduction-specific markers are known in part due to challenges in studying the CCS including large cell type heterogeneity, small cell numbers and complex three-dimensional (3D) anatomy.
- Single-cell RNA-sequencing (scRNA-seq) allows for genome-wide analysis of gene expression at single-cell resolution.

What New Information Does This Article Contribute?

- Using large-scale single-cell RNA-sequencing, we provide the single-cell transcriptome of the entire CCS within developing mice.
- Cell clusters were identified representing each major component of the conduction system, supported by expression enrichment of known conduction genes.
- Differential gene expression analyses between these CCS clusters and other cell types revealed a host of markers not previously associated with the conduction system, a subset of which were validated using fluorescent *in situ* hybridization as well as whole mount immunolabelling with volume imaging (iDISCO+) in three-dimensions on intact mouse hearts.
- Subcluster analysis additionally unveiled distinct conduction cell subtypes, including the clinically-relevant but previously poorly characterized “transitional cells”.

Coordinated contraction of the heart relies on specialized cardiac cells known as the cardiac conduction system. Despite its essential role, the CCS has remained challenging to interrogate due to inherent obstacles including cell type heterogeneity, small cell numbers and their complex 3D anatomy. scRNA-seq allows for genome-wide analysis of gene expression at single-cell resolution. Leveraging the advantages of scRNA-seq, we have overcome these hurdles and successfully isolated conduction cells from the entirety of the murine CCS. Analyses revealed cell clusters consistent with each major component of the CCS and led to the discovery and validation of several novel conduction genes. Further, subcluster analysis unveiled distinct conduction cell subtypes, including the clinically-relevant but poorly characterized “transitional cells”. Our study represents the first comprehensive assessment of transcriptional profiles from the entire CCS at single-cell resolution and provides a gene atlas for facilitating future efforts in conduction cell identification, isolation and characterization.

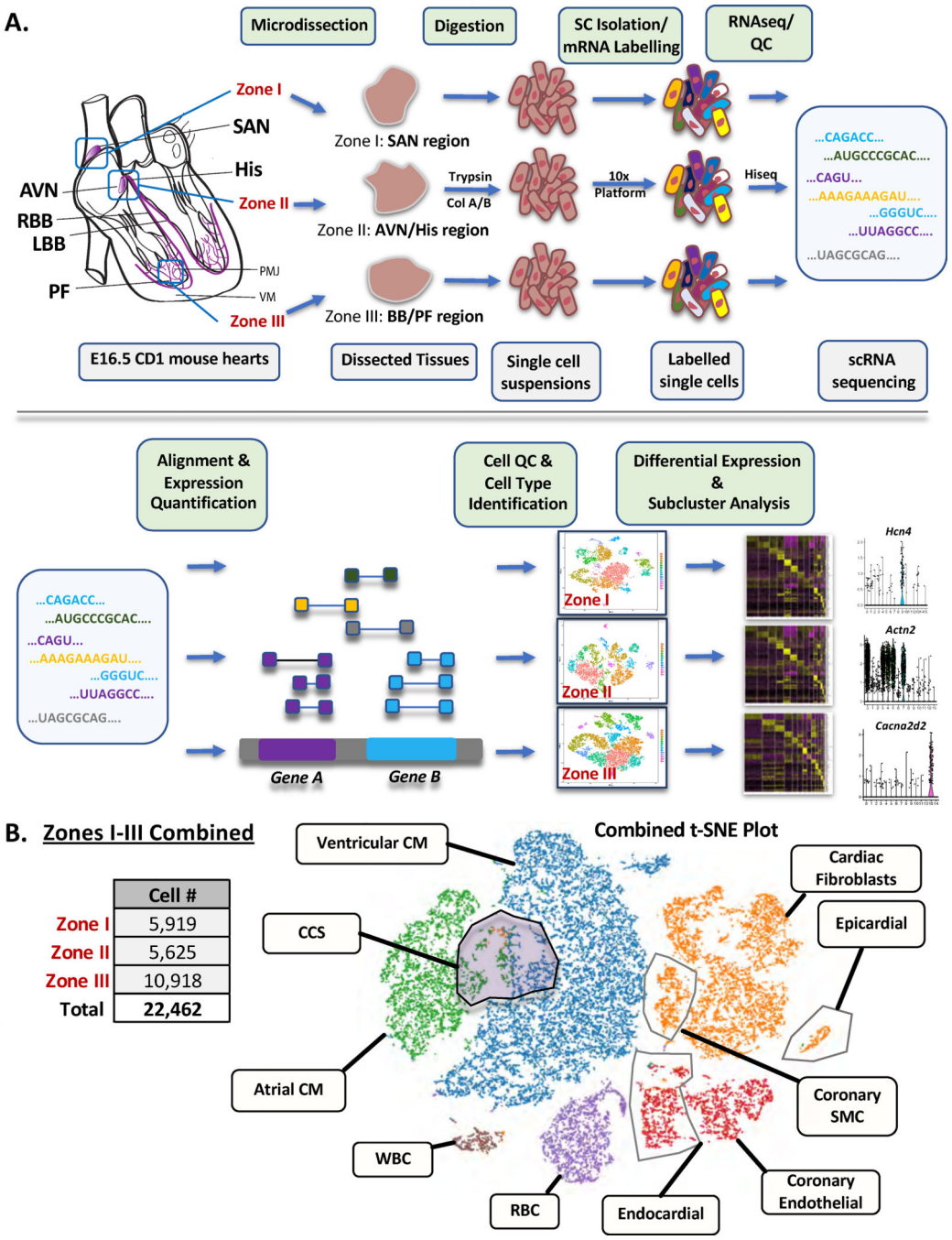


Figure 1. Single-Cell Isolation and Expression Profiling of Murine Cardiac Conduction System Components.

(**A**) Schematic representation of experimental design and workflow. Wild-type, embryonic day 16.5 (E16.5) CD1 mouse hearts were harvested and three zones of microdissection were isolated based on anatomical landmarks and entailed: **Zone I** - Sinoatrial node (SAN) region, **Zone II** - Atrioventricular node (AVN)/His region and **Zone III** - Bundle branch (BB)/Purkinje fiber (PF) region. A minimum of 6 embryonic hearts were pooled per zone. Tissues were digested into a single cell (SC) suspension, isolated via oil droplets, mRNA

labelled with cell-specific barcodes, made into a library and subsequently sequenced, all using the commercial droplet-based cell capture platform from 10x Genomics®. Data was then aligned with gene expression quantified followed by a series quality control (QC) steps as well as differential gene expression and subcluster analysis. (**B**) (Left) Table listing the number of cells isolated from each Zone and total number of cells evaluated. (Right) t-SNE plot of all cells combined with representative cell types identified by cluster. CCS, cardiac conduction system. CM, cardiomyocyte. LBB, left bundle branch. PMJ, Purkinje-myocyte junction. RBC, red blood cell. RBB, right bundle branch. scRNA-seq, single-cell RNA sequencing. SMC, smooth muscle cell. VM, ventricular myocardium. WBC, white blood cell.

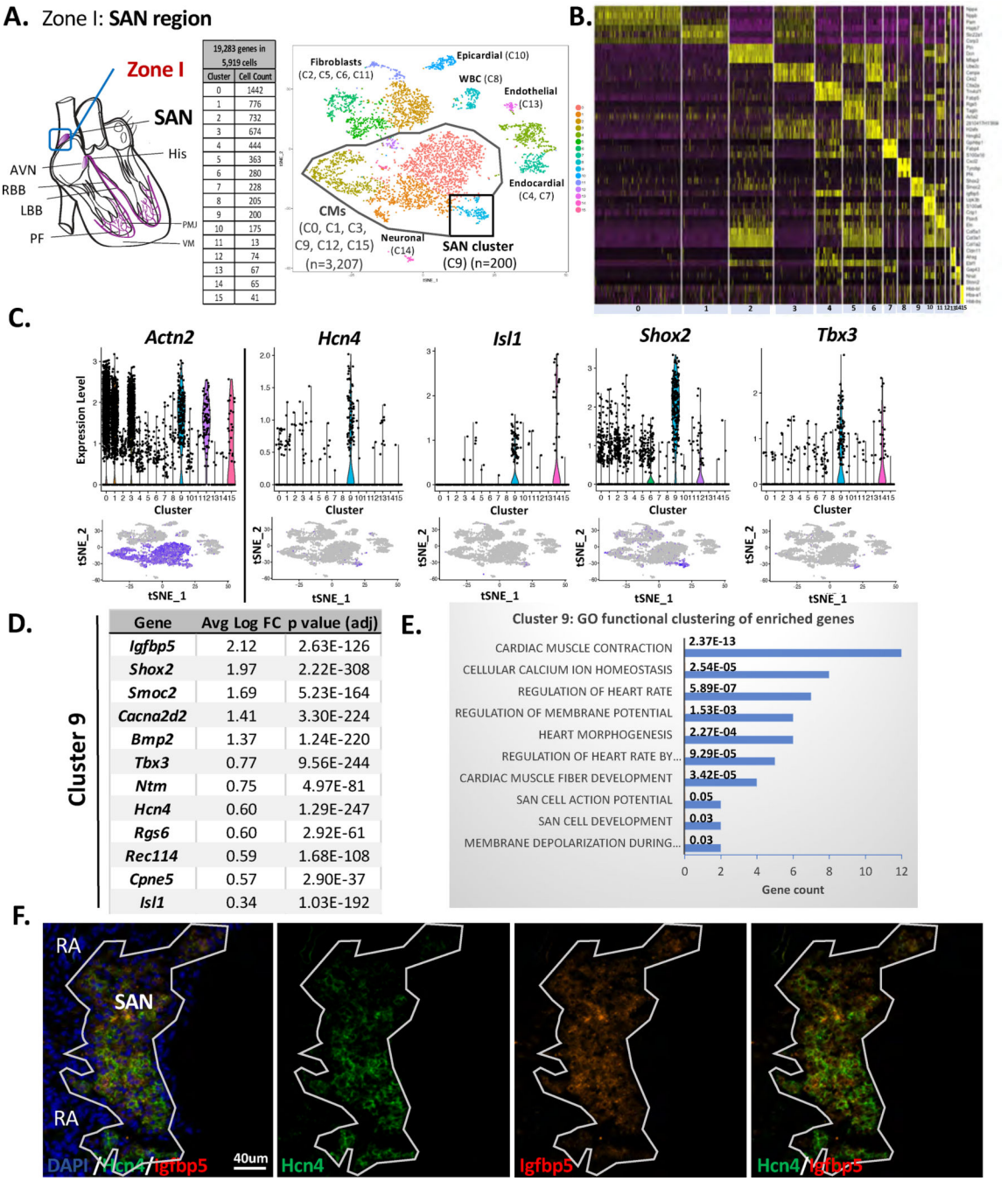


Figure 2. scRNA-seq Analysis of Zone I Revealed a Bona Fide Sino-Atrial Node (SAN) Cell Cluster.

(**A**) Cluster cell numbers and t-SNE plot for Zone I cells. (**B**) Heatmap of differentially expressed genes for each cluster of cells. (**C**) Cardiomyocyte (*Actn2*$^{+}$) and Nodal (*Hcn4*$^{+}$, *Isl1*$^{+}$, *Shox2*$^{+}$, *Tbx3*$^{+}$) signatures visualized by ViolinPlots (top) and FeaturePlots (bottom). (**D**) Table highlighting differentially expressed genes in Cluster 9. Avg log FC, average log fold change. (**E**) Gene ontology (GO) functional cluster analysis of all enriched genes for Cluster 9. (**F**) Immunofluorescence staining of postnatal day 4 (P4) (n=8) wild-type murine

cardiac tissue sections showing a SAN with colocalization of Hcn4 (green) and Igfbp5 (red) protein. Nuclei stained with DAPI (blue). RA, right atrial myocardium.

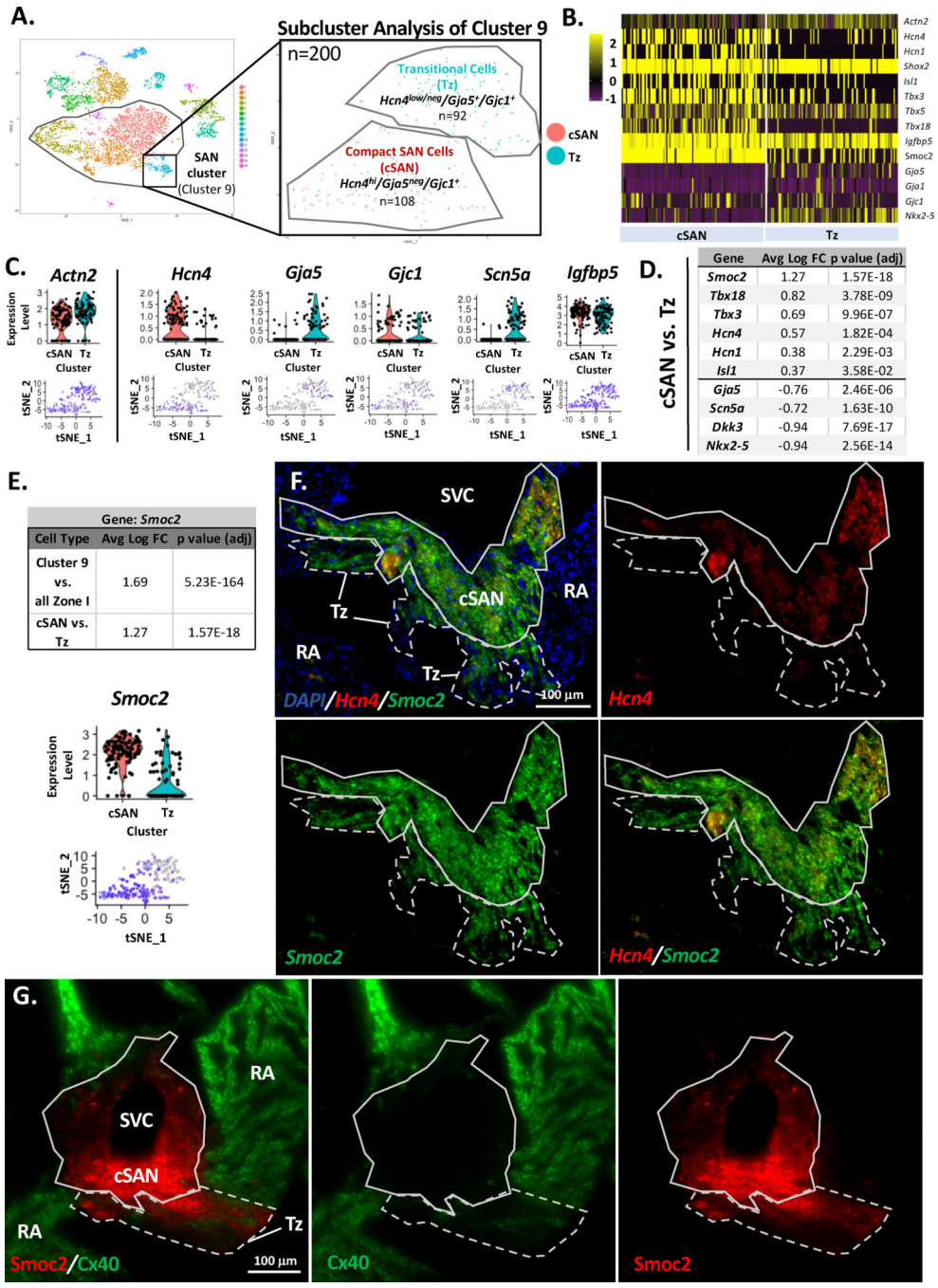


Figure 3. Analysis of Cluster 9 Revealed Unique SAN Subtypes.

(**A**) t-SNE plot of Cluster 9 cells demonstrated two distinct subclusters consistent with compact SA nodal cells (cSAN) (*Hcn4*^{hi}*/Gja5*^{neg}*/Gjc1*$^{+}$) and Transitional cells (Tz) (*Hcn4*$^{low/neg}$*/Gja5*$^{+}$*/Gjc1*$^{+}$). (**B**) Expression heatmap of established SAN and atrial CM genes for each subcluster of cells. (**C**) ViolinPlots (top) and FeaturePlots (bottom) for key cSAN and Tz markers. (**D**) Table highlighting differentially expressed genes between cSAN and Tz subclusters. Avg log FC, average log fold change. (**E**) Gene expression enrichment for *Smoc2* within Cluster 9 and the cSAN cluster as compared to all Zone I cells and the Tz

cluster, respectively. *Smoc2* gene expression within Cluster 9 visualized by ViolinPlot (top) and Featureplot (bottom). (**F**) Fluorescent RNA *in situ* hybridization targeting *Hcn4* (red punctae) and *Smoc2* (green punctae) mRNA within postnatal day 0 (P0) (n=3), wild-type mice. Compact SA node (cSAN), transitional (Tz) and surrounding right atrial (RA) cells. DAPI (blue). Solid line depicting cSAN cells ($Hcn4^{hi}/Smoc2^{+}$) and dashed line showing Tz cells ($Hcn4^{low/neg}/Smoc2^{+}$). Beyond these borders are right atrial cardiomyocytes (RA) ($Hcn4^{neg}/Smoc2^{neg}$). (**G**) Immunofluorescent staining of P0 (n=5) wild-type murine cardiac tissue sections showing a cSAN with expression of Smoc2 protein (red) and negative staining for Cx40 (green). A transitional cell population of cells ($Smoc2^{+}/Cx40^{low}$) emerging from the cSAN is demarcated by dashed border. Surrounding tissue is the RA ($Smoc2^{neg}/Cx40^{+}$).

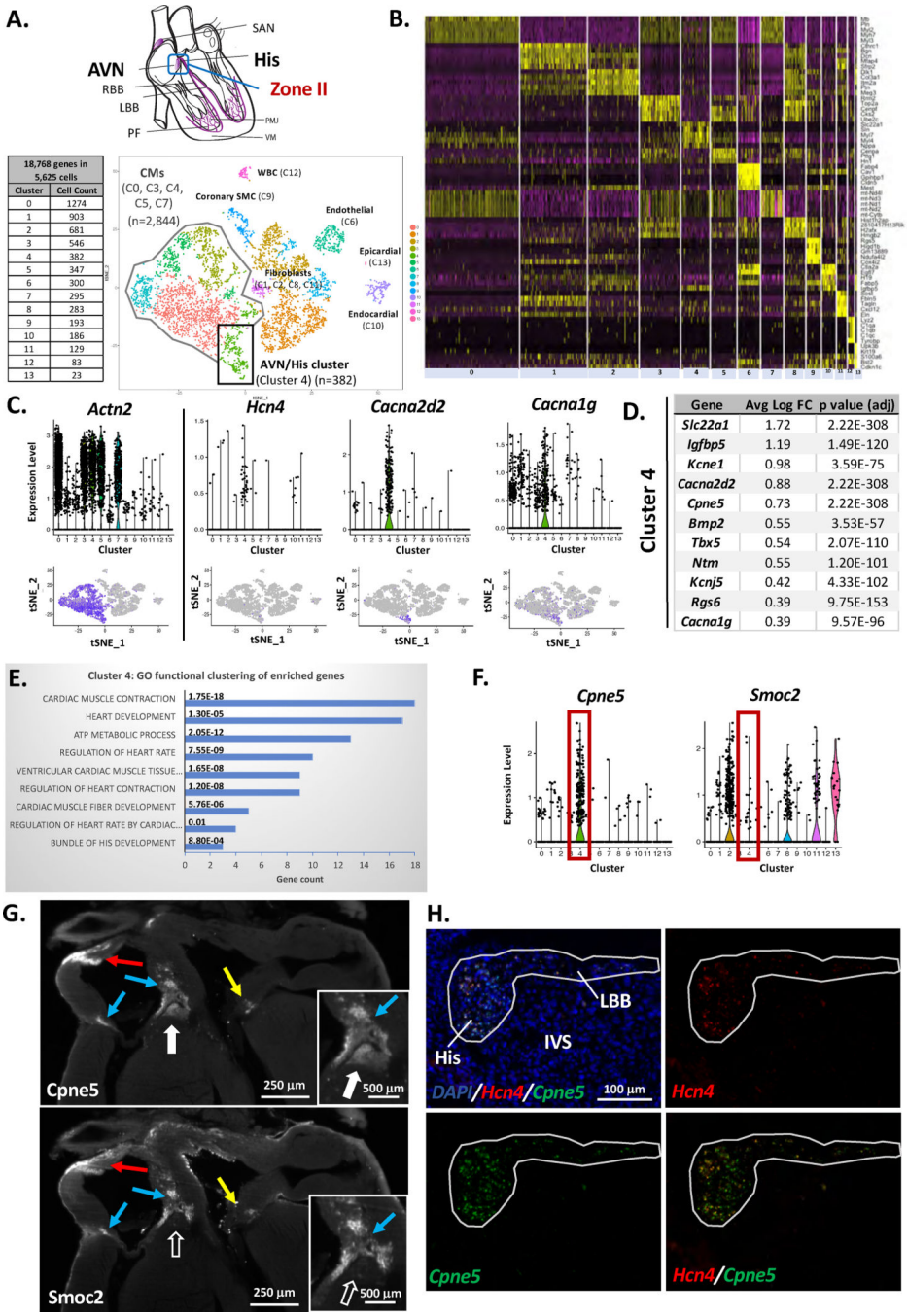


Figure 4. Identification of an AVN/His Cell Cluster Within Zone II.

(**A**) Cluster cell numbers and t-SNE plot of Zone II cells.(**B**) Heatmap of differentially expressed genes by cell cluster. (**C**) Cardiomyocyte (*Actn2*$^{+}$) and AVN (*Hcn4*$^{+}$, *Cacna2d2*$^{+}$, *Cacna1g*$^{+}$) gene signatures visualized by ViolinPlots (top) and FeaturePlots (bottom). (**D**) Table highlighting differentially expressed genes in Cluster 4. Avg log FC, average log fold change. (**E**) Gene ontology (GO) functional cluster analysis of all enriched genes for Cluster 4. (**F**) *Cpne5* and *Smoc2* gene expression within Zone II cells as illustrated by ViolinPlots. *Cpne5* expression is enriched within Cluster 4 while *Smoc2* is not. (**G**) The same postnatal

day 12 (P12) heart co-immunostained for Cpne5 and $Smoc^2$ protein (white); insets show a $Cpne5^+$ (solid white arrow) but $Smoc2^{neg}$ (empty white arrow) AVN. Red arrows, internodal tracks; blue arrows, right AV ring bundle; yellow arrows, left AV ring bundle. (**H**) FISH targeting *Hcn4* (red punctae) and *Cpne5* (green punctae) mRNA within a E16.5 (n=3), wild-type mouse heart section showing the His bundle giving rise to the left bundle branch (LBB) (both $Hcn4^+/Cpne5^+$) and surrounding ventricular cells of the IVS ($Hcn4^{neg}/Cpne5^{neg}$).

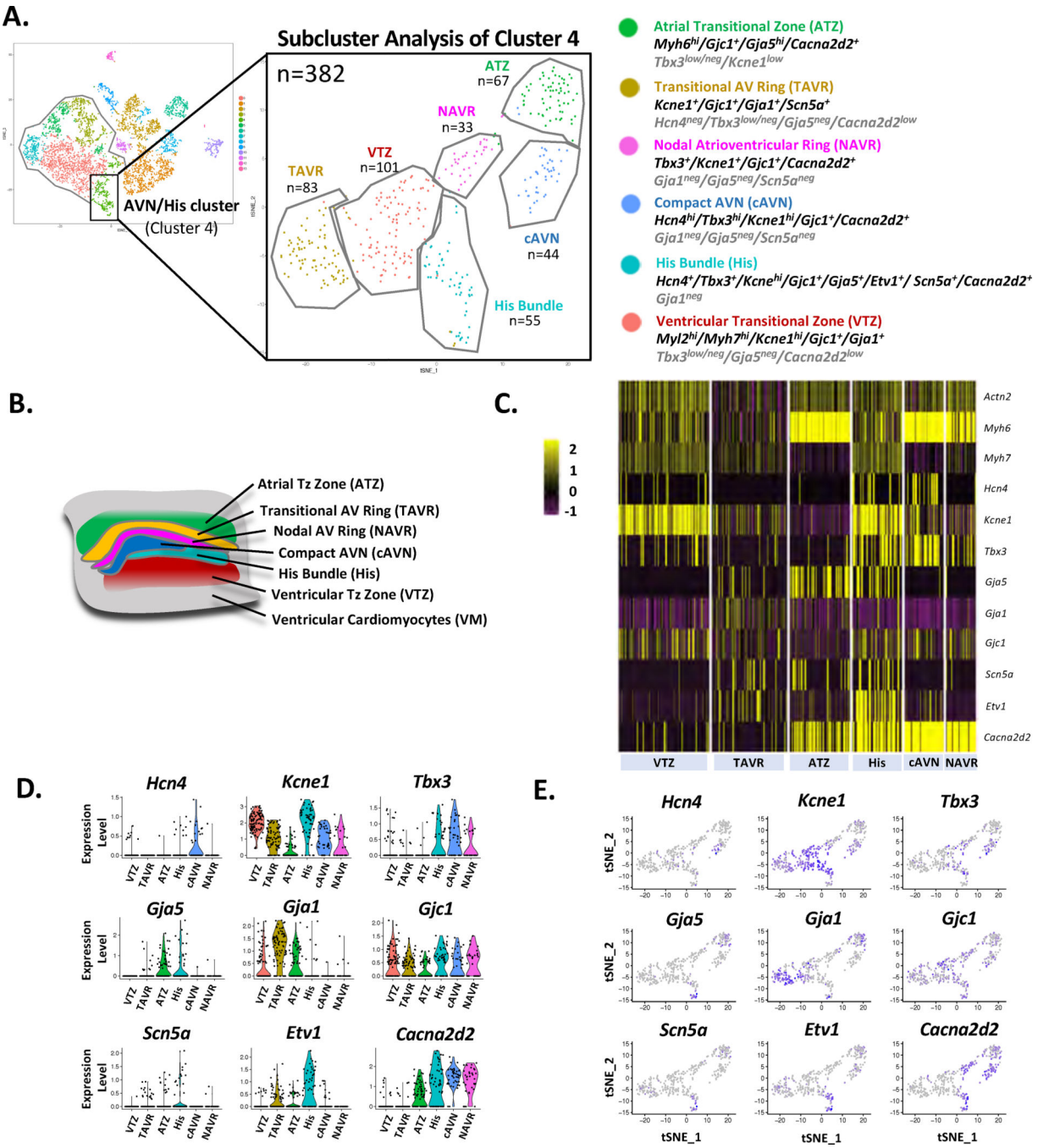


Figure 5. Analysis of Cluster 4 in Zone II Unveiled Distinct AVN and Transitional Cell Subtypes.

(**A**) Cells from the putative AVN/His cluster (Cluster 4 in Zone II) were selected for further analysis. t-SNE plot of Cluster 4 cells demonstrated six distinct subclusters. (**B**) Illustration of AV nodal region subdomains by transcriptional profiling (adapted from Aanhaanen et al. 2010).[6] (**C**) Heatmap, (**D**) ViolinPlots and (**E**) FeaturePlots demonstrating typical nodal, transitional and ventricular CM gene expression signatures for each subcluster of cells. ATZ, Atrial Transitional Zone; cAVN, compact AVN; His, His bundle; NAVR, nodal AV ring;

TAVR, transitional AV ring; VTZ, ventricular transitional zone; VM, ventricular cardiomyocytes.

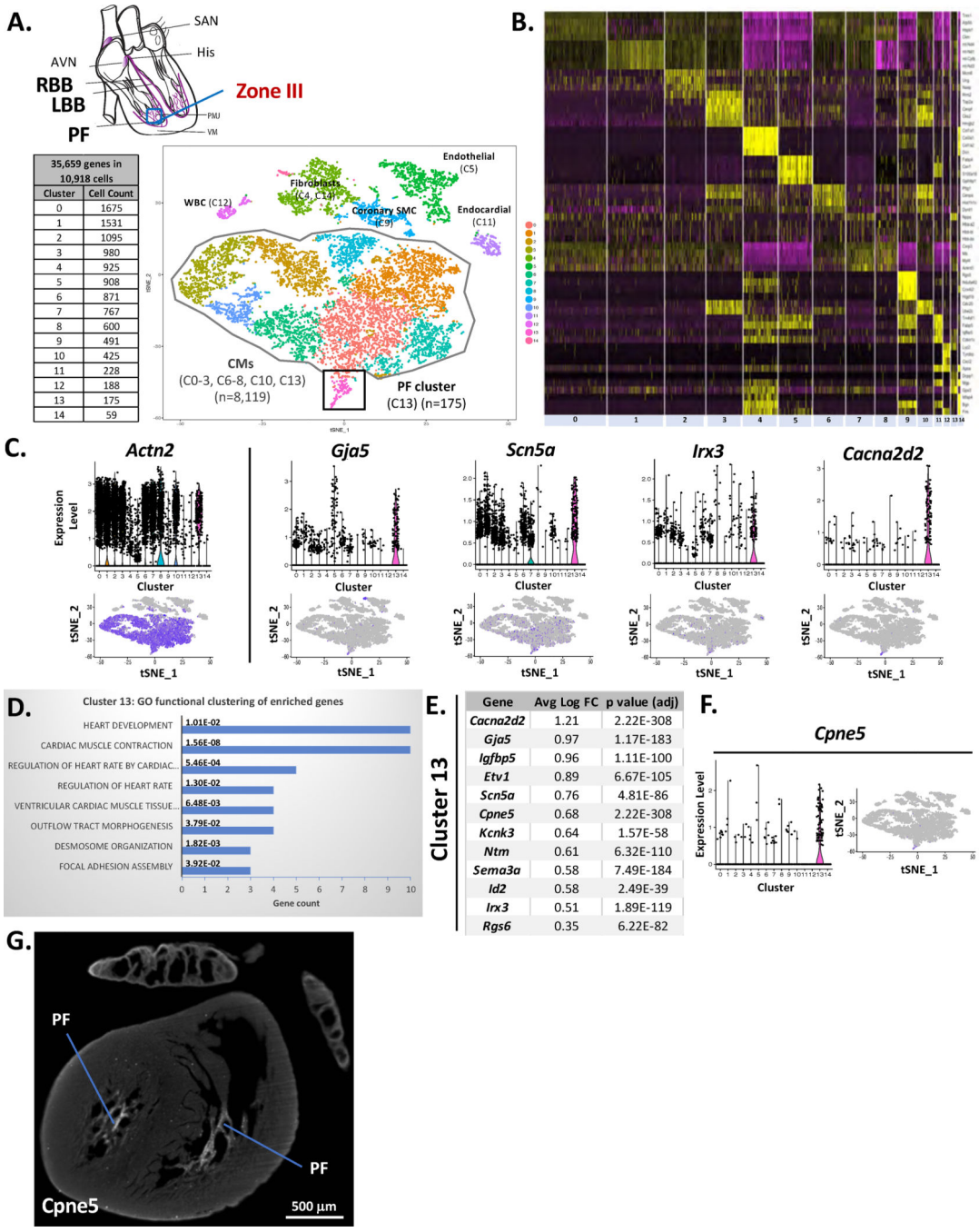


Figure 6. An Immature Purkinje Fiber (PF) Cell Cluster was Detected Within Zone III.

(**A**) Cluster cell numbers and t-SNE plot of Zone III cells. (**B**) Heatmap of differentially expressed genes for each cluster of cells. (**C**) Cardiomyocyte (*Actn2*$^{+}$) and PF (*Gja5*$^{+}$, *Scn5a*$^{+}$, *Irx3*$^{+}$ and *Cacna2d2*$^{+}$) gene signatures visualized by ViolinPlots (top) and FeaturePlots (bottom). (**D**) Gene ontology (GO) functional cluster analysis of all enriched genes for Cluster 13. (**E**) Table highlighting differentially expressed genes in Cluster 13. Avg log FC, average log fold change. (**F**) *Cpne5* gene expression within Zone III cells illustrated by

ViolinPlot (left) and FeaturePlot (right). (**G**) Immunofluorescence staining of an E16.5 (n=6) wild-type murine cardiac tissue section showing Cpne5 (white) of the immature PF network.

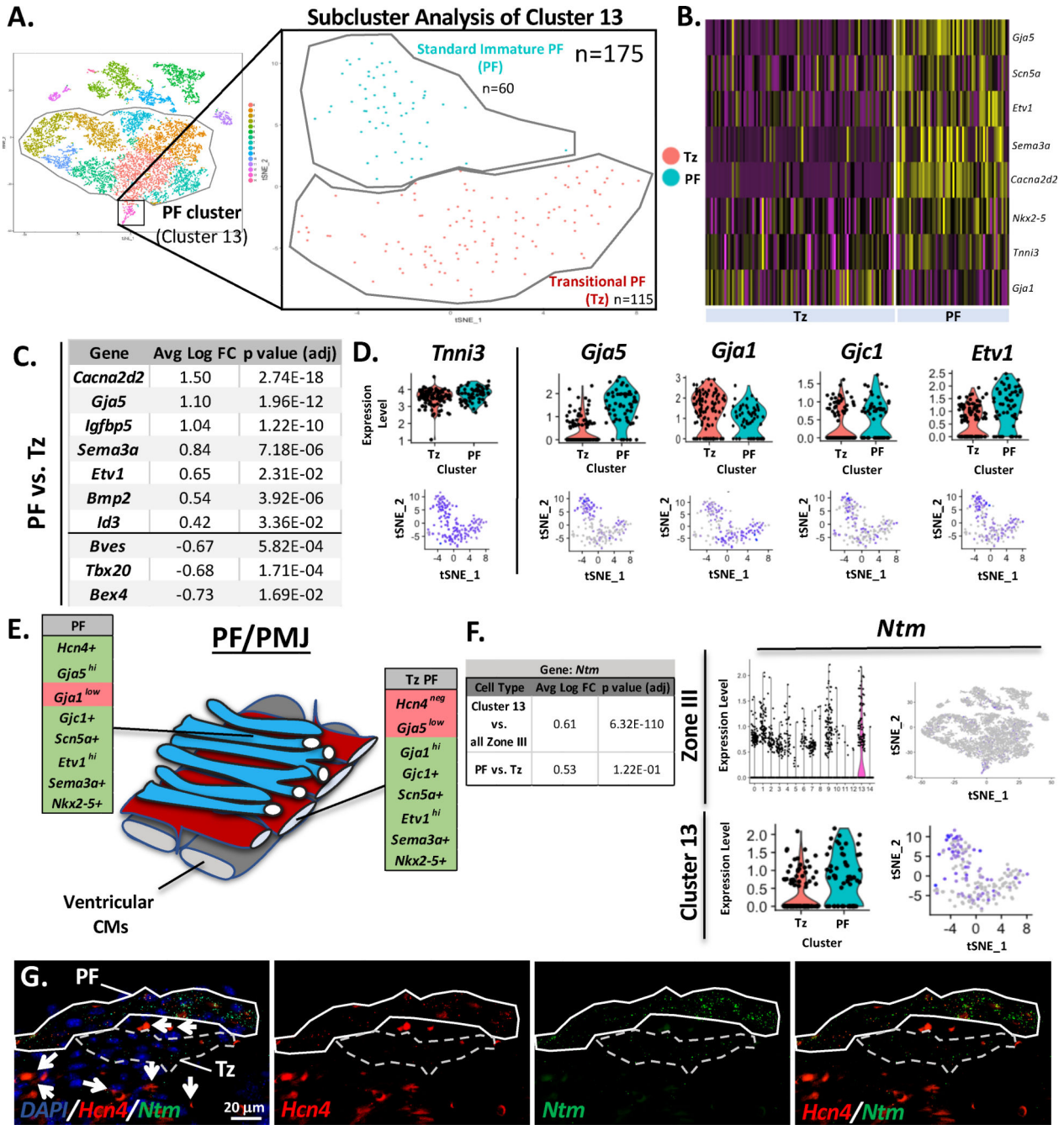


Figure 7. Analysis of Cluster 13 in Zone III Identified Unique Immature PF and Transitional PF Cell Subtypes.

(**A**) t-SNE plot demonstrated two distinct subclusters within Cluster 13, consistent with a standard immature PF cluster ("PF") and a transitional PF cluster ("Tz"). (**B**) Expression heatmap of known PF and ventricular cardiomyocyte genes within the standard PF and Tz clusters. (**C**) Table highlighting differentially expressed genes between the two subclusters, represented by (**D**) ViolinPlots (top) and FeaturePlots (bottom). Avg log FC, average log fold change. (**E**) Illustration of the Purkinje-myocyte junction (PMJ) consisting of standard PF (blue), ventricular CMs (gray) and transitional PFs (red) bridging the two (adapted from

Tranum-Jensen et al. 1991).[9] Summary of enriched (green) and downregulated (red) genes for each cell subtype based on scRNAseq data presented here. (**F**) Gene expression enrichment table for *Ntm* between Cluster 13 and standard PF subcluster as compared to the rest of Zone III cells and Tz subcluster, respectively. *Ntm* gene expression demonstrated in Zone III (upper panels) and within Cluster 13 (lower panels) as visualized by ViolinPlot (left) and Featureplot (right). (**G**) Fluorescent RNA in situ hybridization targeting *Hcn4* (red punctae) and *Ntm* (green punctae) mRNA expression within immature PFs and transitional PF cells within a P0 mouse heart section (n=3). DAPI (blue). Solid line depicting standard immature PF cells (*Hcn4*$^{+}$/*Ntm*$^{+}$) and dashed line showing Tz cells (*Hcn4*$^{low/neg}$/*Ntm*$^{+}$). Beyond these borders are ventricular myocardial cells (*Hcn4*^{neg}/*Ntm*^{neg}). Larger brightly-stained objects (examples noted by white arrows) in both red and green channels represent autofluorescence from red blood cells (yellow in merge photos). True RNAscope fluorescent *in situ* hybridization signal is represented by small punctae in either green (*Ntm*) or red (*Hcn4*).

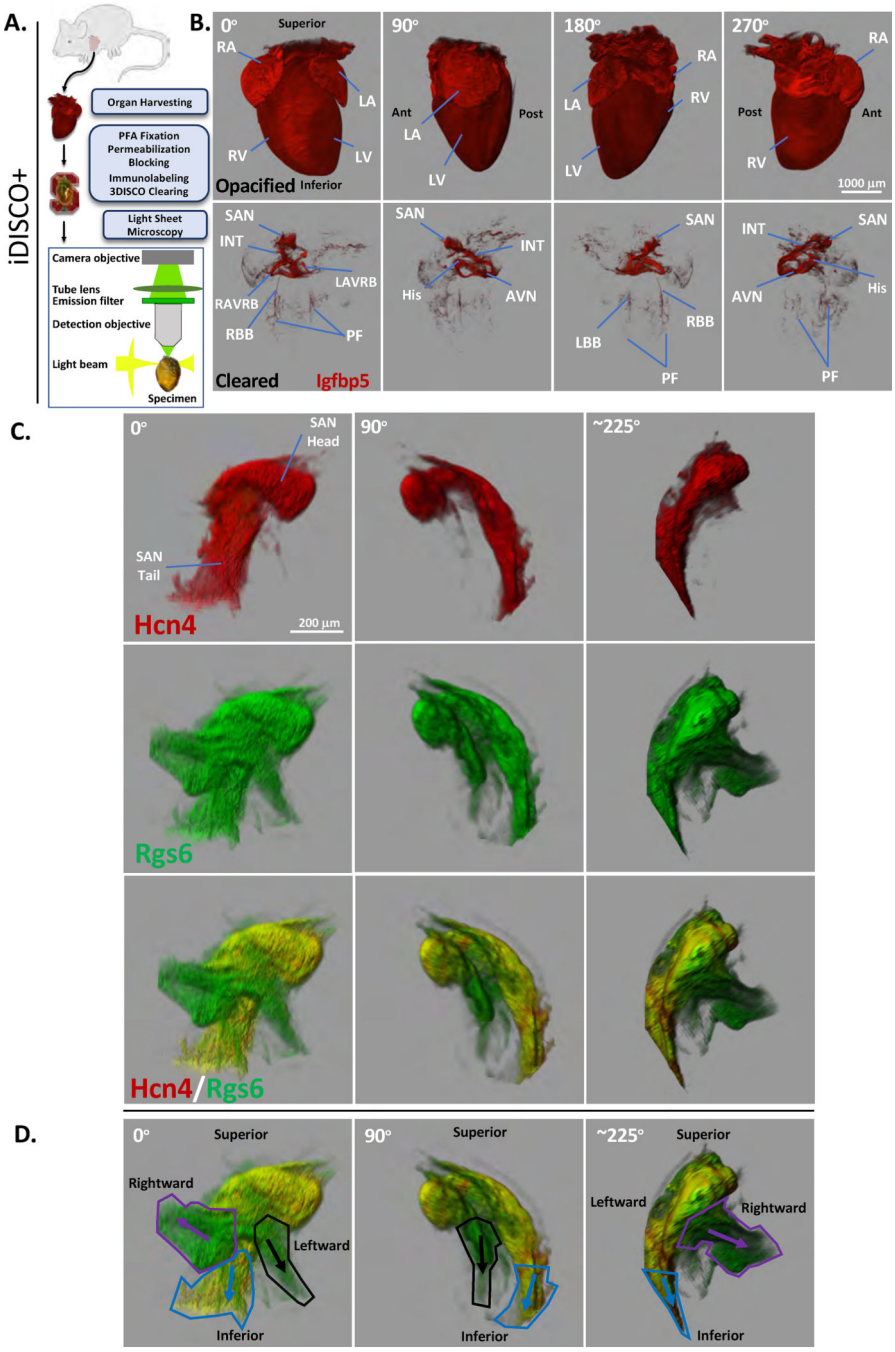


Figure 8. Optical Clearing and 3D Volumetric Analyses Delineate the Architecture of the Entire CCS and SA Nodal Substructure Within Intact Murine Hearts.

(**A**) Schematic representation of workflow for iDISCO+ clearing of intact embryonic and postnatal mouse hearts and visualization using light sheet microscopy. At least 10 biological replicates were used for each immunolabelling. (**B**) A wild-type whole heart from an E16.5 mouse embryo is shown at four angles of view (0, 90, 180 and 270 degrees) at 0.63x magnification. Top and bottom rows are the same optically cleared heart using iDISCO+ where, in the top row, background fluorescence has been saturated to provide a

representation of the opacified heart. Bottom row demonstrates the same tissue-cleared heart, showing Igfbp5 immunostaining which marks the cardiac conduction system in its entirety. (**C**) iDISCO+ cleared wild-type, intact SAN (6.3x magnification) from E16.5 murine heart co-immunolabeled for Hcn4 (red) and Rgs6 (green) protein. SAN shown at three angles of view (0, 90, ~225 degrees). (**D**) Merge image (Hcn4 in red, Rgs6 in green) with two major transitional sinoatrial conduction pathways (SACPs) outlined ($Hcn4^{neg}/Rgs6^{+}$). Purple SACP: from SAN body directed rightward; Black SACP: from the SAN head directed inferiorly and leftward. A third Hcn4+ SACP is also marked by Rgs6 and emerges inferiorly from the tail of the SAN to give rise to the internodal tracks (blue line) not seen here but visible in panel 8B. Ant, Anterior. AVN, atrioventricular node. His, His bundle. INT, internodal tracks. LA/RA, left or right atrium. LAVRB, left AV ring bundle. LBB/RBB, left of right bundle branch. LV/RV, left or right ventricle. PF, Purkinje fiber. Post, Posterior. RAVRB, right AV ring bundle. SAN, sinoatrial node.

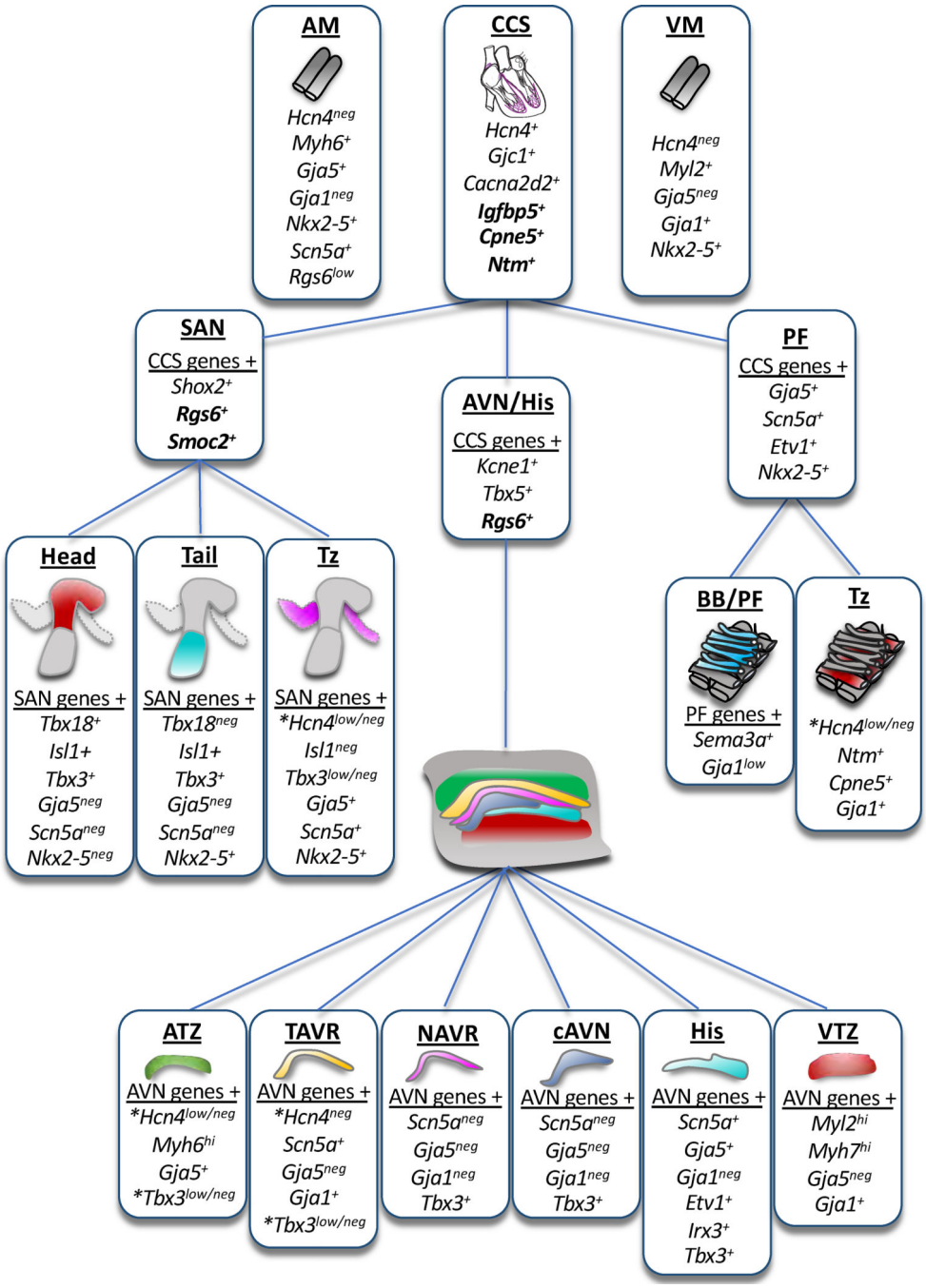


Figure 9. Transcriptional Landscape of the Cardiac Conduction System.

A summary of differential gene expression, both established and novel per the scRNA-seq findings of this study, within all components of the CCS including the sino-atrial node (SAN), atrioventricular node/His bundle (AVN/His) and Purkinje fiber (PF) regions. Asterisk (*) highlights genes with different expression as compared to parental cell type within the hierarchy. AM, atrial myocardium, ATZ, Atrial Transitional Zone; cAVN,

compact AVN; NAVR, nodal AV ring; TAVR, transitioanal AV ring; VTZ, ventricular transitional zone; VM, ventricular cardiomyocytes.