SUPPLEMENTAL MATERIAL

Supplemental Methods

Generation of *Nkx2-5* enhancer Δ226 bp (*Nkx2-5*$^{\Delta enh}$) mouse using the CRISPR-Cas9 system

We used a standard protocol to generate the Δ226-bp mouse line (*Nkx2-5*$^{\Delta enh/\Delta enh}$) using the CRISPR-cas9 system as previously described.[31] In brief, two single guide RNAs (sgRNAs: GATTACTTTATCTTCCCCGG and GGGACATCTTTTGTCTCT) were used to target the cardiac enhancer region located -9435/-8922 from the transcriptional start site of *Nkx2-5*. sgRNAs were synthesized and purified using a MEGAshortscript T7 Transcription Kit (Invitrogen, MA, USA) and a MegaClear Kit (Ambion, CT, USA). Cas9 mRNA (L7206, TriLink, CA, USA) and two sgRNAs were used for microinjections into one-cell mouse embryos to create the Δ226-bp mouse line (Figure 2B). Correct founders were determined by direct PCR and Sanger sequencing with primers (gctccatgtcccagaatctt and atggaactcagaggcacacc). The founders were maintained in a C57BL/6J background. For His-Purkinje system morphology imaging, the Δ226-bp mouse lines were crossed with *Contactin2*$^{EGFP/+}$ line harboring a *Contactin2*-enhanced green fluorescence protein BAC reporter gene. For the survival study, newborns were genotyped and kept with their mothers without further intervention.

Bright field heart imaging and histology

For whole-mount images, neonates at postnatal day 1 (P1) were euthanized and a median sternotomy was performed to visualize the heart and great vessels followed by fixation with 4 % paraformaldehyde (PFA) overnight. After wash with ice-cold phosphate buffered saline (PBS), bright-field images of the thoracic cavity were taken using the Zeiss

Discovery V8 microscope equipped with a Zeiss AxioCam Color camera interfaced with Zeiss Zen 2012 software (ZEISS, Oberkochen, Germany). For subsequent histological assessment, the hearts were excised from the body.

For histological assessment, the excised hearts from P1 neonates were rinsed in ice-cold PBS, and then stored in 70% ethanol. The hearts were then paraffin-embedded, sectioned at 5 μm-thickness, and stained with hematoxylin and eosin according to standard protocols at New York University Experimental Pathology Research Laboratory. All bright field images were acquired using a BZ-X800 microscope (Keyence, Osaka, Japan).

For morphological assessment of embryonic hearts, the thoracic cavity from embryonic day (E) 11.5 embryos were fixed in 4% PFA overnight and processed for paraffin sectioning. Continuous whole embryo sections with 5 μm-thickness were stained with H&E, and scanned in using the Zeiss Discovery V8 microscope. The scanned images were spatially aligned and stacked for volume rendering using MATLAB (Mathworks, Inc., MA, USA) with a custom written routine. For conotruncal cushion volume analysis, 47 sections per animal were volume-rendered for each inferior and superior septal cushions, and manually segmented over the total number of sections.

Immunofluorescence

Embryonic torsos at E12.5 from mating pairs of heterozygous mutants were isolated and fixed in 4% PFA in PBS for 2 hours and soaked in 15% then 30% sucrose. The tissues were embedded in OCT (Sakura Finetek USA, CA, USA) and stored in −80°C until sectioning. Sections of 10 μm-thickness were collected on slides and subjected to

immunostaining. The primary antibodies used were connexin 40 (Cx40-A, 1/400, Alpha Diagnostic International, Inc., TX, USA), and troponin I (ab56357, 1/100, Abcam, MA, USA). Secondary antibodies used were: donkey anti-rabbit 555 (1:500, Thermo Fisher Scientific, A31572); donkey anti-goat 488 (1:500, Thermo Fisher Scientific, A11055). Slides were coverslipped with Vectashield mounting medium with DAPI (Vector Laboratories, CA, USA). Stained sections were visualized with KEYENCE BZ-X800 fluorescence microscope (Keyence, IL, USA).

Scanning electron microscopy (SEM)

A standard procedure for SEM of the embryonic heart was utilized by the Microscopy Core at New York University Grossman School of Medicine. Embryos at E12.5 were assessed. The anterior pericardium was excised to expose the heart of each embryo prior to pinning on dental wax dishes, and fixation in 2.5% glutaraldehyde (G5882, Sigma-Aldrich, Inc., MO, USA) in PBS at 4°C for 48 hours. Samples were washed in PBS, post fixed with 1% Osmium tetroxide aqueous solution (Electron Microscopy Sciences, PA, USA) for 1 hour, then dehydrated in graded series of ethanol prior to drying using a Tousimis autosamdsri 931 critical point dryer (Tousimis, MD, USA). Dried specimens were mounted ventral side up and ion sputtered with gold (Safematic CCU-010 SEM coating system, Rave Scientific, NJ, USA) to ~30 nm thickness before being scanned with a Zeiss Gemini300 FESEM (Carl Zeiss Microscopy, Oberkochen, Germany). SEM photomicrographs were taken in standard orientations and magnifications.

Electrocardiography (ECG)

For ECG measurements, P1 mice were anesthetized with 1% isoflurane using a nose cone. Mice were placed in supine position and warmed to 37.5 °C on a heated platform to maintain a stable body temperature during recordings. Surface ECG signals from standard limb leads were recorded for 3 minutes. Parameters from averaged ECG traces on lead II were analyzed in blinded fashion using LabChart Pro software (v8.1.9, ADInstruments Inc., CO, USA) according to the instructions. Measured parameters included heart rate, P-wave duration, PR interval and QRS interval in 9–10 animals per group. The P-wave duration included both the first positive and negative deflections from baseline. PR interval was measured from the onset of the P wave to the onset of the QRS complex. QRS interval was determined from the initial upward deflection from baseline after the P wave until the maximum deflection in the negative wavefront between S wave and J wave, the additional wave right after the QRS complex representing early repolarization.

Transthoracic Echocardiography

A 30–70 Mhz linear-array probe (MX700) and a Vevo 3100 Imaging System (FUJIFILM VisualSonics, Inc., Toronto, Canada) were used for transthoracic echocardiographic recordings. P1 mice were anesthetized with 1% isoflurane via nose cone. Mice were positioned supine and warmed to 38 °C on a heated platform to maintain a stable body temperature during recordings. Biventricular strain and strain rate were calculated offline in short axis views using analysis routine written in MATLAB (Mathworks, Inc.). Detailed image processing algorithm for biventricular strain is described as follows: (1) A single cardiac cycle is selected by user. (2) Circumferential strain was calculated employing two-

dimensional block matching algorithm for speckle tracking after manually selecting a region of interest in right and left ventricular myocardium in short axis views. (3) Strain is calculated by measuring mechanical deformation of two adjacent tracking points over the single cardiac cycle. For mechanical activation map generation, activation time was measured as time to maximum negative strain (peak systole) from onset of systole in a similar manner to signal processing in a conventional optical mapping analysis. As a last step, a 4 x 4 pixel Gaussian averaging kernel was applied to reduce noise.

Chromatin immunoprecipitation (ChIP) assay in embryonic mouse hearts

ChIP assay was performed in mouse embryonic hearts at E12.5. Embryos of CD-1 mice purchased from Charles River Laboratories (MA, USA) were dissected in ice-cold PBS. Whole hearts including truncal outflow tract (OFT) were pooled and dissociated. The dissociated heart tissues were fixed for 10 min in 1% formaldehyde in cell culture medium, then quenched by 0.125M glycine for 5 min at room temperature. ChIP was performed with the crosslinked samples using ChromaFlash High-Sensitivity ChIP kit (P-2027, EpigenTek Group Inc., NY, USA) according to the manufacturer's protocol. Extracted chromatin was sheared using Diagnode Bioruptor (UCD-200, Diagnode Inc., NJ, USA) for 15 cycles of 30 sec ON and 30 sec OFF at high setting with the Bioruptor water cooler. ChIP-grade rabbit anti-Gata6 (5851S, Cell Signaling Technology, Inc., MA, USA) and rabbit non-immune IgG were used. Immunoprecipitated DNA was used to amplify the region containing GATA binding sites (labeled GS1 and GS2) in the cardiac enhancer region. The primers used were as follows: GS1 (amplicon 98 bp) forward, 5'-CTCTTGACAGGCAGCGTTTG -3' and reverse, 5'- TGTCGGGCCAGGAGTATCA -3';

GS2 (amplicon 109 bp) forward, 5'- CAGAGCAGATTTGGTGTGCG -3' and reverse, 5'- GCAAGAAGTTTTCCTCCGGG -3'. Quantitative PCR (qPCR) was performed using Power SYBR Green PCR Master Mix (4367659, Thermo Fisher Scientific, MA, USA) on a StepOne Real-Time PCR System (Applied Biosystems, MA, USA). Fold enrichment was calculated between specific Ab (GATA6)-immunoprecipitated samples and those immunoprecipitated with the non-immune IgG. The enrichments were analyzed in five biological replicates.

Optical mapping

High-resolution optical mapping experiments were performed with P1 mouse hearts. Hearts were surgically removed with thoracotomy. Excised P1 hearts were cannulated with a 30 gauge blunt tip, hollow bore needle via the aorta, and initially perfused with oxygenated Tyrode's solution (composition [mmol/L]: NaCl 114, $NaHCO_3$ 25, dextrose 10, KCl 4.6, $CaCl_2$ 1.5, Na_2PO_4 1.2, $MgCl_2$ 0.7), to clear blood and stabilize the heart, followed by Tyrode's solution containing 10 microM blebbistatin. Perfusate temperature was maintained at 37°C. Hearts were allowed to recover for 15 minutes and then stained with the voltage-sensitive dye, Di-4-ANEPPS (Molecular Probes Inc., Eugene, OR, USA). Light from green LEDs (530 nm; ThorLabs) was used as an excitation source and the emitted light (620 nm long pass) from the ventricles was detected with a high-resolution CMOS camera (Mi-CAM Ultima-L, SciMedia, Ltd., CA, USA) at 1000 frames/s in bin mode (256 x 256 pixels) with 14-bit resolution. A custom-made software package written in MATLAB was used to analyze optical signals offline. Signal processing algorithm using FIR filter was adapted from an open source package

(https://github.com/optocardiography). In brief, time lapse fluorescence and ECG signals were acquired simultaneously using a single data acquisition system. The pixel resolution of the current optical mapping system is 30 microns per pixel. Detailed image processing algorithm is as follows: (1) Baseline wander and background noise were removed with a bandpass filter (FIR temporal filter with a band pass of 0–100 Hz). (2) For activation map generation, a single cardiac cycle is chosen by user on custom Graphic User Interface. (3) Activation map was generated by measuring upstroke time of optical signals from each pixel. (4) As a final step, a 5 x 5 pixels Gaussian averaging kernel was applied to reduce overall noise. The gradient in activation maps represents the time taken by the activation front to travel a unit distance in the perpendicular direction to the wavefront.

Quantitative RT-PCR

Whole hearts were collected from E12.5 mice, and the atria, right ventricle (RV), left ventricle (LV), and OFT were then removed and stored at −80 °C before RNA extraction. Total RNA was extracted using the RNeasy Plus Mini Kit (74134, Qiagen Inc., MD, USA). For quantification of select gene expression levels of interest, cDNA synthesis was performed using Maxima First Strand cDNA Synthesis Kit for RT-qPCR (K1641, Thermo Fisher Scientific, MA, USA). qPCR was then performed using Power SYBR Green PCR Master Mix (Thermo Fisher Scientific) on a StepOne Real-Time PCR System (Applied Biosystems). All primer pairs for qPCR were purchased from Origene (MD, USA) and used according to the manufacturer's protocol. Target gene expression was normalized

to glyceraldehyde-3-phosphate dehydrogenase (*gapdh*), and fold changes were calculated using the comparative $2^{-\Delta\Delta C_T}$ method.

RNA sequencing

RNA sequencing (RNA-seq) was performed with total RNA extracted from the OFT for transcriptome analysis. For each sample, isolated total RNA was amplified and sequencing libraries were prepared using the Low Input Clontech SMART-Seq RNA Kit (Nugen, CA, USA) with 50 bp paired-ended at 10 to 20 million reads per replicate on an Illumina NextSeq 6000 instrument. Sequencing reads were mapped to the reference genome (mm10) using the STAR aligner (v2.5.0c, http://code.google.com/p/rna-star/). Alignments were guided by a Gene Transfer Format (GTF) file. The mean read insert sizes and their standard deviations were calculated using Picard tools (v.1.126, http://broadinstitute.github.io/picard). The read count tables were generated using HTSeq (v0.6.0, http://www-huber.embl.de/HTSeq), and normalized based on their library size factors using DESeq2 (v3.0, http://www.bioconductor.org/packages/release/bioc/html/DESeq2.html). Using the DESeq2 package, differential expression analysis was performed with Wald test and the p values were corrected by Benjamini-Hochberg method. The Read Per Million (RPM) normalized BigWig files were generated using BEDTools (v2.17.0, http://code.google.com/p/bedtools) and bedGraphToBigWig tool (v4, http://hgdownload.cse.ucsc.edu/admin/exe/linux.x86_64/). To compare the level of similarity between the groups and their replicates, principal component analysis (PCA)

and Euclidean distance analysis were used. All the significant genes between *Nkx2-5*$^{\Delta enh/\Delta enh}$ and *Nkx2-5*$^{+/+}$ were selected based on adjusted p value cut-off of 0.1 (FDR < 0.1). Genes within p value cut-off of 0.05 for gene expression differences were categorized as candidate DEGs. Integrative enrichment analysis was performed with DEGs using an open online source Enrichr-KG (https://maayanlab.cloud/enrichr-kg). Gene Ontology Biological Process and MGI Mammalian Phenotype gene set libraries were selected for the analysis. The following parameters were selected for visualization: 1) top 20 terms in each library, 2) genes with at least 3 links, and 3) terms with at least 3 links. Gene Ontology analysis was performed with candidate DEGs using an open online source Enrichr (https://maayanlab.cloud/Enrichr/). All the downstream statistical analyses and generating plots were performed in R environment (v4.0.3, https://www.r-project.org/). This RNA-seq data have been deposited into NCBI GEO database (GSE227361).

Chromatin immunoprecipitation sequencing (ChIP-seq) and transposase accessible chromatin sequencing (ATAC-seq) data analysis

Raw sequencing data were obtained from the previous publication.[29] The datasets were analyzed using Seq-N-Slide pipeline (https://github.com/igordot/sns). Routes sns-chip and sns-atac were applied to ChIP-seq data and ATAC-seq data, respectively. Sequencing reads were mapped to the reference genome (hg38) using the Bowtie2 (v2.3.4, http://bowtie-bio.sourceforge.net/bowtie2/index.shtml) and duplicate reads were removed by Sambamba (v.0.6.8, http://www.open-bio.org/wiki/Sambamba). The RPM normalized BigWig files were generated using deepTools (v.3.1.0, http://deeptools.ie-freiburg.mpg.de). Transcription factor binding sites (narrow peaks) were detected based

on deduplicated BAMs by MACS (v2, https://macs3-project.github.io/MACS/). Bedtools (v.2.17.0) was used to generate the blacklist-filtered BED files and MACS BigWig files were generated using bedGraphToBigWig tool (v4). To annotate the ChIP/ ATAC peaks, we used the R package ChIPseeker (v.1.24, http://www.bioconductor.org/packages/release/bioc/html/ChIPseeker.html) and DiffBind (v3.0.15, https://bioconductor.org/packages/release/bioc/html/DiffBind.html). All downstream statistical analyses and generating plots were performed in R environment (v.4.0.3). IGV (http://www.broadinstitute.org/igv/) was used to visualize the BigWig files. HOMER (v4.10, http://homer.ucsd.edu/homer/motif/) was used for motif analysis.

Single cell RNA-seq (scRNA-seq) data analysis

Raw sequencing data from E10.5 *Pitx2*$^{hd-/-}$ heart tissues were downloaded from Gene Expression Omnibus (GSE131181).[30] The dataset was analyzed with the following packages in R environment (v4.0.3). Fastq files were mapped to reference genome (mm10), and gene-cell matrices were generated by 10x Genomics Cell Ranger software (https://www.10xgenomics.com/). Expression matrices from different samples were merged and imported into Seurat (v 4.1.1, https://satijalab.org/seurat/). A quality control was then performed on the cells to calculate the number of genes, UMIs and the proportion of mitochondrial genes for each cell and the cells with low number of covered genes (gene-count < 200) and high mitochondrial counts (mt-genes > 0.1) were filtered out. Then log normalization was performed on the filtered matrix. Next, PCA was performed, and top 20 principal components were used as input for Uniform Manifold Approximation and Projection (UMAP) and graph-based clustering. Detection of DEGs

between clusters was performed using the FindAllMarkers function, which is based on the non-parametric Wilcoxon rank sum test, specifying return of only upregulated genes with a log2 fold change cutoff of 0.25. The marker genes were then used to determine the cell types. FindMarkers was used for detecting DEGs between control and *Pitx2*$^{hd-/-}$ in the OFT-CM cluster. Ligand-receptor and ligand-target interactions were assessed using CellChat (v1.5.0, https://github.com/sqjin/CellChat) with default parameters in R environment.

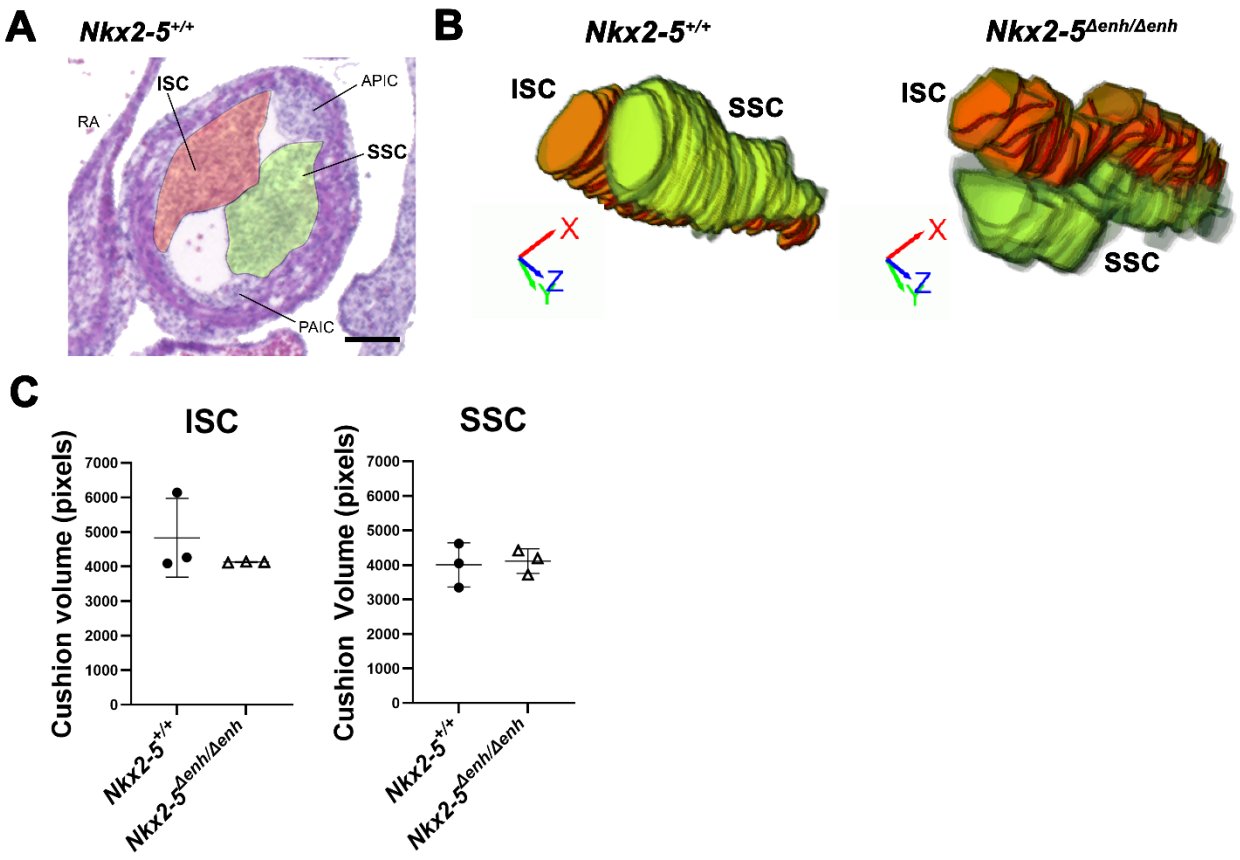
Supplemental Table S1.

Phenotype of outflow tract defects	Number (%)
Persistent truncus arteriosus	8 (44)
Double-outlet right ventricle	3 (17)
Transposition of the great arteries	5 (28)
Hypoplastic pulmonary artery (tetralogy of Fallot like)	2 (11)

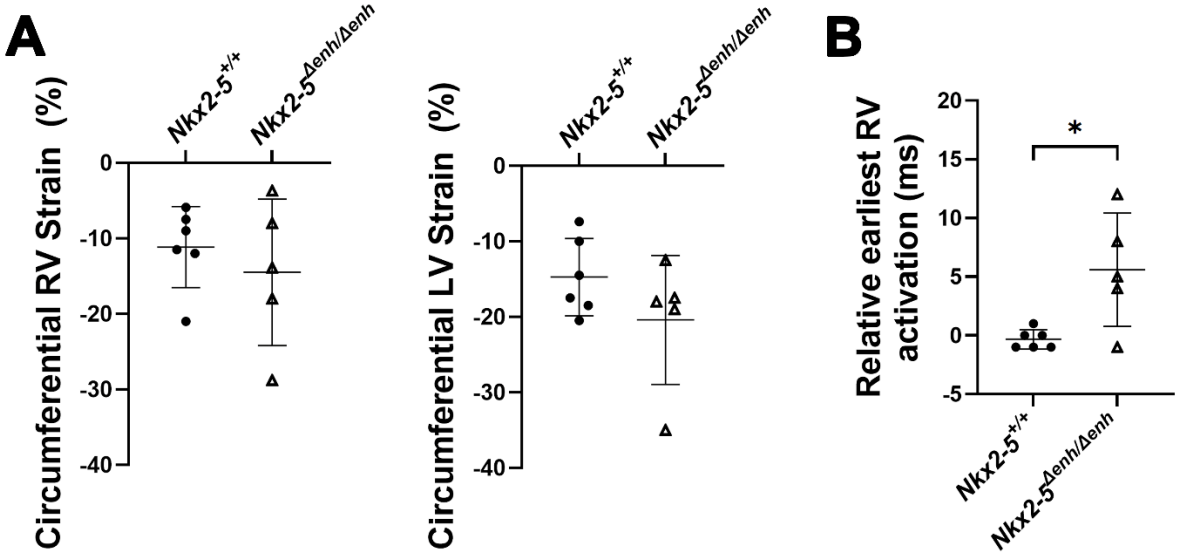
Distribution of outflow tract defects in *Nkx2-5*$^{\Delta enh/\Delta enh}$ mice. n = 18 from 11 consecutive litters. Phenotype was diagnosed by whole-mount image with/without histology.

Supplemental Figures

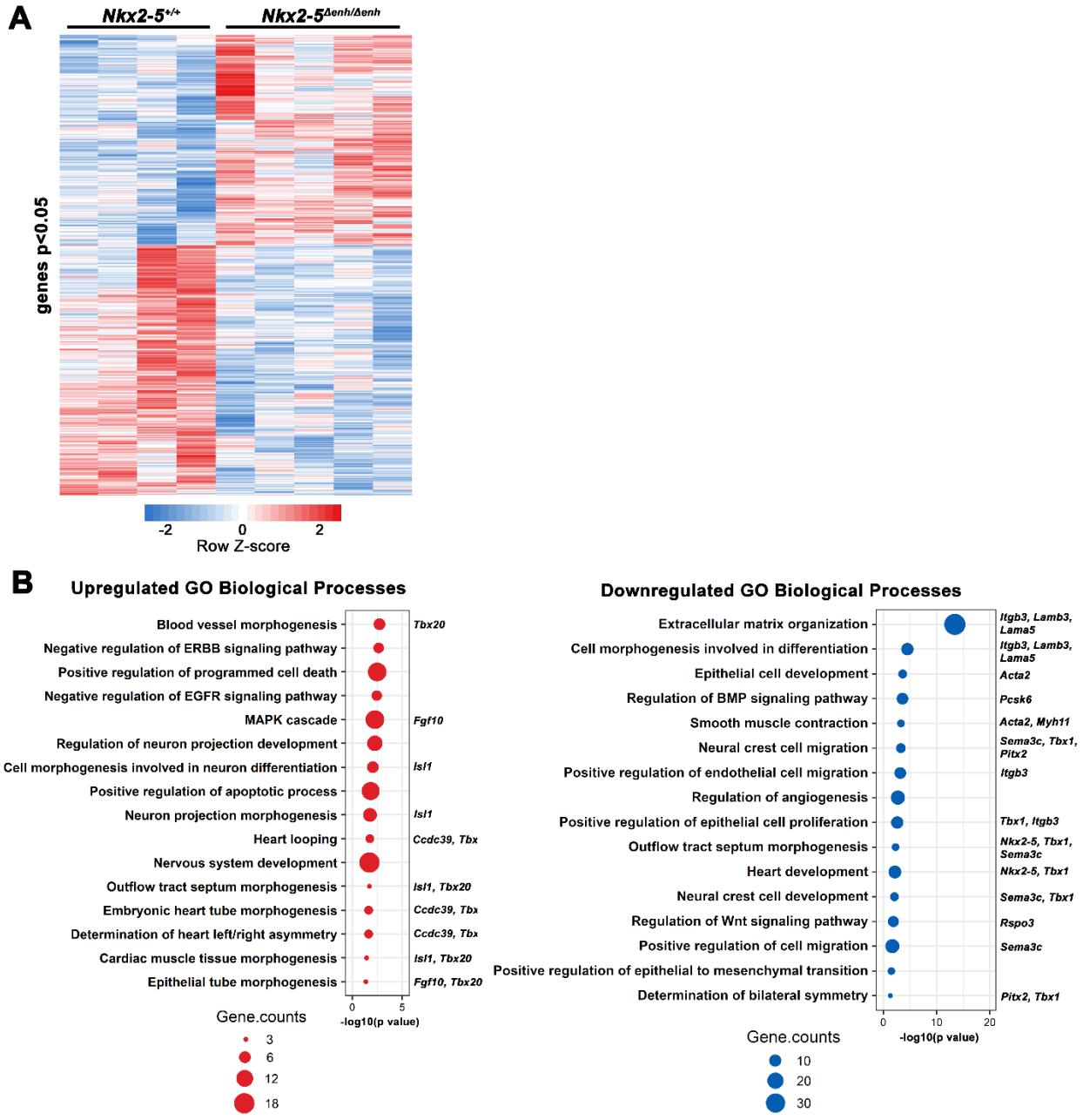
Supplemental Figure S1. Quantification of inferior and superior septal conotruncal cushions in the developing OFT from *Nkx2-5*$^{\Delta enh/\Delta enh}$ and wildtype E11.5 mice. (A) Representative images of conotruncal cushion traces on hematoxylin and eosin stained sections from an E11.5 wildtype mouse. ISC, inferior septal cushion; SSC, superior septal cushion; RA, right atrium; APIC, anterior pulmonary intercalated cushion; PAIC, posterior aortic intercalated cushion. Scale bar: 100 μm. (B) Representative 3D-reconstructed images of conotruncal cushions. The *Nkx2-5*$^{\Delta enh/\Delta enh}$ SSC/ISC axis is under-counterclockwise rotated compared to wildtype littermate control. 3-dimensional axes are X: ventral, Y: left, Z: cranial. (C) Comparisons of cushion volume between *Nkx2-5*$^{\Delta enh/\Delta enh}$ and wildtype OFT shows no significant difference between groups by Student's *t*-test. n = 3 per group. Values show mean ± SD.



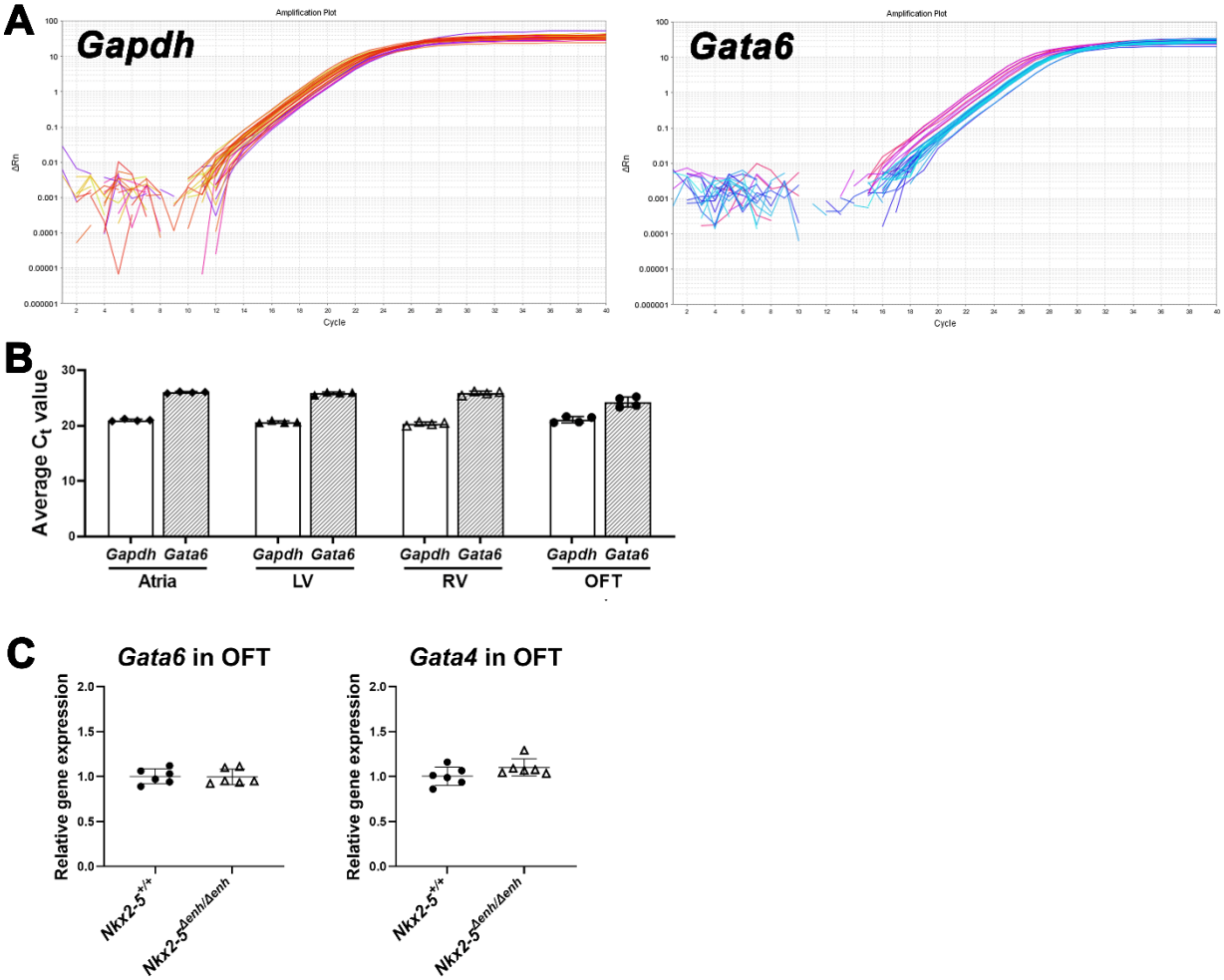
Supplemental Figure S2. Functional assessment of ventricular contraction and electrical activation in the ventricles. (A) Circumferential right (RV) and left ventricular (LV) strain showing no significant difference in systolic function between *Nkx2-5*$^{\Delta enh/\Delta enh}$ and wildtype hearts at postnatal day 1. Student's *t*-test was used for comparison between groups. Values show mean ± SD. n = 5–6 per group. (B) Difference in earliest RV activation relative to earliest LV activation during sinus rhythm obtained from optical mapping at P1. Relative RV initial activation showing delay in *Nkx2-5*$^{\Delta enh/\Delta enh}$ hearts by Student's *t*-test. **P* < 0.05. Values show mean ± SD. n = 5–6 per group.



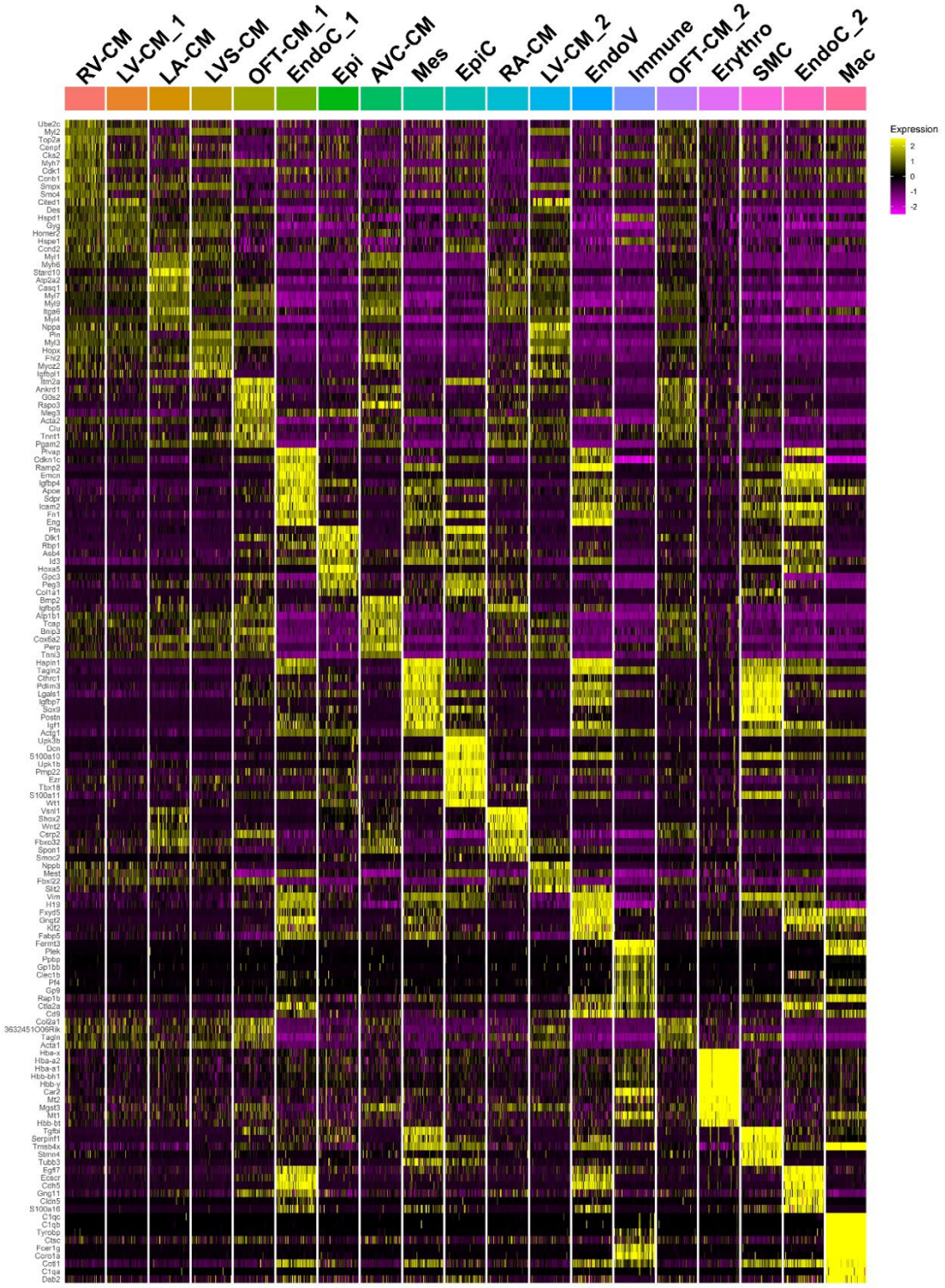
Supplemental Figure S3. Candidate DEGs of embryonic outflow tract (OFT) in *Nkx2-5*$^{\Delta enh/\Delta enh}$ vs. wildtype mice. (A) Heatmap of candidate differentially expressed genes (DEGs, p < 0.05 and basemean > 30) between *Nkx2-5*$^{\Delta enh/\Delta enh}$ vs. wildtype OFT (n = 5 and 4, respectively). (B) Dot plots of curated upregulated (red) and downregulated (blue GO Biological Processes analyzed by Enrichr (p value < 0.05). All candidate DEGs from panel A were included. Genes of interest are shown next to each biological process.



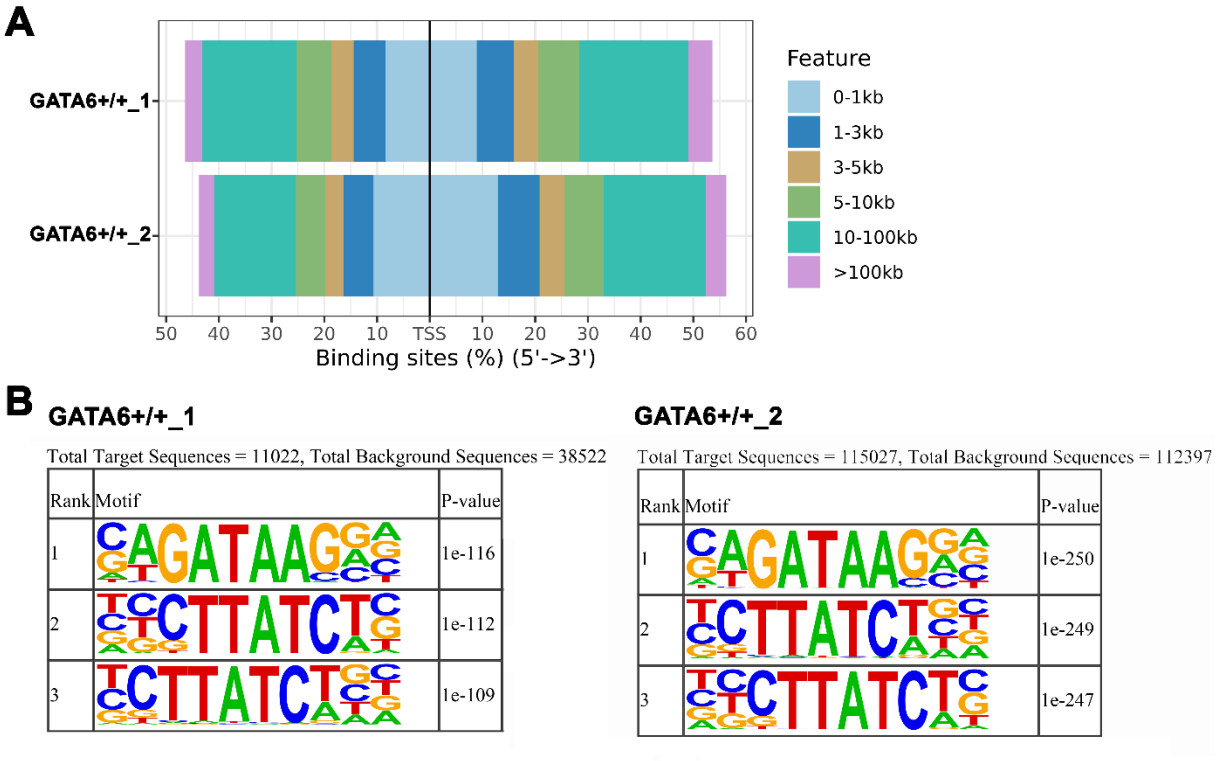
Supplemental Figure S4. RT-qPCR of *Gata6* in regional heart tissues from E12.5 mice. (A) Representative amplification plots of *Gapdh* and *Gata6*. (B) Average C_t values of *Gapdh* and *Gata6* in different cardiac regions from wildtype hearts. n = 4 per group. LV, left ventricle; RV, right ventricle; OFT, outflow tract. (C) *Gata6* and *Gata4* expression in the OFT shows no significant difference between *Nkx2-5*$^{\Delta enh/\Delta enh}$ and wildtype mice. n = 6 per group. Student's *t*-test for *Gata6* and Mann-Whitney U test for *Gata4* expression were performed.



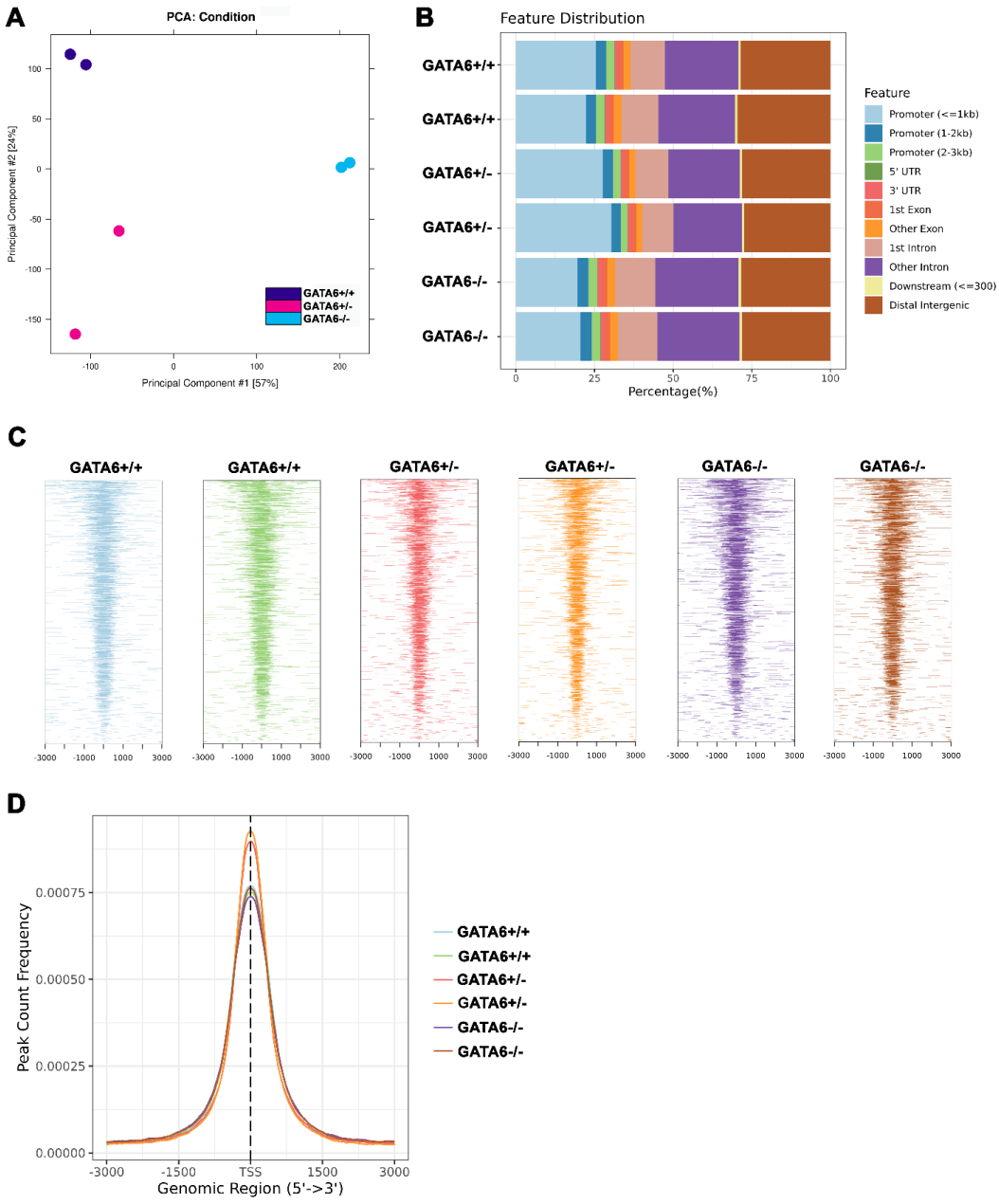
Supplemental Figure S5. Clustering of the scRNA-seq dataset from embryonic *Pitx2* KO hearts. Heatmap of top 10 marker genes for clustering in scRNA-seq from E10.5 control and *Pitx2*$^{hd-/-}$ cardiac tissue reanalyzed from a previously published dataset.[30] 19 subclusters are sorted into 16 clusters in Figure 6. Abbreviations are defined in Figure 6.



Supplemental Figure S6. Overview of GATA6 ChIP-seq dataset for advanced analysis. A previously published dataset was reanalyzed.[29] Day 4 wildtype (GATA6+/+) human induced pluripotent stem cell-derived cardiomyocytes were processed for GATA6 ChIP. (A) Bar plots of the distribution of binding sites upstream and downstream from the transcription start site (TSS) of the nearest genes. (B) HOMER-identified enrichment of known motifs in each sample shows GATA motifs are top-ranked.



Supplemental Figure S7. Overview of ATAC-seq dataset and the accessible chromatin region for advanced analysis. A previously published dataset was reanalyzed.[29] (A) PCA analysis between ATAC-seq samples reveals that biological replicates clustered together based on genotype. (B) Location of ATAC-seq peaks within genomic features. (C) Heatmaps of ATAC-seq from each sample shows peaks within ± 3 kb of proximal promoter regions from transcription start site (TSS). (D) Mean peak count frequency of ATAC-seq peaks at TSS reveals the enrichment around TSS regions.



Supplemental Files

Supplemetal File S1. Downregulated GO biological processes in the *Nkx2-5$^{\Delta enh/\Delta enh}$* OFT

Supplemetal File S2. Upregulated GO biological processes in the *Nkx2-5$^{\Delta enh/\Delta enh}$* OFT

Supplemetal File S3. Gene expression markers for all identity classes in *Pitx2$^{hd-/-}$* scRNA-seq dataset

Supplemetal File S4. Differentially expressed genes in *Pitx2$^{hd-/-}$* versus control in OFT cardiomyocyte cluster