SUPPLEMENTAL MATERIAL
Materials and Methods

Animal studies

All animal studies were approved by and performed under the guidelines of the Institutional Animal Care and Use Committee of the Children's Hospital of Philadelphia (CHOP). Mice were maintained in a temperature- and light-controlled environment with *ad libitum* access to water. Mice in holding cages (after weaning at around 28 days of age) received a standard chow diet (Lab Diet 5L0D); nurturing moms and their pups before weaning received a breeder diet (Lab Diet 5058). All mice were backcrossed at least six generations to and maintained in the C57BL6/J background. The first day we observed new pups born was deemed as P0 and the pups were toe-clipped for identification and genotyping. Both male and female pups were included in the study. The number of mice used in each experiment and number of times experiments are replicated are described in figure legends or related methods section.

Isolation and purification of nuclei from control and KO mice

Control and KO mouse hearts (postnatal 6 or 9-10 days) were rapidly dissected on ice 10am – 12pm of the day. Ventricles were freshly processed or flash frozen in liquid nitrogen and subsequently kept at -80°C before nuclear isolation. Nuclei were isolated and purified as previously described with some modifications (Hu et al. 2017). Specifically, tissues were dounced 15 times with loose pestle and then 10 times with tight pestle. For nuclei purification, samples were ultracentrifuged at 25,000 rpm at 4 °C for 2 hours.

Single-nucleus RNA-Seq library preparation and sequencing

sNucDrop-Seq was performed to make single-nucleus RNA-Seq library (Hu et al. 2017). We
loaded 2.0 pM of library for Nextseq 500 sequencing. The sequencing configuration was 20 bp (Read1), 8 bp (Index1), and 60 bp (Read2). In total, six P10 mouse ventricle samples including 3 control and 3 littermate KO were analyzed with sNucDrop-Seq in three sequencing runs, three P6 control mouse ventricles were analyzed with sNucDrop-Seq in one sequencing run.

*Read mapping, Clustering and Marker gene identification*

Paired-end sequencing reads of sNucDrop-seq were processed as previous described (Hu et al. 2017). Briefly, after mapping the reads to the mouse genome (mm10, Gencode release vM13), both exonic and intronic reads mapped to the predicted strands of annotated genes were retrieved for the cell type classification. Uniquely mapped reads were grouped by cell barcode. To digitally count gene transcripts, a list of UMIs in each gene, within each nucleus, was assembled, and UMIs within ED = 1 were merged together. The total number of unique UMI sequences was counted, and this number was reported as the number of transcripts of that gene for a given nucleus. Raw digital expression matrices were generated for the 3 sequencing runs. The UMI count of mitochondrial genes was considered as contamination and thus was discarded.

The raw digital expression matrices were combined and loaded into the R package Seurat (v 2.0.1) (Butler and Satija 2017). For normalization, UMI counts for all nuclei were scaled by library size (total UMI counts), multiplied by 10,000 and transformed to log space. Only genes found to be expressed in >10 nuclei were retained. Moreover, nuclei with fewer than 500 or more than 4,000 detected genes were omitted. As a result, 4560, 7760 and 7323 nuclei were kept in control P6, control P10 and KO P10 mice, respectively.

For clustering of control mice, the highly variable genes were identified using the function `FindVariableGenes` with the parameters: `x.low.cutoff = 0.0125, x.high.cutoff = 3` and
y.cutoff = 0.8 in Seurat, resulting in 1,952 highly variable genes. Cell cycle effect was assessed by function \textit{CellCycleScoring}. Cell cycle effect and batch effect were regressed out using the function \textit{RegressOut} in Seurat. Before clustering, we identify the 38 animal-specific genes among 3 biology replicates by function \textit{FindMarkers} with the \textit{thresh.use} parameter setting to 0.5. Then 12 ribosomal protein genes and 14 animal-specific genes were excluded from highly variable gene set, resulting in remaining 1,926 genes as input for principal component (PC) analysis (PCA). The expression level of 1,926 genes in the nuclei was scaled and centered along each gene, and was conducted to the PCA. We selected different cut-offs of the number of PCs and empirically found that downstream clustering analyses were optimized when using 20-PCs cutoff. First 20 PCs were selected and used for two-dimension t-distributed stochastic neighbor embedding (tSNE), implemented by the Seurat software with default parameters. Based on the tSNE map, 14 clusters were identified using the function \textit{FindCluster} in Seurat with the resolution parameter of 1.0. As a result of these steps, we were able to assign 7,760 nuclei into 14 cell type clusters. To perform clustering analysis on 7,323 KO nuclei, we identified 1,831 highly variable genes. After excluding 12 ribosomal protein genes and 32 animal-specific genes, the remaining 1,787 highly variable genes were loaded to the PCA. The 7,323 KO nuclei were classified into 13 clusters by the first 25 PCs with the resolution parameter set to 1.6. For clustering analysis on 4,560 control P6 nuclei, 2,698 highly variable genes were loaded to the PCA. The 4,560 P6 nuclei were classified into 10 clusters by the first 15 PCs with the resolution parameter set to 1.

To identify marker genes, differential expression analysis was performed by the function \textit{FindAllMarkers} in Seurat with likelihood-ratio test. Differentially expressed genes that were expressed in at least 10% cells within the cluster and with a fold change more than 0.5 (log scale)
were considered to be marker genes. In total, 2,334 marker genes were identified for 14 control clusters, while 1,738 marker genes were identified for 13 KO clusters. The cluster dendrogram of clusters and heat map of marker genes were generated as previous described (Hu et al. 2017).

**Comparison of control cluster to KO clusters**

To examine how closely the cell types identified in KO nuclei matched the corresponding cell types defined in control nuclei, we computed the pairwise Pearson correlation coefficients between each pair of cell types in control and KO. We first combined control and KO dataset and load them to the Seurat package. Then function `FindVariableGenes` was used to identify top 2,000 highly variable genes within the combined dataset. For each cell type cluster, the average natural log scaled UMI counts of these top 2,000 highly variable genes were used to generate the cell-type-specific gene expression matrix. R function cor.test was used to calculate the pairwise Pearson correlation coefficients.

**Identification of differentially expressed genes between control and KO nuclei**

Differential gene expression analysis within cell type between control and KO nuclei was performed using the function `FindMarkers` in Seurat, using a likelihood ratio test. Genes with a fold-change of more than 0.25 (log(e) scale) and a $P$-value less than 0.01 were considered to be differentially expressed.

**Identification of enriched transcriptional regulators in cell type-specific Gdf15+ nuclei**

For the differential expression test of transcriptional regulators between $Gdf15^+$ nuclei and $Gdf15^-$ nuclei, only exonic reads were used to generate the digital expression matrix of nUMI
count. Nuclei with more than 300 detected genes, more than 400 nUMI, and less than 4,000 detected genes were kept for the downstream analysis. To this end, 6,108 control nuclei and 7,300 KO nuclei were passed and loaded into the Seurat. Differential gene expression analysis of transcriptional regulator within each cell type between control and KO nuclei was performed using the function FindMarkers in Seurat, using Standard AUC classifier test with the setting: test.use=“roc”. We performed two comparisons to identify the enriched transcriptional regulators in Gdf15+ nuclei of a specific cell type: KO Gdf15+ nuclei versus KO Gdf15- nuclei; and KO Gdf15+ nuclei versus control Gdf15- nuclei. Transcriptional regulators with a higher expression level in Gdf15+ nuclei (AUC>0.5) in both comparisons were kept and were subjected into gene regulatory network construction. Protein-protein interaction among transcriptional regulators was retrieved from the literature based on the TRRUST database(Han et al. 2018). The illustration of network was constructed using Cytoscape software (v 3.5.1) (Shannon et al. 2003). To further evaluate the co-expressed pattern between transcriptional regulators and Gdf15, we calculate the Pearson correlation coefficients between a transcriptional regulator and Gdf15 in the nuclei expressing both the transcriptional regulator and Gdf15. Transcription factor binding motifs in the upstream sequences of Gdf15 were identified using FIMO (Grant et al. 2011) which is available as part of MEME Suite of motif-based sequence analysis tools (http://meme.nbcr.net). H3K4me1 and H3K27Ac Chip-seq data of mouse heart (postnatal 0 day) was downloaded from Encode (Consortium 2012).

*GO/KEGG enrichment analysis*
The GO annotations were downloaded from the Ensembl database and KEGG annotations were retrieved by KEGG API. An enrichment analysis was performed via a hypergeometric test. The P value was calculated using the following formula:

\[ P = 1 - \sum_{i=0}^{m-1} \frac{\binom{M}{i}\binom{N-M}{n-i}}{\binom{N}{n}}, \]

where \( N \) is the total number of genes, \( n \) is the total number of marker genes, \( M \) is the number of genes annotated to a certain GO term or KEGG pathway, and \( i \) is the number of marker genes annotated to a certain GO term or KEGG pathway. P value was corrected by function `p.adjust` with false discovery rate (FDR) correction in R. GO terms or KEGG pathways with FDR below 0.05 were considered enriched. Statistical calculations were performed in R. Alternatively pathway analysis was performed using metascape.org (July 2018).

**Whole heart RNA-Seq**

RNA-Seq of hearts from 16-day-old littermate control and KO was as previously described (Wang et al. 2017). The data were deposited in the GEO database (accession number GSE88761).

**Gene expression analysis**

We isolated total RNA from mouse tissues or HL-1 cells using RNAzol RT (Molecular Research Center) following the manufacturer’s instructions. We synthesized cDNA from 500 ng total RNA using iScript cDNA synthesis kit (Bio-Rad) and quantified mRNA levels by real-time qPCR using SYBR Green (Bio-Rad). We ran samples in technical triplicates and calculated relative mRNA levels using a standard curve and normalized to \( 36b4 \) mRNA level in the same sample.
**Immunohistochemistry, immunofluorescence and quantification of GDF15+ cells**

PECAM1 (Santa Cruz sc-46694), NPR3 (Santa Cruz sc-515449) and GDF15 (Abcam ab189358) immunohistochemistry were performed on paraffin sections of 10-day-old littermate control and KO mouse heart ventricles as previously described (Wang et al. 2017). Total number of nuclei in the whole-heart cross section was counted using Image J. GDF15+ cells were counted manually under microscope. TNNI3 (Abcam ab47003), PECAM1 (BD Biosciences 550274), NPR3 (Santa Cruz sc-515449) and Ki67 (Vector Laboratories vp-RM04) immunofluorescent staining were performed on cryosections of 6 or 10-day-old WT mice. Fluorescent 2\(^{nd}\) antibodies were from Jackson ImmunoResearch Laboratories, and Hoechst 33342 was from Cayman Chemical (15547).

**Sirius red staining**

Sirius red staining was done in 10-day-old littermate control and KO hearts using the Picrosirius Red Stain Kit (Polysciences) following the manufacture’s instruction, and quantified using Image J.

**Adenovirus construction and injection**

Recombinant *Gata4* adenovirus was constructed using the AdEasy Adenoviral Vector System (Agilent) as previously described (Pei et al. 2006; Pei et al. 2011). Pericardial injection of adenovirus was performed in 3-day-old C57BL6/J WT mice using a Hamilton syringe as previously described (Wang et al. 2017). Mice were randomized to receive $5 \times 10^7$ infectious units Ad-GFP (control) or Ad-Gata4. Hearts and blood were collected 5 days after injection for
further analysis. As quality control and based upon pre-established criteria, mice that showed no cardiac Gfp or Gata4 overexpression (presumably due to unsuccessful injection or ineffective adenovirus infection) were excluded from the study.

ChIP and transfection

ChIP was performed in HL1 cardiomyocytes as previously described (Wang et al. 2015; Zhao et al. 2018). The GATA4 antibody used was from Santa Cruz (sc-25310x). The ChIP primers used were: GATTTCCCTCTCCCTCCATTCC and ATGACTGGATGCTCACACG. The mouse Gdf15 promoter region (−261→+303, relative to the transcription start site) containing a GATA binding site (−109) was inserted into pGL4.23 luciferase reporter using KpnI and XhoI sites. The GATA binding site was mutated from GATA to CCTA as previously described (Pei et al. 2006). Transfection and luciferase assays were performed as previously described (Wang et al. 2015).

Statistical analysis

Fishers’ exact test, Chi-square test, Correlation test and Student’s t-test were performed in R to determine the statistical significance, with p<0.05 deemed as statistically significant.
References

Butler A, Satija R. 2017. Integrated analysis of single cell transcriptomic data across conditions, technologies, and species. bioRxiv.

Consortium EP. 2012. An integrated encyclopedia of DNA elements in the human genome. Nature 489: 57-74.

Grant CE, Bailey TL, Noble WS. 2011. FIMO: scanning for occurrences of a given motif. Bioinformatics 27: 1017-1018.

Han H, Cho JW, Lee S, Yun A, Kim H, Bae D, Yang S, Kim CY, Lee M, Kim E et al. 2018. TRRUST v2: an expanded reference database of human and mouse transcriptional regulatory interactions. Nucleic Acids Res 46: D380-D386.

Hu P, Fabyanic E, Kwon DY, Tang S, Zhou Z, Wu H. 2017. Dissecting Cell-Type Composition and Activity-Dependent Transcriptional State in Mammalian Brains by Massively Parallel Single-Nucleus RNA-Seq. Mol Cell 68: 1006-1015 e1007.

Pei L, Leblanc M, Barish G, Atkins A, Nofsinger R, Whyte J, Gold D, He M, Kawamura K, Li HR et al. 2011. Thyroid hormone receptor repression is linked to type I pneumocyte-associated respiratory distress syndrome. Nat Med 17: 1466-1472.

Pei L, Waki H, Vaitheesvaran B, Wilpitz DC, Kurland JJ, Tontonoz P. 2006. NR4A orphan nuclear receptors are transcriptional regulators of hepatic glucose metabolism. Nat Med 12: 1048-1055.

Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. 2003. Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. Genome Research 13: 2498-2504.
Wang T, Liu J, McDonald C, Lupino K, Zhai X, Wilkins BJ, Hakonarson H, Pei L. 2017.
GDF15 is a heart-derived hormone that regulates body growth. *EMBO Mol Med* **9**: 1150-1164.

Wang T, McDonald C, Petrenko NB, Leblanc M, Giguere V, Evans RM, Patel VV, Pei L. 2015.
Estrogen-related receptor alpha (ERRalpha) and ERRgamma are essential coordinators of cardiac metabolism and function. *Mol Cell Biol* **35**: 1281-1298.

Zhao J, Lupino K, Wilkins BJ, Qiu C, Liu J, Omura Y, Allred AL, McDonald C, Susztak K, Barish GD et al. 2018. Genomic integration of ERRgamma-HNF1beta regulates renal bioenergetics and prevents chronic kidney disease. *Proc Natl Acad Sci U S A* **115**: E4910-E4919.
Figure S1. Validation of sNucDrop-Seq for mouse hearts.
(A) Brief work flow chart of the sNucDrop-Seq procedure for mouse heart snRNA-Seq.
(B) Violin plots illustrating number of transcripts (UMIs) and unique genes detected by sNucDrop-Seq from P10 control and KO hearts. Center line: median; circle: mean; limits: first and third quartile; whiskers, ±1.5 Inter quartile range (IQR).
(C) Reproducibility among biological replicates. Left panel shows gene expression correlation between samples with p value labeled. Right panel shows clustering results based on transcriptome similarity.
(D) Scatter plot showing the high correlation of expression levels of KO/control between sNucDrop-Seq and bulk RNA-Seq.
(E) Scatter plot showing Nppa and Cox7a1 expression at the single-nucleus resolution. Relative mean expression value and percentage of Nppa+ or Cox7a1+ nuclei are labeled.
Figure S2. Clusters and their transcriptional signatures in control hearts.
(A) Numbers of detected genes (nGene) and transcripts (nUMIs) in all clusters.
(B) Distribution of nGene and nUMI in the tSNE plot.
(C) Heatmap showing cluster-specific non-coding RNA in control hearts.
(D) Violin plot showing cluster-specific non-coding RNA in control hearts.
mCM: mature cardiomyocytes; pCM: proliferating cardiomyocytes; dCM: developing cardiomyocytes; LEC: lymphatic endothelial cells; Fb: fibroblasts; Epi: epicardial cells; BC: blood cells; EC: endothelial cells; PC/SMC: pericytes/smooth muscle cells.
Figure S3. Cellular heterogeneity in the postnatal developing heart. 
(A) Violin plot showing EC1 or EC2 enriched genes. 
(B) Immunostaining of PECAM1 and NPR3 in ventricles (n=2). Arrows: endocardium/trabeculae; Asterisks: major arteries or veins. Scale bar is 100 µm. 
(C) Confocal microscopy of PECAM1 (green) and NRP3 (red) immunostaining. 
(D) Confocal microscopy of PECAM1 (green) and TNNI3 (red) immunostaining. 
(E) Violin plot showing epicardial specific genes. 
(F) Violin plots show that Piezo2, Shank3 and reelin are highly and specifically expressed in lymphatic endothelial cells. 
mCM: mature cardiomyocytes; pCM: proliferating cardiomyocytes; dCM: developing cardiomyocytes; LEC: lymphatic endothelial cells; Fb: fibroblasts; Epi: epicardial cells; BC: blood cells; EC: endothelial cells; PC/SMC: pericytes/smooth muscle cells.
Figure S4. Cluster-specific non-coding RNA in KO hearts.

(A) Numbers of detected genes and transcripts (UMIs) in all clusters.
(B) Heatmap showing expression of cluster-specific non-coding RNA in KO hearts.
(C) Violin plot showing expression of cluster-specific non-coding RNA in KO hearts.

mCM: mature cardiomyocytes; dCM: developing cardiomyocytes; aFb: activated myofibroblasts; Fb: fibroblasts; Epi: epicardial cells; BC: blood cells; EC: endothelial cells; PC/SMC: pericytes/smooth muscle cells.
Figure S5. EC1 cell markers in control and KO hearts.
(A) Immunostain of pan-EC (PECAM1) and EC1-specific markers (NPR3) in control and KO hearts. Scale bar is 100 µm.
(B) Violin plot showing enriched expression of Npr3, Tie1 and Tie2 in EC1.
Figure S6. Expression correlation between Gdf15 and cell type-specific transcriptional regulators. Y axis is the relative expression of Gdf15. X axis is the relative expression of the transcriptional regulator specified.