

METHODS AND RESOURCES

Integrated transcriptomics and epigenomics reveal chamber-specific and species-specific characteristics of human and mouse hearts

Junpeng Gao^{1,2}, Yuxuan Zheng^{1,2,3}, Lin Li⁴, Minjie Lu⁵, Xiangjian Chen⁶, Yu Wang⁷, Yanna Li⁸, Xiaomeng Liu⁹, Yun Gao^{1,2}, Yunuo Mao^{1,2}, Peng Zhao¹⁰, Jinan Zhang⁶, Fuchou Tang^{1,2,3}, Lei Song^{5,11}*, Lu Wen^{1,2}*, Jizheng Wang⁵*

1 Beijing Advanced Innovation Center for Genomics, School of Life Sciences, Peking University, Beijing, China, **2** Biomedical Pioneering Innovation Center, Ministry of Education Key Laboratory of Cell Proliferation and Differentiation, Beijing, China, **3** Peking-Tsinghua Center for Life Sciences, Academy for Advanced Interdisciplinary Studies, Peking University, Beijing, China, **4** Guangdong Provincial Key Laboratory of Proteomics, Department of Pathophysiology, School of Basic Medical Sciences, Southern Medical University, Guangzhou, China, **5** State Key Laboratory of Cardiovascular Disease, Fuwai Hospital, National Center for Cardiovascular Diseases, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China, **6** Department of Cardiology, First Affiliated Hospital of Nanjing Medical University, Nanjing, China, **7** Department of Ultrasound, Xiangyang No. 1 People's Hospital, Hubei University of Medicine, Xiangyang, China, **8** Department of Obstetric and Gynecology, Beijing Anzhen Hospital, Capital Medical University, Beijing, China, **9** Sichuan Cancer Hospital & Institute, Sichuan Cancer Center, School of Medicine, University of Electronic Science and Technology of China, Radiation Oncology Key Laboratory Of Sichuan Province, Chengdu, China, **10** Department of Pathology, The Affiliated Hospital of Qingdao University, Qingdao, China, **11** Clinical Research Center for Cardiovascular Diseases, Fuwai Hospital, National Center for Cardiovascular Diseases, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China

☞ These authors contributed equally to this work.

* songlqd@126.com (LS); wenlu@pku.edu.cn (LW); jzwang@hotmail.com (JW)



OPEN ACCESS

Citation: Gao J, Zheng Y, Li L, Lu M, Chen X, Wang Y, et al. (2021) Integrated transcriptomics and epigenomics reveal chamber-specific and species-specific characteristics of human and mouse hearts. *PLoS Biol* 19(5): e3001229. <https://doi.org/10.1371/journal.pbio.3001229>

Academic Editor: Cecilia W. Lo, University of Pittsburgh, UNITED STATES

Received: September 10, 2020

Accepted: April 12, 2021

Published: May 18, 2021

Copyright: © 2021 Gao et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: The authors confirm that all data underlying the findings are fully available without restriction. The raw NOME-seq and RNA-seq data of human have been deposited in Genome Sequence Archive (GSA) for Human under the accession code HRA000365. The raw NOME-seq and RNA-seq data of mouse have been deposited in GSA under CRA003798. The processed NOME-seq data of human have been deposited in OMix (<https://bigd.big.ac.cn/omix/>) under the ID OMIX247, and the processed RNA-seq data of human and BAM files of 176 novel

Abstract

DNA methylation, chromatin accessibility, and gene expression represent different levels information in biological process, but a comprehensive multiomics analysis of the mammalian heart is lacking. Here, we applied nucleosome occupancy and methylome sequencing, which detected DNA methylation and chromatin accessibility simultaneously, as well as RNA-seq, for multiomics analysis of the 4 chambers of adult and fetal human hearts, and adult mouse hearts. Our results showed conserved region-specific patterns in the mammalian heart at transcriptome and DNA methylation level. Adult and fetal human hearts showed distinct features in DNA methylome, chromatin accessibility, and transcriptome. Novel long noncoding RNAs were identified in the human heart, and the gene expression profiles of major cardiovascular diseases associated genes were displayed. Furthermore, cross-species comparisons revealed human-specific and mouse-specific differentially expressed genes between the atria and ventricles. We also reported the relationship among multiomics and found there was a bell-shaped relationship between gene-body methylation and expression in the human heart. In general, our study provided comprehensive spatiotemporal and evolutionary insights into the regulation of gene expression in the heart.

lncRNAs have been deposited in OMix under OMIX243. The processed NOME-seq and RNA-seq data of mouse have been deposited in OMix under OMIX245 and OMIX242, respectively. All data mentioned above can be accessed in GSA under BioProject ID PRJCA003611.

Funding: This work was supported by grants from the National Key R&D Program of China (2017YFA0103402 to L.W.), the CAMS Innovation Fund for Medical Sciences (CAMS-I2M, 2016-I2M-1-015 to J.W.), the National Natural Science Foundation of China (81870286 to L.S.) and the Beijing Municipal Natural Science Foundation (7181008 to L.S.). This work was also supported by the Beijing Advanced Innovation Center for Genomics at Peking University. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Abbreviations: AF, atrial fibrillation; ATAC-seq, assay for transposase-accessible chromatin using sequencing; CGI, CpG island; CHD, congenital heart disease; CPC, Coding Potential Calculator; DEG, differentially expressed gene; DMR, differentially methylated region; DMW, differentially methylated window; DNase-Seq, DNase sequencing; FDR, false discovery rate; GO, Gene Ontology; HCM, hypertrophic cardiomyopathy; LA, left atrium; lncRNA, long noncoding RNA; LV, the left ventricle; MNase-seq, micrococcal nuclease sequencing; NDR, nucleosome-depleted region; NOME-seq, nucleosome occupancy and methylome sequencing; PBAT, post-bisulfite adaptor tagging; PCA, principal component analysis; PVDF, polyvinylidene fluoride; qRT-PCR, quantitative reverse transcription PCR; RA, right atrium; RPKM, reads per kilobase per million; RRBS, reduced representation bisulfite sequencing; RV, right ventricle; TBS, Tris-buffered saline; TES, transcription end site; TF, transcription factor; TSS, transcription start site; WGBS, whole genome bisulfite sequencing.

Introduction

The heart is the first organ formed during mammalian embryogenesis [1] and is composed of 4 chambers: the left atrium (LA), the right atrium (RA), the left ventricle (LV), and the right ventricle (RV) [2]. The atria and the ventricles play different roles, and all 4 chambers exhibit distinct properties with respect to structure, biochemistry, and electrochemistry [3,4]. With the rapid development of high-throughput sequencing technologies, genomic studies have emerged to reveal the molecular characteristics of the mammalian heart [5–13]. For example, region-resolved transcriptional and quantitative proteomic maps of the human heart have been reported in recent years [14–19]. However, all these studies have focused on individual omics, especially transcriptomics. DNA methylation and other epigenetic aspects, which are essential for cardiogenesis and cardiovascular disease, have not been sufficiently investigated. In addition to this, a few studies seek to resolve cardiac physiology and pathology at the multiomics level [20,21]. But a comprehensive multiomics analysis comparing DNA methylation, chromatin accessibility, and transcriptome at one time has not been reported for 4 chambers of mammalian hearts.

Previous studies have revealed the fetal origins of cardiovascular disease [22,23]. Heart development is a complex process that is controlled by interacting molecular pathways [24]. The formation of all 4 chambers is completed before the ninth week of human embryogenesis [25]. Current research has mainly concentrated on the process of embryonic heart development before birth or on regional differences in the adult human heart [26]. Because of their similarity to humans in physiology and pathology, mice are the most commonly used model organisms in human biology and disease research [27]. However, the chamber-specific characteristics and the species differences between human and mouse hearts have not yet been comprehensively investigated, although these data could support decision-making in drug discovery.

In the past few years, a large number of techniques have arisen to map the DNA methylome and chromatin accessibility at the whole-genome scale. For example, whole genome bisulfite sequencing (WGBS), post-bisulfite adaptor tagging (PBAT), and reduced representation bisulfite sequencing (RRBS) have been widely used to investigate the state of DNA methylation at single-base resolution [28–30]. DNase sequencing (DNase-Seq), micrococcal nuclease sequencing (MNase-seq), and assay for transposase-accessible chromatin using sequencing (ATAC-seq) are several ways to identify chromatin organization [31–33]. Here, we performed multiomics analyses of the 4 chambers of both human and mouse hearts using nucleosome occupancy and methylome sequencing (NOME-seq), which simultaneously detects DNA methylation and chromatin accessibility [34], as well as RNA-seq. Our results revealed chamber-specific and species-specific characteristics of gene expression, DNA methylation, and chromatin accessibility of human and mouse hearts. The resulting multiomics map of healthy human and mouse hearts can be used as a reference to identify novel biomarkers or drug targets when compared with malfunctioning hearts.

Results

Multiomics analyses of human and mouse hearts

To obtain an integrated multiomics map of mammalian hearts, we collected 11 adult human hearts from healthy donors who contributed for other reasons instead of cardiovascular disease and 3 fetal human hearts from aborted fetuses. We also harvested hearts from 3 groups of male C57BL/6 mice (10 weeks old) with 5 mice in each group. Four chambers, i.e., the LA, RA, LV, and RV, were separately analyzed for each sample.

We performed a modified version of NOME-seq for the heart samples with 2 replicates for each sample. Ribosomal RNA-depleted RNA-seq was also applied for profiling the transcriptome (Fig 1A, see Methods). We sequenced an average of 34.26 Gb raw data for each NOME-seq library and 8.48 Gb raw data for each RNA-seq library (S1 Table). In total, 4.14 Tb sequencing data were obtained. The results of the technical replicates were highly reproducible, as shown by the DNA methylation patterns around the gene-body region and the chromatin accessibility patterns around the transcription start site (TSS) (S1A and S1B Fig). In human and mouse samples, the median efficiencies of GpC methyltransferase and bisulfite conversion were 92.65% and 95.93%, respectively, as assessed by spiked lambda DNA. These results together indicated that we successfully applied NOME-seq to the heart tissues, and the data were reliable for subsequent analysis.

First, the total levels of DNA methylation and chromatin accessibility were compared among the 4 chambers of human and mouse hearts. The DNA methylation and chromatin accessibility were calculated using the methylation levels of cytosine in the WCG (ACG/TCG) and GCH (GCA/GCT/GCC) sites, respectively. The results showed that the total DNA methylation levels were comparable among the different chambers of adult and fetal human hearts and adult mouse hearts (Fig 1B–1E, S1C and S1D Fig). The average methylation levels were 74% for humans and 69% for mice, consistent with previous reports [35,36]. The total chromatin accessibility levels were also largely similar among the different chambers of human and mouse hearts (Fig 1E, S1C and S1D Fig). In addition, both the DNA methylation levels and the chromatin accessibility levels were similar between fetal and adult human hearts (Fig 1B and 1C).

Comparative analysis of the four chambers of the adult human heart

We first analyzed the transcriptomes of the 4 chambers of the adult human heart. Notably, principal component analysis (PCA) clearly distinguished between the atria and the ventricles (Fig 2A). However, individual atrium (LA, RA) samples or ventricle (LV, RV) samples were almost indistinguishable. A substantial number of protein-coding genes were differentially expressed in the human atria and ventricles (S2 Table). A total of 808 differentially expressed genes (DEGs) highly expressed in the atria were mainly relevant to the Gene Ontology (GO) terms including cell adhesion, extracellular matrix organization, or ion transport. In addition, the 381 DEGs highly expressed in the ventricles were enriched in the GO terms including oxidation–reduction process, blood circulation, and heart contraction (Fig 2B). Long noncoding RNAs (lncRNAs) have emerged as important regulators of cardiac development and homeostasis [37–40]. Our results showed that PCA based on lncRNA transcriptome data also distinguished between the atria and the ventricles (Fig 2C). We identified 124 and 63 lncRNAs with significantly higher expression in the atria and ventricles, respectively (S2A Fig, S2 Table). Pearson analysis found a statistically significant correlation between these lncRNAs and protein-coding genes compared with random pairs (Fig 2D). Then, we identified 19 and 5 gene pairs that showed *cis*-regulatory relationships between lncRNAs and protein-coding genes in the atria and the ventricles, respectively (Fig 2E). Among these gene pairs, *NKX2-6* is known to play a key role in cardiac development, and its mutation has been linked to familial atrial fibrillation [41,42]; *KCNJ3* is a causative gene of hereditary bradyarrhythmias [43]; *MYL2* has been reported to be connected to hypertrophic cardiomyopathy [44,45]; and *HEY2* participates in left ventricular maturation [46]. Among the lncRNAs in these gene pairs, *CRNDE*, *CHL1-AS2*, and *TBX5-AS1* have been associated with different types of cancers [47–49]. The roles of the lncRNAs in these gene pairs in cardiac development and homeostasis remain to be further investigated.

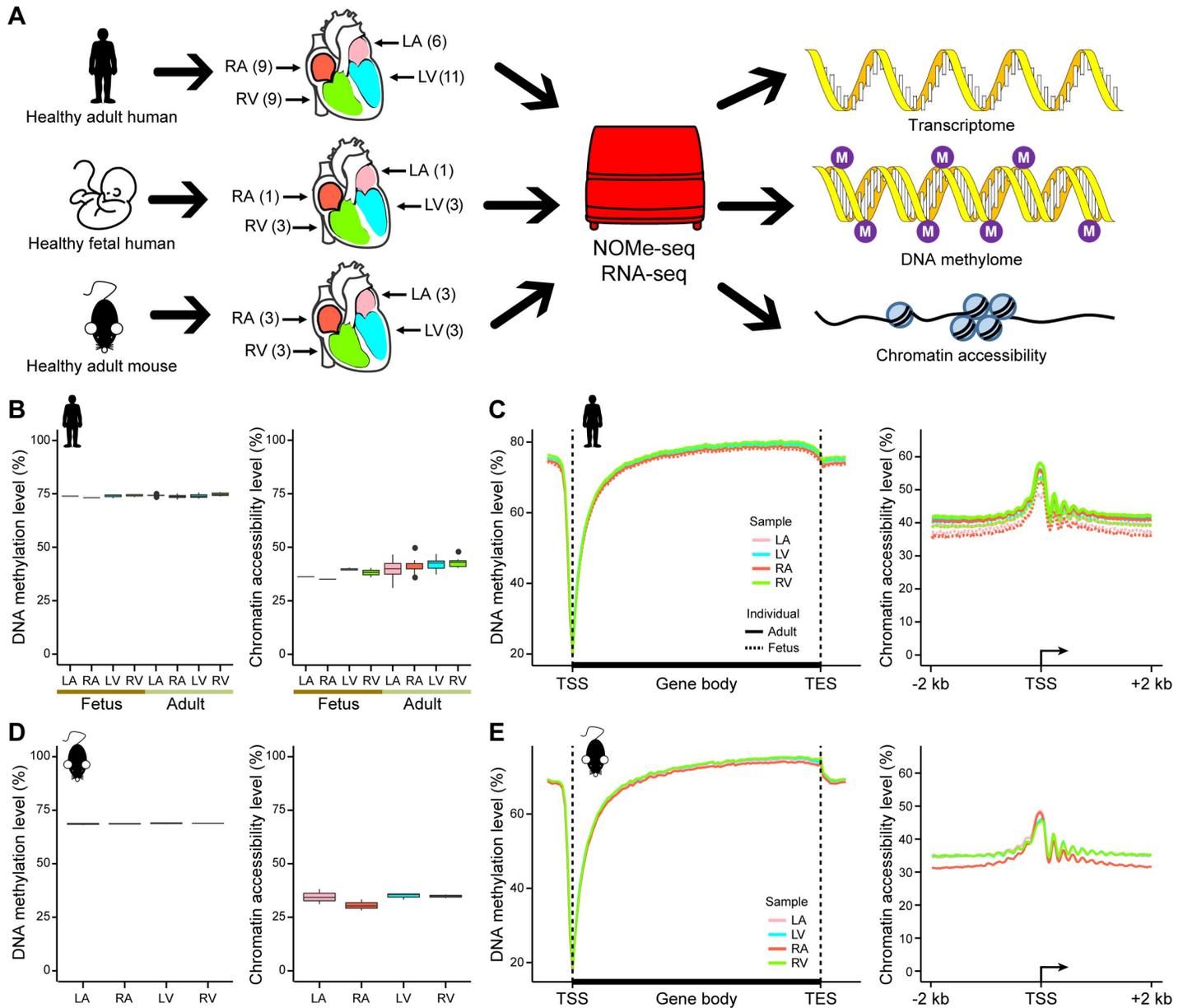


Fig 1. Global features of DNA methylation and chromatin accessibility in human and mouse hearts. (A) Experimental design, including heart sources and schematic depiction of the multiomics analysis. Numbers in brackets refer to the numbers of hearts used in the study. (B) Boxplots showing the endogenous DNA methylation level (left) and chromatin accessibility levels (right) in the human heart. The endogenous DNA methylation level was calculated using the DNA methylation level of WCG sites (ACG and TCG trinucleotides), and the chromatin accessibility level was calculated using the DNA methylation level of GCH sites (GCA, GCT, and GCC trinucleotides). (C) Line plots showing the average endogenous DNA methylation level around the gene body ± 5 kb region (left) and the average chromatin accessibility levels around the TSS ± 2 kb region (right) in the human heart. Different colors indicate different anatomic regions, and the solid line and dashed line indicate adult samples and fetal samples, respectively. (D) Boxplots showing the endogenous DNA methylation level (left) and the chromatin accessibility levels (right) in the mouse heart. (E) Line plots showing the average endogenous DNA methylation level around the gene body ± 5 kb regions (left) and the average chromatin accessibility levels around the TSS ± 2 kb regions (right) in the mouse heart. Different colors indicate different anatomic regions. The raw data for B–E can be found in [S1 Data](#). LA, left atrium; LV, left ventricle; NOMe-seq, nucleosome occupancy and methylome sequencing; RA, right atrium; RV, right ventricle; TES, transcription end site; TSS, transcription start site.

<https://doi.org/10.1371/journal.pbio.3001229.g001>

In our study, to assess the methylation levels of over 16 million WCG sites, we divided the whole genome into 1-kb bins and, notably, found that the atria and the ventricles could be clearly distinguished (Fig 2F). Then, differentially methylated regions (DMRs) between the

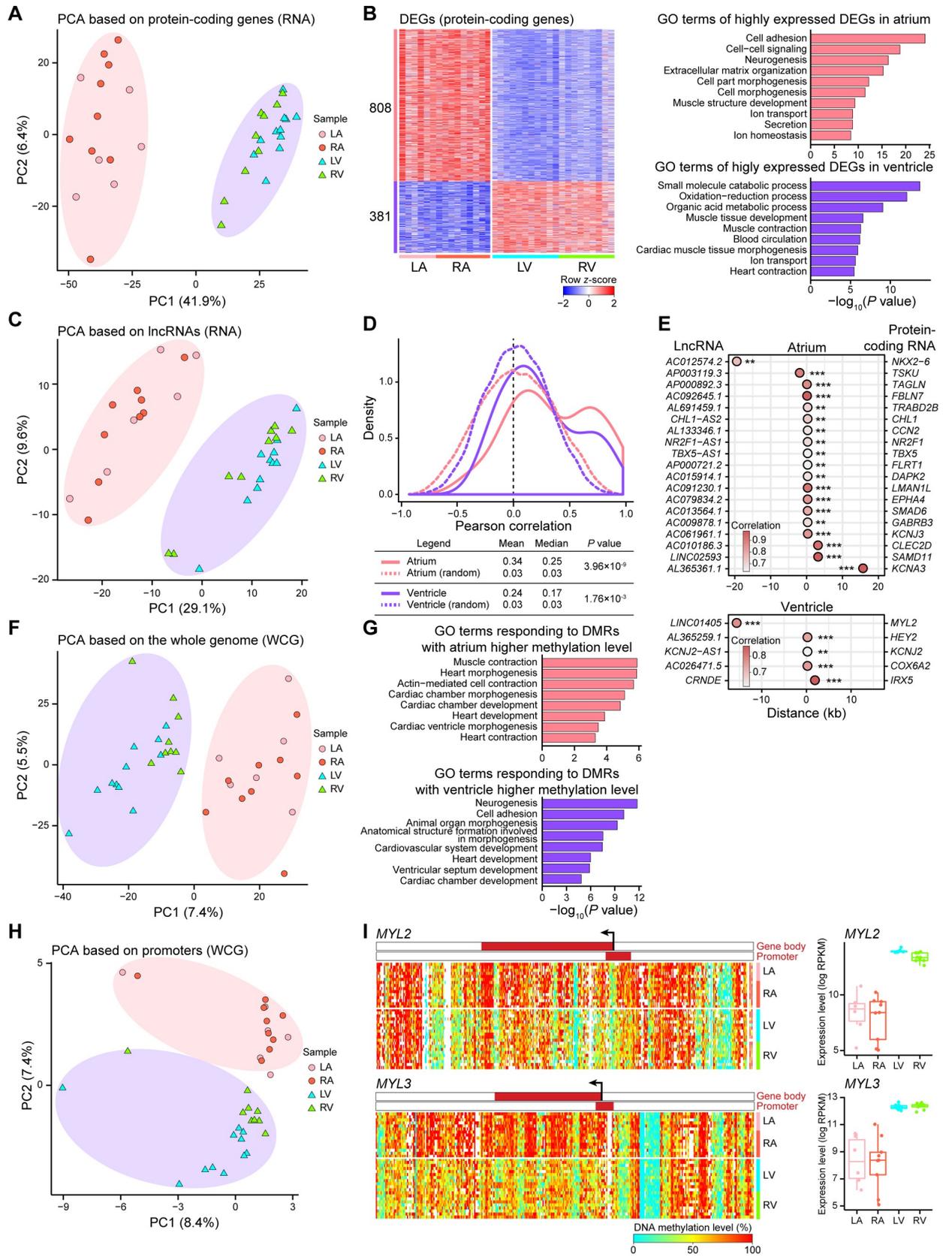


Fig 2. Regional differences in transcriptome and endogenous DNA methylation profiles in the adult human heart. (A) PCA plot showing the transcriptome pattern of protein-coding genes in the adult human heart. Circles indicate atrium samples, and triangles indicate ventricle samples. The variation values of PC1 and PC2 were 41.9% and 6.4%, respectively. (B) Heatmap showing row z -score scaled gene expression levels of DEGs (protein-coding genes) between atria and ventricles (left), and corresponding GO terms are shown (right). The number of DEGs is reported on the left. (C) PCA plot showing the transcriptome pattern of lncRNAs in the adult human heart. Circles indicate atrium samples, and triangles indicate ventricle samples. The variation values of PC1 and PC2 were 29.1% and 9.6%, respectively. (D) Line plot showing the density distribution of the Pearson correlation coefficient between the paired protein-coding genes and lncRNAs (top). Solid lines indicate the real distribution, dashed lines indicate the simulated distribution, and pink lines and purple lines correspond to the atria and ventricles, respectively. The mean and median Pearson correlation coefficients are reported, and two-tailed Student t test P values between the real and simulated distributions are indicated (bottom). (E) Dot plots showing gene pairs between lncRNAs (left) and protein-coding genes (right) in *cis*-regulation analysis in the atria (top) and the ventricles (bottom). The x-axis indicates the linear distance between the lncRNA and the protein-coding genes. Dot colors indicate the Pearson correlation coefficient, and stars indicate the correlation significance. ** P value ≤ 0.01 and *** P value ≤ 0.001 . (F) PCA plot showing the endogenous DNA methylation pattern of the whole genome in the adult human heart. Circles indicate atrium samples, and triangles indicate ventricle samples. The variation values of PC1 and PC2 were 7.4% and 5.5%, respectively. (G) GO terms of genes corresponding to gene bodies in which DMRs were located. Top, GO terms corresponding to DMRs with higher endogenous DNA methylation levels in the atria; bottom, GO terms corresponding to DMRs with higher endogenous DNA methylation levels in the ventricles. (H) PCA plot showing the endogenous DNA methylation pattern of promoters in the adult human heart. Circles indicate atrium samples, and triangles indicate ventricle samples. The variation values of PC1 and PC2 were 8.4% and 7.4%, respectively. (I) Heatmaps showing endogenous DNA methylation levels around the gene body ± 10 kb regions (left) and boxplots showing the gene expression levels (right) of *MYL2* and *MYL3* in the adult human heart. The color bars in the heatmaps indicate the gene-body regions and the promoter regions (from 1 kb upstream of the TSS to 0.5 kb downstream of the TSS). Gene expression levels were quantified with $\log_2(\text{RPKM} + 1)$. The raw data for A–I can be found in [S1 Data](#). DEGs, differentially expressed genes; DMRs, differentially methylated regions; GO, Gene Ontology; LA, left atrium; lncRNAs, long noncoding RNAs; LV, left ventricle; PCA, principal component analysis; RA, right atrium; RPKM, reads per kilobase per million; RV, right ventricle; TSS, transcription start site.

<https://doi.org/10.1371/journal.pbio.3001229.g002>

human atria and ventricles were identified ([S3 Table](#)). There were 278 DMRs hypermethylated in the atria, and GO analysis showed that the genes associated with these DMRs were enriched in the GO terms such as contraction and cardiac chamber morphogenesis; 1,199 DMRs were hypermethylated in the ventricles, and they were highly enriched in the GO terms such as cell adhesion and heart development ([Fig 2G](#)). While these DMRs tended to occur in the intergenic regions, a number of DMRs were located in promoter and CpG island (CGI) regions ([S2B Fig](#)). PCA based on DNA methylation levels in both the promoter and CGI regions distinguished the atria and ventricles ([Fig 2H](#), [S2C Fig](#)). Interestingly, the methylation levels in both the promoters and gene-body regions of *MYL2* and *MYL3* were notably lower in the ventricles than in the atria ([Fig 2I](#)). The gene body of *MYH7* also showed lower methylation levels in the ventricles ([S2D Fig](#)). These 3 genes are all typical ventricle marker genes, and their mutations can lead to hypertrophic cardiomyopathy [[45,50,51](#)]. Three genes highly expressed in the atria, *GJA5*, *SMAD6*, and *SMAD7*, showed an opposite methylation pattern, with lower DNA methylation levels in their gene bodies in the atria than in the ventricles ([S2D Fig](#)). These genes also contribute to heart development or functions [[52–54](#)]. Thus, our results identified a negative correlation between gene expression and gene-body methylation in several key heart genes.

Chromatin accessibility is a prerequisite for transcription factor (TF) binding and can regulate gene transcription [[55](#)]. To map open chromatin regions, we identified 766,486 distal nucleosome-depleted regions (NDRs) and 33,546 proximal NDRs in the human heart. Based on the chromatin accessibility level of the distal or proximal NDRs, the 4 chambers of the human heart could not be distinguished from each other ([S2E Fig](#)), suggesting that there were no significant differences in the overall chromatin accessibility in the 4 chambers of the heart.

Comparison of fetal and adult human hearts

Then, we compared the fetal and adult human heart. PCA using transcriptome data of both protein-coding genes and known lncRNAs showed that the PC1 axis separated the samples by developmental stage and the PC2 axis distinguished between the atria and the ventricles ([Fig 3A](#), [S3A Fig](#)). We identified 176 novel lncRNAs in the human heart ([S4 Table](#), see [Methods](#)).

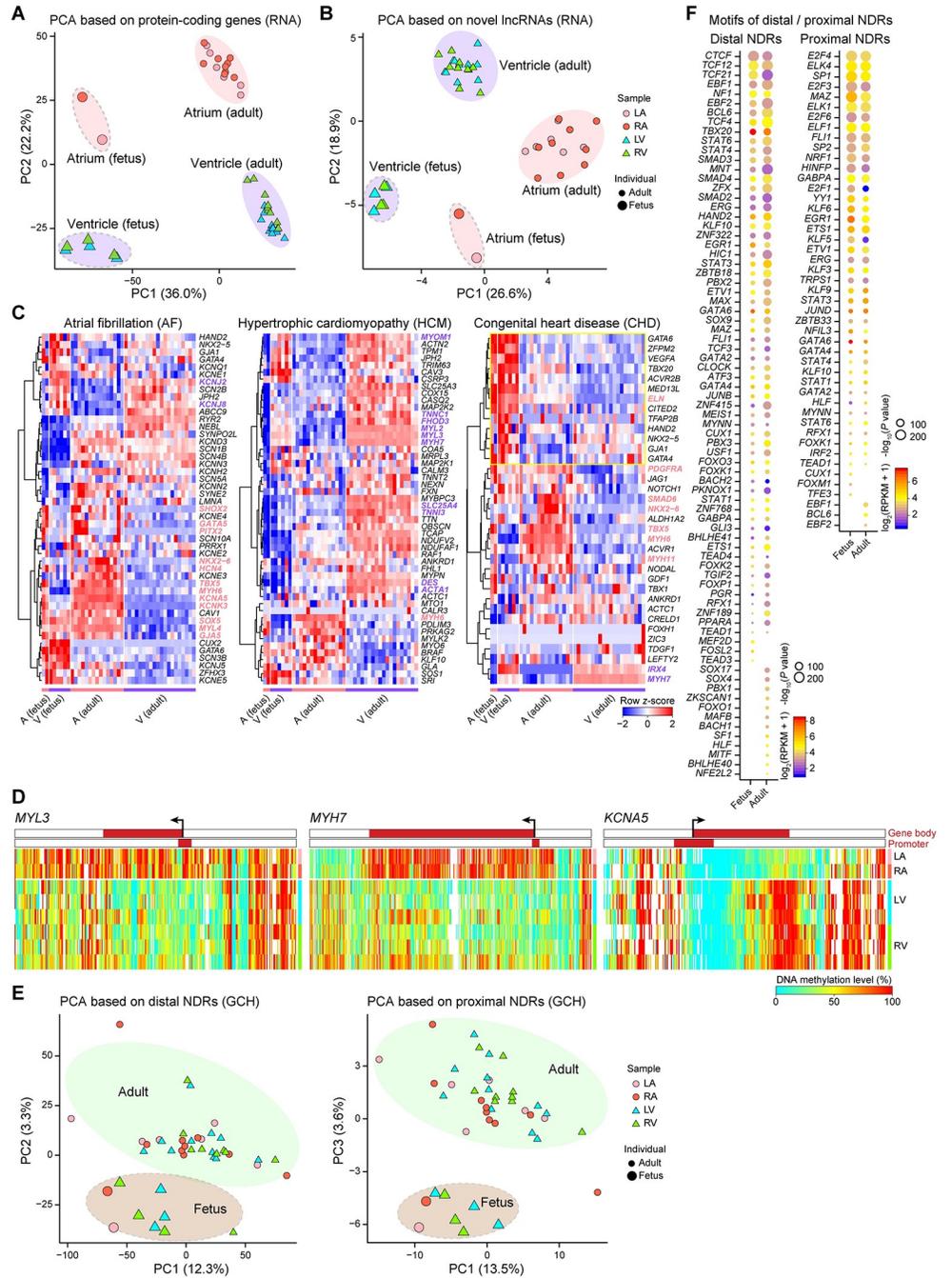


Fig 3. Distinct features of the transcriptome, endogenous DNA methylation, and chromatin accessibility in the human heart between fetuses and adults. (A) PCA plot showing the transcriptome pattern of protein-coding genes in the human heart. Circles and triangles indicate the atria and ventricles, respectively; point sizes indicate the adult heart and fetal heart. The variation values of PC1 and PC2 were 36.0% and 22.2%, respectively. (B) PCA plot showing the transcriptome pattern of novel lncRNAs in the human heart. Circles and triangles indicate the atria and ventricles, respectively; point sizes indicate the adult heart and fetal heart. The variation values of PC1 and PC2 were 26.6% and 18.9%, respectively. (C) Heatmaps showing z-score scaled expression levels of genes associated with heart diseases, such as atrial fibrillation (left), hypertrophic cardiomyopathy (middle), and congenital heart disease (right) in the rows. DEGs between the atria and ventricles among these genes are indicated by pink and purple. (D) Heatmaps showing endogenous DNA methylation levels in the gene body \pm 10 kb regions of *MYL3* (left), *MYH7* (middle), and *KCNA5* (right) in the human fetal heart. The color bars in the heatmaps indicate the gene-body regions and the promoter regions (from 1 kb upstream of the TSS to 0.5 kb downstream of the TSS). (E) PCA plot showing the chromatin accessibility pattern of distal (left) and proximal (right) NDRs in the human heart. Circles and triangles indicate the

atria and ventricles, respectively; point sizes indicate the adult heart and fetal heart. (F) Motif enrichment analysis of distal (left) and proximal (right) NDRs in the human heart. Colors indicate average expression levels, and sizes indicate the P values ($P \leq 10^{-10}$) of the corresponding transcription factors. The raw data for A–F can be found in [S1 Data](#). A, atrium; AF, atrial fibrillation; CHD, congenital heart disease; DEGs, differentially expressed genes; HCM, hypertrophic cardiomyopathy; LA, left atrium; lncRNAs, long noncoding RNAs; LV, left ventricle; NDRs, nucleosome-depleted regions; PCA, principal component analysis; RA, right atrium; RPKM, reads per kilobase per million; RV, right ventricle; TSS, transcription start site; V, ventricle.

<https://doi.org/10.1371/journal.pbio.3001229.g003>

The distance from these novel lncRNAs to the nearest known transcripts was 33 kb, and their median nearest distance from each other was 283 kb ([S3B Fig](#)). Interestingly, PCA using transcriptome data of these 176 novel lncRNAs could distinguish between the atria and the ventricles of both fetal and adult human hearts, suggesting that they may play roles in heart development ([Fig 3B](#)).

Cardiovascular disease is the top killer of humans [56]. Atrial fibrillation (AF), hypertrophic cardiomyopathy (HCM), and congenital heart disease (CHD) are common cardiovascular diseases, and genes associated with these diseases have been uncovered in recent years [57–59]. Here, we displayed the expression patterns of AF-associated genes (47 genes), HCM-associated genes (50 genes), and CHD-associated genes (35 genes) in fetal and adult human hearts ([Fig 3C](#)). AF is characterized by abnormal electrical activity of the atrium, and HCM is characterized by thickening of the left ventricular wall [60,61]. We then investigated the overlap distribution between highly expressed DEGs in atria/ventricles and disease-associated genes. We found that there were 14 DEGs overlapped with AF-associated genes, including 12 (85.7%) DEGs were highly expressed in the atria (i.e., *SHOX2*, *GATA2*, and *PITX2*). And there were 11 DEGs overlapped with HCM-associated genes, including 10 (90.9%) DEGs were highly expressed in the ventricles (i.e., *MYOM1*, *TNNC1*, and *DES*). The percentage of DEGs up-regulated in the corresponding chambers implied that chamber-specific expression patterns contributed to disease occurrence and progression ([S3C Fig](#)). CHD is a congenital cardiovascular disease related to abnormal heart structure and function [62]. Thirteen CHD-associated genes showed relatively higher expression in fetal hearts (yellow box highlight in [Fig 3C](#)). The regional and temporal specificity of cardiovascular disease-associated genes demonstrated that the physiology and pathology of the human heart are under precise spatiotemporal expression and regulation. To reveal potential core genes in cardiovascular diseases, we also constructed coexpression networks of disease-associated genes ([S3D Fig](#)). Coexpression networks showed several key genes in these cardiovascular diseases, for example, *KCNK3*, *TNNI3*, and *ELN*.

Considering the DNA methylome, the adult and fetal hearts could be clearly distinguished by the whole genome, promoters, or CGI data. The DNA methylation levels of the whole genome, but not those of promoters or CGI, could distinguish between the atria and the ventricles of fetal hearts ([S3E Fig](#)). The gene-body methylation levels of *MYL2*, *MYL3*, and *MYH7* were lower in the fetal ventricles than in the fetal atria, similar to those of the adult heart ([Fig 3D](#)). The gene-body methylation level of *KCNA5* was significantly higher in the ventricles than in the atria of fetal hearts, while no difference was found in the adult hearts ([Fig 3D](#)).

We identified 886,426 distal NDRs and 34,700 proximal NDRs by combining the adult and fetal human heart data. Notably, PCA using the chromatin accessibility of the proximal or distal NDRs clearly distinguished between fetal and adult hearts ([Fig 3E](#)). The enriched TF motifs of the distal and proximal NDRs of fetal and adult hearts were compared ([S5 Table](#)). The fetal and adult human heart shared most of the enriched distal motifs, such as the GATA factors (*GATA2*, *GATA4*, and *GATA6*), STAT proteins (*STAT1*, *STAT3*, *STAT4*, and *STAT6*), and *PBX2/3*, though not *PBX1* ([Fig 3F](#)). The binding motifs of a few TFs, such as *MITF*, which has been shown to play a role in cardiac growth and hypertrophy [63], displayed adult-specific

enrichment (Fig 3F). The proximal NDRs in both fetal and adult hearts were strongly enriched for the motifs of TFs related to basic transcription regulation.

Comparison of human and mouse hearts

Next, we performed a comparative analysis of human and mouse hearts. DEGs between the atria and the ventricles of the murine heart were enriched in GO terms such as cell adhesion, heart development, and cardiac muscle contraction, which were highly similar to those of the human heart (S4A Fig, S6 Table). The human and mouse hearts shared 230 DEGs with higher expression levels in the atria, including *NPPA* and *ROR2*, and 98 DEGs with higher expression levels in the ventricles, including *MYL2* and *MYL3* (Fig 4A, S6 Table). We further investigated species-specific DEGs that showed over 5-fold changes in average expression levels and were not detected (reads per kilobase per million (RPKM) < 1) in the other species. In total, 199 human-specific DEGs were identified and found to be enriched in the GO terms such as synaptic signaling and neurogenesis (Fig 4B, S6 Table), while 150 mouse-specific DEGs were revealed and shown to be related to ion transport. *STAT1* and *STAT3* have been shown to play roles in ischemic heart disease [64,65], and in our study, we found that a gene in the STAT family (*STAT4*) was a human-specific DEG (Fig 4C). *ANKRD2* was detected in the adult human heart and has been determined to be up-regulated in human dilated cardiomyopathy [66]. The expression patterns of these genes were confirmed by real-time quantitative reverse transcription PCR (qRT-PCR) and western blotting analysis (Fig 4D and 4E). Compared with single-cell RNA-seq data sets of human and mouse hearts [16,67], we found that both human-specific and mouse-specific ventricular DEGs (with higher expression levels in the ventricles) were mainly expressed in cardiomyocyte (S4B and S4C Fig). In the meantime, those species-specific atrial DEGs (with higher expression levels in the atria) were expressed in several cell types, such as fibroblast, neuron, endothelial cell, and cardiomyocyte (S4B and S4C Fig). In mouse hearts, PCA based on DNA methylation of the whole genome, only the promoters or only the CGI regions distinguished the atria from the ventricles, similar to the results in human hearts (S4D Fig). The gene bodies of *Myl2* and *Myl3* were also hypermethylated in the mouse atria compared with the ventricles (Fig 4F). In addition, *Epha4* showed a higher methylation level in the gene body in the ventricles, and this phenomenon was conserved in both mouse and human hearts (Fig 4F, S4E Fig).

TF binding motif enrichment analysis was performed on the NDRs to establish the relationship between TF regulation and open chromatin. To compare motif enrichment related to chromatin accessibility between human and mouse hearts, we identified 313,517 distal NDRs and 25,966 proximal NDRs in the mouse heart to seek motifs for TF binding. The results showed that the distal NDRs of both species were enriched for the binding motifs of TFs such as GATAs (*GATA2*, *GATA4*, and *GATA6*), STATs (*STAT1* and *STAT6*), and PBX1/2/3 (Fig 4G, S7 Table). GATA factors play important roles in heart development and cardiac diseases, and overexpression of *Gata4* or *Gata6* in mouse hearts could lead to cardiac hypertrophy [68,69]. Pbx proteins are necessary for myocardium differentiation [70]. Tcf21 functions in specification of the cardiac fibroblast lineage as a bHLH TF [71] and showed higher expression in the atria. These key TFs formed conserved regulatory networks in both human and mouse hearts. We also identified binding motifs for a few TFs that were enriched solely in either human or mouse hearts. *Mef2* is an early marker of the myocardium lineage [72]. Interestingly, although *MEF2A* (mean expression level RPKM = 46), *MEF2C* (RPKM = 13), and *MEF2D* (RPKM = 29) were expressed in human hearts, the binding motifs of *Mef2a*, *Mef2c*, and *Mef2d* were enriched in only mouse hearts. The proximal NDRs of both human and mouse hearts were enriched in the motifs of TFs such as Sp proteins (*SP1* and *SP2*), the E2F

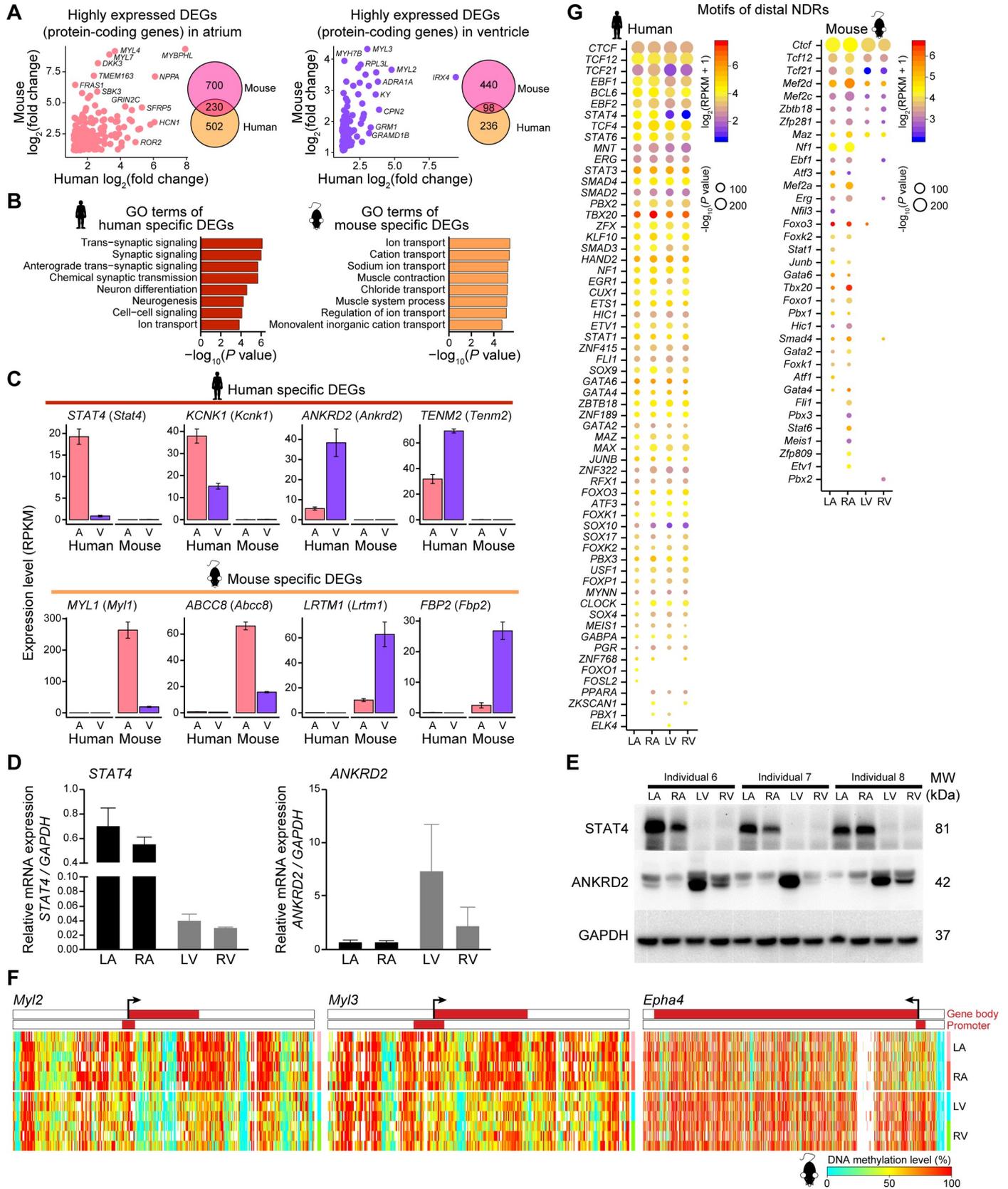


Fig 4. Comparisons of the transcriptome and regulatory elements in the adult human heart and mouse heart. (A) Scatter plots showing fold change of the overlapping DEGs (protein-coding genes) in cross-species comparisons. Venn diagrams showing the number of overlapping DEGs (protein-coding genes) in cross-species comparisons. DEGs were defined as differentially expressed genes with statistical significance between the atria and ventricles. Left, highly expressed DEGs (protein-coding genes) in the atria; right, highly expressed DEGs (protein-coding genes) in the ventricles. (B) GO terms of human- (left) and mouse- (right) specific DEGs. DEGs were defined as in Fig 4A. (C) Bar plots showing the expression levels of human- (top) and mouse- (bottom) specific DEGs. DEGs were defined as in Fig 4A. Data are shown as the mean \pm SEM. (D) Relative mRNA levels of *STAT4* and *ANKRD2* in the 4 chambers of the human heart using quantitative RT-PCR. (E) Protein abundances of *STAT4* and *ANKRD2* in the 4 chambers of the human heart by western blot analysis. (F) Heatmaps showing endogenous DNA methylation levels in the gene body \pm 10 kb regions of *Myl2* (left), *Myl3* (middle), and *Epha4* (right) in mouse hearts. The color bars in the heatmaps indicate the gene-body regions and the promoter regions (from 1 kb upstream of the TSS to 0.5 kb downstream of the TSS). (G) Motif enrichment analysis of distal NDRs in the adult human heart (left) and mouse heart (right). Colors indicate average expression levels, and sizes indicate the P values ($P \leq 10^{-10}$) of the corresponding transcription factors. The raw data for A–D, F, and G can be found in [S1 Data](#). The raw images for E can be found in [S1 Raw images](#). A, atrium; DEGs, differentially expressed genes; GO, Gene Ontology; LA, left atrium; LV, left ventricle; MW, molecular weight; NDRs, nucleosome-depleted regions; RA, right atrium; RPKM, reads per kilobase per million; RT-PCR, reverse transcription PCR; RV, right ventricle; TSS, transcription start site; V, ventricle.

<https://doi.org/10.1371/journal.pbio.3001229.g004>

TF family (E2F4 and E2F6), and Kruppel-like factors (KLF3, KLF6, and KLF9), which are mainly involved in basic transcription regulation ([S4F Fig](#), [S7 Table](#)).

Correlations among DNA methylation, chromatin accessibility, and gene expression

To explore how DNA methylation and chromatin accessibility are correlated with gene expression, all the protein-coding genes were divided into 4 groups: high- (RPKM > 10), intermediate- ($1 < \text{RPKM} \leq 10$), and low-expression genes ($0.1 < \text{RPKM} \leq 1$) and silenced genes ($\text{RPKM} \leq 0.1$). We first observed the average chromatin accessibility levels around the TSS and found that chromatin accessibility was positively correlated with gene expression levels. The high-expression genes showed the highest level of accessibility, and the silenced genes were nearly inaccessible ([Fig 5A](#)). The high-expression lncRNAs also showed higher accessibility around the TSS than the low-expression lncRNAs ([S5A Fig](#)).

Then, we investigated the average DNA methylation levels across the genes. Consistent with previous reports, DNA methylation around the TSS was negatively correlated with gene expression ([Fig 5B](#), [S5B Fig](#)). In promoter regions, we found a clear positive correlation (R1) between the expression levels of protein-coding genes and the corresponding chromatin accessibility level ([S5C Fig](#)). In contrast, a negative correlation (R2) between expression and DNA methylation level was observed.

Interestingly, the results showed that the gene bodies of intermediate-expression genes exhibited the highest methylation level ([Fig 5B](#)). Indeed, gene-body methylation displayed a bell-shaped relationship with gene expression, with high-expression genes showing prominently lower methylation levels than intermediate-expression genes ([Fig 5C](#)). Notably, gene-body chromatin accessibility revealed a reverse bell-shaped relationship with gene expression, indicating that the gene-body chromatin of high-expression genes was generally open in the human heart ([Fig 5D](#)). These patterns of correlations between gene expression and DNA methylation or chromatin accessibility were conserved in mice ([S5D and S5E Fig](#)).

Discussion

This study, to our knowledge, represents the first comprehensive multiomics analysis of all 4 chambers of adult human and mouse hearts and fetal human hearts. The mouse is the most widely used model for studying heart development and cardiac diseases. Although human and mouse hearts are similar in anatomy, they are quite different in size and electrophysiology [[2,73–76](#)]. Additionally, studies have shown that human and mouse hearts respond differently to cardiovascular drugs [[77,78](#)]. Our data showed that the NDRs of human and mouse hearts share many TF motifs, indicating that the main regulatory networks are conserved between

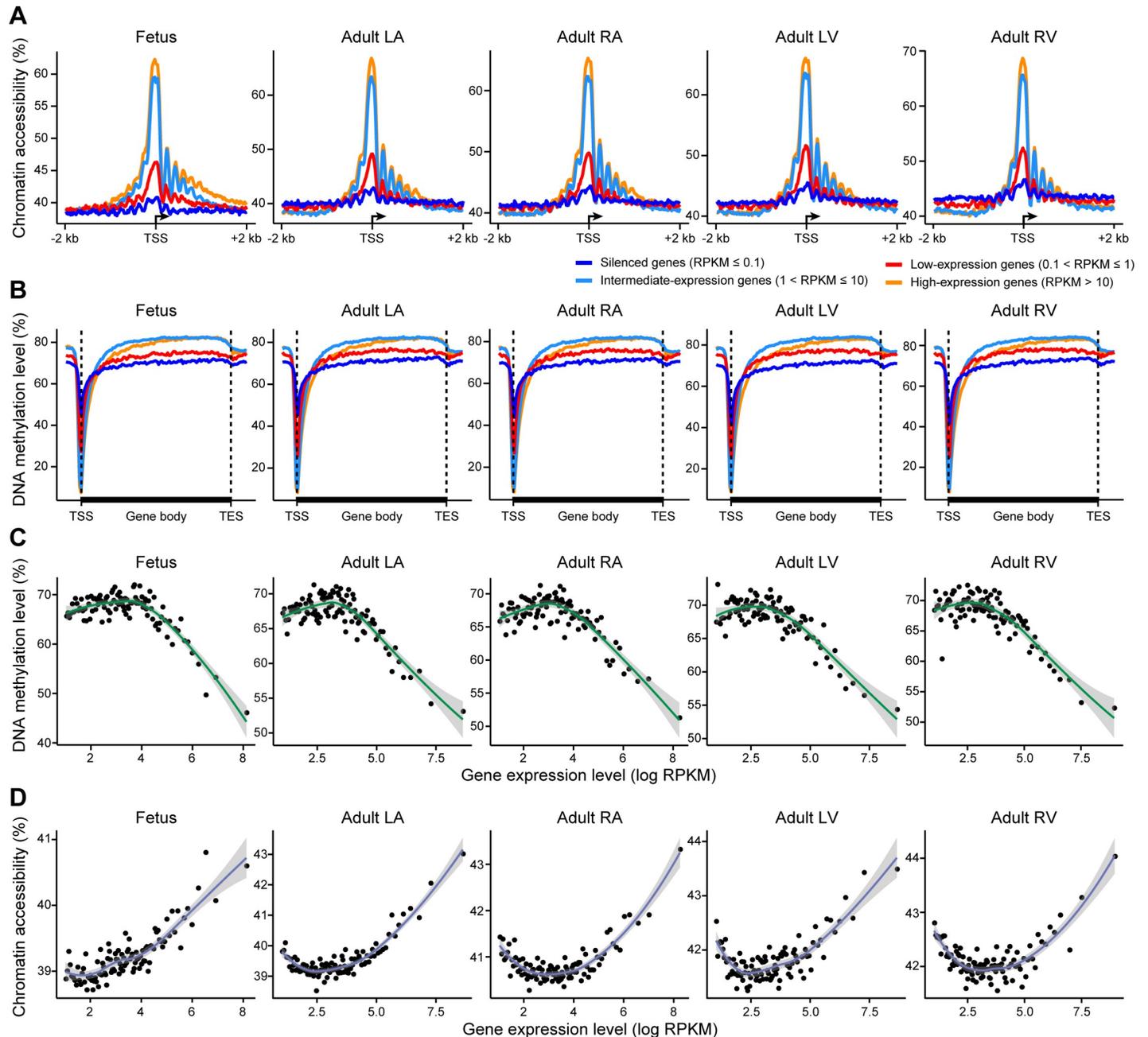


Fig 5. Multiomics relationship patterns in the human heart. (A and B) Line plots showing the average chromatin accessibility level in the TSS \pm 2 kb regions (A) and the average endogenous DNA methylation level in the gene body \pm 5 kb regions (B) of protein-coding genes in the human heart. Different colors indicate 4 gene groups classified according to expression levels. (C and D) Expressed genes (RPKM > 0) were ranked by expression level and arranged into 100 bins, and the x-axis from left to right represents genes with increased expression levels. Each point represents a bin. The line plots show the average endogenous DNA methylation level (C) and the average chromatin accessibility level (D) of the gene bodies in each bin. The shadow indicates the 0.95 confidence interval around smooth. The raw data for A–D can be found in [S1 Data](#). LA, left atrium; LV, left ventricle; RA, right atrium; RPKM, reads per kilobase per million; RV, right ventricle; TES, transcription end site; TSS, transcription start site.

<https://doi.org/10.1371/journal.pbio.3001229.g005>

these 2 species. On the other hand, the species-specific DEGs and motifs identified in our data provide insights into the species differences between human and mouse hearts that could potentially accelerate the development of cardiovascular drugs. A recent study revealed

transcriptional and cellular diversity of the nonfailing human heart at single-cell resolution and identified a cluster of neuronal cells which were likely to derived from the intrinsic cardiac autonomic network [16]. The identified human-specific DEGs in our study related to synaptic signaling showed relatively higher expression level in neuronal cells than other types of cells.

Our study revealed significant differences among 4 chambers of the adult human heart and between fetal and adult human hearts. We identified novel chamber-specific and stage-specific lncRNAs, as well as lncRNAs that were paired with important genes, suggesting that they are worth further research. The epigenomes of the different chambers and developmental stages were distinguishable, reflecting the roles of epigenetic modifications in precise spatiotemporal regulation of heart development and cardiac homeostasis. Some key ventricle marker genes, such as *MYL2*, *MYL3*, and *MYH7*, showed gene-body hypomethylation in the ventricles in both the fetal and adult stages, suggesting that there are critical regulatory elements in these regions.

The complex interactions characterized by multiomics guide diverse cells and organs to play their roles. A negative correlation between gene expression and DNA methylation with positively correlated gene expression and chromatin accessibility has been reported in mouse embryonic stem cells using a single-cell multiomics method [79]. While our data are largely consistent with this study, we interestingly found that the gene bodies of intermediate-expression genes in the human heart had the highest level of DNA methylation and lowest level of chromatin accessibility. A previous study also showed a bell-shaped correlation between gene expression and gene-body DNA methylation in cell lines [80]. It is possible that the relatively high transcription rate of the intermediate-expression genes combined with the presence of few intragenic regulatory elements contributes to this pattern. The detailed mechanism underlying the bell-shaped relationship between expression and gene-body DNA methylation or chromatin accessibility needs further research in the future.

In summary, our study provides integrated transcriptome, DNA methylation, and chromatin accessibility maps for understanding the spatiotemporal characteristics of both the human and mouse hearts.

Methods

Heart sample collection

Samples of normal adult human hearts were obtained from donors who died for reasons other than cardiovascular disease ($n = 11$; 10 male; mean age, 29.4 ± 7.1 years). Human fetal hearts aged 16 ($n = 1$) and 20 weeks ($n = 2$) of gestation were collected after elective abortion. The demographics of the donors are summarized in [S1 Table](#). The LA, RA, LV, and RV were collected and immediately frozen in liquid nitrogen for future use as soon as adult or fetal hearts were separated. This study was performed in accordance with the principle of the Helsinki Declaration and approved by the Ethics Committees of Fuwai Hospital and Xiangyang No.1 People's Hospital. The approval numbers for adult and fetal human samples were 2018–1065 and XYFH20170124, respectively. Written informed consent was obtained from all donors or parents. Samples of all 4 cardiac chambers from C57BL/6 mice aged 10 weeks were collected for cross-species analysis, in conformity with the Chinese National Regulation on the Administration of Laboratory Animals (2017 Revision) and approved by the Fuwai Hospital Committee for Laboratory Animal Use (approval number: FW-2019-0002). To avoid gender-related effects in the cardiovascular system of mice [81–84], the mice used in this study were all male.

RNA extraction and RNA-seq library preparation

Following sample homogenization by grinding in liquid nitrogen and subsequent use of QIAshredder (QIAGEN, Dusseldorf, North Rhine-Westphalia, Germany, Cat. 79656), total

RNA from human and mouse hearts was extracted using the RNeasy Fibrous Tissue Mini Kit (QIAGEN, Cat. 74704) according to the manufacturer's recommendation. Then, rRNAs were removed from the total RNA by the NEBNext rRNA Depletion Kit (NEB, Ipswich, Massachusetts, USA, Cat. E7755X). First-strand cDNAs were synthesized by the NEBNext RNA First Strand Synthesis Module (NEB, Cat. E7525L), and second-strand synthesis was performed by the NEBNext Ultra II Non-Directional RNA Second Strand Synthesis Module (NEB, Cat. E6111L). Finally, the RNA-seq library was prepared using KAPA Hyper Prep Kits (KAPA Biosystems, Wilmington, Massachusetts, USA, Cat. KK8504).

NOMe-seq library preparation and sequencing

To optimize the NOMe-seq procedure for frozen heart tissue samples, we modified the protocol based on previous studies [29,34,85,86]. Approximately 30 mg tissues were used for NOMe-seq in this study. In brief, the heart tissues were thoroughly ground at low temperatures. The tissue pellets were resuspended in 500 μ l ice-cold lysis buffer with Protease Inhibitor Cocktail: 10 mM Tris-HCl (pH 7.5), 10 mM NaCl, 3 mM MgCl₂, 0.1 mM EDTA, and 0.5% NP-40. The tubes were incubated for 60 min with vortexing every 10 min to release nuclei. The nuclei were washed twice with cold DPBS, and the lysates were incubated in 60 units of M.CviPI GpC Methyltransferase (NEB, Cat. M0227L) for 1 h, with the addition of 3 ng unmodified lambda DNA (Thermo Fisher Scientific, Waltham, Massachusetts, Cat. SD0021). Twenty units of M.CviPI and 0.75 μ l of 200 \times SAM were added for another 1 h incubation. The reactions were then stopped by adding EDTA and proteinase K to digest the proteins overnight. Genomic DNAs were purified by phenol:chloroform:isoamyl alcohol extraction and ethanol precipitation. Next, bisulfite conversion of genomic DNAs was conducted using EZ-96 DNA Methylation-Direct MagPrep (Zymo Research, Irvine, California, USA, Cat. D5044). Oligo1 (5'-biotin-CTACACGACGCTCTTCCGATCTNNNNNNNNN-3') and Oligo2 (AGACGTGTGCTCTTCCGATCTNNNNNNNNN) were used to synthesize the first and second strands, respectively. Finally, the NOMe-seq libraries were amplified with 12 cycles of PCR using KAPA HiFi Hot Start Ready Mix (KAPA Biosystems, Cat. KK2602). The NOMe-seq libraries and RNA-seq libraries were sequenced on the HiSeq 4000 platform (Novogene, Beijing, China).

Quantitative RT-PCR

Total RNA was isolated using TRIzol (Thermo Fisher Scientific, Cat. 15596018) according to the manufacturer's protocol, and reverse transcription was performed with the PrimeScript RT reagent Kit (Takara, Beijing, China, Cat. RR047A). Quantitative PCR was performed in an ABI 7500 using TB Green Premix Ex Taq (Takara, Cat. RR820B). After an initial denaturation at 95°C for 30 s, the reaction conditions (5 s at 95°C, 34 s at 60°C, 34 s at 72°C) were applied for 40 cycles, and each qRT-PCR was repeated thrice for every sample as technical replicates. Relative mRNA expression was calculated by the $2^{-\Delta\Delta CT}$ method and normalized using *GAPDH* as an internal control. The specific primers used were as follows: for human *STAT4*, forward primer 5'-ATCCTCCACCTGCCACATTG-3' and reverse primer 5'-TTCCTTGCAAGAACTGGCCC-3'; for human *ANKRD2*, forward primer 5'-AGACCTTCCTGAAAGCTGCG-3' and reverse primer 5'-TCCACAGTGGCCCCATTATC-3'; for human *GAPDH*, forward primer 5'-ACAACCTTGGTATCGTGGAAGG-3' and reverse primer 5'-GCCATCACGCCACAGTTTC-3'.

Western blot analysis

Total proteins were harvested using RIPA buffer supplemented with proteinase and phosphatase inhibitors (Roche, Basel, Switzerland, Cat. 05892791001, Cat. 04906845001) according to

standard protocols. The protein concentrations were measured by the BCA (Thermo Fisher Scientific, Cat. 23227) assay. The protein samples were separated by 4% to 12% SDS-PAGE gels (Invitrogen, Carlsbad, California, USA, Cat. NP0323BOX) followed by electrotransfer to polyvinylidene fluoride (PVDF) membranes (Merck Millipore, Burlington, Massachusetts, USA, Cat. ISEQ00010). The membranes were blocked in 5% milk in Tris-buffered saline (TBS)-Tween (Applygen, Beijing, China, Cat. B1009) for 1 h at room temperature and then incubated overnight at 4°C in primary antibodies against STAT4 (CST, Danvers, Massachusetts, USA, Cat. 2653, 1:1,000), ANKRD2 (Proteintech, Rosemont, Illinois, USA, Cat. 11821, 1:1,000), or GAPDH (Proteintech, Cat. 60004, 1:20,000). After the membranes were washed 5 times in TBS-T for 5 min at room temperature, they were incubated with secondary antibodies conjugated to HRP (ZSGB-BIO, Beijing, China, Cat. ZB-2301, Cat. ZB-2305) for 1 h at room temperature. After washing 5 times in TBS-T for 5 min, protein bands were visualized using the SuperSignal West Femto Chemiluminescent Substrate Kit (Thermo Fisher Scientific, Cat. 34096) and imaged using the ProteinSimple FluorChem M System (ProteinSimple, San Jose, California, USA).

Processing NOME-seq data

The NOME-seq raw sequencing reads were processed by removing the first 9 bases containing random primer sequences, adaptors, and low-quality bases using *trim_galore* (version: 0.1.3) with the parameters '*-quality 20-stringency 3-length 50-clip_R1 9-clip_R2 9-paired-trim1-phred33*'. The clean reads were then mapped into the UCSC human genome (hg19) or mouse genome (mm10) using *Bismark* (version: 0.7.6) with the parameters '*-fastq-non-directional-unmapped-phred33-quals*' in the paired-end nondirectional alignment mode [87]. Then, the unmapped reads were realigned into the same reference genome in a single-end alignment mode to improve the number of mapped reads. After alignment, PCR-duplicated reads were removed using *SAMtools* (version: 0.1.18) [88], and the final BAM file was obtained. Cytosine sites with at least 3× coverage were used to detect technological replicates. After that, the final BAM files of the same technological replicate were merged using *SAMtools*, and 3× coverage was used as the read depth cutoff in the downstream analysis.

The methylation level of each cytosine site was calculated as the number of methylated reads 'C' divided by the number of methylated and unmethylated reads ('C+T'). We used WCG (ACG and TCG trinucleotides) for DNA methylation analysis and GCH (GCA, GCT, and GCC trinucleotides) for chromatin accessibility analysis.

Processing RNA-seq data

RNA-seq raw sequencing reads were first processed by discarding reads contaminated with adaptors and reads with low-quality bases using custom scripts. Then, the clean reads were mapped to the GENCODE human genome (hg19) or mouse genome (mm10) using *HISAT2* (version: 2.1.0) [89] with the default parameters. We generated splice site files for mapping with the *hisat2_extract_splice_sites.py* script in *HISAT2*. In each sample, the transcripts were then directly assembled and quantified using the subcommand '*-e -B*' in *StringTie* (version: 1.3.6) [90] with read alignments and the reference annotation. Finally, the gene expression levels were quantified as reads per kilobase per million mapped reads (RPKM).

Identification of differentially expressed genes

DEGs were identified with the RNA-seq read count data using the R package *DESeq2* (version: 1.20.0) [91]. The false discovery rate (FDR, Benjamini and Hochberg) was calculated to obtain the statistical significance of DEGs. Only genes matching the following 2 criteria were

considered DEGs: (1) the log₂-transformed fold change was greater than 1 with a P value ≤ 0.05 and $FDR \leq 0.05$; (2) the average expression level (RPKM) in the corresponding cluster was greater than 1 (for example, the average expression level of a highly expressed DEG in the atria should be greater than 1 in the atria). GO analysis of the DEGs was performed with ToppGene [92].

Identification of correlations between long noncoding RNAs and protein-coding RNAs

The correlation between lncRNAs and protein-coding RNAs was calculated with DEGs in the atria and ventricles. There were 808 protein-coding RNAs and 124 lncRNAs in the atria and 381 protein-coding RNAs and 63 lncRNAs in the ventricles. We calculated the Pearson correlation between each lncRNA and any protein-coding RNAs within 500 kb, and there were 67 and 30 such pairs in the atria and ventricles, respectively.

lncRNAs and protein-coding RNAs were then randomly arranged throughout the genome to generate random gene pairs, but the gene length was kept fixed. The Pearson correlation between lncRNAs and protein-coding RNAs within 500 kb in the random distribution was calculated. This simulation process was repeated 1,000 times.

Identification of *cis*-regulatory relationships between long noncoding RNAs and protein-coding RNAs

Pearson correlations between lncRNAs and protein-coding RNAs were calculated to detect coexpressed gene pairs, and only DEGs in the atria and ventricles were considered. Coexpressed gene pairs were chosen only if the Pearson correlation coefficient was greater than 0.6 and the correlation P value was ≤ 0.05 . If the linear distance between the lncRNA and protein-coding RNA in the coexpressed gene pair was less than 100 kb, we considered this gene pair to have a *cis*-regulatory relationship.

Identification of species-specifically expressed genes

We first examined gene homology correspondences from the human and mouse annotation references [93] and discarded genes with one-to-many correspondences. Then, genes with the following 3 features were identified as species-specifically expressed genes (in short, taking human-specifically expressed genes as an example): (1) the average difference in the gene expression level between humans and mice was more than 5 times; (2) at least 2 human samples expressed this gene ($RPKM \geq 0.1$) with an average expression level in all human samples of more than 1; and (3) the average expression level in all mouse samples was less than 1. Finally, 768 human- and 399 mouse-specifically expressed genes were identified. We then overlapped these species-specifically expressed genes with the DEGs in the atria or ventricles to obtain species-specific DEGs.

Identification of novel transcripts in human hearts

To investigate novel transcripts in humans, the clean RNA-seq sequencing data were aligned with *HISAT2* as described above, but the alignment reads were not directly assembled and quantified using the subcommand '*-e -B*' in *StringTie* with the reference annotation. After the alignment, the transcripts were initially assembled and quantified with *StringTie* in each sample, and then, we merged the transcripts in each human sample with the subcommand '*merge*' in *StringTie*, which created a uniform transcript set for all human samples. We then processed the alignment reads and the merged transcripts using the subcommand '*-e -B*' in *StringTie*.

To exclude known transcripts and identify novel transcripts, annotation files from UCSC, Ensembl, and GENCODE were collected together to form a set of “known transcript sites.” Any transcript located within 10 kb upstream and 10 kb downstream from a known transcript was not defined as a novel transcript. Transcripts not mapped to assembled chromosomes were excluded from the downstream analysis.

After that, 2,424 novel transcripts were obtained, and their coding potential was calculated with the *CPC* (Coding Potential Calculator) online tool [94], which separated these transcripts into 2 groups, coding transcripts and noncoding transcripts. Furthermore, we filtered the novel transcripts with the following 3 criteria: (1) the length of a novel transcript was more than 1 kb; (2) the number of exons in a novel transcript was more than 2 (in other words, there was an exon–exon junction in the transcript); and (3) at least 1 human sample expressed this novel transcript. Finally, 51 coding RNAs and 176 noncoding RNAs were detected. Since the lengths of these filtered noncoding RNAs were more than 1 kb, we directly considered these RNAs as lncRNAs in this study.

Establishing the coexpression networks with cardiovascular diseases associated genes

To identify the potential key genes in cardiovascular diseases, the coexpression networks analysis was performed only with associated genes. Based on our RNA-seq data quantified with RPKM, the Pearson correlation coefficient was calculated, and only gene pairs with correlation coefficient greater than 0.5 were considered in the following analysis. The networks were visualized with software *Cytoscape*, and the point size indicated the number of connections for a given gene, and the line thickness indicated the correlation coefficient for a given gene pair.

Estimation of DNA methylation and chromatin accessibility levels

DNA methylation levels (WCG) and chromatin accessibility levels (GCH) were estimated in several different genomic regions, and only regions with at least 3 WCG/GCH sites were considered.

The whole genome (human or mouse) was separated into bins with a size of 1 kb. The gene body was the region from the TSS to the transcription end site (TES). When we observed WCG levels around gene bodies (Fig 1C and 1E, S1A and S1B Fig), we extended the gene-body region (5 kb upstream from TSS and 5 kb downstream from TES), each gene body was separated into 100 bins with different sizes, and each extended region was separated into 10 bins with a size of 500 bp. When we observed GCH levels around the TSS (Fig 1C and 1E, S1A and S1B Fig), we extended the window 2 kb upstream and 2 kb downstream from the TSS, and each region was separated into 200 bins with a size of 20 bp.

The promoter was defined as the region from 1 kb upstream of the TSS to 0.5 kb downstream of the TSS. The human annotation information identifying the genomic elements (exons, introns, and CpG islands) was downloaded from the UCSC Genome Browser (hg19), and all repetitive elements and their subfamilies (LINE, SINE, LTR, L1, L2, Alu, MIR, ERV1, ERVK, ERVL, and ERVL-MaLR) were collected from the “RepeatMasker” track from the UCSC Genome Browser (hg19). Intragenic regions were considered gene-body regions, and intergenic regions were considered the complement of the intragenic regions throughout the genome. Human enhancer information was collected from a previous study [95].

Identification of differentially methylated regions

DMRs were identified with the 3× coverage WCG sites. The whole genome was separated into several 300 bp fixed windows to calculate the average WCG level in each sample, and windows

with at least 3 WCG sites were considered. Differentially methylated windows (DMWs) between 2 group samples were first selected only if the average difference was more than 20% with two-tailed Student *t* test $P \leq 0.05$ and FDR ≤ 0.05 (Benjamini and Hochberg method). Neighboring DMWs within 300 bp were then joined as DMRs.

Identification of nucleosome-depleted regions

NDRs were identified with 3× coverage GCH sites. NDRs were defined as regions with significantly higher GCH levels than those in the background. Similar to the previous study [96,97], 100 bp windows with 20 bp sliding steps were used to call NDRs, and only regions matched following 3 criteria were considered NDRs: (1) the average GCH level in the region was significantly higher than that in the whole-genome background with chi-squared test $P \leq 10^{-10}$; (2) the number of GCH sites in the region was more than 5; and (3) the length of the region was longer than 140 bp.

According to the distance from NDRs to TSS, NDRs were grouped into 2 types: proximal NDRs and distal NDRs. NDRs located within promoters (1 kb upstream and 0.5 kb downstream of the TSS) were defined as proximal NDRs, and all others were defined as distal NDRs.

Identification of transcription factor motifs corresponding to open chromatin

To establish the relationship between TF regulation and open chromatin, TF motif enrichment in NDRs was investigated using the *findMotifsGenome.pl* script in HOMER (version: 4.10.4) [98] with the parameters ‘-size 2000 -len 8 -S 100’. Only motifs with $P \leq 10^{-10}$ and RPKM ≥ 5 in at least 1 subject were kept in the downstream analysis. In the figures of this study, any motifs without statistical significance are not shown.

Supporting information

S1 Fig. Global patterns of DNA methylation and chromatin accessibility in human and mouse hearts. (A) Line plots showing the average endogenous DNA methylation level around the gene body ± 5 kb region (left) and the average chromatin accessibility levels around the TSS ± 2 kb region (right) in the adult human heart NC_LV_ind1. Different colors indicate different technology replications. (B) Line plots showing the average endogenous DNA methylation level around the gene body ± 5 kb region (left) and the average chromatin accessibility levels around the TSS ± 2 kb region (right) in the mouse heart NM_LA_m1to5. Different colors indicate different technology replications. (C) Bar plots showing the average endogenous DNA methylation levels (left) and the average chromatin accessibility levels (right) in the whole genome in the human heart. (D) Bar plots showing the average endogenous DNA methylation levels (left) and the average chromatin accessibility levels (right) in the whole genome in the mouse heart. The raw data for A–D can be found in [S2 Data](#). LA, left atrium; LV, left ventricle; TES, transcription end site; TSS, transcription start site. (TIF)

S2 Fig. Multiomics differences (differences in the transcriptome and endogenous DNA methylation) in the adult human heart. (A) Heatmap showing row *z*-score scaled gene expression levels of DEGs (lncRNAs) between the atria and ventricles in the adult human heart. The number of DEGs is reported on the left. (B) Bar plots showing the relative enrichment of DMRs located in different genome elements. (C) PCA plot showing the endogenous DNA methylation pattern of CpG islands in the adult human heart. Circles indicate atrium

samples, and triangles indicate ventricle samples. The variation values of PC1 and PC2 were 10.2% and 7.8%, respectively. (D) Heatmaps showing endogenous DNA methylation levels around the gene body \pm 10 kb region of *MYH7* (top), *GJA5*, *SMAD6*, and *SMAD7* (bottom) in the adult human heart. The color bars in the heatmaps indicate the gene-body region and the promoter region (from TSS upstream 1 kb to TSS downstream 0.5 kb). Gene expression levels were quantified with $\log_2(\text{RPKM} + 1)$. (E) PCA plot showing the chromatin accessibility pattern of distal (left) and proximal (right) NDRs in the adult human heart. Circles and triangles indicate atrium and ventricle samples, respectively. The raw data for A–E can be found in [S2 Data](#). CGI, CpG island; DEGs, differentially expressed genes; DMRs, differentially methylated regions; LA, left atrium; LINE, long interspersed nuclear elements; lncRNAs, long noncoding RNAs; LTR, long terminal repeat; LV, left ventricle; MIR, mammalian-wide interspersed repeat; NDRs, nucleosome-depleted regions; PCA, principal component analysis; RA, right atrium; RPKM, reads per kilobase per million; RV, right ventricle; SINE, short interspersed nuclear elements; TSS, transcription start site.

(TIF)

S3 Fig. Multiomics comparisons (comparisons of the transcriptome, endogenous DNA methylation, and chromatin accessibility) in the human heart between fetuses and adults.

(A) PCA plot showing the transcriptome pattern of known lncRNAs in the human heart. Circles and triangles indicate atrium and ventricle samples, respectively; point sizes indicate the adult heart and fetal heart. The variation values of PC1 and PC2 were 24.6% and 16.3%, respectively. (B) Left, line plot showing the density distribution of length from novel transcripts to the nearest known transcripts; the zoom-out visualization of the dashed region is shown. Right, line plot showing the density distribution of the nearest length between each novel transcript. (C) Bar plot showing the distribution of human DEGs in heart disease-associated gene sets. These DEGs were identified with comparisons between the adult atria and ventricles. (D) Coexpression networks of cardiovascular diseases (AF, HCM, and CHD) associated genes. The point size indicated the number of connections for a given gene, and the line thickness indicated the correlation coefficient for a given gene pair. (E) PCA plot showing the endogenous DNA methylation pattern of the whole genome (left), promoters (middle), and CpG islands (right) in the human heart. Rounds and triangles indicate atrium and ventricle samples, respectively; point sizes indicate the adult heart and fetal heart. The raw data for A–E can be found in [S2 Data](#). AF, atrial fibrillation; CHD, congenital heart disease; DEGs, differentially expressed genes; HCM, hypertrophic cardiomyopathy; LA, left atrium; lncRNAs, long noncoding RNAs; LV, left ventricle; PCA, principal component analysis; RA, right atrium; RV, right ventricle.

(TIF)

S4 Fig. Comparisons of the transcriptome and regulatory elements in the adult human heart and mouse heart.

(A) Heatmap showing row *z*-score scaled gene expression levels of DEGs (protein-coding genes) between the atria and ventricles in the mouse heart (left), and corresponding GO terms are shown (right). The number of DEGs is reported on the left. (B) Heatmaps showing *z*-score scaled expression levels of human-specific DEGs among different cell types from single-cell RNA-seq data. Left: human-specific DEGs with higher expression levels in the atria; right, human-specific DEGs with higher expression levels in the ventricles. (C) Heatmaps showing *z*-score scaled expression levels of mouse-specific DEGs among different cell types from single-cell RNA-seq data. Left: mouse-specific DEGs with higher expression levels in the atria; right, mouse-specific DEGs with higher expression levels in the ventricles. (D) PCA plots showing the endogenous DNA methylation pattern of the whole genome, promoters, and CpG islands in the mouse heart. Circles and triangles indicate atrium and

ventricle samples, respectively. (E) Heatmaps showing endogenous DNA methylation levels in the gene body \pm 10 kb regions of *EPHA4* in human hearts. The color bars in the heatmaps indicate the gene-body regions and the promoter regions (from 1 kb upstream of the TSS to 0.5 kb downstream of the TSS). (F) Motif enrichment analysis of proximal NDRs in the adult human heart (left) and mouse heart (right). Colors indicate average expression levels, and sizes indicate P values ($P \leq 10^{-10}$) of the corresponding transcription factor. The raw data for A–F can be found in [S2 Data](#). DEGs, differentially expressed genes; GO, Gene Ontology; LA, left atrium; LV, left ventricle; NDRs, nucleosome-depleted regions; PCA, principal component analysis; RA, right atrium; RPKM, reads per kilobase per million; RV, right ventricle; TSS, transcription start site.

(TIF)

S5 Fig. Multiomics relationship patterns in human and mouse hearts. (A and B) Line plots showing the average chromatin accessibility level around the TSS \pm 2 kb region (A) and the average endogenous DNA methylation level around the gene body \pm 5 kb region (B) of lncRNAs in the human heart. Different colors indicate 2 lncRNA groups classified according to expression levels. (C) Line plots showing the average endogenous DNA methylation level (green) and the average chromatin accessibility level (blue) in promoters of corresponding genes. Genes were ranked by expression levels, and the x-axis from left to right represented genes with increased expression levels. Spearman correlation coefficients between chromatin accessibility levels and gene expression levels (R1) and Spearman correlation coefficients between endogenous DNA methylation levels and gene expression levels (R2) are indicated. (D) Line plots showing the average chromatin accessibility level around the TSS \pm 2 kb region (left) and the average endogenous DNA methylation level around the gene body \pm 5 kb region (right) of protein-coding genes in the mouse heart LA. Different colors indicate 4 gene groups classified according to expression levels. (E) Line plots showing the average chromatin accessibility level around the TSS \pm 2 kb region (left) and the average endogenous DNA methylation level around the gene body \pm 5 kb region (right) of lncRNAs in the mouse heart LA. Different colors indicate 2 lncRNA groups classified according to expression levels. The raw data for A–E can be found in [S2 Data](#). LA, left atrium; lncRNAs, long noncoding RNAs; LV, left ventricle; RA, right atrium; RPKM, reads per kilobase per million; RV, right ventricle; TES, transcription end site; TSS, transcription start site.

(TIF)

S1 Table. Sequencing information of NOME-seq and RNA-seq in human and mouse hearts.
(XLSX)

S2 Table. DEGs (protein-coding genes and lncRNAs) between the human atria and ventricles.

(XLSX)

S3 Table. DMRs between the human atria and ventricles.

(XLSX)

S4 Table. Novel lncRNAs identified in human hearts.

(XLSX)

S5 Table. Motifs enriched in distal and proximal NDRs in fetal and adult human hearts.

(XLSX)

S6 Table. Conserved and species-specific DEGs between the atria and ventricles in human and mouse hearts.

(XLSX)

S7 Table. Motifs enriched in distal and proximal NDRs in human and mouse hearts.

(XLSX)

S1 Data. The individual numerical values in Figs 1B–1E; 2A–2I; 3A–3F; 4A–4D, 4F, 4G; 5A–5D.

(XLSX)

S2 Data. The individual numerical values in S1A–S1D; S2A–S2E; S3A–S3E; S4A–S4F; S5A–S5E Figs.

(XLSX)

S1 Raw images. Original gel and blot images for Fig 4E.

(PDF)

Author Contributions

Conceptualization: Junpeng Gao, Fuchou Tang.

Data curation: Junpeng Gao, Yuxuan Zheng, Jizheng Wang.

Formal analysis: Junpeng Gao, Yuxuan Zheng.

Funding acquisition: Fuchou Tang, Lei Song, Lu Wen, Jizheng Wang.

Investigation: Junpeng Gao, Yuxuan Zheng, Lin Li, Minjie Lu, Xiangjian Chen, Yu Wang, Yanna Li, Xiaomeng Liu, Yun Gao, Yunuo Mao, Peng Zhao, Jinan Zhang, Fuchou Tang, Lei Song, Lu Wen, Jizheng Wang.

Methodology: Junpeng Gao.

Project administration: Junpeng Gao, Fuchou Tang.

Supervision: Lei Song, Jizheng Wang.

Validation: Junpeng Gao.

Visualization: Junpeng Gao, Yuxuan Zheng.

Writing – original draft: Junpeng Gao.

Writing – review & editing: Junpeng Gao, Yuxuan Zheng, Lei Song, Lu Wen, Jizheng Wang.

References

1. Buckingham M, Meilhac S, Zaffran S. Building the mammalian heart from two sources of myocardial cells. *Nat Rev Genet.* 2005; 6(11):826–35. Epub 2005/11/24. <https://doi.org/10.1038/nrg1710> PMID: 16304598.
2. Wessels A, Sedmera D. Developmental anatomy of the heart: a tale of mice and man. *Physiol Genomics.* 2003; 15(3):165–76. Epub 2003/11/13. <https://doi.org/10.1152/physiolgenomics.00033.2003> PMID: 14612588.
3. Xin M, Olson EN, Bassel-Duby R. Mending broken hearts: cardiac development as a basis for adult heart regeneration and repair. *Nat Rev Mol Cell Biol.* 2013; 14(8):529–41. Epub 2013/07/11. <https://doi.org/10.1038/nrm3619> PMID: 23839576; PubMed Central PMCID: PMC3757945.
4. Ho SY, Anderson RH, Sánchez-Quintana D. Atrial structure and fibres: morphologic bases of atrial conduction. *Cardiovasc Res.* 2002; 54(2):325–36. [https://doi.org/10.1016/s0008-6363\(02\)00226-2](https://doi.org/10.1016/s0008-6363(02)00226-2) PMID: 12062338

5. Lindskog C, Linne J, Fagerberg L, Hallstrom BM, Sundberg CJ, Lindholm M, et al. The human cardiac and skeletal muscle proteomes defined by transcriptomics and antibody-based profiling. *BMC Genomics*. 2015; 16:475. Epub 2015/06/26. <https://doi.org/10.1186/s12864-015-1686-y> PMID: 26109061; PubMed Central PMCID: PMC4479346.
6. Lu ZQ, Sinha A, Sharma P, Kislinger T, Gramolini AO. Proteomic analysis of human fetal atria and ventricle. *J Proteome Res*. 2014; 13(12):5869–78. Epub 2014/10/18. <https://doi.org/10.1021/pr5007685> PMID: 25323733.
7. Zhao X-S, Gallardo TD, Lin L, Schageman JJ, Shohet RV. Transcriptional mapping and genomic analysis of the cardiac atria and ventricles. *Physiol Genomics*. 2002; 12(1):53–60. <https://doi.org/10.1152/physiolgenomics.00086.2002> PMID: 12502795
8. Tabibiazar R, Wagner RA, Liao A, Quertermous T. Transcriptional profiling of the heart reveals chamber-specific gene expression patterns. *Circ Res*. 2003; 93(12):1193–201. <https://doi.org/10.1161/01.RES.0000103171.42654.DD> PMID: 14576202
9. Hu HL, Kang Y, Zeng Y, Zhang M, Liao Q, Rong MQ, et al. Region-resolved proteomics profiling of monkey heart. *J Cell Physiol*. 2019; 234(8):13720–34. Epub 2019/01/16. <https://doi.org/10.1002/jcp.28052> PMID: 30644093.
10. Zhi Y, Cao Z, Li QH, Li XL, Sun Y, Zhang T, et al. Transcriptional analysis of atrial and ventricular muscles from rats. *Genet Mol Res*. 2016; 15(1). Epub 2016/02/26. <https://doi.org/10.4238/gmr.15017330> PMID: 26909933.
11. Kakimoto Y, Tanaka M, Kamiguchi H, Hayashi H, Ochiai E, Osawa M. MicroRNA deep sequencing reveals chamber-specific miR-208 family expression patterns in the human heart. *Int J Cardiol*. 2016; 211:43–8. Epub 2016/03/15. <https://doi.org/10.1016/j.ijcard.2016.02.145> PMID: 26974694.
12. Li G, Xu A, Sim S, Priest JR, Tian X, Khan T, et al. Transcriptomic Profiling Maps Anatomically Patterned Subpopulations among Single Embryonic Cardiac Cells. *Dev Cell*. 2016; 39(4):491–507. <https://doi.org/10.1016/j.devcel.2016.10.014> PMID: 27840109; PubMed Central PMCID: PMC5130110.
13. DeLaughter DM, Bick AG, Wakimoto H, McKean D, Gorham JM, Kathiriyas IS, et al. Single-Cell Resolution of Temporal Gene Expression during Heart Development. *Dev Cell*. 2016; 39(4):480–90. <https://doi.org/10.1016/j.devcel.2016.10.001> PMID: 27840107; PubMed Central PMCID: PMC5198784.
14. Zhou Y, Zhang J. Single-Cell Transcriptomics: New Insights in Heart Research. *Circulation*. 2020; 141(21):1720–3. Epub 2020/05/27. <https://doi.org/10.1161/CIRCULATIONAHA.120.046043> PMID: 32453661.
15. Wang L, Yu P, Zhou B, Song J, Li Z, Zhang M, et al. Single-cell reconstruction of the adult human heart during heart failure and recovery reveals the cellular landscape underlying cardiac function. *Nat Cell Biol*. 2020; 22(1):108–19. Epub 2020/01/10. <https://doi.org/10.1038/s41556-019-0446-7> PMID: 31915373.
16. Tucker NR, Chaffin M, Fleming SJ, Hall AW, Parsons VA, Bedi KC, et al. Transcriptional and Cellular Diversity of the Human Heart. *Circulation*. 2020; 142(5):466–82. <https://doi.org/10.1161/CIRCULATIONAHA.119.045401> PMID: 32403949
17. Johnson EK, Matkovich SJ, Nerbonne JM. Regional Differences in mRNA and lncRNA Expression Profiles in Non-Failing Human Atria and Ventricles. *Sci Rep*. 2018; 8(1):13919. Epub 2018/09/19. <https://doi.org/10.1038/s41598-018-32154-2> PMID: 30224797; PubMed Central PMCID: PMC6141608.
18. Doll S, Dressen M, Geyer PE, Itzhak DN, Braun C, Doppler SA, et al. Region and cell-type resolved quantitative proteomic map of the human heart. *Nat Commun*. 2017; 8(1):1469. Epub 2017/11/15. <https://doi.org/10.1038/s41467-017-01747-2> PMID: 29133944; PubMed Central PMCID: PMC5684139.
19. Barth AS, Merk S, Arnoldi E, Zwermann L, Kloos P, Gebauer M, et al. Functional profiling of human atrial and ventricular gene expression. *Pflugers Arch*. 2005; 450(4):201–8. Epub 2005/05/07. <https://doi.org/10.1007/s00424-005-1404-8> PMID: 15877233.
20. Nothjunge S, Nührenberg TG, Grüning BA, Doppler SA, Preissl S, Schwaderer M, et al. DNA methylation signatures follow preformed chromatin compartments in cardiac myocytes. *Nat Commun*. 2017; 8(1). <https://doi.org/10.1038/s41467-017-01724-9> PMID: 29162810
21. Grunert M, Dorn C, Cui H, Dunkel I, Schulz K, Schoenhals S, et al. Comparative DNA methylation and gene expression analysis identifies novel genes for structural congenital heart diseases. *Cardiovasc Res*. 2016; 112(1):464–77. Epub 2016/08/09. <https://doi.org/10.1093/cvr/cvw195> PMID: 27496870.
22. Palinski W, Napoli C. The fetal origins of atherosclerosis: maternal hypercholesterolemia, and cholesterol-lowering or antioxidant treatment during pregnancy influence in utero programming and postnatal susceptibility to atherogenesis. *FASEB J*. 2002; 16(11):1348–60. Epub 2002/09/03. <https://doi.org/10.1096/fj.02-0226rev> PMID: 12205026.
23. Barker DJ. Fetal origins of cardiovascular disease. *Ann Med*. 1999; 31 Suppl 1:3–6. Epub 1999/05/26. PMID: 10342493.

24. Olson EN, Srivastava D. Molecular pathways controlling heart development. *Science*. 1996; 272(5262):671–6. <https://doi.org/10.1126/science.272.5262.671> PMID: 8614825
25. Krishnan A, Samtani R, Dhanantwari P, Lee E, Yamada S, Shiota K, et al. A detailed comparison of mouse and human cardiac development. *Pediatr Res*. 2014; 76(6):500–7. Epub 2014/08/29. <https://doi.org/10.1038/pr.2014.128> PMID: 25167202; PubMed Central PMCID: PMC4233008.
26. Cui Y, Zheng Y, Liu X, Yan L, Fan X, Yong J, et al. Single-Cell Transcriptome Analysis Maps the Developmental Track of the Human Heart. *Cell Rep*. 2019; 26(7):1934–50e5. Epub 2019/02/14. <https://doi.org/10.1016/j.celrep.2019.01.079> PMID: 30759401.
27. Rosenthal N, Brown S. The mouse ascending: perspectives for human-disease models. *Nat Cell Biol*. 2007; 9(9):993–9. Epub 2007/09/01. <https://doi.org/10.1038/ncb437> PMID: 17762889.
28. Ulrich MA, Nery JR, Lister R, Schmitz RJ, Ecker JR. MethylC-seq library preparation for base-resolution whole-genome bisulfite sequencing. *Nat Protoc*. 2015; 10(3):475–83. <https://doi.org/10.1038/nprot.2014.114> PMID: 25692984; PubMed Central PMCID: PMC4465251.
29. Miura F, Enomoto Y, Dairiki R, Ito T. Amplification-free whole-genome bisulfite sequencing by post-bisulfite adaptor tagging. *Nucleic Acids Res*. 2012; 40(17):e136. <https://doi.org/10.1093/nar/gks454> PMID: 22649061; PubMed Central PMCID: PMC3458524.
30. Meissner A, Mikkelsen TS, Gu H, Wernig M, Hanna J, Sivachenko A, et al. Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature*. 2008; 454(7205):766–70. Epub 2008/07/05. <https://doi.org/10.1038/nature07107> PMID: 18600261; PubMed Central PMCID: PMC2896277.
31. Boyle AP, Davis S, Shulha HP, Meltzer P, Margulies EH, Weng Z, et al. High-resolution mapping and characterization of open chromatin across the genome. *Cell*. 2008; 132(2):311–22. Epub 2008/02/05. <https://doi.org/10.1016/j.cell.2007.12.014> PMID: 18243105; PubMed Central PMCID: PMC2669738.
32. Schones DE, Cui K, Cuddapah S, Roh TY, Barski A, Wang Z, et al. Dynamic regulation of nucleosome positioning in the human genome. *Cell*. 2008; 132(5):887–98. Epub 2008/03/11. <https://doi.org/10.1016/j.cell.2008.02.022> PMID: 18329373.
33. Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat Methods*. 2013; 10(12):1213–8. Epub 2013/10/08. <https://doi.org/10.1038/nmeth.2688> PMID: 24097267; PubMed Central PMCID: PMC3959825.
34. Kelly TK, Liu Y, Lay FD, Liang G, Berman BP, Jones PA. Genome-wide mapping of nucleosome positioning and DNA methylation within individual DNA molecules. *Genome Res*. 2012; 22(12):2497–506. <https://doi.org/10.1101/gr.143008.112> PMID: 22960375; PubMed Central PMCID: PMC3514679.
35. Xie L, Weichel B, Ohm JE, Zhang K. An integrative analysis of DNA methylation and RNA-Seq data for human heart, kidney and liver. *BMC Syst Biol*. 2011; 5(3):S4. <https://doi.org/10.1186/1752-0509-5-S3-S4> PMID: 22784623
36. Hon GC, Rajagopal N, Shen Y, McCleary DF, Yue F, Dang MD, et al. Epigenetic memory at embryonic enhancers identified in DNA methylation maps from adult mouse tissues. *Nat Genet*. 2013; 45(10):1198–206. <https://doi.org/10.1038/ng.2746> PMID: 23995138; PubMed Central PMCID: PMC4095776.
37. Frank S, Aguirre A, Hescheler J, Kurian L. A lncRNA Perspective into (Re)Building the Heart. *Front Cell Dev Biol*. 2016; 4:128. Epub 2016/11/25. <https://doi.org/10.3389/fcell.2016.00128> PMID: 27882316; PubMed Central PMCID: PMC5101577.
38. Devaux Y, Zangrando J, Schroen B, Creemers EE, Pedrazzini T, Chang CP, et al. Long noncoding RNAs in cardiac development and ageing. *Nat Rev Cardiol*. 2015; 12(7):415–25. Epub 2015/04/10. <https://doi.org/10.1038/nrcardio.2015.55> PMID: 25855606.
39. Scheuermann JC, Boyer LA. Getting to the heart of the matter: long non-coding RNAs in cardiac development and disease. *EMBO J*. 2013; 32(13):1805–16. Epub 2013/06/13. <https://doi.org/10.1038/emboj.2013.134> PMID: 23756463; PubMed Central PMCID: PMC3981183.
40. Anderson KM, Anderson DM, McAnally JR, Shelton JM, Bassel-Duby R, Olson EN. Transcription of the non-coding RNA upperhand controls Hand2 expression and heart development. *Nature*. 2016; 539(7629):433–6. Epub 2016/11/04. <https://doi.org/10.1038/nature20128> PMID: 27783597; PubMed Central PMCID: PMC5261552.
41. Biben C, Hatzistavrou T, Harvey R. Expression of NK-2 class homeobox gene Nkx2–6 in foregut endoderm and heart. *Mech Dev*. 1998; 73(1):125–7. [https://doi.org/10.1016/s0925-4773\(98\)00037-9](https://doi.org/10.1016/s0925-4773(98)00037-9) PMID: 9545560
42. Wang J, Zhang D-F, Sun Y-M, Li R-G, Qiu X-B, Qu X-K, et al. NKX2-6 mutation predisposes to familial atrial fibrillation. *Int J Mol Med*. 2014; 34(6):1581–90. <https://doi.org/10.3892/ijmm.2014.1971> PMID: 25319568

43. Yamada N, Asano Y, Fujita M, Yamazaki S, Inanobe A, Matsuura N, et al. Mutant KCNJ3 and KCNJ5 Potassium Channels as Novel Molecular Targets in Bradyarrhythmias and Atrial Fibrillation. *Circulation*. 2019; 139(18):2157–69. Epub 2019/02/16. <https://doi.org/10.1161/CIRCULATIONAHA.118.036761> PMID: 30764634.
44. Flavigny J, Richard P, Isnard R, Carrier L, Charron P, Bonne G, et al. Identification of two novel mutations in the ventricular regulatory myosin light chain gene (MYL2) associated with familial and classical forms of hypertrophic cardiomyopathy. *J Mol Med (Berl)*. 1998; 76(3–4):208–14. Epub 1998/04/16. <https://doi.org/10.1007/s001090050210> PMID: 9535554.
45. Morita H, Rehm HL, Menesses A, McDonough B, Roberts AE, Kucherlapati R, et al. Shared genetic causes of cardiac hypertrophy in children and adults. *N Engl J Med*. 2008; 358(18):1899–908. Epub 2008/04/12. <https://doi.org/10.1056/NEJMoa075463> PMID: 18403758; PubMed Central PMCID: PMC2752150.
46. Koibuchi N, Chin MT. CHF1/Hey2 plays a pivotal role in left ventricular maturation through suppression of ectopic atrial gene expression. *Circ Res*. 2007; 100(6):850–5. Epub 2007/03/03. <https://doi.org/10.1161/01.RES.0000261693.13269.bf> PMID: 17332425.
47. Liu T, Zhang X, Yang YM, Du LT, Wang CX. Increased expression of the long noncoding RNA CRNDE-h indicates a poor prognosis in colorectal cancer, and is positively correlated with IRX5 mRNA expression. *Onco Targets Ther*. 2016; 9:1437–48. Epub 2016/04/05. <https://doi.org/10.2147/OTT.S98268> PMID: 27042112; PubMed Central PMCID: PMC4795576.
48. Zhang C, Wu W, Ye X, Ma R, Luo J, Zhu H, et al. Aberrant expression of CHL1 gene and long non-coding RNA CHL1-AS1, CHL1-AS2 in ovarian endometriosis. *Eur J Obstet Gynecol Reprod Biol*. 2019; 236:177–82. Epub 2019/04/04. <https://doi.org/10.1016/j.ejogrb.2019.03.020> PMID: 30943448.
49. Qiao F, Li N, Li W. Integrative Bioinformatics Analysis Reveals Potential Long Non-Coding RNA Biomarkers and Analysis of Function in Non-Smoking Females with Lung Cancer. *Med Sci Monit*. 2018; 24:5771–8. Epub 2018/08/19. <https://doi.org/10.12659/MSM.908884> PMID: 30120911; PubMed Central PMCID: PMC6110140.
50. Maron BJ, Maron MS, Semsarian C. Double or compound sarcomere mutations in hypertrophic cardiomyopathy: a potential link to sudden death in the absence of conventional risk factors. *Heart Rhythm*. 2012; 9(1):57–63. Epub 2011/08/16. <https://doi.org/10.1016/j.hrthm.2011.08.009> PMID: 21839045.
51. Sabater-Molina M, Perez-Sanchez I, Hernandez Del Rincon JP, Gimeno JR. Genetics of hypertrophic cardiomyopathy: A review of current state. *Clin Genet*. 2018; 93(1):3–14. Epub 2017/04/04. <https://doi.org/10.1111/cge.13027> PMID: 28369730.
52. Gollob MH, Jones DL, Krahn AD, Danis L, Gong X-Q, Shao Q, et al. Somatic mutations in the connexin 40 gene (GJA5) in atrial fibrillation. *N Engl J Med*. 2006; 354(25):2677–88. <https://doi.org/10.1056/NEJMoa052800> PMID: 16790700
53. Chen Q, Chen H, Zheng D, Kuang C, Fang H, Zou B, et al. Smad7 is required for the development and function of the heart. *J Biol Chem*. 2009; 284(1):292–300. Epub 2008/10/28. <https://doi.org/10.1074/jbc.M807233200> PMID: 18952608; PubMed Central PMCID: PMC2610499.
54. Galvin KM, Donovan MJ, Lynch CA, Meyer RI, Paul RJ, Lorenz JN, et al. A role for smad6 in development and homeostasis of the cardiovascular system. *Nat Genet*. 2000; 24(2):171–4. Epub 2000/02/02. <https://doi.org/10.1038/72835> 10655064. PMID: 10655064
55. McKinsey TA, Vondriska TM, Wang Y. Epigenomic regulation of heart failure: integrating histone marks, long noncoding RNAs, and chromatin architecture. *F1000Res*. 2018; 7. Epub 2018/11/13. <https://doi.org/10.12688/f1000research.15797.1> PMID: 30416708; PubMed Central PMCID: PMC6206605.
56. Roth GA, Abate D, Abate KH, Abay SM, Abbafati C, Abbasi N, et al. Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980–2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet*. 2018; 392 (10159):1736–88. [https://doi.org/10.1016/S0140-6736\(18\)32203-7](https://doi.org/10.1016/S0140-6736(18)32203-7) PMID: 30496103
57. Fatkin D, Santiago CF, Huttner IG, Lubitz SA, Ellinor PT. Genetics of Atrial Fibrillation: State of the Art in 2017. *Heart Lung Circ*. 2017; 26(9):894–901. Epub 2017/06/12. <https://doi.org/10.1016/j.hlc.2017.04.008> PMID: 28601532.
58. Roma-Rodrigues C, Fernandes AR. Genetics of hypertrophic cardiomyopathy: advances and pitfalls in molecular diagnosis and therapy. *Appl Clin Genet*. 2014; 7:195–208. Epub 2014/10/21. <https://doi.org/10.2147/TACG.S49126> PMID: 25328416; PubMed Central PMCID: PMC4199654.
59. Fahed AC, Gelb BD, Seidman JG, Seidman CE. Genetics of congenital heart disease: the glass half empty. *Circ Res*. 2013; 112(4):707–20. Epub 2013/02/16. <https://doi.org/10.1161/CIRCRESAHA.112.300853> PMID: 23410880; PubMed Central PMCID: PMC3827691.
60. Linde C, Steffel J. The year in cardiology 2017: arrhythmias and cardiac devices. *Eur Heart J*. 2018; 39(6):434–41. Epub 2018/01/05. <https://doi.org/10.1093/eurheartj/ehx765> PMID: 29300864.

61. Nakamura M, Sadoshima J. Mechanisms of physiological and pathological cardiac hypertrophy. *Nat Rev Cardiol*. 2018; 15(7):387–407. Epub 2018/04/21. <https://doi.org/10.1038/s41569-018-0007-y> PMID: 29674714.
62. Bruneau BG. The developmental genetics of congenital heart disease. *Nature*. 2008; 451(7181):943–8. Epub 2008/02/22. <https://doi.org/10.1038/nature06801> PMID: 18288184.
63. Tshori S, Gilon D, Beerli R, Nechushtan H, Kaluzhny D, Pikarsky E, et al. Transcription factor MITF regulates cardiac growth and hypertrophy. *J Clin Invest*. 2006; 116(10):2673–81. Epub 2006/09/26. <https://doi.org/10.1172/JCI27643> PMID: 16998588; PubMed Central PMCID: PMC1570375.
64. Kishore R, Verma SK. Roles of STATs signaling in cardiovascular diseases. *JAKSTAT*. 2012; 1(2):118–24. Epub 2012/04/01. <https://doi.org/10.4161/jkst.20115> PMID: 24058760; PubMed Central PMCID: PMC3670291.
65. Knight RA, Scarabelli TM, Stephanou A. STAT transcription in the ischemic heart. *JAKSTAT*. 2012; 1(2):111–7. Epub 2012/04/01. <https://doi.org/10.4161/jkst.20078> PMID: 24058759; PubMed Central PMCID: PMC3670290.
66. Nagueh SF, Shah G, Wu Y, Torre-Amione G, King NM, Lahmers S, et al. Altered titin expression, myocardial stiffness, and left ventricular function in patients with dilated cardiomyopathy. *Circulation*. 2004; 110(2):155–62. Epub 2004/07/09. <https://doi.org/10.1161/01.CIR.0000135591.37759.AF> PMID: 15238456.
67. Ren Z, Yu P, Li D, Li Z, Liao Y, Wang Y, et al. Single-Cell Reconstruction of Progression Trajectory Reveals Intervention Principles in Pathological Cardiac Hypertrophy. *Circulation*. 2020; 141(21):1704–19. Epub 2020/02/27. <https://doi.org/10.1161/CIRCULATIONAHA.119.043053> PMID: 32098504.
68. Brewer A, Pizzey J. GATA factors in vertebrate heart development and disease. *Expert Rev Mol Med*. 2006; 8(22):1–20. Epub 2006/09/22. <https://doi.org/10.1017/S1462399406000093> PMID: 16987437.
69. van Berlo JH, Aronow BJ, Molkentin JD. Parsing the roles of the transcription factors GATA-4 and GATA-6 in the adult cardiac hypertrophic response. *PLoS ONE*. 2013; 8(12):e84591. Epub 2014/01/07. <https://doi.org/10.1371/journal.pone.0084591> PMID: 24391969; PubMed Central PMCID: PMC3877334.
70. Maves L, Tyler A, Moens CB, Tapscott SJ. Pbx acts with Hand2 in early myocardial differentiation. *Dev Biol*. 2009; 333(2):409–18. Epub 2009/07/18. <https://doi.org/10.1016/j.ydbio.2009.07.004> PMID: 19607825; PubMed Central PMCID: PMC2752274.
71. Acharya A, Baek ST, Huang G, Eskiocak B, Goetsch S, Sung CY, et al. The bHLH transcription factor Tcf21 is required for lineage-specific EMT of cardiac fibroblast progenitors. *Development*. 2012; 139(12):2139–49. Epub 2012/05/11. <https://doi.org/10.1242/dev.079970> PMID: 22573622; PubMed Central PMCID: PMC3357908.
72. Edmondson DG, Lyons GE, Martin JF, Olson EN. Mef2 gene expression marks the cardiac and skeletal muscle lineages during mouse embryogenesis. *Development*. 1994; 120(5):1251–63. PMID: 8026334
73. Killeen MJ, Thomas G, Gurung I, Goddard C, Fraser J, Mahaut-Smith M, et al. Arrhythmogenic mechanisms in the isolated perfused hypokalaemic murine heart. *Acta Physiol*. 2007; 189(1):33–46. <https://doi.org/10.1111/j.1748-1716.2006.01643.x> PMID: 17280555
74. Baker LC, London B, Choi B-R, Koren G, Salama G. Enhanced dispersion of repolarization and refractoriness in transgenic mouse hearts promotes reentrant ventricular tachycardia. *Circ Res*. 2000; 86(4):396–407. <https://doi.org/10.1161/01.res.86.4.396> PMID: 10700444
75. London B. Cardiac arrhythmias: from (transgenic) mice to men. *J Cardiovasc Electrophysiol*. 2001; 12(9):1089–91. <https://doi.org/10.1046/j.1540-8167.2001.01089.x> PMID: 11573703
76. Killeen MJ, Thomas G, Sabir IN, Grace AA, Huang CL. Mouse models of human arrhythmia syndromes. *Acta Physiol (Oxf)*. 2008; 192(4):455–69. Epub 2007/11/30. <https://doi.org/10.1111/j.1748-1716.2007.01822.x> PMID: 18045245.
77. Fedorov VV, Glukhov AV, Ambrosi CM, Kosteki G, Chang R, Janks D, et al. Effects of KATP channel openers diazoxide and pinacidil in coronary-perfused atria and ventricles from failing and non-failing human hearts. *J Mol Cell Cardiol*. 2011; 51(2):215–25. Epub 2011/05/19. <https://doi.org/10.1016/j.yjmcc.2011.04.016> PMID: 21586291; PubMed Central PMCID: PMC3124600.
78. Glukhov AV, Flagg TP, Fedorov VV, Efimov IR, Nichols CG. Differential K(ATP) channel pharmacology in intact mouse heart. *J Mol Cell Cardiol*. 2010; 48(1):152–60. Epub 2009/09/12. <https://doi.org/10.1016/j.yjmcc.2009.08.026> PMID: 19744493; PubMed Central PMCID: PMC2813353.
79. Clark SJ, Argelaguet R, Kapourani CA, Stubbs TM, Lee HJ, Alda-Catalinas C, et al. scNMT-seq enables joint profiling of chromatin accessibility DNA methylation and transcription in single cells. *Nat Commun*. 2018; 9(1):781. Epub 2018/02/24. <https://doi.org/10.1038/s41467-018-03149-4> PMID: 29472610; PubMed Central PMCID: PMC5823944.

80. Jjingo D, Conley AB, Yi SV, Lunyak VV, Jordan IK. On the presence and role of human gene-body DNA methylation. *Oncotarget*. 2012; 3(4):462–74. Epub 2012/05/12. <https://doi.org/10.18632/oncotarget.497> PMID: 22577155; PubMed Central PMCID: PMC3380580.
81. Saito T, Ciobotaru A, Bopassa JC, Toro L, Stefani E, Eghbali M. Estrogen contributes to gender differences in mouse ventricular repolarization. *Circ Res*. 2009; 105(4):343–52. Epub 2009/07/18. <https://doi.org/10.1161/CIRCRESAHA.108.190041> PMID: 19608983; PubMed Central PMCID: PMC2921935.
82. Babiker FA, De Windt LJ, van Eickels M, Thijssen V, Bronsauer RJ, Grohe C, et al. 17beta-estradiol antagonizes cardiomyocyte hypertrophy by autocrine/paracrine stimulation of a guanylyl cyclase A receptor-cyclic guanosine monophosphate-dependent protein kinase pathway. *Circulation*. 2004; 109(2):269–76. Epub 2004/01/14. <https://doi.org/10.1161/01.CIR.0000105682.85732.BD> PMID: 14718400.
83. Condorelli G, Morisco C, Latronico MV, Claudio PP, Dent P, Tsichlis P, et al. TNF-alpha signal transduction in rat neonatal cardiac myocytes: definition of pathways generating from the TNF-alpha receptor. *FASEB J*. 2002; 16(13):1732–7. Epub 2002/11/01. <https://doi.org/10.1096/fj.02-0419com> PMID: 12409315.
84. Condorelli G, Roncarati R, Ross J Jr., Pisani A, Stassi G, Todaro M, et al. Heart-targeted overexpression of caspase3 in mice increases infarct size and depresses cardiac function. *Proc Natl Acad Sci U S A*. 2001; 98(17):9977–82. Epub 2001/08/09. <https://doi.org/10.1073/pnas.161120198> PMID: 11493678; PubMed Central PMCID: PMC55563.
85. Guo H, Hu B, Yan L, Yong J, Wu Y, Gao Y, et al. DNA methylation and chromatin accessibility profiling of mouse and human fetal germ cells. *Cell Res*. 2017; 27(2):165–83. <https://doi.org/10.1038/cr.2016.128> PMID: 27824029; PubMed Central PMCID: PMC5339845.
86. Guo F, Li L, Li J, Wu X, Hu B, Zhu P, et al. Single-cell multi-omics sequencing of mouse early embryos and embryonic stem cells. *Cell Res*. 2017; 27(8):967–88. <https://doi.org/10.1038/cr.2017.82> PMID: 28621329
87. Krueger F, Andrews SR. Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics*. 2011; 27(11):1571–2. <https://doi.org/10.1093/bioinformatics/btr167> PMID: 21493656
88. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The sequence alignment/map format and SAMtools. *Bioinformatics*. 2009; 25(16):2078–9. <https://doi.org/10.1093/bioinformatics/btp352> PMID: 19505943
89. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. *Nat Methods*. 2015; 12(4):357–60. <https://doi.org/10.1038/nmeth.3317> PMID: 25751142
90. Pertea M, Pertea G, Antonescu C, Chang TC, Mendell JT, Salzberg SL. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat Biotechnol*. 2015; 33(3):290–5. <https://doi.org/10.1038/nbt.3122> PMID: 25690850
91. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014; 15(12):550. <https://doi.org/10.1186/s13059-014-0550-8> PMID: 25516281
92. Chen J, Bardes EE, Aronow BJ, Jegga AG. ToppGene Suite for gene list enrichment analysis and candidate gene prioritization. *Nucleic Acids Res*. 2009; 37:305–11. <https://doi.org/10.1093/nar/gkp427> PMID: 19465376
93. Wheeler DL, Church DM, Edgar R, Federhen S, Helmberg W, Madden TL, et al. Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res*. 2004; 32(1):13–6. <https://doi.org/10.1093/nar/gkh073> PMID: 14681353
94. Kong L, Zhang Y, Ye Z, Liu X, Zhao S, Wei L, et al. CPC: assess the protein-coding potential of transcripts using sequence features and support vector machine. *Nucleic Acids Res*. 2007; 35:345–9. <https://doi.org/10.1093/nar/gkm391> PMID: 17631615
95. Guo H, Zhu P, Yan L, Li R, Hu B, Lian Y, et al. The DNA methylation landscape of human early embryos. *Nature*. 2014; 511(7511):606. <https://doi.org/10.1038/nature13544> PMID: 25079557
96. Li L, Guo F, Gao Y, Ren Y, Yuan P, Yan L, et al. Single-cell multi-omics sequencing of human early embryos. *Nat Cell Biol*. 2018; 20(7):847–58. <https://doi.org/10.1038/s41556-018-0123-2> PMID: 29915357
97. Chen Y, Lyu R, Rong B, Zheng Y, Lin Z, Dai R, et al. Refined spatial temporal epigenomic profiling reveals intrinsic connection between PRDM9-mediated H3K4me3 and the fate of double-stranded breaks. *Cell Res*. 2020; 30(3):256–68. <https://doi.org/10.1038/s41422-020-0281-1> PMID: 32047271
98. Heinz S, Benner C, Spann NJ, Bertolino E, Lin YC, Laslo P, et al. Simple Combinations of Lineage-Determining Transcription Factors Prime cis-Regulatory Elements Required for Macrophage and B Cell Identities. *Mol Cell*. 2010; 38(4):576–89. <https://doi.org/10.1016/j.molcel.2010.05.004> PMID: 20513432