Original Article

The Genetic Association between CDKN1A and Heart Failure: Genome-Wide Exploration of m^{6}A-SNPs and Mendelian Randomization*

Ziyi Yang[1], Zhennan Lin[1], Xiaotong Ning[1], Xingbo Mo[2], Laiyuan Wang[1], Xiangfeng Lu[1], and Shufeng Chen[1,#]

1. Key Laboratory of Cardiovascular Epidemiology, Department of Epidemiology, Fuwai Hospital, National Center for Cardiovascular Diseases, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100037, China; 2. Center for Genetic Epidemiology and Genomics, Department of Epidemiology, Jiangsu Key Laboratory of Preventive and Translational Medicine for Geriatric Diseases, School of Public Health, Medical College of Soochow University, Suzhou 215123, Jiangsu, China

Abstract

Objective N6-methyladenosine (m^{6}A) is a common epigenetic modification in eukaryotes. In this study, we explore the potential impact of m^{6}A-associated single nucleotide polymorphisms (m^{6}A-SNPs) on heart failure (HF).

Methods Data from genome-wide association studies (GWAS) investigating HF in humans and from m^{6}A-SNPs datasets were used to identify HF-associated m^{6}A-SNPs. Their functions were explored using expression quantitative trait locus (eQTL), gene expression, and gene enrichment analyses. Mediation protein quantitative trait locus (pQTL)-Mendelian randomization (MR) was used to investigate the potential mechanism between critical protein levels and risk factors for HF.

Results We screened 44 HF-associated m^{6}A-SNPs, including 10 m^{6}A-SNPs that showed eQTL signals and differential expressions in HF. The SNP rs1801270 in CDKN1A showed the strongest association with HF (P = 7.75 × 10^{-6}). Additionally, MR verified the genetic association between the CDKN1A protein and HF, as well as the mediating effect of blood pressure (BP) in this pathway. Higher circulating level of CDKN1A was associated with a lower risk of HF (odds ratio [*OR*] = 0.82, 95% confidence interval [*CI*]: 0.69 to 0.99). The proportions of hypertension, systolic BP, and diastolic BP were 48.10%, 28.94%, and 18.02%, respectively. Associations of PDIA6 (P = 1.30 × 10^{-2}) and SMAD3 (P = 4.80 × 10^{-2}) with HF were also detected.

Conclusion Multiple HF-related m^{6}A-SNPs were identified in this study. Genetic associations of CDKN1A and other proteins with HF and its risk factors were demonstrated, providing new ideas for further exploration of the molecular mechanisms of HF.

Key words: Heart failure; N6-methyladenosine modification; Genome-wide association study; Expression quantitative trait locus; Mendelian randomization

Biomed Environ Sci, 2024; 37(12): 1397-1408 doi: 10.3967/bes2024.172 ISSN: 0895-3988

www.besjournal.com (full text) CN: 11-2816/Q

*This study was supported by the National Natural Science Foundation of China [82070473, 82170480, 82030102] and Chinese Academy of Medical Sciences Innovation Fund for Medical Sciences [2021-I2M-1-010].

#Correspondence should be addressed to Shufeng Chen, PhD, Tel: 86-10-60866589, E-mail: chenshufeng@fuwai.cams.cn

Biographical note of the first author: Ziyi Yang, female, born in 2000, MS Candidate, majoring in epidemiology and health statistics.

INTRODUCTION

Heart failure (HF) is a clinical syndrome caused by structural and/or functional cardiac abnormalities. HF is characterized by significant morbidity, mortality and a high economic burden, with the prevalence and years lived with disability rates of 711.90 and 63.92 per 100,000 population in 2019[1], affecting approximately 64 million patients. Furthermore, the prevalence of HF is estimated to increase due to an aging population[2].

HF is induced by several aspects of risk factors such as traditional cardiometabolic factors, lifestyle factors, genetic and epigenetic factors, etc. HF has substantial heritability of approximately 15%–35%[3,4]. More than 50 susceptibility loci for HF have been identified in genome-wide association studies (GWASs) in different populations[5–7]. In a GWAS of over 90,000 HF cases and more than one million control individuals, seven proteins were identified as potential targets for preventative interventions[7]. The next challenge is to determine the roles of these variants in RNA regulation. Distinct regions of single nucleotide polymorphisms (SNP) affect RNA modifications in various ways, such as by changing modifiable nucleotides at modified positions or altering nucleotides around the modified sites, thereby indirectly influencing RNA modifications and their functions[8].

In the last few decades, an increasing body of evidence has shown that RNA translation is significantly influenced by epigenetic modifications and transcription factors. Indeed, N6-methyladenosine (m^6A) is one of the most common post-transcriptional modifications in eukaryotes. More than 170 RNA modifications were identified, among which m^6A was found to mediate over 80% of RNA methylation[9], and play roles in mRNA, rRNA, snRNA, and lncRNA regulation[10]. The progression of m^6A modification is dynamic and reversible, and three protein regulators participate in this process: m^6A methyltransferases (writers), m^6A binding proteins (readers), and demethylases (erasers). These affect mRNA stability[11], splicing[12], translation[13], nuclear export[14], and other processes involved in various stages of the RNA lifecycle. Consequently, m^6A modification plays roles in the occurrence and development of various diseases.

Recent evidences have shown that m^6A is a candidate biomarker and therapeutic target for cardiovascular diseases. Additionally, the landscape of m^6A has been observed to be altered in HF[15], and multiple studies have indicated that m^6A influences a wide range of diseases in its progression, such as atherosclerosis, cardiac hypertrophy, and myocardial infarction[16]. For example, methyltransferase-like 14 (METTL14) promotes forkhead box O1 (FOXO1) expression by enhancing its m^6A modification, inducing an endothelial cell inflammatory response and atherosclerotic plaque formation[17]. Moreover, the association between m^6A modifications and gene expression in whole-transcriptome m^6A confirmed that genetic variants influence m^6A by changing the RNA sequences at the modifiable position or in the flanking region[8]. However, whether m^6A-SNPs affect HF remains unclear. Therefore, in this study, we aimed to identify and annotate HF-associated m^6A-SNPs. First, we explored HF-associated m^6A-SNPs in GWAS and m^6A datasets. We then evaluated the effects of these variants on gene expression using expression quantitative trait locus (eQTL) and gene expression analyses, which further supported the regulatory effect of the mutation sites. Finally, we conducted Mendelian randomization (MR) analyses to clarify the genetic relationships between these genes and HF risk factors.

METHODS

Determination of HF-associated m^6A-SNPs

To explore the predominant m^6A-SNPs associated with HF, we overlapped the published GWAS and m^6A-SNP data (Figure 1). A total of 352,014 genetic variants from different populations (mainly European and Asian) were downloaded from the m^6AVar database (m6avar.renlab.org/download.html)[18], which included functional variants involved in RNA modifications. The m^6A-SNPs were divided into three confidence levels: 1) high confidence, with variants located near the RNA modification sites and verified by immunoprecipitation (mi CLIP) or photo-crosslinking-assisted m^6A sequencing (PA-m^6A-seq) experiments; 2) medium confidence, with modification sites generated from methylated RNA immunoprecipitation (MeRIP-Seq) experiments; and 3) low medium confidence, with predictions based on the random forest algorithm performed from sequences around all variants from Single Nucleotide Polymorphism Database (dbSNP) and The Cancer Genome Atlas (TCGA).

Three GWAS summary datasets were included in this study: (1) a 2019 HF GWAS from the BioBank Japan Project (BBJ) dataset, which contains 9,413 chronic HF cases and 203,040 controls from East

Asia[19]; (2) a 2021 HF GWAS from a meta GWAS dataset (BBJ, UK biobank, and Finngene), which contains 10,540 congestive HF cases and 168,186 controls from a cross-population (East Asian and European)[20]; (3) a 2022 HF GWAS from a meta GWAS dataset (HERMES, Penn Medicine Biobank, etc), which contains 115,150 chronic HF or all-cause HF cases and 1,550,331 controls from a cross-population (mainly European)[6]. The SNPs were screened at a threshold of 1.00×10^{-4}. Manhattan plots were generated using the 'CMplot' R package[21]. Enrichment analyses of m^{6}A-SNPs passing the threshold were performed using Metascape (http://metascape.org/) to categorize and investigate the functional pathways of these genes. Metascape integrates comprehensive gene list annotations, including Kyoto Encyclopedia of Genes and Genomes (KEGG) and Gene Ontology (GO)[22].

eQTL Analysis

To investigate whether HF-associated m^{6}A-SNPs affect transcriptional regulation and gene expression at the genome-wide level, we performed eQTL analysis using HaploReg v4.2 (https://pubs.broadinstitute.org/mammals/haploreg/haploreg.php) and GTEx V8 (https://www.gtexportal.org/home/). HaploReg is an efficient tool for exploring the annotations of variants from the 1000 Genomes Project, which was designed to develop mechanistic hypotheses on the impact of variants on phenotypes[23]. HaploReg also aggregates eQTL data from numerous studies conducted on different human cell lines and tissues. Additionally, GTEx was used to study tissue-specific gene expression and regulation. We searched for possible functions of the m^{6}A-SNPs in the aorta, coronary artery, right atrial appendage, and left ventricle using the eQTL calculator in GTEx. Finally, we annotated important m^{6}A-SNPs in the UCSC browser (http://www.genome.ucsc.edu/), which provides a comprehensive and interactive visualization of genomic data, such as DNA sequences, gene structures, and regulatory elements in Encyclopedia of DNA Elements (ENCODE). The eQTL analysis revealed potential functional interactions between m^{6}A-SNPs and RNA.

Gene Expression Analysis

For the m^{6}A-SNPs that showed eQTL signals, we further explored whether the genes were differentially expressed in HF compared with healthy controls. We retrieved three expression profiles from the Gene Expression Omnibus (GEO), GSE168742, GSE141910, and GSE95140, containing 1,355 HF cases and 316 controls. The GSE168742 expression profile comprises integrated single-cell

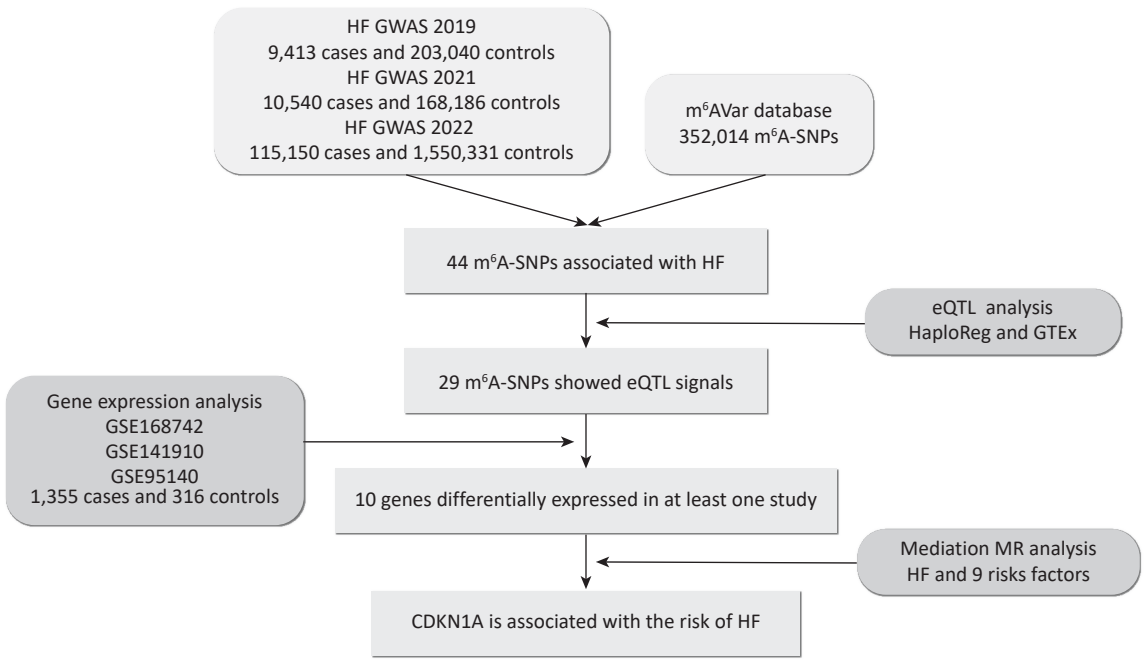

Figure 1. Flow chart illustrating the design and main results of the study. eQTL, expression quantitative trait locus; GWAS, genome-wide association study; HF, heart failure; m^{6}A, N6-methyladenosine; MR, Mendelian randomization; SNP, single nucleotide polymorphism.

RNA-seq and bulk RNA-seq data of human cardiomyocytes and cardiac fibroblasts from 678 cases with HF and 84 controls[24]. The GSE141910 expression profile comprises human cardiac tissue from 189 cases with HF and 161 healthy donors[25], of which the HF cases consisted of 161 cases with dilated cardiomyopathy and 28 cases with hypertrophic cardiomyopathy. The GSE95140 expression profile comprises human cardiomyocytes from 488 HF cases with dilated cardiomyopathy and 71 controls[26]. In R, we used the 'limma' package[27] to analyze bulk RNA-seq gene expression data. The threshold was set as follows to recognize the differences in expression between HF cases and controls: P_{adj} value < 0.05, and an absolute value of log2 fold change (FC) (|logFC|) > 1.

pQTL-MR and Mediation Analysis

Genes affected by m^{6}A-SNPs may further influence HF at the protein level via RNA translation. To explore how these genes influence HF, we selected m^{6}A-SNPs with eQTL signals to perform MR analysis. Four genes with available protein quantitative trait locus (pQTL) data were selected for subsequent analyses: CDKN1A, SMAD3, PDIA6, and TTN. To identify the potential functional mechanisms of these four critical genes, we performed a two-step MR to explore the possible pathways involved. We first conducted a two-sample MR to investigate the causal association between plasma proteins and HF. A two-sample bidirectional MR was also performed to evaluate the mutual causality of HF on critical proteins. We then performed a two-sample MR to identify whether risk factors for HF, including hypertension, blood pressure (BP), obesity, body mass index (BMI), diabetes, coronary heart disease (CHD), cardiomyopathy, and myocardial infarction (MI), may mediate the pathways from plasma protein to HF. Benjamini–Hochberg false discovery rate (FDR) adjustments were used for multiple tests. The effects of proteins on HF outcomes were designated as the total effect (β_0). The effects of proteins on each mediator were estimated as β_1, and the effects of mediators on HF outcomes were estimated as β_2. The mediation proportion of mediators in the causal association between proteins and HF outcomes was calculated as the product of β_1 and β_2 divided by β_0[28], and the 95% confidence intervals (*CI*) of the mediation proportions were calculated using the delta method[29].

The pQTL data were obtained from the Decode study containing 100,000 individuals and 4,970 proteins[30]. Summary-level GWAS outcome data were downloaded from the MRC Integrative Epidemiology Unit (IEU) Open GWAS database (https://gwas.mrcieu.ac.uk/)[31] (Supplementary Table S1, available in www.besjournal.com).

All MR analyses were performed using the 'TwosampleMR'[31] R package. For the two-sample MR, we selected significant ($P < 5.00 \times 10^{-6}$) and independent (linkage disequilibrium $r^2 < 0.01$, > 10,000 kb) SNPs as instrumental variables. Five MR methods were used: inverse-variance weighted, weighted median, MR-Egger, simple mode, and weighted mode. For multiple testing-adjusted significance thresholds, we selected a liberal FDR threshold of 20% to explore biological characteristics. For sensitivity analyses, heterogeneity and pleiotropy tests were performed using the MR-Egger and inverse-variance weighted methods. The MR-Egger method was used to provide a credible estimate when there is potential horizontal pleiotropy[32]. Calculations using the delta method were conducted using the 'RMediation' R package[33], with the default algorithm.

RESULTS

HF-associated m^{6}A-SNPs

We found 44 m^{6}A-SNPs associated with HF (Figure 2A). Among those SNPs, 29 were located in coding sequences (CDS), 11 in 3' untranslated regions (UTR'3), 3 in 5' untranslated regions (UTR'5), and 1 in an intron (Figure 2B). Regarding the confidence level, 1 SNP was high (mi CLIP/PA-m^{6}A-Seq), 4 were medium (MeRIP-Seq), and 39 were low (transcriptome-wide prediction). The most significant SNP, rs1801270, which is located in a CDS, and has a medium confidence level, was significant in both GWAS2019 ($P = 7.75 \times 10^{-6}$) and GWAS2022 ($P = 4.89 \times 10^{-4}$). Enrichment analyses indicated that these genes were predominantly enriched in the positive regulation of protein import into the nucleus, transition of the mitotic cell cycle, and establishment of protein localization to organelles (Figure 2C).

The Impact of m^{6}A-SNPs on Gene Expression

The effects of m^{6}A modification predominantly act through the translation or degradation of mRNA, and m^{6}A-SNPs therefore likely affect RNA expression. Among 44 m^{6}A-SNPs, 29 SNPs showed eQTL signals in the HaploReg browser. Among these signals, rs1048414 hits the highest count with 196 records

and was associated with the mRNA levels of HLA-DQA1 and HLA-DQA2 in most cells and tissues. The SNP rs3734264 hit 65 records and was associated with the mRNA levels of bridge-like lipid transfer protein family member 3A (BLTP3A, also known as UHRF1BP1) and small nuclear ribonucleoprotein polypeptide C (SNRPC) in most cells and tissues. Most SNPs regulate RNA translation through histone marks, protein binding, and motif changes.

We further searched for eQTL signals in specific HF-associated tissues using the GTEx V8 eQTL calculator and found that 8 SNPs showed significant signals in the related tissues. Both rs1731259 and rs1731260 were associated with KCNK3 expression in the right atrial appendage. The rs4233729 was associated with TRMT61B expression levels (P = 7.40×10^{-34}, 3.00×10^{-17}, 1.80×10^{-27}, and 9.10×10^{-15}), and rs3734264 with UHRF1BP1 expression levels (P = 2.30×10^{-55}, 7.60×10^{-35}, 1.00×10^{-14}, and 2.40×10^{-24}), in the aorta, coronary artery, right atrial appendage, and left ventricle, respectively (Figure 3).

Target Genes Differentially Expressed in HF

Gene expression may be alerted in HF and related diseases. Cardiomyopathies encompass a diverse array of heart muscle disorders that are considered as significant risk factors contributing to the development of HF. To investigate whether the m^{6}A-SNPs related genes can alter mRNA levels, we also conducted gene expression analyses in HF with different etiologies. The GSE168742, GSE141910, and GSE95140 datasets were used, and included 1,355 HF cases and 316 controls. Of the 28 genes, 10 were differentially expressed between the cases and controls (P_{adj} < 0.05, |log2FC| > 1).

In GSE168742, TTN (P = 1.76×10^{-36}, logFC = –3.10), GIMAP7 (P = 5.01×10^{-5}, logFC = 2.37), CDKN1A (P = 2.06×10^{-11}, logFC = 3.32), PPP1CB (P = 2.07×10^{-6}, logFC = 1.64), and WFS1 (P = 8.94×10^{-8}, logFC = –1.86) showed different levels of RNA expression between cases with HF and controls. In GSE141910, HLA-DQA1 (P = 8.30×10^{-23}, logFC = –1.31) was differentially expressed between cases with HF and controls. In GSE95140, KDM3A (P = 2.57×10^{-4}, logFC = –1.34), SHPRH (P = 2.31×10^{-9}, logFC = –1.83), GIMAP7 (P = 3.36×10^{-10}, logFC = 3.74), TRIM27 (P = 3.29×10^{-3}, logFC = –1.19), CDKN1A (P = 8.63×10^{-3}, logFC = 1.52), UHRF1BP1 (P = 4.45×10^{-7}, logFC = –1.30), and TTN (P = 4.94×10^{-6}, logFC = –1.19) were found to be differently expressed between cases with HF and controls. These results are shown in Figure 4.

Integrative Analyses Indicate Several SNPs may Potentially Affect m^{6}A Modification

By integrating all evidence (Supplementary Table S2, available in www.besjournal.com), we identified 10 SNPs with eQTL signals, and the associated genes were differentially expressed in HF (Table 1). The SNP rs1801270, a missense variant from C to A, causes amino acid transversion of serine (Ser) to arginine (Arg). The SNP rs1801270 is located in the exon region of the CDKN1A gene on chromosome 6. It was associated with chronic HF in both the cross-

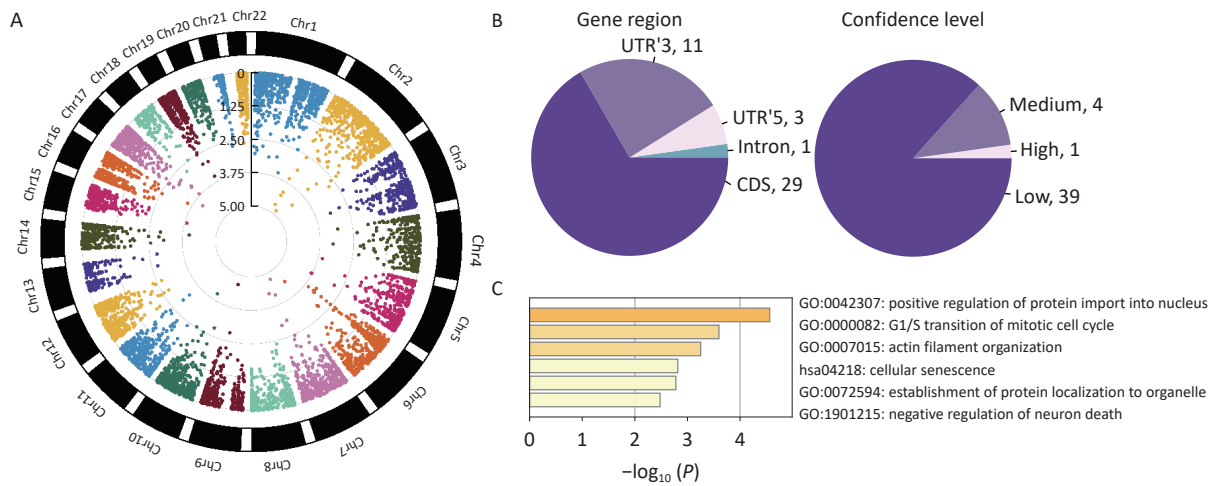

Figure 2. Characteristics of identified HF-associated m^{6}A-SNPs. (A) Circle Manhattan plot of GWAS2022 (which overlapped most SNPs with the m^{6}A-SNP dataset). (B) Information on m^{6}A-SNPs associated with heart failure. (C) Gene ontology terms of genes at which m^{6}A-SNPs are placed (Metascape: http://metascape.org/).

ancestry GWAS (P = 4.89 × 10^{-4}) and the East Asian GWAS (P = 7.75 × 10^{-6}), and its m^6A-SNP confidence level was medium (resulting from MeRIP-seq data). The SNP rs1801270 affected the expression of RAB44 in fibroblasts, affecting only one eQTL. Additionally, this SNP was found to affect four motifs: 2 protein-bound areas (POL2 and ZBTB), DNase in 8 tissues, and enhancer histone marks in 16 tissues. CDKN1A expression was higher in cases with HF than in controls in both GSE168742 (P = 2.06 × 10^{-11}) and GSE95140 (P = 8.63 × 10^{-3}).

The SNP rs3735081 was associated with chronic HF in a cross-ancestry GWAS (P = 5.48 × 10^{-5}) and showed 6 eQTL signals. Differential expression analyses showed that GIMAP7 expression was higher in patients with HF than in controls in GSE168742 and GSE95140 (P = 5.01 × 10^{-5} and P = 3.36 × 10^{-10}, respectively). The RNA immunoprecipitation and microarray analysis (RIP-chip) indicated a potential interaction between rs3735081 and PABPC1 (Supplementary Figure S1A, available in www.besjournal.com). The SNP rs3734264 was associated with chronic HF in a cross-ancestry GWAS (P = 4.98 × 10^{-4}) and showed 65 eQTL signals. The GTEx eQTL calculator also indicated its effect on UHRF1BP1 expression in the aorta, coronary artery, right atrial appendage, and left ventricle. Different expression analyses showed that UHRF1BP1 expression was lower in patients with HF than in the controls in GSE95140 (P = 4.45 × 10^{-7}). Furthermore, RIP-chip analysis indicated a potential interaction between rs3734264 and IGF2BP1 (Supplementary Figure S1B).

MR Analysis Indicated the Association of Critical Proteins with Heart Failure and Its Risk Factors

First, we tested whether four proteins were associated with HF. Indeed, MR results showed that CDKN1A (P = 4.90 × 10^{-2}), PDIA6 (P = 1.30 × 10^{-2}), and SMAD3 (P = 4.80 × 10^{-2}) were associated with HF, whereas the association between TTN and HF was not statistically significant (Figure 5). Higher circulating levels of both CDKN1A (*OR* = 0.825, 95% *CI*, 0.681–0.999) and PDIA6 (*OR* = 0.755, 95% *CI*, 0.605–0.941) were associated with a lower risk of HF, while higher circulating levels of SMAD3 (*OR* = 1.235, 95% *CI*, 1.002–1.521) were associated with a higher risk of HF. In further bidirectional MR analyses, no reverse causality was observed for genetically predicted HF on CDKN1A (P = 0.20),

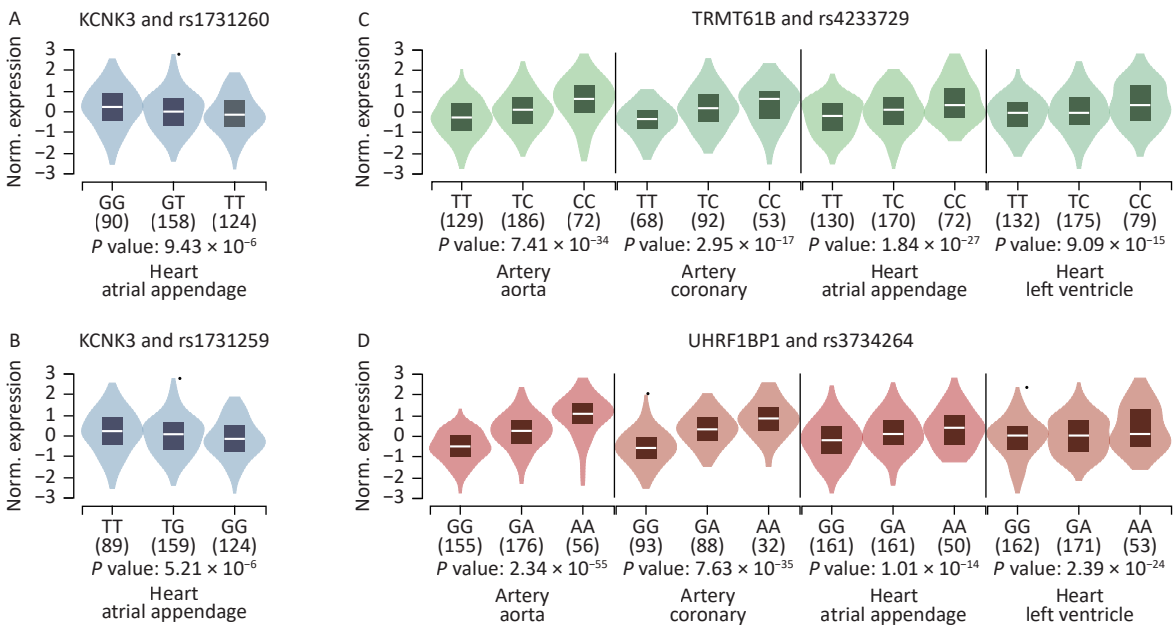

Figure 3. Violin plots of the eQTLs of some significant SNPs. *P* values were calculated using the eQTL calculator in GTEx V8 (GTEx Portal). (A) The effect allele of rs1731260 was associated with KCNK3 expression in the right atrial appendage. (B) The effect allele of rs1731259 was associated with KCNK3 expression in the right atrial appendage. (C) The effect allele of rs4233729 was associated with TRMT61B expression in the aorta, coronary artery, right atrial appendage, and left ventricle. (D) The effect allele of rs3734264 was associated with UHRF1BP1 expression in the aorta, coronary artery, right atrial appendage, and left ventricle.

PDIA6 (P = 0.10), or SMAD3 levels (P = 0.24) using the inverse variance weighted (IVW) method (Supplementary Table S3, available in www.besjournal.com).

We then searched for causal associations between these proteins and a range of risk factors for HF. Among all risk factors assessed, the circulating level of CDKN1A was inversely associated with hypertension (*OR* = 0.985, 95% *CI*: 0.971 to 0.999), systolic BP (SBP) (β = −0.11, 95% *CI*: −0.197 to −0.023), and diastolic BP (DBP) (β = −0.12, 95% *CI*: −0.209 to −0.03). Circulating PDIA6 level was inversely associated with CHD (*OR* = 0.991, 95% *CI*, 0.985–0.997) and MI (*OR* = 0.993, 95% *CI*, 0.988–0.999). Circulating SMAD3 level was positively associated with CHD (*OR* = 1.008, 95% *CI*, 1.002–1.015). The circulating level of TTN was positively associated with DBP (β = 0.062, 95% *CI*: 0.006 to 0.117) and BMI (β = 0.111, 95% *CI*: 0.046 to 0.176). The aforementioned effects were significant for at least one method, and all five MR methods showed consistent associations. Statistically significant associations of CDKN1A, PDIA6 and SMAD3 with HF and its risk factors are shown in Supplementary Tables S4.1, 4.2, and 4.3 (available in www.besjournal.com).

Mediator Analysis Showed CDKN1A and Other Proteins may Affect HF by Risk Factor Pathways

We further evaluated the potential mechanisms by which the significantly associated risk factors may influence the relationship between proteins and HF outcomes. The SNP rs1801270, which is located in the exon region of CDKN1A and has a medium confidence level, was associated with chronic HF in both the cross-ancestry GWAS ($P = 4.89 \times 10^{-4}$) and the East Asian GWAS ($P = 7.75 \times 10^{-6}$). Differential expression analyses showed that CDN1KA was highly expressed in both dilated cardiomyopathy and HF samples when compared to controls. We analyzed hypertension and BP as mediators of the pathway from CDKN1A to HF, as shown in Figure 6A. The association between CDKN1A and HF risk is likely mediated by hypertension and high BP. The total and mediation effects of the three mediators are shown in Figure 6B–D. The mediation effects ($\beta_1 \times \beta_2$) of SBP, DBP, and hypertension were −0.056 (95% *CI*: −0.111 to −0.011), −0.035 (95% *CI*: −0.076 to −0.005)

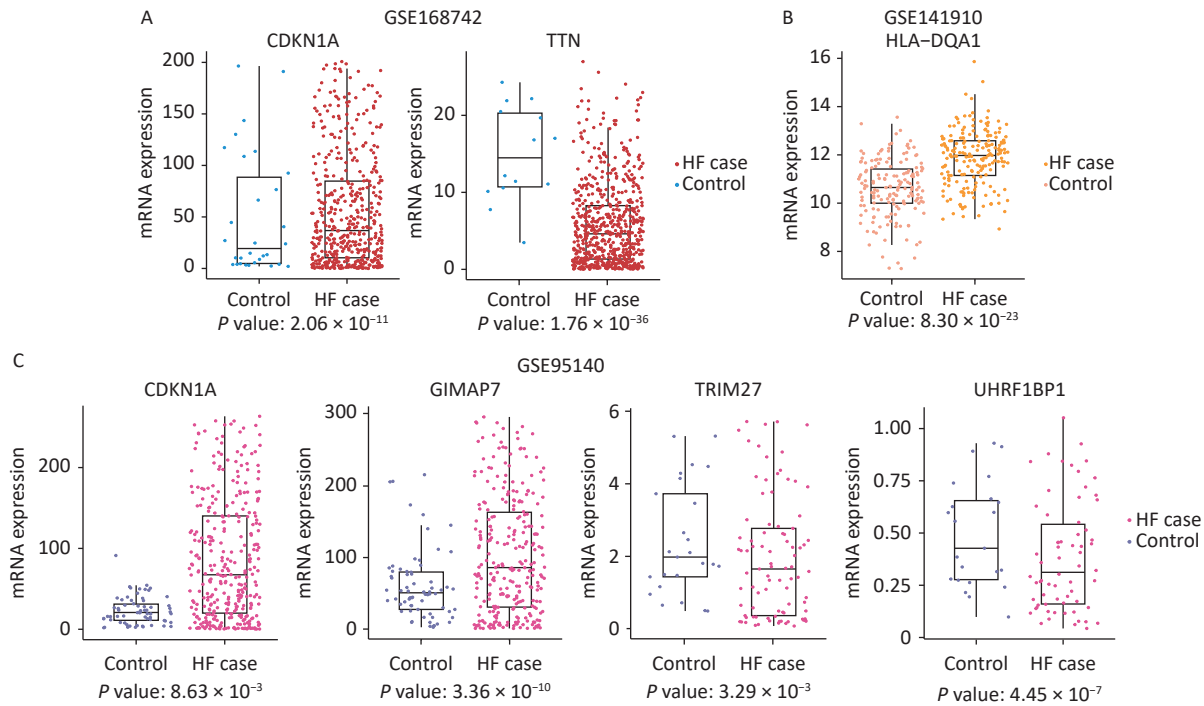

Figure 4. Box plots depicting the gene expression of the three datasets. (A) Differentially expressed genes found in GSE168742: expression of CDKN1A and TTN in HF cases and controls. (B) Differentially expressed genes found in GSE141910: expression of HLA−DQA1 in HF cases and controls. (C) Differentially expressed genes found in GSE95140: expression of CDKN1A, GIMAP7, TRIM27, and UHRF1BP1 in HF cases and controls. HF: heart failure.

and −0.047 (95% *CI*: −0.099 to −0.003), respectively. Hypertension described 48.10% (95% *CI*: 1.36%–51.58%) of the total effect of CDKN1A on HF, followed by SBP (28.94%, 95% *CI*: 5.61%–57.32%) and DBP (18.02%, 95% *CI*: 2.84%–39.48%) (Figure 7). Overall, this suggests that CDKN1A may protect against HF by inhibiting the BP pathway.

Following the same design, we subsequently analyzed the role of CHD and MI as mediators in the pathway from PDIA6 to HF and CHD as a mediator in the pathway from SMAD3 to HF. The total and mediation effects of the mediators are shown in Figure 6E–G. CHD explained 21.20% (95% *CI*: 4.80%–37.60%) of the total effect of PDIA6 on HF, followed by MI (18.17%, 95% *CI*: 2.07%–34.27%). CHD explained 26.55% (95% *CI*: 4.52%–48.57%) of the total effect of SMAD3 on HF (Figure 7).

DISCUSSION

In this study, we identified HF-associated m^{6}A-SNPs by exploring large-scale GWAS data and performing eQTL analysis. Some genes, including CDKN1A, showed differential expression levels in cases with HF when compared to controls. The genetic association between plasma CDKN1A protein

Table 1. The significant m^{6}A-SNPs for heart failure

rsID	Chr	Effect allele frequency	β	*P*	Gene	Gene region	Confidence level	HaploReg	GTEx	GSE168742	GSE95140	GSE141910
rs1801270*	6	0.068	0.015	7.75×10^{-6}	CDKN1A	CDS	Medium	1	N	3.315	1.525	–
rs2030259	2	0.225	0.027	1.41×10^{-5}	KDM3A	CDS	Low	2	N	–	−1.344	–
rs6900607	6	0.742	−0.101	4.21×10^{-5}	SHPRH	CDS	Low	2	N	–	−1.833	–
rs3735081	7	0.330	0.022	5.48×10^{-5}	GIMAP7	CDS	Low	6	N	2.367	3.742	–
rs1048414	6	0.604	0.051	2.49×10^{-4}	HLA−DQA1	CDS	Low	196	N	–	–	−1.314
rs11545587	6	0.105	0.029	4.74×10^{-4}	TRIM27	UTR5	Medium	2	N	–	−1.187	–
rs1801270	6	0.238	0.029	4.89×10^{-4}	CDKN1A	CDS	Medium	1	N	3.315	1.525	–
rs3734264	6	0.353	0.019	4.98×10^{-4}	UHRF1BP1 (BLTP3A)	CDS	Low	65	Y	–	−1.305	–
rs1128416	2	0.443	−0.018	7.14×10^{-4}	PPP1CB	CDS	Low	57	N	1.641	–	–
rs66523653	2	0.046	−0.192	8.44×10^{-4}	TTN	CDS	Low	1	N	−3.098	−1.187	–
rs1801206	4	0.585	0.018	8.65×10^{-4}	WFS1	CDS	Low	8	N	−1.858	–	–

Note. *This record is the result of the BioBankJapan Project (BBJ) genome−wide association studies (GWAS) dataset (GWAS 2019), and the rest are from the meta−database (GWAS 2022). rsID: number of SNP; Chr: chromosome; HaploReg: number of eQTL signals from HaploReg v4.2; GTEx: eQTL signals from GTEx V8 (in artery aorta, artery coronary, heart atrial appendage, and heart left ventricle); GSE16874: logFC between the heart failure group and control group; GSE95140: logFC between the HF group and control group; GSE141910: logFC between the HF group and control group. The "–" in these three columns means no significant expression. The "Y" in GTEx means the SNP showed an eQTL effect in the specific tissue described above, and the "N" means the SNP didn't show eQTL effect in a specific tissue.

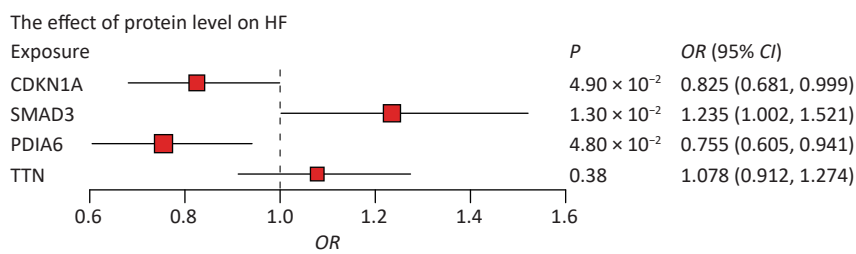

Figure 5. Associations of protein levels with HF. The point estimates are represented by a square along with the 95% confidence intervals. The size of the squares represented the standard error. *CI*, confidence interval; HF, heart failure; *OR*, odds ratio.

and HF was verified by MR analysis. Mediation analysis indicated that CDKN1A may influence HF through BP regulation. Consequently, we revealed the role of CDKN1A in HF development through its contribution to RNA modification, gene expression, and protein action, thereby providing new ideas for selecting therapeutic targets.

Recent studies have examined the relationships of m^{6}A-SNP with BP[34] and coronary artery disease[35]. Indeed, GWAS data has revealed associations between SNPs located in CDKN1A and HF-associated risk traits, such as left ventricular systolic function (rs4135240)[5], pulse pressure (rs1801270)[36], hypertrophic cardiomyopathy (rs3176326)[37], and cardiovascular biological trait (rs2376620)[38]. In another HF GWAS, CDKN1A (rs4135240) was found to be associated with HF and reduced left ventricular systolic function[5]. In the present study, we examined HF-associated m^{6}A-SNPs and found that 10 SNPs in CDKN1A and other genes showed eQTL effects in various tissues and cells. Susceptibility genes were significantly expressed in HF cases. Enrichment analysis revealed that these genes were enriched in processes related to positive

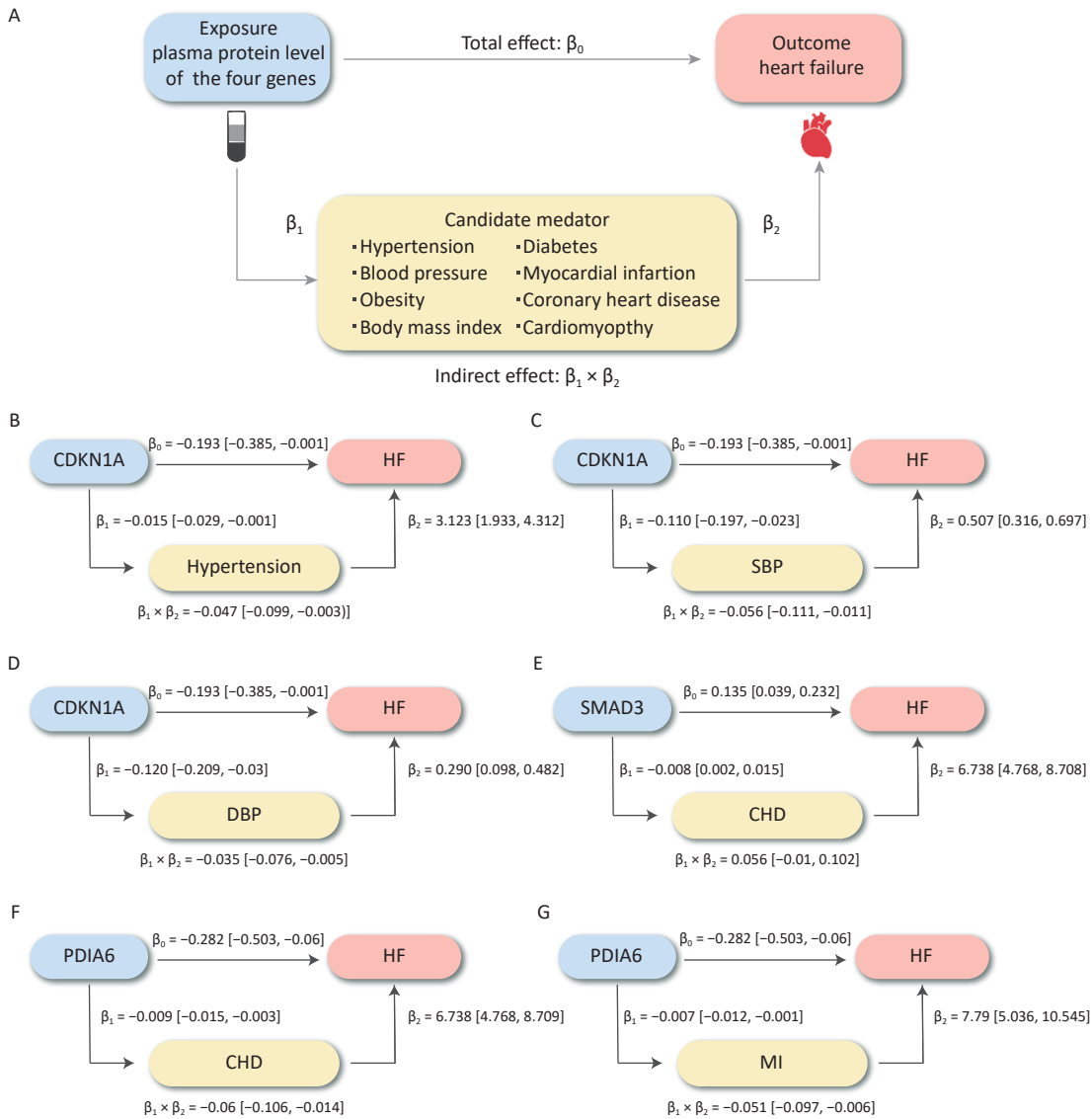

Figure 6. The causal effect of plasma protein on HF risk mediated by risk factors. (A) The overall design of mediation MR. (B–G) The effects of three proteins on HF, hypertension, SBP, DBP, CHD, and MI and the mediation proportion of three mediators. CHD, coronary heart disease; DBP, diastolic blood pressure; HF, heart failure; MI, myocardial infarction; MR, Mendelian randomization; SBP, systolic blood pressure.

regulation of protein import, transition of mitotic cell cycle, and establishment of protein localization to organelles. These results indicate that these genes may play roles in the regulation of molecular functions. Moreover, CDKN1A was identified as a critical gene and subsequently included in the MR analysis. The role of CDKN1A in cardiovascular diseases has been previously reported and validated through recent experiments on the regulation of various cardiovascular risk factors. CDKN1A encodes p21, a potent cell cycle inhibitor of postnatal cardiomyocyte cell cycle arrest[39], and regulates growth-related processes in cardiac hypertrophy[40]. Research has indicated that CDKN1A is repressed by hypoxia-inducible factor 1α (HIF1α), thereby affecting the proliferation of hypoxic fetal cardiomyocytes[41]. Whether the m6A-SNPs in CDKN1A affect the development of HF through these processes needs to be verified.

Both PDIA6 and SMAD3 were found to be associated with HF in an MR study. PDIA6 encodes a member of the disulfide isomerase family that catalyzes protein disulfide bond formation. It also protects cardiomyocytes from ischemia/reperfusion-induced death by promoting disulfide bond formation and enhancing ER protein folding[42]. In a mouse model of myocardial ischemia/reperfusion, PDIA6 was experimentally validated as a gene regulated by alternative splicing[43]. SMAD3 directly upregulates programmed cell death 5 (PDCD5) during cardiac fibrosis. SMAD3 has been shown to broadly regulate cardiac fibrosis through various pathways. For example, SMAD3 upregulates PDCD5 during myocardial fibrosis, which subsequently mitigates progressive fibrosis and cardiac dysfunction[44]. Additionally, SMAD3 critically regulates the function of infarcted macrophages by facilitating the phagocytic phenotype to an anti-inflammatory transition, thereby protecting the infarcted heart from adverse remodeling[45].

Our study is the first to report an association between the m^{6}A-SNPs and the risk of HF. We have provided detailed annotations for these m^{6}A-SNPs, which show that m^{6}A-SNPs are critical for regulating mRNA in HF. Moreover, the strengths of our study include the relatively large sample size of the GWAS and the gene expression data used. Furthermore, we conducted extensive tests on the mediation of a broad range of risk factors in regulating HF progression, utilizing the largest available pQTL dataset to date.

Our study also has some limitations. First, the m^{6}A-SNPs we screened out did not reach genome-wide significance. Second, experiments on animal models and cell lines were not conducted to further verify causal associations. Third, although the results of the sensitivity analyses passed the tests, the causal associations should be interpreted with caution, as several assumptions of the methods are untestable and the heterogeneity of the instrumental variables and residual horizontal pleiotropy might still distort some results. Fourth, we utilized GWAS summary data rather than individual data and could not further explore the differences in subgroups, such as sex and age. Hence, the potential genes and m^{6}A-SNPs identified in this study require further verification to provide solid evidence for clinical diagnosis and prevention. Finally, as an exploratory study, we selected a relatively liberal FDR threshold to identify potentially meaningful results. This also means accepting a higher false positive rate, and future validation of the identified SNP loci and genes is necessary.

In conclusion, this study identified HF-related m^{6}A-SNPs, especially in the CDKN1A gene. CDKN1A showed a significant association with HF in MR

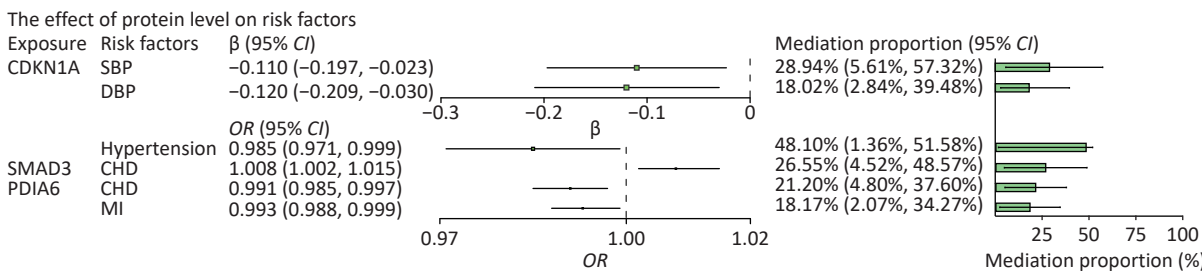

Figure 7. Effect of protein levels on mediators and mediating role of each mediator in associations of protein levels with heart failure. The point estimates were represented by a square along with the 95% confidence intervals. The size of the squares represented the standard error. CHD, coronary heart disease; *CI*, confidence interval; DBP, diastolic blood pressure; HF, heart failure; MI, myocardial infarction; *OR*, odds ratio; SBP, systolic blood pressure.

analysis, and the effects were mediated by BP. These findings, if confirmed in larger studies, may have further implications for the development and prevention of HF.

DATA SHARING STATEMENT

The data analyzed and reported in this manuscript are publicly available and can be accessed from various databases.

CONFLICT OF INTEREST

The authors have no competing financial interests or conflicts of interest to disclose.

AUTHOR CONTRIBUTIONS

Concept and design: Ziyi Yang, Xingbo Mo, and Shufeng Chen; Statistical analysis and interpretation of data: Ziyi Yang, Xiaotong Ning, and Zhennan Lin; Drafting the manuscript: Ziyi Yang; Critical revision of the manuscript: Xingbo Mo, Laiyuan WANG, Xiangfeng Lu, and Shufeng Chen; Funding: Laiyuan Wang, Xiangfeng Lu, and Shufeng Chen; Supervision: Shufeng Chen. All the authors reviewed the final manuscript.

Received: April 1, 2024;
Accepted: September 3, 2024

REFERENCES

1. Yan T, Zhu SJ, Yin XJ, et al. Burden, trends, and inequalities of heart failure globally, 1990 to 2019: a secondary analysis based on the global burden of disease 2019 study. J Am Heart Assoc, 2023; 12, e027852.
2. Bui AL, Horwich TB, Fonarow GC. Epidemiology and risk profile of heart failure. Nat Rev Cardiol, 2011; 8, 30–41.
3. Lindgren MP, Pirouzifard M, Smith JG, et al. A Swedish nationwide adoption study of the heritability of heart failure. JAMA Cardiol, 2018; 3, 703–10.
4. Reza N, Owens AT. Advances in the genetics and genomics of heart failure. Curr Cardiol Rep, 2020; 22, 132.
5. Shah S, Henry A, Roselli C, et al. Genome-wide association and Mendelian randomisation analysis provide insights into the pathogenesis of heart failure. Nat Commun, 2020; 11, 163.
6. Levin MG, Tsao NL, Singhal P, et al. Genome-wide association and multi-trait analyses characterize the common genetic architecture of heart failure. Nat Commun, 2022; 13, 6914.
7. Rasooly D, Peloso GM, Pereira AC, et al. Genome-wide association analysis and Mendelian randomization proteomics identify drug targets for heart failure. Nat Commun, 2023; 14, 3826.
8. Luo XT, Li HQ, Liang JQ, et al. RMVar: an updated database of functional variants involved in RNA modifications. Nucleic Acids Res, 2021; 49, D1405–12.
9. Wiener D, Schwartz S. The epitranscriptome beyond m6A. Nat Rev Genet, 2021; 22, 119–31.
10. Gomes CPC, Schroen B, Kuster GM, et al. Regulatory RNAs in heart failure. Circulation, 2020; 141, 313–28.
11. Wang X, Lu ZK, Gomez A, et al. N6-methyladenosine-dependent regulation of messenger RNA stability. Nature, 2014; 505, 117–20.
12. Xiao W, Adhikari S, Dahal U, et al. Nuclear m^{6}A reader YTHDC1 regulates mRNA splicing. Mol Cell, 2016; 61, 507–19.
13. Wang X, Zhao BS, Roundtree IA, et al. N6-methyladenosine modulates messenger RNA translation efficiency. Cell, 2015; 161, 1388–99.
14. Fustin JM, Doi M, Yamaguchi Y, et al. RNA-methylation-dependent RNA processing controls the speed of the circadian clock. Cell, 2013; 155, 793–806.
15. Berulava T, Buchholz E, Elerdashvili V, et al. Changes in m^{6}A RNA methylation contribute to heart failure progression by modulating translation. Eur J Heart Fail, 2020; 22, 54–66.
16. Kumari R, Ranjan P, Suleiman ZG, et al. mRNA modifications in cardiovascular biology and disease: with a focus on m^{6}A modification. Cardiovasc Res, 2022; 118, 1680–92.
17. Jian DD, Wang Y, Jian LG, et al. METTL14 aggravates endothelial inflammation and atherosclerosis by increasing FOXO1 N6-methyladeosine modifications. Theranostics, 2020; 10, 8939–56.
18. Zheng YY, Nie P, Peng D, et al. m6AVar: a database of functional variants involved in m^{6}A modification. Nucleic Acids Res, 2018; 46, D139–45.
19. Ishigaki K, Akiyama M, Kanai M, et al. Large-scale genome-wide association study in a Japanese population identifies novel susceptibility loci across different diseases. Nat Genet, 2020; 52, 669–79.
20. Sakaue S, Kanai M, Tanigawa Y, et al. A cross-population atlas of genetic associations for 220 human phenotypes. Nat Genet, 2021; 53, 1415–24.
21. Yin LL, Zhang HH, Tang ZS, et al. rMVP: a memory-efficient, visualization-enhanced, and parallel-accelerated tool for genome-wide association study. Genomics Proteomics Bioinformatics, 2021; 19, 619–28.
22. Zhou YY, Zhou B, Pache L, et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. Nat Commun, 2019; 10, 1523.
23. Ward LD, Kellis M. HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. Nucleic Acids Res, 2012; 40, D930–4.
24. Ko T, Nomura S, Yamada S, et al. Cardiac fibroblasts regulate the development of heart failure via Htra3-TGF-β-IGFBP7 axis. Nat Commun, 2022; 13, 3275.
25. Flam E, Jang C, Murashige D, et al. Integrated landscape of cardiac metabolism in end-stage human nonischemic dilated cardiomyopathy. Nat Cardiovasc Res, 2022; 1, 817–29.
26. Nomura S, Satoh M, Fujita T, et al. Cardiomyocyte gene programs encoding morphological and functional signatures in cardiac hypertrophy and failure. Nat Commun, 2018; 9, 4435.
27. Ritchie ME, Phipson B, Wu D, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res, 2015; 43, e47.
28. Carter AR, Sanderson E, Hammerton G, et al. Mendelian randomisation for mediation analysis: current methods and challenges for implementation. Eur J Epidemiol, 2021; 36, 465–78.
29. MacKinnon DP, Lockwood CM, Hoffman JM, et al. A comparison of methods to test mediation and other intervening variable effects. Psychol Methods, 2002; 7, 83–104.

30. Ferkingstad E, Sulem P, Atlason BA, et al. Large-scale integration of the plasma proteome with genetics and disease. Nat Genet, 2021; 53, 1712−21.
31. Hemani G, Zheng J, Elsworth B, et al. The MR-Base platform supports systematic causal inference across the human phenome. Elife, 2018; 7, e34408.
32. Sproviero W, Winchester L, Newby D, et al. High blood pressure and risk of dementia: a two-sample mendelian randomization study in the UK biobank. Biol Psychiatry, 2021; 89, 817−24.
33. Tofighi D, MacKinnon DP. RMediation: an R package for mediation analysis confidence intervals. Behav Res Methods, 2011; 43, 692−700.
34. Mo XB, Lei SF, Zhang YH, et al. Examination of the associations between m6A-associated single-nucleotide polymorphisms and blood pressure. Hypertens Res, 2019; 42, 1582−9.
35. Mo XB, Lei SF, Zhang YH, et al. Detection of M6A-associated SNPs as potential functional variants for coronary artery disease. Epigenomics, 2018; 10, 1279−87.
36. Surendran P, Feofanova EV, Lahrouchi N, et al. Discovery of rare variants associated with blood pressure regulation through meta-analysis of 1.3 million individuals. Nat Genet, 2020; 52, 1314−32.
37. Harper AR, Goel A, Grace C, et al. Common genetic variants and modifiable risk factors underpin hypertrophic cardiomyopathy susceptibility and expressivity. Nat Genet, 2021; 53, 135−42.
38. Yu MY, Harper AR, Aguirre M, et al. Genetic determinants of the interventricular septum are linked to ventricular septal defects and hypertrophic cardiomyopathy. Circ Genom Precis Med, 2023; 16, 207−15.
39. Tane S, Ikenishi A, Okayama H, et al. CDK inhibitors, p21Cip1 and p27Kip1, participate in cell cycle exit of mammalian cardiomyocytes. Biochem Biophys Res Commun, 2014; 443, 1105−9.
40. Hauck L, Grothe D, Billia F. p21CIP1/WAF1-dependent inhibition of cardiac hypertrophy in response to Angiotensin II involves Akt/Myc and pRb signaling. Peptides, 2016; 83, 38−48.
41. Guimarães-Camboa N, Stowe J, Aneas I, et al. HIF1α represses cell stress pathways to allow proliferation of hypoxic fetal cardiomyocytes. Dev Cell, 2015; 33, 507−21.
42. Vekich JA, Belmont PJ, Thuerauf DJ, et al. Protein disulfide isomerase-associated 6 is an ATF6-inducible ER stress response protein that protects cardiac myocytes from ischemia/reperfusion-mediated cell death. J Mol Cell Cardiol, 2012; 53, 259−67.
43. Ma N, Xu H, Zhang WH, et al. Genome-wide analysis revealed the dysregulation of RNA binding protein-correlated alternative splicing events in myocardial ischemia reperfusion injury. BMC Med Genomics, 2023; 16, 251.
44. Weng L, Ye JJ, Yang FH, et al. TGF-β1/SMAD3 regulates programmed cell death 5 that suppresses cardiac fibrosis post-myocardial infarction by inhibiting HDAC3. Circ Res, 2023; 133, 237−51.
45. Chen BJ, Huang SB, Su Y, et al. Macrophage Smad3 protects the infarcted heart, stimulating phagocytosis and regulating inflammation. Circ Res, 2019; 125, 55−70.