Clinical Epigenetics

RESEARCH

Open Access

Mendelian randomization provides a multi-omics perspective on the regulation of genes involved in ribosome biogenesis in relation to cardiac structure and function

Shuxu Wei[1,2†], Ronghuai Shen[1†], Xiaojia Lu[1], Xinyi Li[1,2], Lingbin He[1], Youti Zhang[3], Jiahang Yang[4], Zhouwu Shu[1*] and Xianxi Huang[1,2,3*]

Abstract

Background Ribosome biogenesis (RiboSis) is a complex process for generating ribosomes, the cellular machinery responsible for protein synthesis. Dysfunctional RiboSis can disrupt cardiac structure and function, contributing to cardiovascular diseases. This study employed a Mendelian randomization (MR) approach, integrating multi-omics data, to investigate the relationship between RiboSis-related genes and standard cardiac structure and function.

Methods We utilized summary stats for methylation, RNA splicing, and gene expression, and UK Biobank cardiopulm MRI genetic associations ($N = 41{,}135$). MR evaluated RiboSis gene features against traits, complemented by hypothesis prioritization for multi-trait colocalization (HyPrColoc) and colocalization. Composite scores ranked RiboSis genes, and phenome-wide association study (PheWAS) with scQTLbase instrumental variables (IVs) confirmed results.

Results We identified 15 RiboSis-related genes: HEATR1, SENP3, ERI1, ERCC2, TSR1, UTP11, DDX17, SMARCB1, NIP7, ERAL1, NOP56, RPL10A, EIF6, EXOSC9, and NOP58. Notably, HEATR1 and SENP3 were ranked in the top quartile (Q1), scoring 25. In validation cohort, 12 genes associated with cardiac structures, functions, diseases. Only ERAL1, TSR1, and NIP7 lacked significant associations with cardiac traits.

Conclusion Our multi-omics MR analysis identified 15 RiboSis-related genes associated with cardiac risk, with 12 further validated through gene set enrichment analysis. These findings suggest a link between RiboSis and cardiac health, enhancing understanding of cardiac disease mechanisms.

Keywords Ribosome biogenesis, Cardiac structure and function, Mendelian randomization, Multi-omics

†Shuxu Wei and Ronghuai Shen have contributed equally to this work.

*Correspondence:
Zhouwu Shu
shuzhw@163.com
Xianxi Huang
xianxihuang@sina.com
Full list of author information is available at the end of the article

© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

Introduction

Ribosome biogenesis (RiboSis) is a complex process essential for producing the ribosomes necessary for protein synthesis during cellular growth and proliferation [1, 2]. This tightly coordinated process involves three RNA polymerases, approximately 80 ribosomal proteins, and around 200 non-ribosomal trans-acting factors [3, 4]. RiboSis encompasses rRNA transcription, cleavage, modification, ribosome assembly, and ribosomal preparticle [5].

The heart, having evolved as a tubular organ over hundreds of millions of years, exhibits distinct left and right chambers originating from different progenitor cell populations and operating under disparate pressure regimes [6]. The left heart encounters high afterload, while the right heart typically faces relatively lower afterload [6]. Variations in left–right cardiac structure can contribute to various pathologies, including arrhythmogenic right ventricular cardiomyopathy (ARVC) and pulmonary hypertension [7, 8]. Conversely, cardiac diseases can also induce structural and functional abnormalities. For example, right or left ventricular dysfunction can arise from intrinsic myocardial diseases, coronary artery disease, or pulmonary or systemic hypertension [9].

Many studies have indicated that ribosomal dysfunction can lead to various cardiovascular diseases [10, 11], including myocardial hypertrophy [12], myocardial infarction [13–15], and atherosclerosis [16]. For example, Wang et al. [17] analyzed plasma samples from patients with familial hypercholesterolemia and found that RPL9, RPL35, RPS7, and RPL23 exhibited a downward trend. The expressed genes are mainly involved in the ribosome and oxidative phosphorylation pathways. Studies indicate that during myocardial infarction, the protein levels of Nucleolin (Ncl, the biogenesis of ribosomes begins with Ncl, where different rRNA subunits are transcribed by RNA polymerase I into precursor transcripts) significantly decrease. At the same time, overexpression of Ncl reduces the infarct area and cell death in rat hearts [18]. In murine models, dysregulation of nucleolar protein nucleostemin (NS), a component of RiboSis, has been shown to increase the vulnerability of diabetic hearts to ischemia/reperfusion injury [19]. Furthermore, a high-throughput phenotypic platform has identified ribosomal protein RPL13 as associated with congenital heart disease (CHD) [20]. Ribosomal proteins have also been implicated in the pathogenesis of transplant-associated coronary artery disease (TxCAD) following cardiac transplantation [21]. Moreover, aberrant regulation of certain ribosomal genes has been observed in various pathological cardiac conditions, including ischemia, diabetic cardiomyopathy, and hypertrophic or dilated cardiomyopathy [22]. The occurrence and development of various cardiac diseases are closely associated with abnormalities in the structure and function of the left and right cardiac systems. Therefore, in this study, we analyzed relevant cardiac structure and function features.

In recent years, Mendelian randomization (MR) studies have been widely applied in disease etiology. Without randomized controlled trials, MR is the most compelling strategy to explore causal relationships between exposure and outcome [23]. MR uses exposure-related single-nucleotide polymorphisms (SNPs) as instrumental variables (IVs) [24], and this IVs substitution method mimics randomized controlled trials, as SNPs are randomly assigned to offspring at conception, largely avoiding confounding, with individual characteristics such as gender and age less likely to bias causal effects [25]. Similarly, errors in reverse causation in MR studies are also less likely as genotype is formed before disease onset [26]. The increasing availability of large-scale genome-wide association study (GWAS) and molecular quantitative trait loci (QTL) data enables us to explore the causal relationships between RiboSi-related gene regulation and cardiac structure and function from the perspectives of methylation, RNA splicing, and expression levels. Methylation QTL (mQTL) refers to a genomic location where genetic variation affects the DNA methylation pattern of a specific site or region. Methylation is an epigenetic modification that regulates gene expression by adding a methyl group (a group comprising one carbon atom and three hydrogen atoms) to the DNA molecule without altering the DNA sequence itself. Splicing QTL (sQTL) refers to specific genomic loci where genetic variation is associated with changes in the RNA splicing process, leading to the production of different mRNA splice isoforms by affecting the choice of splicing sites or splicing modes. Such genetic variation may affect the binding sites of splicing factors or the splicing mechanism itself, altering the structure and function of proteins and ultimately impacting the phenotype or disease risk. Expression QTL (eQTL) refers to loci controlling gene expression levels associated with quantitative traits. Here, we utilize MR to investigate the potential associations between methylation, RNA splicing, and expression of RiboSi genes and 22 common structural and functional variations in the heart.

Methods

The current MR analysis is based on publicly available datasets. In this study, we conducted summary data-based MR (SMR) analysis [27] separately at the levels of methylation, RNA splicing, and expression with 22 common cardiac structural and functional variations. To strengthen causal inference, we employed the HEIDI test [27], colocalization analysis [28], and the hypothesis

prioritization for multi-trait colocalization (HyPrColoc) [29]. Finally, we integrated the results of the 22 cardiac-related variations across the three gene regulation levels for comprehensive evaluation. Causal candidate genes were identified by integrating results from these three different levels of MR analysis. There was no overlap between the sample populations of the exposure and outcome groups. Following the analyses above, we leveraged human single-cell eQTL (sc-eQTL) data to identify all SNPs associated with RiboSis-related genes identified in our study. Subsequently, we further performed individual phenome-wide association studies (PheWAS) for each SNP using FinnGen data to validate these genes' association with cardiac phenotypes. Our research flowchart is shown in Fig. 1.

Sources of methylation, RNA splicing, and expression data

The original mQTL data were generated in two cohorts, BSGS ($n=614$) and LBC ($n=1366$), from peripheral blood [30]. This study utilized Whole Blood methylation measured on the Illumina HumanMethylation450 chip in the Brisbane Systems Genetics Study ($n=614$ from 177 families) and the Lothian Birth Cohorts of 1921 and 1936 (merged $n=1366$). Specific data are available for download from the YangLab website (https://yanglab.westlake.edu.cn/software/smr/#mQTLsummarydata). Blood eQTL data were extracted from the eQTLGen Consortium (https://eqtlgen.org/), which includes 31,684 individuals [30]. Our sQTL data were obtained from the Genotype-Tissue Expression (GTEx) project, where the latest GTEx v8 dataset extensively characterizes genetic associations, gene expression, and splicing in 838 individuals across 49 tissues [31]. We downloaded all sQTL data from these 49 tissues for analysis, and these data are available for download from the YangLab website (https://yanglab.westlake.edu.cn/software/smr/#sQTLsummarydata).

RiboSis-associated genes were sourced from Zang et al.'s study [32], which defined a gene set of 331 ribonucleic acid-related genes based on GO terms from MSigDB and characterization from Nerurkar et al. [5, 32]. The gene set did not provide Ensemble IDs for all genes. Hence, manual annotation was performed by accessing the Ensembl Genome Database (see supplementary Table S1 for details).

Cardiac common 22 structural and functional outcome dataset

Our outcome data encompass 22 common cardiac variations, including four left ventricular characteristics: left ventricular stroke volume (LVSV), left ventricular ejection fraction (LVEF), maximal left ventricle (MAX_LV), and minimum left ventricle (MIN_LV); one aortic characteristic: maximum aortic diameter (max_aa_diam); seven right ventricular characteristics: right ventricular ejection fraction (RVEF), right ventricular stroke volume (RVSV), right ventricular end-diastolic volume (RVEDV), right ventricular end-systolic volume (RVESV), maximum right atrial area (ramax area), minimum right atrial area (ramin area), and right atrial fractional area change (RAFAC); four pulmonary artery characteristics: magnetic resonance short-axis image of pulmonary artery root diameter (short_axis_cm_root), MRI short-axis image of proximal pulmonary artery diameter (short_axis_cm_pa), MRI short-axis image of proximal pulmonary artery diameter during diastole (short_axis_cm_pa_diastole), and short-axis magnetic resonance imaging pulmonary artery strain (short_axis_cm_pa_strain); two pulmonary artery to aorta ratio characteristics: ratio of pulmonary artery to aorta (PA_AO) and diastolic pulmonary artery to aorta ratio (PA_AO_diastole); and four left-to-right ventricular ratio characteristics: right ventricular stroke volume/left ventricular stroke volume (RVSV/LVSV), right ventricular end-diastolic volume/left ventricular end-diastolic volume (RVEDV/LVEDV), right ventricular end-systolic volume/left ventricular end-systolic volume (RVESV/LVESV), and RVEF/LVEF. These data were derived from a deep learning model-based study utilizing cardiovascular magnetic resonance imaging in the UK Biobank imaging sub-study, encompassing 45,504 individuals [6]. Of these, 41,135 individuals participated in at least one genome-wide association study after genotyping quality control and exclusion of prevalent diseases (http://www.broadcvdi.org).

The human sc-eQTL data for our validation cohort and PheWAS were retrieved from scQTLbase (http://bioinfo.szbl.ac.cn/scQTLbase/), a comprehensive sc-eQTL portal encompassing data from 304 datasets spanning 57 cell types and 95 cell states. This resource contains approximately 16 million SNPs significantly associated with gene expression in specific cell types or states, including ~ 690,000 disease-associated sc-eQTLs mapped to 3,333 traits/diseases. We retrieved all SNPs associated with RiboSis-related genes identified in our study and performed a PheWAS and gene enrichment analysis using FinnGen R10 data (https://www.finngen.fi/en/access_results). FinnGen is a large-scale biobank project encompassing genomic and health data from over 500,000 FinnGen individuals ($n=453,733$; 254,618 females and 199,115 males) to investigate disease mechanisms and susceptibility [33].

SMR analysis

We employed SMR to estimate the relationship between methylation, expression, and RNA splicing of

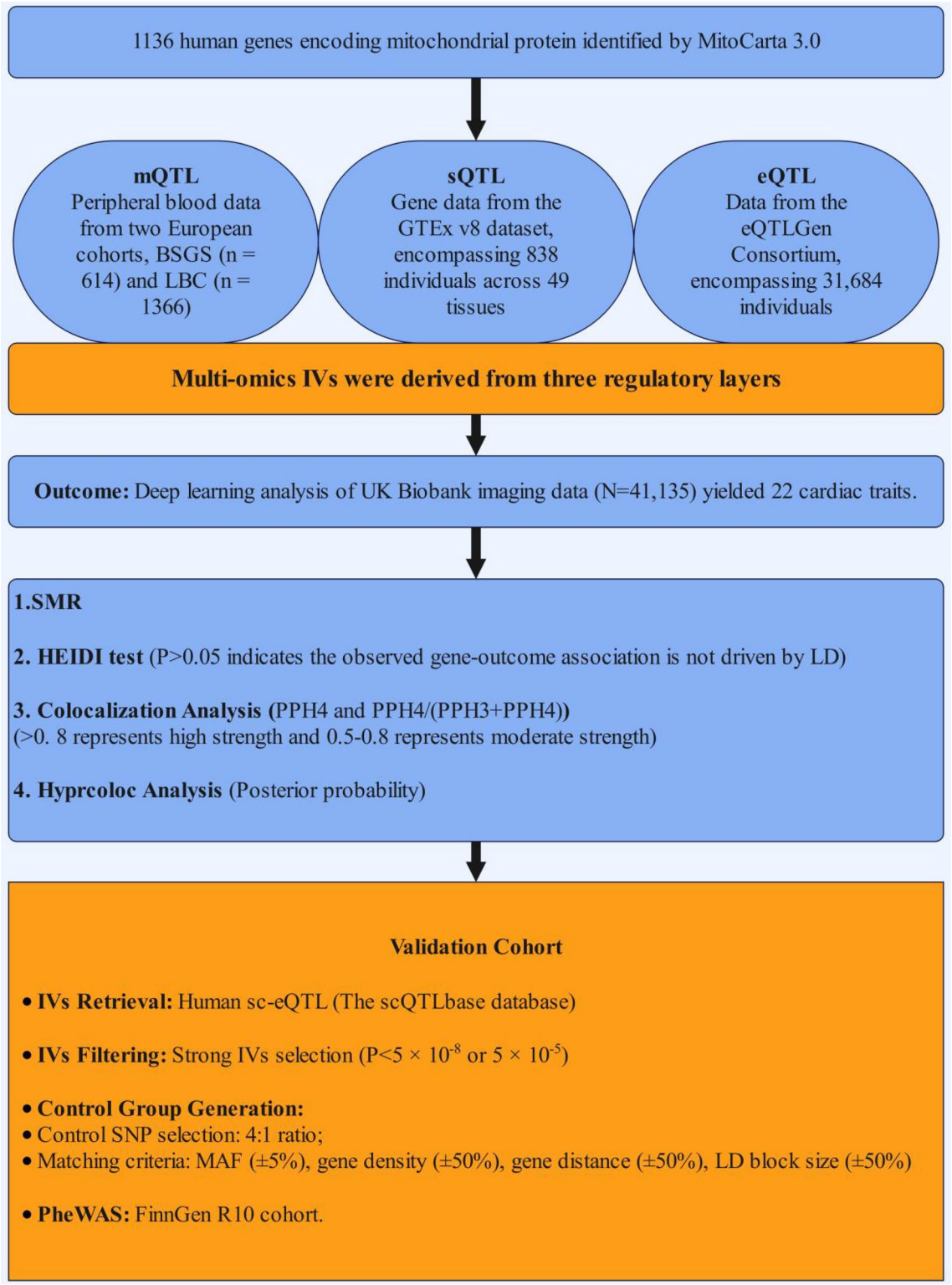

Fig. 1 Study design and analysis workflow. QTL: Quantitative trait locr; 1Vs: instrumental variables; Hypreoloe: the hypothesis prioritization for multi-trait colocalization; MAF: minor allele frequency; LD: linkage disequilibrium; SNP: single-nucleotide polymorphism; PheWAS: phenome-wide association study

RiboSi-associated genes and common variations in cardiac structure and function. SMR generally has much higher statistical power than traditional MR analysis when exposures and outcomes are derived from two independent large sample sets [27]. We considered windows centered around corresponding genes (± 2000 kb) and selected the most relevant cis-QTLs by a p value threshold of 5.0×10^{-8}. SNPs with allele frequency differences more significant than a specified threshold (set at 0.3 in this study) between any pairwise datasets (including LD reference samples, QTL summary data, and outcome summary data) were excluded. The HEIDI heterogeneity test was used to distinguish pleiotropy and linkage, with P-HEIDI < 0.05 indicative of pleiotropy and therefore excluded from the analysis [27]. SMR and HEIDI tests were conducted using the SMR software tool. Adjustments were made to p values in consideration of possible correlations between various phenotypes, as Bonferroni thresholds may be deemed overly conservative [34]. Thus, false discovery rate (FDR) correction was utilized. FDR control, as described by Benjamini and Hochberg (1995) [35], was employed. The procedure controls the expected proportion of false positives among all rejected null hypotheses (FDR = FP/(FP + TP), where FP and TP represent false and true positives, respectively) [35]. The FDR was maintained below a threshold of 0.05, a conventionally accepted level for statistical significance [35]. Given that our cardiac-related structures and functions are measured indicators rather than tissue specimen data, the tissue specificity of GTEx v8 sQTL summary data for this outcome may be relatively weak; therefore, a comprehensive scan of data from all 49 tissues was conducted. We initially performed a comprehensive R-based SMR analysis of sQTL data using the SMR analysis package in R software to ensure robust results. Subsequently, we conducted SMR and HEIDI tests using the SMR software tool based on analysis results with FDR p value < 0.05. Finally, colocalization and HyPrColoc analyses were performed using FDR-corrected p values < 0.05. The above analyses were conducted using R software (R-4.3.1) and the SMR software tool (smr-1.3.1).

Integrated analysis

After completing the analysis of each cardiac structural and functional variation feature with mQTL, sQTL, and eQTL, we merged these three gene regulatory layers with RiboSi-associated genes based on the respective variation features to identify positive genes and subsequently conducted colocalization analysis and HyPrColoc analysis.

Colocalization analysis and HyPrColoc analysis

We conducted colocalization analysis and HyPrColoc analysis to detect shared causal variants for each common cardiac structural and functional variation feature and identified mitochondrial-related mQTL, sQTL, and eQTL using the coloc R package. Bayesian-based colocalization analysis is an essential post-GWAS approach designed to identify genetic variants associated with phenotypes based on GWAS results [28]. This method involves four hypotheses: **H0:** no significant association between SNPs in the genomic region and either phenotype; **H1/H2:** significant association between SNPs in the genomic region and either phenotype one or phenotype 2; **H3:** significant association between SNPs in the genomic region and both phenotypes, driven by different causal variants; **H4:** significant association between SNPs in the genomic region and both phenotypes, driven by the same causal variant [28]. The analysis tests the posterior probability of the fourth hypothesis (H4). Based on these hypotheses, a higher statistical probability of H4 (PH4) better explains the influence of significant signal loci on the phenotype. Some studies suggest that a PH4 > 0.8 indicates strong colocalization support, while 0.8 > PH4 > 0.5 indicates moderate colocalization support [36, 37]. HyPrColoc analysis, also a Bayesian method, uses a novel branch and bound algorithm to identify subsets of features displaying colocalization at distinct causal variants within a genetic locus (called clusters) [29]. It is based on three key assumptions: **(1)** for non-independent studies, GWAS results are from the same underlying population, i.e., the LD patterns are identical across studies; **(2)** each trait has at most one causal variant per genomic region; **(3)** these causal variants are either directly genotyped or well imputed in all GWAS datasets. The support probability for clustering evidence is similar to that for colocalization analysis.

Integrating results from the multi-omics level of evidence

We integrated results from three distinct gene regulatory layers to comprehensively understand the associations between RiboSis-related gene regulation at different levels and the 20 cardiac variation features. Subsequently, we synthesized and evaluated RiboSis genes associated with these 20 cardiac variation features. Recognizing the equal importance of these three gene regulatory layers, we categorized them into two classes: Class I Recommendation: Genes associated with outcome features in more than two gene regulatory layers (FDR-corrected SMR analysis result < 0.05); Class II Recommendation: Genes associated with outcome features in one gene regulatory layer (FDR-corrected SMR analysis result < 0.05).

Given that we performed colocalization and HEIDI tests for genes in each regulatory layer, and HyPrColoc analysis for genes associated with outcome features in more than two layers, we further classified the recommendations as follows: For Class II recommendations, genes with at least one significant sensitivity analysis (HEIDI test > 0.05 or colocalization analysis probability > 0.8) were classified as Grade A; genes with only one significant sensitivity analysis (colocalization analysis probability between 0.5 and 0.8) were classified as Grade B; genes with no significant sensitivity analyses (HEIDI test < 0.05 and colocalization analysis probability < 0.5) were classified as Grade C.

Based on the classifications for Class II, we categorized Class I recommendations as follows: Grade A, where each gene regulatory layer had at least one significant sensitivity analysis (HEIDI test > 0.05 or colocalization analysis probability > 0.8); Grade B, where each layer had only one significant sensitivity analysis (colocalization analysis probability between 0.5 and 0.8), or only one layer had a HEIDI test < 0.05 and colocalization analysis probability < 0.5, with at least one other layer showing HEIDI test > 0.05 and colocalization analysis probability > 0.8; Grade C, where no layer had significant sensitivity analyses (HEIDI test < 0.05 and colocalization analysis probability < 0.5), or the conditions for Grade B were not met.

For the comprehensive evaluation, we applied a scoring system. We assigned an initial scores 10 for Class I and 5 for Class II recommendations. Within Class II, weights were 1 for Grade A, 0.5 for Grade B, and 0.1 for Grade C. The weights for the three grades in Class I were similar to those in Class II, with additional considerations for HyPrColoc clustering: 1 for clustering across all regulatory layers and outcome features, 0.5 for clustering across gene regulatory layers only, 0.25 for clustering between a single gene regulatory layer and outcome features, and 0 for no clustering.

The final score, the gene hypothesis influence factor based on MR analysis (GHIF-MR), was calculated by summing the scores for each gene across all cardiac features and dividing by the number of appearances of the gene. We presented the evaluation results for each of the 22 cardiac features and the comprehensive evaluation.

Validation and PheWAS

Following identifying RiboSis-related genes, we queried scQTLbase to retrieve all associated SNPs for each gene, encompassing data from various single-cell studies and large-scale projects. The principles for phenotype selection are as follows: (1) Exclude phecode with less than 200 cases [38]. (2) Binary phenotypes with fewer than 100 cases and continuous and categorical ordered phenotypes with sample sizes of fewer than 10,000 were excluded [39]. (3) International Classification of Diseases (ICD)-coded binary phenotypes without principal ICD codes or with external causes (codes as Z00–Z99) were also excluded [40]. We then selected strongly associated SNPs for each gene using two thresholds: a stringent threshold of 5×10^{-8} and, in cases with limited SNPs meeting this criterion, a more lenient threshold of 5×10^{-5} [41, 42]. Subsequently, we utilized an R package to generate a control group of SNPs for each selected SNP via random sampling. These control SNPs exhibited comparable allele frequencies, linkage disequilibrium (LD) patterns, distances to the nearest gene, and gene densities as the investigated SNPs. Based on previous research suggesting an optimal ratio of four control SNPs per SNP of interest [43], we implemented specific matching parameters: minor allele frequency(MAF) within ± 5%, gene density within ± 50%, distance to the nearest gene within ± 50%, and the number of other SNPs with $R^2 \geq 0.50$ (representing haplotype block size) within ± 50% [44]. Finally, we performed gene set enrichment analysis using FinnGen R10 data. This analysis aimed to determine whether the instrumental variables represented by the strongly associated SNPs for each gene were significantly enriched in specific traits (e.g., cardiac structural abnormalities or diseases) compared to the control SNP set, inferring a potential association between the gene and the respective trait.

Results

Gene methylation and cardiac structural function of RIBOSIS-associated genes

After FDR correction, we identified four unique RiboSis-related genes with eight proximal CpG sites exhibiting causal effects on five cardiac structural and functional traits. We employed two probability approaches for colocalization analysis [28]: calculating PPH4 and calculating the ratio PPH4/(PPH4 + PPH3). While the former generally sufficed, the latter was employed when PPH4 values were low, maintaining the same interpretation. Our analysis revealed the following significant colocalization signals (see Supplementary Tables S2 and S3 for details. The forest map is shown in Fig. 2): SMARCB1 and LVEF(cg08219923, $P_{adj}=0.0020$, $P_{HEIDI}=2.88\times10^{-5}$, $PPH4=2.17\times10^{-5}$); ERI1 and max_aa_diam(cg10039590, $P_{adj}=0.0036$, $P_{HEIDI}=0.0621$, $PPH4=1.28\times10^{-3}$; cg04756296, $P_{adj}=0.0151$, $P_{HEIDI}=0.1828$, $PPH4=8.08\times10^{-3}$; cg18237548, $P_{adj}=0.0055$, $P_{HEIDI}=0.3233$, $PPH4=9.60\times10^{-3}$); TSR1 and max_aa_diam (cg18334121, $P_{adj}=0.0003$, $P_{HEIDI}=0.0045$, $PPH4=4.90\times10^{-10}$); SMARCB1 and MIN_LV(cg08219923, $P_{adj}=0.0036$, $P_{HEIDI}=0.0002$, $PPH4=2.68\times10^{-4}$); ERCC2 and ramin area (cg27517897, $P_{adj}=0.0089$, $P_{HEIDI}=0.2289$,

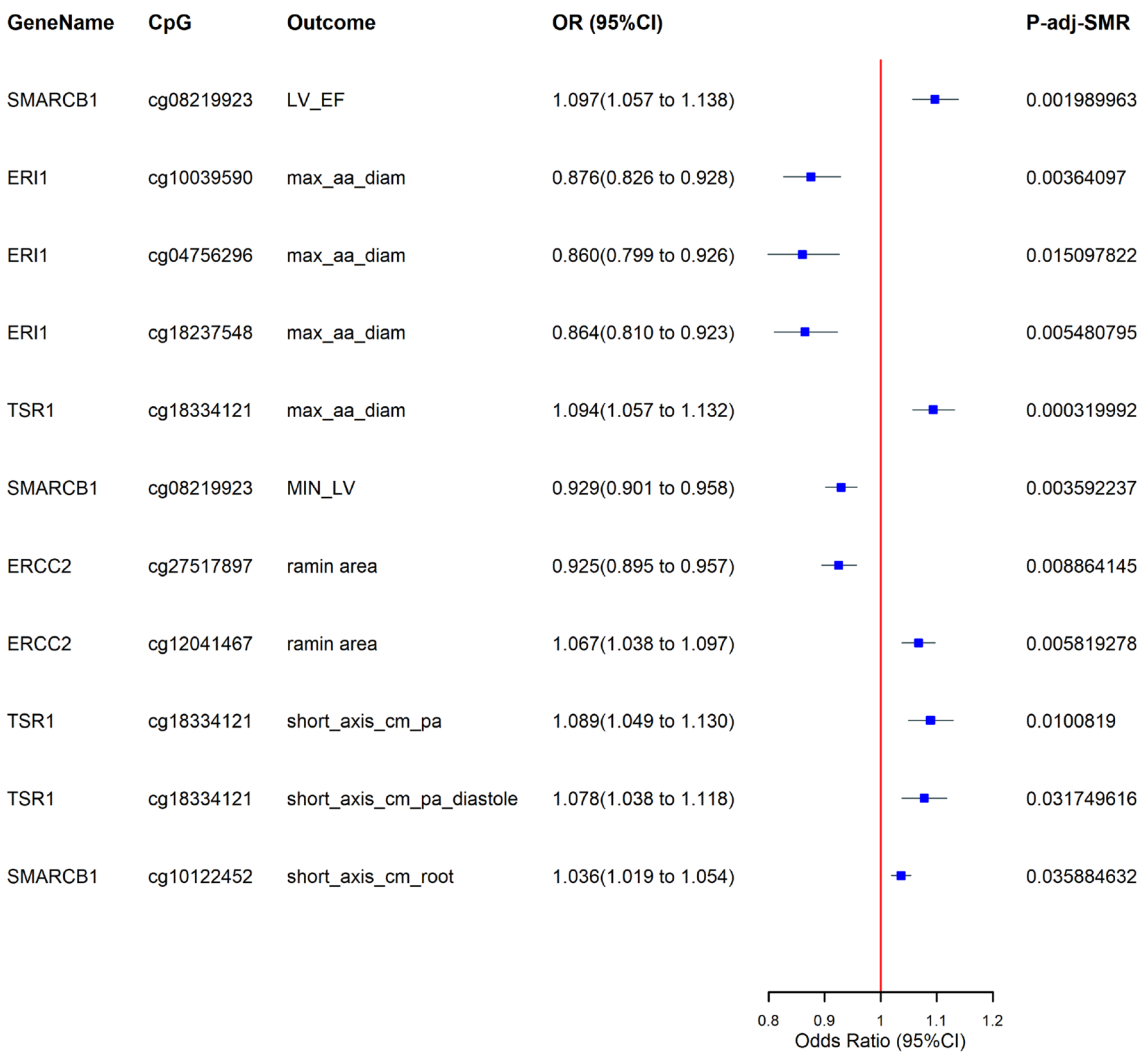

Fig. 2 Forest map of methylation levels and SMR analysis results of cardiac structure–function. OR, odds ratio; CI, confidence interval; P-adj-SMR: The FDR-corrected *p* value in SMR analysis, where SMR analysis refers to summary data-based Mendelian randomization, utilizing summary-level data from GWAS and quantitative trait loci (QTL) studies to test for pleiotropic associations between gene expression levels and the complex traits of interest. FDR correction stands for false discovery rate correction, which is the expected proportion of falsely rejected (i.e., rejecting the true null hypothesis) null hypotheses among all the rejected null hypotheses. This correction is performed using the R programming language

PPH4 = 5.27×10^{-9}; cg12041467, $P_{adj}=0.0058$, $P_{HEIDI}=0.4335$, PPH4 = 5.81×10^{-9}); TSR1 and short_axis_cm_pa(cg18334121, $P_{adj}=0.0101$, $P_{HEIDI}=0.0090$, PPH4 = 4.13×10^{-6}); TSR1 and short_axis_cm_pa_diastole(cg18334121, $P_{adj}=0.0317$, $P_{HEIDI}=0.0468$, PPH4 = 3.12×10^{-5}); SMARCB1 and short_axis_cm_root (cg10122452, $P_{adj}=0.0359$, $P_{HEIDI}=0.0468$, PPH4 = 0.885).

RNA splicing and cardiac structural function of RIBOSIS-associated genes

Following FDR multiple testing correction, we identified five unique RiboSis-related genes significantly associated with cardiac traits across nine tissues (Supplementary Table S4 and Fig. 3). A complete list of analyzed GTEx v8 tissues and sample sizes is provided in Supplementary Table S5. Specifically, the results are as follows: ERI1 exhibited a significant association with max_aa_diam across five tissues(Thyroid, $P_{adj}=0.0375$; Whole Blood, $P_{adj}=0.0248$, Lung, $P_{adj}=0.0368$; Pancreas, $P_{adj}=0.0302$; Spleen, $P_{adj}=0.0264$); a significant association was observed between TSR1 and max_aa_diam in Skin Sun Exposed Lower leg ($P_{adj}=0.0235$); HEATR1 was significantly associated with MIN_LV across two tissues: Skin Not Sun Exposed Suprapubic ($P_{adj}=0.0481$) and Esophagus Muscularis ($P_{adj}=0.0147$); EIF6 was significantly associated with RVEDV in Esophagus Muscularis

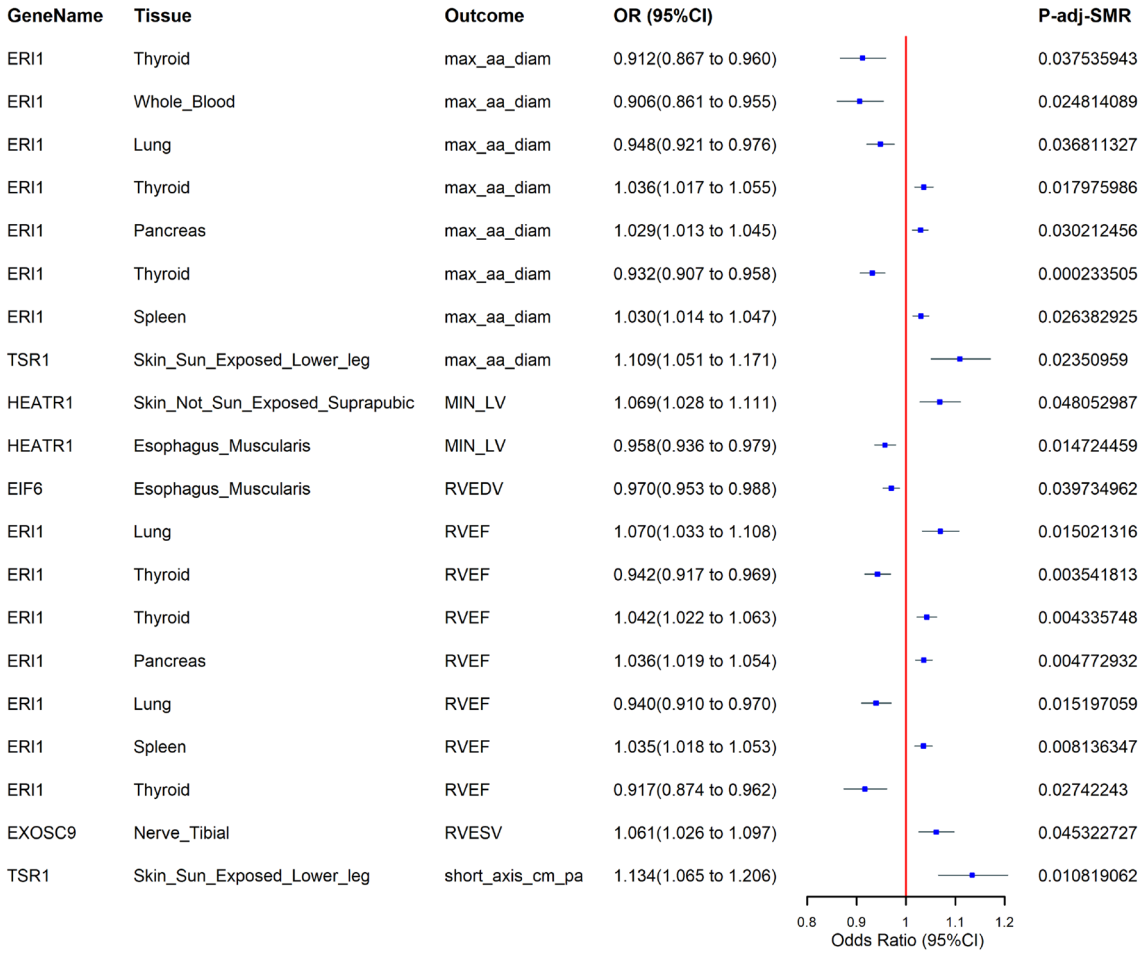

Fig. 3 Forest map of RNA splicing and SMR analysis results of cardiac structure–function. OR, odds ratio; CI, confidence interval; P-adj-SMR: The FDR-corrected *p* value in SMR analysis, where SMR analysis refers to summary data-based Mendelian randomization, utilizing summary-level data from GWAS and quantitative trait loci (QTL) studies to test for pleiotropic associations between gene expression levels and the complex traits of interest. FDR correction stands for false discovery rate correction, which is the expected proportion of falsely rejected (i.e., rejecting the true null hypothesis) null hypotheses among all the rejected null hypotheses. This correction is performed using the R programming language. Within the ERI1 gene, multiple independent linkage disequilibrium (LD) blocks harboring distinct top SNPs were identified in Thyroid and Lung tissues. Independent analyses of these SNPs yielded significant, albeit varying, results

($P_{adj=}$0.0397); ERI1 exhibited a significant association with RVEF across four tissues(Thyroid, $P_{adj}=0.0035$; Lung, $P_{adj}=0.0150$; Pancreas, $P_{adj}=0.0047$; Spleen, $P_{adj}=0.0081$); EXOSC9 was significantly associated with RVESV in Nerve Tibial ($P_{adj}=0.0453$), while TSR1 showed a significant association with short_axis_cm_pa in Skin Sun Exposed Lower leg ($P_{adj}=0.0108$). Of note, multiple independent and LD blocks containing different top SNPs (SNPs with the smallest *P* values in the original GWAS data used as instrumental variables for SMR analysis) within the ERI1 gene were identified in both Thyroid and Lung tissues. Analyses of these SNPs individually yielded different yet significant ($P_{adj}<0.05$) results. For brevity, only one representative *P* value is presented here, with detailed results provided in Supplementary Table S4. Furthermore, the direction of effect estimates for the same gene on the same cardiac structure or function was inconsistent. For instance, the odds ratio (OR) for ERI1 on max_aa_diam was less than 1 in Whole Blood (OR = 0.9063, 95% CI: 0.8605–0.9545) and Lung (OR = 0.9483, 95% CI: 0.9209–0.9765) but greater than 1 in Pancreas (OR = 1.0292, 95% CI: 1.0134–1.0452) and Spleen (OR = 1.0301, 95% CI: 1.0140–1.0465). Different top SNPs displayed divergent directions of effect estimates even within the same tissue, such as the ERI1 and max_aa_diam association in the Thyroid. The results of colocalization analysis are shown in Supplementary Table S3.

Gene expression and cardiac structural function of RIBOSIS-associated genes

Following FDR correction, we identified 11 RiboSis-related genes significantly associated with 12 cardiac structure and function traits (Supplementary Table S6 and Fig. 4). Colocalization results for these genes are presented in Supplementary Table S3. Five causal effect estimates—SMARCB1 on LVEF ($P_{HEIDI}=2.26\times10^{-6}$, OR = 0.8969, 95% CI: 0.8653–0.9297), SMARCB1 on MIN_LV ($P_{HEIDI}=8.48\times10^{-7}$, OR = 1.0909, 95% CI: 1.0591–1.1236), NOP58 on RVESV/LVESV ($P_{HEIDI}=0.0007$, OR = 0.8981, 95% CI: 0.8463–0.9531), SMARCB1 on RVESV/LVESV ($P_{HEIDI}=0.0004$, OR = 0.9332, 95% CI: 0.8983–0.9695), and UTP1 on short_axis_cm_root ($P_{HEIDI}=0.0440$, OR = 1.1122, 95% CI: 1.0584–1.1688)—yielded nonsignificant HEIDI test

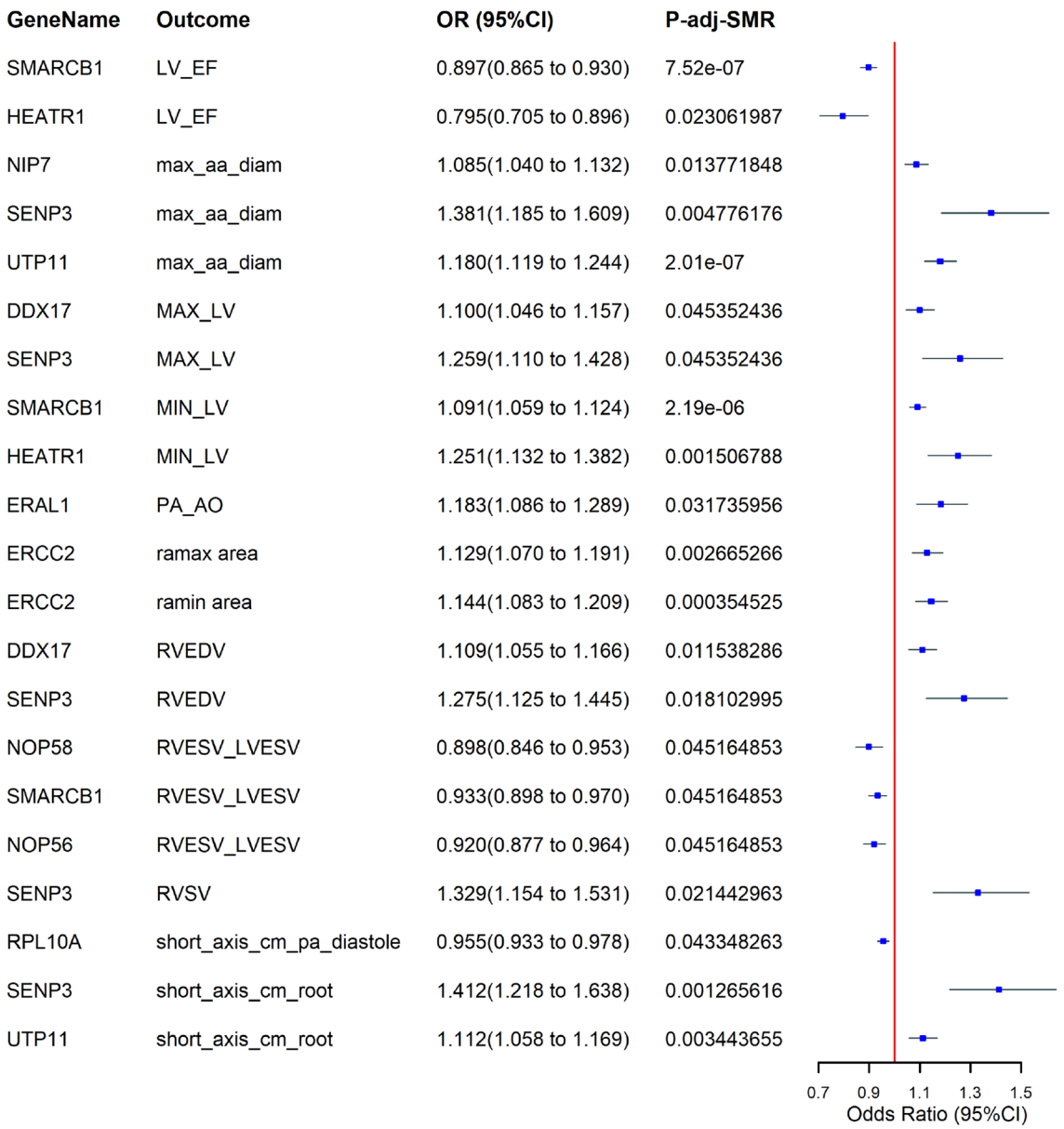

Fig. 4 Forest map of gene expression and SMR analysis results of cardiac structure–function. OR, odds ratio; CI, confidence interval; P-adj-SMR: The FDR-corrected *p* value in SMR analysis, where SMR analysis refers to summary data-based Mendelian randomization, utilizing summary-level data from GWAS and quantitative trait loci (QTL) studies to test for pleiotropic associations between gene expression levels and the complex traits of interest. FDR correction stands for false discovery rate correction, which is the expected proportion of falsely rejected (i.e., rejecting the true null hypothesis) null hypotheses among all the rejected null hypotheses. This correction is performed using the R programming language

results. All other associations exhibited P-HEIDI > 0.05. Notably, except SMARCB1, the predicted expression of the remaining genes exhibited consistent directional effects across different cardiac traits. For example, SENP3 consistently showed ORs more remarkable than 1 for max_aa_diam (OR = 1.3806, 95% CI: 1.1846–1.6090), MAX_LV (OR = 1.2586, 95% CI: 1.1096–1.4276), RVEDV (OR = 1.2751, 95% CI: 1.1253–1.4448), RVSV (OR = 1.3291, 95% CI: 1.1537–1.5313), and short_axis_cm_root (OR = 1.4123, 95% CI: 1.21780–1.6376). Additionally, our findings suggest potential cardioprotective roles for HEATR1 (LVEF, OR = 0.7949, 95% CI: 0.7051–0.8962), NOP56 (RVESV/LVESV, OR = 0.9195, 95% CI: 0.8770–0.9641), and RPL10A (short_axis_cm_pa_diastole, OR = 0.9553, 95% CI: 0.9328–0.9783), in addition to SMARCB1 and NOP58.

Integrating multi-omics evidence

We prioritized RiboSis-related genes by integrating the above results with colocalization and HyPrColoc analyses (Supplementary Table S3, Supplementary Table S7, and Supplementary Table S8). This prioritization scheme informed a composite scoring system, reflected in the summed and averaged GHIF-MR scores (Table 1). Most genes only demonstrated causal effects on one of the 22 cardiac traits, so their summed and averaged scores were frequently identical. However, some genes exhibited pleiotropic effects on multiple cardiac traits. SENP3 and HEATR1 achieved the highest summed GHIF-MR score [25], while NOP58 had the lowest (0.5). HEATR1 exhibited the highest average GHIF-MR score (12.5), with NOP58 again demonstrating the lowest (0.5). We then categorized these 15 genes into quartiles based on their summed GHIF-MR score: Quartile 1 (Q1) represented the top 25%, Quartile 2 (Q2) encompassed 25–50%, Quartile 3 (Q3) spanned 50–75%, and Quartile 4 (Q4) constituted the bottom 25%. This stratification yielded the following classifications for the summed GHIF-MR score: Q1—HEATR1 and SENP3; Q2—ERI1 and ERCC2; Q3—TSR1, UTP11, DDX17, and SMARCB1; and Q4—NIP7, ERAL1, NOP56, RPL10A, EIF6, EXOSC9, and NOP58. Due to the predominance of genes impacting a single cardiac trait, only HEATR1 could be confidently classified as Q1 based on the average GHIF-MR score, precluding further quartile distinctions using this metric.

Table 1 Score results of RIBOSIS-related genes related to cardiac structure and function based on multi-omics evidence, where GHIF-MR indicates the gene hypothesis influence factor based on MR analysis

RiboSis-related gene	Total GHIF-MR	Average GHIF-MR
SMARCB1	7.5	1.875
HEATR1	25	12.5
NIP7	5	5
SENP3	25	5
UTP11	10	5
DDX17	10	5
ERAL1	5	5
ERCC2	15	7.5
NOP58	0.5	0.5
NOP56	5	5
RPL10A	5	5
ERI1	20	10
TSR1	11.5	3.83
EIF6	5	5
EXOSC9	5	5

Validation and PheWAS

To investigate the phenotypic associations of 15 RiboSis-related genes, we first searched the scQTLbase database to identify all available IVs for each gene. Subsequently, we applied stringent thresholds to select strongly associated IVs. Specifically, a threshold of $P < 5 \times 10^{-5}$ was employed for ERAL1, ERCC2, NIP7, NOP58, SENP3, TSR1, and UTP11, while a more stringent threshold of $P < 5 \times 10^{-8}$ was used for DDX17, EXOSC9, ERI1, HEATR1, NOP56, RPL10A, EIF6, and SMARCB1 (Supplementary Table S9). Notably, ERAL1 exhibited only one strongly associated IV (rs12948254). We then utilized the R package to generate a matched SNP set (Supplementary Table S10) for gene set enrichment analysis leveraging the FinnGen R10 data. Following FDR correction, we identified a range of traits significantly associated with these genes.

Our analysis revealed compelling associations between HEATR1 and both "Cardiac arrest" ($P_{adj} = 5.14 \times 10^{-20}$) and "Heart failure and bmi 25plus" ($P_{adj} = 7.95 \times 10^{-5}$). SENP3 showed strong associations with "Hypertension, essential" ($P_{adj} = 2.75 \times 10^{-43}$) and "Cardiovascular diseases (excluding rheumatic etc.)" ($P_{adj} = 1.44 \times 10^{-13}$). Notably, UTP11 exhibited significant associations with multiple cardiac-related traits, including "Angina pectoris" ($P_{adj} = 1.20 \times 10^{-16}$, "Unstable angina pectoris" ($P_{adj} = 2.99 \times 10^{-16}$), "Major coronary heart disease event" ($P_{adj} = 5.83 \times 10^{-16}$), "Ischaemic heart disease, wide definition" ($P_{adj} = 5.83 \times 10^{-16}$), and "Myocardial infarction, strict" ($P_{adj} = 7.10 \times 10^{-15}$).

Furthermore, ERCC2 was associated with "Congenital malformations of cardiac chambers and connections" ($P_{adj} = 2.61 \times 10^{-10}$), "Secondary hypertension" ($P_{adj} = 0.0014$), "Heart failure and antihypertensive medication" ($P_{adj} = 0.0022$), and "Heart failure and bmi 25plus" ($P_{adj} = 0.0112$). NOP58

displayed strong associations with "Statin medication" ($P_{adj} = 9.47 \times 10^{-75}$) and "Disorders of lipoprotein metabolism and other lipidaemias" ($P_{adj} = 2.68 \times 10^{-70}$), while NOP56 was associated with "Thoracic aortic aneurysm" ($P_{adj} = 1.17 \times 10^{-11}$), "Aortic aneurysm" ($P_{adj} = 2.89 \times 10^{-10}$), and "Pre-eclampsia or poor fetal growth" ($P_{adj} = 8.20 \times 10^{-5}$).

Our analysis also revealed associations between RPL10A and "Thoracic aortic aneurysm" ($P_{adj} = 1.10 \times 10^{-7}$), "Atrial septal defect" ($P_{adj} = 0.0024$), "Coronary atherosclerosis" ($P_{adj} = 0.0038$), and "Hypertensive Heart Disease" ($P_{adj} = 0.0168$). ERI1 exhibited significant associations with "Hypertension, essential" ($P_{adj} = 2.96 \times 10^{-35}$), "Hypertension" ($P_{adj} = 3.23 \times 10^{-27}$), "Antihypertensive medication—note that there are other indications" ($P_{adj} = 1.23 \times 10^{-26}$), and "Hypertensive Heart Disease" ($P_{adj} = 0.0009$). Similarly, EIF6 was associated with "Hypertensive Heart Disease" ($P_{adj} = 3.58 \times 10^{-17}$), "Hypertension" ($P_{adj} = 3.26 \times 10^{-16}$), and "Hypertension, essential" ($P_{adj} = 2.56 \times 10^{-15}$).

Furthermore, DDX17 showed a strong association with "STROKE" ($P_{adj} = 3.15 \times 10^{-7}$), while EXOSC9 exhibited significant associations with numerous cardiac-related traits, including "Coronary angioplasty" ($P_{adj} = 4.86 \times 10^{-108}$), "Coronary atherosclerosis" ($P_{adj} = 5.70 \times 10^{-99}$), "Unstable angina pectoris" ($P_{adj} = 6.95 \times 10^{-88}$), "Major coronary heart disease event" ($P_{adj} = 3.81 \times 10^{-87}$), "Atrial fibrillation and flutter with reimbursement" ($P_{adj} = 2.05 \times 10^{-75}$), "Ischaemic heart disease, wide definition" ($P_{adj} = 4.49 \times 10^{-71}$), "Myocardial infarction, strict" ($P_{adj} = 9.66 \times 10^{-62}$), and "Heart failure and coronary heart disease" ($P_{adj} = 9.55 \times 10^{-58}$). Lastly, SMARCB1 exhibited significant associations with "Nonischemic cardiomyopathy" ($P_{adj} = 4.43 \times 10^{-129}$), "Hypertensive Heart Disease" ($P_{adj} = 2.74 \times 10^{-114}$), "Cardiac arrhythmias" ($P_{adj} = 1.82 \times 10^{-106}$), "Aortic aneurysm" ($P_{adj} = 1.38 \times 10^{-74}$), "Thoracic aortic aneurysm" ($P_{adj} = 0.0002$), and "Atrial fibrillation and flutter" ($P_{adj} = 0.0006$).

These findings suggest that the identified genes are significantly associated with various cardiac and cardiovascular traits, implicating their potential roles in cardiac development, function, and disease. We acknowledge that ERAL1 exhibited only a strongly associated IV, and no significant enrichment for cardiac-related traits was observed. Further investigation using the LDlink (https://ldlink.nih.gov/?tab=home) portal to explore the phenotypic associations of rs12948254 yielded no significant findings. Similarly, no significant associations with cardiac-related traits were observed for TSR1 and NIP7. For details, see Supplementary Table S11.

Discussion

In this study, we integrated SMR, colocalization, and HyPrColoc analyses to investigate the relationships between genetically predicted methylation, splicing, and expression levels of RiboSis-related genes and cardiac structure and function. A composite scoring system and PheWAS validation strengthened our findings. We identified 15 RiboSis-related genes, 12 of which were supported by PheWAS replication. This study represents a novel investigation into the role of ribosomal factors in cardiac health using a multi-omics approach, potentially unveiling new avenues for understanding disease mechanisms and therapeutic interventions.

HEATR1, identified in our study as the highest-ranking gene, encodes a ribosome biogenesis factor essential for pre-18S rRNA processing in the nucleolus. This crucial role in ribosome biogenesis optimizes pre-rRNA transcription by RNA polymerase I [45]. According to GeneCards (https://www.genecards.org/cgi-bin/carddisp.pl?gene=HEATR1), HEATR1 is potentially implicated in pancreatic ductal adenocarcinoma and Diamond–Blackfan anemia. Despite limited research directly linking HEATR1 to cardiac function, existing evidence suggests a potential mechanism involving the p53 signaling pathway [46]. HEATR1 has been shown to modulate and activate p53 signaling, and extensive research implicates p53 dysregulation in various cardiovascular events [47, 48], including heart failure. Preclinical studies in rats have demonstrated an association between p53 expression and both ischemia/reperfusion-induced cardiomyocyte apoptosis and the incidence of coronary restenosis following percutaneous coronary intervention [49]. Conversely, a murine model demonstrated that p53 acetylation induced coronary microvascular dysfunction, whereas a lack of p53 acetylation improved coronary flow reserve and rescued impaired cardiac function [50]. Furthermore, p53 expression has been observed in cardiac tissue from patients with congestive heart failure [49]. Therefore, HEATR1 may indirectly influence cardiac structure and function, potentially contributing to cardiovascular events via the p53 pathway.

Another gene ranked within the Q1 is SENP3 (SUMO Specific Peptidase 3), encoding a protein responsible for the reversible posttranslational modification of proteins via small ubiquitin-like modifier (SUMO) conjugation. A previous study utilizing SENP3 knockout mice demonstrated that SENP3 regulates Nodal, a gene critical for establishing left–right asymmetry during heart development. This work identified a SENP3-Nodal signaling axis that modulates cardiac aging and autophagy homeostasis, ultimately influencing heart development and potentially contributing to coronary heart disease [51]. Mechanistic studies in mice have shown that SENP3

silencing ameliorates cardiomyocyte apoptosis primarily by suppressing endoplasmic reticulum (ER) stress and mitochondrial-mediated apoptosis pathways, while simultaneously promoting mitochondrial translocation of dynamin-related protein 1 (Drp1) in reperfused myocardium [52]. These findings provide direct evidence that SENP3 upregulation plays a crucial, Drp1-dependent role in myocardial ischemia/reperfusion (MIR) injury, suggesting that SENP3 inhibition may offer a therapeutic avenue for attenuating MIR damage [52]. Conversely, another study reported a cardioprotective role for SENP3 during ischemia–reperfusion injury [53]. In contrast to this latter finding, our study revealed that at the gene expression level, SENP3 is positively associated with an increased risk for multiple cardiac structural and functional traits, suggesting a potential role in promoting heart disease. These contrasting findings highlight the need for further basic research to elucidate the precise role of SENP3 in cardiac health and disease.

ERI1 (Ribonuclease Exonuclease 1), a protein-coding gene located within the Q2 partition of our scoring system, binds to and degrades the 3'-ends of histone mRNAs, suggesting a critical role in post-replicative histone mRNA decay [54]. A GWAS investigating the circulating proteome linked ERI1 to both hypertension and spontaneous coronary artery dissection (SCAD), both significant risk factors for aortic dilatation [55]. This finding aligns with our study, which demonstrated a protective effect of ERI1 methylation levels at CpG sites cg10039590, cg04756296, and cg18237548 on maximum aortic diameter. However, ERI1 has also been shown to interact with the PB2, PB1, and NP components of viral ribonucleoproteins, thereby enhancing viral transcription [56]. This suggests a potential detrimental role for ERI1 in the context of viral infection within the heart. A case report described a homozygous microdeletion in the ERI1 gene, resulting in intellectual disability, limb malformations, and cardiac defects [57]. The authors hypothesized that aberrant ERI1 transcripts may lead to abnormal expression of HOXC8, a transcription factor crucial for body axis formation during embryogenesis [57]. This suggests a potential role for ERI1 in the development of cardiac structure and function, potentially contributing to congenital heart disease.

According to the GeneCards database, ERCC2 (ERCC Excision Repair 2, TFIIH Core Complex Helicase Subunit) encodes a protein crucial for transcription-coupled nucleotide excision repair. This protein is an integral component of the basal transcription factor BTF2/TFIIH complex and exhibits ATP-dependent DNA helicase activity. A real-time polymerase chain reaction (RT-PCR) study for genotyping peripheral blood DNA samples revealed that homozygous mutations in ERCC2 may be a risk factor for coronary artery disease (CAD). At the same time, the wild-type allele may confer a decreased risk [58]. ERCC2 has been implicated in oxidative stress, a process known to contribute to cardiovascular disease development [59]. Bioinformatic analysis identified ERCC2 as a candidate gene associated with patent ductus arteriosus (PDA) in patients, suggesting a potential role in congenital cardiac structural and functional abnormalities [60]. Our study revealed a trend toward increased risk for both maximum and minimum heart areas with elevated ERCC2 gene expression. Conversely, at the methylation level, CpG sites cg27517897 and cg12041467 exhibited a protective effect against these traits. We hypothesize that higher methylation levels at these CpG sites may reduce ERCC2 expression, thereby conferring cardioprotection. However, further investigation is warranted to confirm this proposed mechanism.

Additionally, our study identified several other genes of interest. For instance, UTP11 (UTP11 Small Subunit Processome Component), which exhibits RNA-binding activity and participates in neurodevelopment and positive regulation of apoptotic processes, may influence cardiac events via modulation of the p53 pathway [61]. Similarly, DDX17 (DEAD-Box Helicase 17) could impact cardiac health by regulating mitochondrial homeostasis and lipid accumulation [62, 63]. SMARCB1 (SWI/SNF Related, Matrix Associated, Actin-Dependent Regulator Of Chromatin, Subfamily B, Member 1) encodes a protein core component of a complex responsible for remodeling repressive chromatin structures. This remodeling facilitates enhanced access to transcriptional machinery to target genes. SMARCB1 has been implicated in cardiomyopathy risk, increased left ventricular wall thickness, and TGF-β1-mediated myocardial fibrosis [64]. Expression of TSR1 (TSR1 ribosome maturation factor) has been linked to spontaneous coronary artery dissection (SCAD) [65].

Our study's primary strength lies in its integrative approach, leveraging multi-omics evidence to investigate the association between candidate genes and 22 common cardiac structural and functional traits. This approach and an aggregated scoring system enable a clear interpretation of results. Furthermore, we employed the large-scale FinnGen R10 dataset for PheWAS analyses, providing independent validation and strengthening our conclusions. However, our study is limited to using cardiac traits derived from a single, albeit significant, study. This lack of independent replication cohorts restricts our ability to establish causal relationships between RiboSis-related genes and cardiac outcomes. Another limitation concerns the interpretation of colocalization analyses. While a low posterior probability of colocalization (PPH4) might suggest a lack of shared

genetic effects, caution is warranted when PPH3 (posterior probability of association with only one trait) is also low due to limited statistical power. Although we calculated PPH4/(PPH3 + PPH4) to mitigate this issue, it does not fully address this inherent limitation. Furthermore, PheWAS analyses, particularly those utilizing summary-level data, typically lack the explicit adjustment for confounders characteristic of traditional regression-based GWAS. This limitation stems from the unavailability of individual-level data necessary for confounder control. Consequently, observed associations may be influenced by unmeasured or unadjusted factors, such as age, sex, lifestyle, and environmental exposures, which may affect both the genetic variant and the phenotype. Another limitation of this study is that three genes (ERAL1, TSR1, and NIP7) identified in the PheWAS validation analysis lacked support from gene set enrichment analysis. This was due to limited strongly associated SNPs for these genes in the scQTLbase, precluding robust screening. In addition, utilizing the GTEx v8 dataset, derived from postmortem samples, introduces a potential bias. Gene expression and abundance alterations due to death may influence our findings, particularly in colocalization and HyPrColoc analyses. Meanwhile, our validation analysis was conducted solely within the FinnGen R10 cohort. Future studies should consider expanding validation to additional cohorts, such as the UK Biobank (UKB), to enhance the generalizability and robustness of our findings. Finally, the identified genes warrant further investigation as potential drug targets. This includes database searches for potential biomarkers and therapeutic targets, and bioinformatic analyses to elucidate functional relationships and pathways, ultimately informing future clinical translation efforts. Translation of the 15 identified genes to clinical application requires further investigation. This includes initial in vitro validation (e.g., single-cell sequencing and animal models). Furthermore, the absence of proteomics data in our multi-omics approach limits our ability to definitively assess the potential of these genes as drug targets or key pathogenic drivers, given the central role of proteins in biological processes. Subsequent in vivo studies and clinical trials will necessitate careful consideration of demographic factors such as sex and ethnicity, which are critical for clinical translation.

Conclusion

In summary, our study identified 15 RiboSis-related genes: HEATR1, SENP3, ERI1, ERCC2, TSR1, UTP11, DDX17, SMARCB1, NIP7, ERAL1, NOP56, RPL10A, EIF6, EXOSC9, and NOP58. Notably, 12 of these genes were further supported by gene set enrichment analyses. These findings provide valuable insights into the intricate relationship between ribosome biogenesis and cardiac health, potentially informing novel therapeutic strategies for cardiac diseases.

Abbreviations

RiboSis	Ribosome biogenesis
MR	Mendelian randomization
HyPrColoc	Hypothesis prioritization for multi-trait colocalization
PheWAS	Phenome-wide association study
ARVC	Arrhythmogenic right ventricular cardiomyopathy
SNPs	Single-nucleotide polymorphisms
IVs	Instrumental variables
QTL	Quantitative trait loci
SMR	Summary data-based MR
sc-eQTL	Single-cell eQTL
GWAS	Genome-wide association study
PH4	Statistical probability of H4
FDR	False discovery rate
GHIF-MR	The gene hypothesis influence factor based on MR analysis
CHD	Congenital heart disease
TxCAD	Transplant-associated coronary artery disease
ICD	International classification of diseases
ER	Endoplasmic reticulum
Drp1	Dynamin-related protein 1
MIR	Myocardial ischemia/reperfusion
PDA	Patent ductus arteriosus

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13148-025-01850-y.

Additional file1 (XLSX 223 KB)

Acknowledgements

The authors thank all individuals who provided the original GWAS data and catalog databases mentioned in this study. The authors would also like to acknowledge the use of MR methods and R software packages provided by Medical Pea Studio.

Authors' contributions

S.X.W and R.H.S were involved in data curation, formal analysis, methodology, resources, software, validation, visualization, writing—original draft. X.Y.L, X.J.L, L.B.H, and Y.T.Z helped in conceptualization, data curation, formal analysis, resources, writing—review & editing. J.H.Y contributed to writing—review & editing. Corresponding authors X.X.H and Z.W.S were involved in writing—review & editing, supervision, validation, conceptualization. All authors reviewed the manuscript.

Funding

The authors received no funding for this study.

Availability of data and materials

Individual-level data are not available. We do not report any raw code. Sources for all raw data used in this study are cited within the manuscript. The SMR analysis method is described in the Methods section; the SMR software (SMR-1.3.1) is available from the Yang Lab (https://yanglab.westlake.edu.cn/software/smr/#ExecutableFiles(version1.3.1)), including parameter selection and usage instructions (https://yanglab.westlake.edu.cn/software/smr/#SMR&HEIDIanalysis). Colocalization analysis employed the "colo" package (version 5.2.3; https://github.com/chr1swallace/coloc), and HyPrColoc used the "hyprcoloc" package (version 0.0.2; https://github.com/jrs95/hyprcoloc). PheWAS analysis was performed using the "mrasst" package (version 0.2.0) in R. All SMR results and GWAS/QTL associations for the selected SNPs are provided in Supplementary Table. Interested readers are encouraged to obtain data from their respective sources.

Declarations

Ethics approval and consent to participate

The dataset used in this study is publicly available, and ethical approval and informed consent were obtained before implementation. Therefore, our study requires no additional informed consent or ethical approval.

Competing interests

The authors declare no competing interests.

Author details

[1]Department of Cardiology, The First Affiliated Hospital of Shantou University Medical College, No.57, Changping Road, Shantou 515041, China. [2]Laboratory of Molecular Cardiology, The First Affiliated Hospital of Shantou University Medical College, No.57, Changping Road, Shantou 515041, China. [3]Department of Cardiology, Jie xi's People Hospital, Jieyang, Guangdong, China. [4]Guangdong Pharmaceutical University, Guangzhou, Guangdong, China.

Received: 25 November 2024 Accepted: 19 February 2025
Published online: 05 March 2025

References

1. Khatter H, Myasnikov AG, Natchiar SK, Klaholz BP. Structure of the human 80S ribosome. Nature. 2015;520(7549):640–5.
2. Shen PS, Park J, Qin Y, Li X, Parsawar K, Larson MH, et al. Rqc2p and 60 S ribosomal subunits mediate mRNA-independent elongation of nascent chains. Science. 2015;347(6217):75–8.
3. Pena C, Hurt E, Panse VG. Eukaryotic ribosome assembly, transport and quality control. Nat Struct Mol Biol. 2017;24(9):689–99.
4. Ferreira-Cerca S, Poll G, Gleizes PE, Tschochner H, Milkereit P. Roles of eukaryotic ribosomal proteins in maturation and transport of pre-18S rRNA and ribosome function. Mol Cell. 2005;20(2):263–75.
5. Nerurkar P, Altvater M, Gerhardy S, Schutz S, Fischer U, Weirich C, et al. Eukaryotic ribosome assembly and nuclear export. Int Rev Cell Mol Biol. 2015;319:107–40.
6. Pirruccello JP, Di Achille P, Nauffal V, Nekoui M, Friedman SF, Klarqvist MDR, et al. Genetic analysis of right heart structure and function in 40,000 people. Nat Genet. 2022;54(6):792–803.
7. McKoy G, Protonotarios N, Crosby A, Tsatsopoulou A, Anastasakis A, Coonar A, et al. Identification of a deletion in plakoglobin in arrhythmogenic right ventricular cardiomyopathy with palmoplantar keratoderma and woolly hair (Naxos disease). Lancet. 2000;355(9221):2119–24.
8. Gotschy A, Saguner AM, Niemann M, Hamada S, Akdis D, Yoon J-N, et al. Right ventricular outflow tract dimensions in arrhythmogenic right ventricular cardiomyopathy/dysplasia—a multicentre study comparing echocardiography and cardiovascular magnetic resonance. European Heart Journal - Cardiovascular Imaging. 2018;19(5):516–23.
9. Schmidt AF, Bourfiss M, Alasiri A, Puyol-Anton E, Chopade S, van Vugt M, van der Laan SW, Gross C, Clarkson C, Henry A, Lumbers TR. Druggable proteins influencing cardiac structure and function: implications for heart failure therapies and cancer cardiotoxicity. Sci Adv. 2023;9(17):eadd4984.
10. Fan Y, Cheng Y, Li Y, Chen B, Wang Z, Wei T, et al. Phosphoproteomic analysis of neonatal regenerative myocardium revealed important roles of checkpoint kinase 1 via activating mammalian target of rapamycin C1/ribosomal protein S6 kinase b-1 pathway. Circulation. 2020;141(19):1554–69.
11. Domitrovic T, Moreira MH, Carneiro RL, Ribeiro-Alves M, Palhano FL. Natural variation of the cardiac transcriptome in humans. RNA Biol. 2020;18(10):1374–81.
12. Hannan RD, Luyken J, Rothblum LI. Regulation of rDNA transcription factors during cardiomyocyte hypertrophy induced by adrenergic agents. J Biol Chem. 1995;270(14):8290–7.
13. Tran QK. Reciprocality between estrogen biology and calcium signaling in the cardiovascular system. Front Endocrinol (Lausanne). 2020;11:568203.
14. Deng H-f, Zou J, Wang N, Ma H, Zhu L-l, Liu K, et al. Nicorandil alleviates cardiac remodeling and dysfunction post-infarction by up-regulating the nucleolin/autophagy axis. Cell Signal. 2022. https://doi.org/10.1016/j.cellsig.2022.110272.
15. Tong Z, Tang Y, Jiang B, Wu Y, Liu Y, Li Y, et al. Phosphorylation of nucleolin is indispensable to upregulate miR-21 and inhibit apoptosis in cardiomyocytes. J Cell Physiol. 2018;234(4):4044–53.
16. Holdt LM, Stahringer A, Sass K, Pichler G, Kulak NA, Wilfert W, et al. Circular non-coding RNA ANRIL modulates ribosomal RNA maturation and atherosclerosis in humans. Nat Commun. 2016;7:12429.
17. Wang HX, Zhao YX. Prediction of genetic risk factors of atherosclerosis using various bioinformatic tools. Genet Mol Res. 2016. https://doi.org/10.4238/gmr.15027347.
18. Jiang B, Zhang B, Liang P, Chen G, Zhou B, Lv C, et al. Nucleolin protects the heart from ischaemia–reperfusion injury by up-regulating heat shock protein 32. Cardiovasc Res. 2013;99(1):92–101.
19. Zhao S, Xia Y, Zhang F, Xiong Z, Li Y, Yan W, et al. Nucleostemin dysregulation contributes to ischemic vulnerability of diabetic hearts: role of ribosomal biogenesis. J Mol Cell Cardiol. 2017;108:106–13.
20. Schroeder AM, Allahyari M, Vogler G, Missinato MA, Nielsen T, Yu MS, et al. Model system identification of novel congenital heart disease gene candidates: focus on RPL13. Hum Mol Genet. 2019;28(23):3954–69.
21. Linke AT, Marchant B, Marsh P, Frampton G, Murphy J, Rose ML. Screening of a HUVEC cDNA library with transplant-associated coronary artery disease sera identifies RPL7 as a candidate autoantigen associated with this disease. Clin Exp Immunol. 2001;126(1):173–9.
22. Marrocco V, Bogomolovas J, Ehler E, Dos Remedios CG, Yu J, Gao C, et al. PKC and PKN in heart disease. J Mol Cell Cardiol. 2019;128:212–26.
23. Yun Z, Guo Z, Li X, Shen Y, Nan M, Dong Q, et al. Genetically predicted 486 blood metabolites in relation to risk of colorectal cancer: a Mendelian randomization study. Cancer Med. 2023;12(12):13784–99.
24. Zuccolo L, Holmes MV. Commentary: mendelian randomization-inspired causal inference in the absence of genetic data. Int J Epidemiol. 2017;46(3):962–5.
25. Richmond RC, Davey Smith G. Mendelian randomization: concepts and scope. Cold Spring Harb Perspect Med. 2022. https://doi.org/10.1101/cshperspect.a040501.
26. Burgess S, Butterworth A, Thompson SG. Mendelian randomization analysis with multiple genetic variants using summarized data. Genet Epidemiol. 2013;37(7):658–65.
27. Zhu Z, Zhang F, Hu H, Bakshi A, Robinson MR, Powell JE, et al. Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. Nat Genet. 2016;48(5):481–7.
28. Zhao SS, Yiu ZZN, Barton A, Bowes J. Association of lipid-lowering drugs with risk of psoriasis: a Mendelian randomization study. JAMA Dermatol. 2023;159(3):275–80.
29. Foley CN, Staley JR, Breen PG, Sun BB, Kirk PDW, Burgess S, et al. A fast and efficient colocalization algorithm for identifying shared genetic risk factors across multiple traits. Nat Commun. 2021. https://doi.org/10.1038/s41467-020-20885-8.
30. McRae AF, Marioni RE, Shah S, Yang J, Powell JE, Harris SE, et al. Identification of 55,000 replicated DNA methylation QTL. Sci Rep. 2018;8(1):17605.
31. Gt C. The GTEx consortium atlas of genetic regulatory effects across human tissues. Science. 2020;369(6509):1318–30.
32. Zang Y, Ran X, Yuan J, Wu H, Wang Y, Li H, et al. Genomic hallmarks and therapeutic targets of ribosome biogenesis in cancer. Brief Bioinform. 2024. https://doi.org/10.1093/bib/bbae023.
33. Kurki MI, Karjalainen J, Palta P, Sipila TP, Kristiansson K, Donner KM, et al. FinnGen provides genetic insights from a well-phenotyped isolated population. Nature. 2023;613(7944):508–18.
34. Jain P, Miller-Fleming T, Topaloudi A, Yu D, Drineas P, Georgitsi M, et al. Polygenic risk score-based phenome-wide association study identifies novel associations for Tourette syndrome. Transl Psychiatry. 2023;13(1):69.
35. YaDY B. The control of the false discovery rate in multiple testing under dependency. Ann Stat. 2001;29:1165–88.
36. Yuan S, Xu F, Zhang H, Chen J, Ruan X, Li Y, Burgess S, Åkesson A, Li X, Gill D, Larsson SC. Proteomic insights into modifiable risk of venous thromboembolism and cardiovascular comorbidities. J Thromb Haemost. 2024;22(3):738–48.
37. Yuan S, Xu F, Li X, Chen J, Zheng J, Mantzoros CS, et al. Plasma proteins and onset of type 2 diabetes and diabetic complications: proteome-wide mendelian randomization and colocalization analyses. Cell Rep Med. 2023;4(9): 101174.

38. Verma A, Bradford Y, Dudek S, Lucas AM, Verma SS, Pendergrass SA, et al. A simulation study investigating power estimates in phenome-wide association studies. BMC Bioinformatics. 2018;19(1):120.
39. Delongchamp R, Faramawi MF, Feingold E, Chung D, Abouelenein S. The association between SNPs and a quantitative trait: power calculation. Eur J Environ Public Health. 2018. https://doi.org/10.20897/ejeph/3925.
40. Li S, Schooling CM. A phenome-wide association study of ABO blood groups. BMC Med. 2020;18(1):334.
41. Kurilshikov A, Medina-Gomez C, Bacigalupe R, Radjabzadeh D, Wang J, Demirkan A, et al. Large-scale association analyses identify host factors influencing human gut microbiome composition. Nat Genet. 2021;53(2):156–65.
42. Li P, Wang H, Guo L, Gou X, Chen G, Lin D, et al. Association between gut microbiota and preeclampsia-eclampsia: a two-sample Mendelian randomization study. BMC Med. 2022;20(1):443.
43. Hennessy S, Bilker WB, Berlin JA, Strom BL. Factors influencing the optimal control-to-case ratio in matched case-control studies. Am J Epidemiol. 1999;149(2):195–7.
44. Khaire AS, Wimberly CE, Semmes EC, Hurst JH, Walsh KM. An integrated genome and phenome-wide association study approach to understanding Alzheimer's disease predisposition. Neurobiol Aging. 2022;118:117–23.
45. Prieto JL, McStay B. Recruitment of factors linking transcription and processing of pre-rRNA to NOR chromatin is UBF-dependent and occurs independent of transcription in human cells. Genes Dev. 2007;21(16):2041–54.
46. He S, Ma X, Ye Y, Zhang M, Zhuang J, Song Y, et al. HEATR1 modulates cell survival in non-small cell lung cancer via activation of the p53/PUMA signaling pathway. Onco Targets Ther. 2019;12:4001–11.
47. Lorenzana-Carrillo MA, Tejay S, Nanoa J, Huang G, Liu Y, Haromy A, et al. TRIM35 Monoubiquitinates H2B in Cardiac Cells. Implications for Heart Failure Circ Res. 2024;135(2):301–13.
48. Fang T, Wang J, Sun S, Deng X, Xue M, Han F, et al. JinLiDa granules alleviates cardiac hypertrophy and inflammation in diabetic cardiomyopathy by regulating TP53. Phytomedicine. 2024;130: 155659.
49. Song H, Shuxun R, Fen X, Ruiwen Z, Chi-Ming W. Presence of P53 and transforming growth factor beta-1 in human cardiac tissue with congestive heart failure. Microsc Microanal. 1997;3(S2):149–50.
50. He X, Cantrell AC, Williams QA, Gu W, Chen Y, Chen JX, et al. p53 acetylation exerts critical roles in pressure overload-induced coronary microvascular dysfunction and heart failure in mice. Arterioscler Thromb Vasc Biol. 2024;44(4):826–42.
51. Chen X, Su Q, Ling X, Yang Y, Liu Y, Zhu X, He A, Wu H, Qi Y. SENP3-regulated Nodal signaling plays a potential role in cardiac left-right asymmetry development. Int J Biol Macromol. 2024;274(Pt 2):133294.
52. Gao L, Zhao Y, He J, Yan Y, Xu L, Lin N, et al. The desumoylating enzyme sentrin-specific protease 3 contributes to myocardial ischemia reperfusion injury. J Genet Genomics. 2018;45(3):125–35.
53. Rawlings N, Lee L, Nakamura Y, Wilkinson KA, Henley JM. Protective role of the deSUMOylating enzyme SENP3 in myocardial ischemia-reperfusion injury. PLoS ONE. 2019;14(4): e0213331.
54. Kupsco JM, Wu MJ, Marzluff WF, Thapar R, Duronio RJ. Genetic and biochemical characterization of Drosophila Snipper: A promiscuous member of the metazoan 3'hExo/ERI-1 family of 3' to 5' exonucleases. RNA. 2006;12(12):2103–17.
55. Chai T, Tian M, Yang X, Qiu Z, Lin X, Chen L. Association of Circulating Cathepsin B Levels With Blood Pressure and Aortic Dilation. Front Cardiovasc Med. 2022;9: 762468.
56. Declercq M, Biquand E, Karim M, Pietrosemoli N, Jacob Y, Demeret C, et al. Influenza A virus co-opts ERI1 exonuclease bound to histone mRNA to promote viral transcription. Nucleic Acids Res. 2020;48(18):10428–40.
57. Choucair N, Rajab M, Megarbane A, Chouery E. Homozygous microdeletion of the ERI1 and MFHAS1 genes in a patient with intellectual disability, limb abnormalities, and cardiac malformation. Am J Med Genet A. 2017;173(7):1955–60.
58. Altinkilic EM, Isbir S, Gormus U, Yilmaz SG, Dalan AB, Duman S, Isbir T. RRM1, RRM2 and ERCC2 gene polymorphisms in coronary artery disease. In Vivo. 2016;30(5):611–5.
59. Foret MK, Orciani C, Welikovitch LA, Huang C, Cuello AC, Do CS. Early oxidative stress and DNA damage in Abeta-burdened hippocampal neurons in an Alzheimer's-like transgenic rat model. Commun Biol. 2024;7(1):861.
60. Chen B, Hou A, Zhao L, Liu Y, Shi X, Du B, et al. Next generation sequencing identify rare copy number variants in non-syndromic patent ductus arteriosus. Front Genet. 2020;11: 600787.
61. Gan Y, Deng J, Hao Q, Huang Y, Han T, Xu JG, et al. UTP11 deficiency suppresses cancer development via nucleolar stress and ferroptosis. Redox Biol. 2023;62: 102705.
62. Yan M, Gao J, Lan M, Wang Q, Cao Y, Zheng Y, et al. DEAD-box helicase 17 (DDX17) protects cardiac function by promoting mitochondrial homeostasis in heart failure. Signal Transduct Target Ther. 2024;9(1):127.
63. Ning D, Jin J, Fang Y, Du P, Yuan C, Chen J, et al. DEAD-Box Helicase 17 exacerbates non-alcoholic steatohepatitis via transcriptional repression of cyp2c29, inducing hepatic lipid metabolism disorder and eliciting the activation of M1 macrophages. Clin Transl Med. 2024;14(2): e1529.
64. Shabani M, Wang M, Jenkins GD, Rotter JI, Rich SS, Batzler A, et al. Myocardial fibrosis and cardiomyopathy risk: a genetic link in the MESA. Circ Heart Fail. 2023;16(9): e010262.
65. Sun Y, Chen Y, Li Y, Li Z, Li C, Yu T, et al. Association of TSR1 variants and spontaneous coronary artery dissection. J Am Coll Cardiol. 2019;74(2):167–76.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.