Supplementary Materials for

Leveraging the shared and opposing genetic mechanisms in the heritable cardiomyopathies

Daria R. Kramarenko[1,2,3], Poeya Haydarlou[1], George J. Powell[4], Joel T. Rämö[5,6,7], Riyad Janan[4], Claire Prince[4], Dominic S. Zimmerman[1], Pantazis Theotokis[4], Prisca K. Thami[4], Jan Haas[8,3], Sophie Garnier[9,10,3], Frank Rühle[11,12,3], Edwin Poel[1,3], Amand F. Schmidt[13,3,14,15,16], Sharlene Day[17,18], Adam Helms[19,18], Rachel Lampert[20,18], Victoria Parikh[21,18], Jodie Ingles[22,18], Iacopo Olivotto[23,18], Neal Lakdawala[24,18], Anjali Owens[17,18], Sara Saberi[19,18], John Stendhal[20,18], Euan Ashley[21,18], Belinda Gray[25,18], Mark W. Russell[19,18], Thomas D. Ryan[26,18], Joseph W. Rossano[27,18], Dominic Abrams[28,18], Erin Miller[26,18], Kimberly Lin[29,18], Niccolo Maurizi[30,18], Alessia Argiro[23,18], Colin Berry[31,32], Rob Cooper[33,32], Andrew S. Flett[34,32], Roy S. Gardner[31,35,32], John P. Greenwood[36,37,38,32], Brian P. Halliday[39,40,32], David Hutchings[41,32], Masliza Mahmod[42,32], Gerry P. McCann[43,32], Stephen P. Page[37,32], Charles Peebles[44,32], Betty Raman[42,32], Peter Swoboda[38,32], Amanda Varnava[45,32], David Wright[33,32], Sanjay Prasad[39,32], Stuart Cook[46,32], Upsala (Paz) Tayal[39,40,32], Rachel Buchan[40,32], Roddy Walsh[47,3], Arthur A. M. Wilde[1,2,13], Benjamin Meder[8,3,48], Philippe Charron[49,10,50], Anuj Goel[51], Ahmad S. Amin[1,2,13,3], Patrick T. Ellinor[6,52], Krishna G. Aragam[6,52], Rafik Tadros[53], Yigal M. Pinto[1,13,2,3], Carolyn Y. Ho[24,18], Hugh Watkins[51], James S. Ware[4,54,18,3], Connie R. Bezzina *[1,3], Sean J. Jurgens *[1,5,13,3]

1 Department of Experimental Cardiology, Amsterdam Cardiovascular Sciences, Heart Failure & Arrhythmias, Amsterdam UMC location University of Amsterdam, Amsterdam, the Netherlands

2 European Reference Network for rare low-prevalence and complex diseases of the heart: ERN GUARD-Heart

3 DCM-NEXT

4 National Heart and Lung Institute & MRC Laboratory of Medical Sciences, Imperial College London, London, UK

5 Cardiovascular Disease Initiative, Broad Institute of MIT and Harvard, Cambridge, MA, USA

6 Cardiovascular Research Center, Massachusetts General Hospital, Boston, MA, USA

7 Institute for Molecular Medicine Finland (FIMM), Helsinki Institute of Life Science (HiLIFE), University of Helsinki, Helsinki, Finland

8 Department of Medicine III, Institute for Cardiomyopathies Heidelberg (ICH), University Hospital Heidelberg, Heidelberg, Germany

9 Research Unit on Cardiovascular Disorders, Metabolism and Nutrition, Team Genomics and

Pathophysiology of Cardiovascular Disease, Sorbone Université, INSERM, Paris, France
10 ICAN Institute for Cardiometabolism and Nutrition, Paris, France
11 Bioinformatics Core Facility, Institute of Molecular Biology gGmbH (IMB), Mainz, Germany
12 Department of Genetic Epidemiology, Institute of Human Genetics, University of Münster, Münster, Germany
13 Department of Clinical Cardiology, Amsterdam UMC location University of Amsterdam, Amsterdam, the Netherlands
14 Institute of Cardiovascular Science, Faculty of Population Health, University College London, London, UK
15 University College London British Heart Foundation Research Accelerator, London, UK
16 Department of Cardiology, Division Heart and Lungs, University Medical Center Utrecht, Utrecht University, Utrecht, the Netherlands
17 University of Pennsylvania, Philadelphia, PA, USA
18 SHaRe
19 Division of Pediatric Cardiology, University of Michigan, Ann Arbor, MI
20 Yale School of Medicine, New Haven, CT, USA
21 Stanford School of Medicine, Stanford, CA, USA
22 Garvan Institute of Medical Research, Sydney, NSW, Australia
23 University of Florence, Florence, Italy
24 Cardiovascular Division, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA
25 University of Sydney, Sydney, NSW, Australia
26 Cincinnati Children’s Hospital Medical Center, Cincinnati, OH, USA
27 Children’s Hospital of Philadelphia & University of Pennsylvania, Philadelphia, PA, USA
28 Boston Children’s Hospital, Harvard Medical School, Boston, MA, USA
29 Children’s Hospital of Philadelphia, Philadelphia, PA, USA
30 University Hospital Lausanne (CHUV), Lausanne, Switzerland
31 Golden Jubilee National Hospital, Clydebank, Scotland, UK
32 GoDCM
33 University of Liverpool / Liverpool Heart and Chest Hospital NHS Foundation Trust, Liverpool, UK
34 University Hospital Southampton NHS Foundation Trust, Southampton, UK
35 University of Glasgow, Glasgow, UK
36 Baker Heart and Diabetes Institute, Melbourne, Australia
37 Leeds Teaching Hospitals NHS Trust, Leeds, UK
38 The University of Leeds, Leeds, UK
39 National Heart and Lung Institute, Imperial College London, London, UK
40 Royal Brompton and Harefield Hospitals, Guy’s and St. Thomas NHS Foundation Trust, London, UK
41 University of Manchester, UK
42 University of Oxford, Oxford, UK

43 Division of Cardiovascular Sciences, University of Leicester, and the NIHR Leicester Biomedical Research Centre, Glenfield Hospital, Leicester, UK
44 University of Southampton, Southampton, UK
45 Imperial College Healthcare NHS Trust, London, UK
46 MRC Laboratory of Medical Sciences, Imperial College London, London, UK
47 Cardiovascular and Genomics Research Institute, St George's University of London, London, UK
48 Site Heidelberg/Mannheim, DZHK, Heidelberg, Germany
49 Sorbonne Université, Inserm, Unité de recherche sur les maladies cardiovasculaires et métaboliques, ICAN, F-75013 Paris, France
50 APHP, Cardiology and Genetics Departments, Pitié-Salpêtrière Hospital, Paris, France
51 Division of Cardiovascular Medicine, Radcliffe Department of Medicine, University of Oxford, John Radcliffe Hospital, Oxford, UK
52 Cardiovascular Genomics Initiative, Cleveland Clinic, Cleveland, OH, USA
53 Electrophysiology Service and Adult Congenital Heart Disease Center, Montreal Heart Institute, Université de Montréal, Montreal, Canada.
54 Program in Medical and Population Genetics, Broad Institute of Harvard and MIT, Cambridge, MA, USA

* contributed equally

Correspondence to:
s.j.jurgens@amsterdamumc.nl
c.r.bezzina@amsterdamumc.nl

Table of contents:

Supplementary Note

Data sources

To investigate the genetic similarity between DCM and HCM, we leveraged large-scale genome-wide association (GWAS) data from two recent studies. Data for DCM were obtained from a large GWAS meta-analysis by Jurgens *et al.* (2024)[1]. In this study, case-control GWAS data were assembled from 6 European-ancestry datasets, including clinical case-control datasets (4,343 clinically ascertained DCM cases) and biobank datasets (5,022 DCM cases defined by billing-codes). This GWAS included Using a genome-wide significance threshold ($P< 5x10^{-8}$), 38 distinct loci were reported. To maximize discovery, the GWAS data were subsequently integrated into a multi-trait GWAS (MTAG)[2] with GWAS data for MRI-derived left ventricular (LV) traits (global circumferential strain, indexed left ventricle end-systolic volume (LVESVi), and ejection fraction ; N=36,083; ref.[3]). From this MTAG, 65 significant loci were reported ($P< 5x10^{-8}$). Of note, further details on the MTAG methodology are described below.

Data for HCM were retrieved from a recent GWAS meta-analysis by Tadros *et al.* (2025)[3]. This study included a total of 5,900 clinically-ascertained HCM cases and 68,359 controls of European genetic ancestry. At genome-wide significance ($P< 5x10^{-8}$), 34 distinct loci were reported. Similar to the DCM study, an MTAG approach was used to boost discovery: HCM GWAS was integrated with GWAS data for three LV traits (global circumferential strain, indexed left ventricle end-systolic volume (LVESVi), and LV concentricity). From this MTAG, 68 significant loci were reported ($P< 5x10^{-8}$).
In this study, we processed and utilized both the single-trait GWAS and MTAG summary statistics, for both DCM and HCM.

Processing of GWAS summary statistics

We processed the DCM and HCM GWAS summary statistics, aligning all datasets to genome build GRCh37. We then aimed to remove variants driven by disproportionately small sample sizes. To this end, for DCM GWAS, we restricted to variants with at least 70% of the total case number contributing to the meta-analysis. In the HCM GWAS summary statistics only the total sample size was available, and

therefore variants were restricted to those with at least 96% of the total sample size contributing to the meta-analysis. Of note, the broad *MYBPC3* locus in HCM GWAS is known to tag rare founder variants[4]; we therefore removed the extended region surrounding this locus (chr11: 29,978,453–80,288,956) from the summary statistics[5]. These filters left 6,635,031 variants for DCM GWAS, and 6,035,750 variants for HCM GWAS (**Supplementary table 1**).

Processing of MTAG summary statistics

DCM and HCM MTAG summary statistics were also reprocessed to ensure consistent loci annotation and gene prioritization across studies and approaches. The datasets were aligned to genome build GRCh37. For DCM MTAG, to remove variants with disproportionately small contributing sample size, we removed variants with effective sample size <70% of the maximum effective sample size. For HCM MTAG, we restricted to variants that passed all filters in the filtered HCM GWAS summary statistics above. Finally, we removed the extended *MYBPC3* region from both datasets. These filters left 5,513,180 variants for DCM MTAG, and 5,117,470 variants for HCM MTAG.

Genetic correlation with LV traits

To assess shared genetic architecture across the cardiomyopathy spectrum and quantitative cardiac traits, we estimated genetic correlations using bivariate LD score regression [6]**(Methods)**. Specifically, we computed genetic correlations between CC-GWAS summary statistics and GWAS results for ten left ventricular (LV) traits relevant to cardiomyopathy, measured in 36,083 participants from the UK Biobank (UKB). [7] The strongest correlations were observed with LVESVi ($r_{g,global}$ = 0.624), global circumferential strain (Ecc) ($r_{g,global}$ = 0.705), and LV concentricity (LVconc) ($r_{g,global}$ = -0.575), indicating substantial genetic overlap. **(Methods; Supplementary Figure 2**; **Supplementary Table 7)** Interestingly, these correlations were notably stronger than those observed for previously published single trait case–control GWAS of DCM and HCM: the correlation between DCM GWAS and LVESVi was $r_{g,global}$= 0.7, DCM GWAS and Ecc was $r_{g,global}$= 0.747 and between HCM GWAS and LVconc was $r_{g,global}$= 0.61(**Supplementary Table 3**). These findings suggest that CC-GWAS captures a better genetic spectrum of cardiac traits as compared to the traditional approach.

Novel regions identified by LAVA

Three novel genomic regions exhibited significant regional genetic correlation between DCM and HCM. Neither of these regions were captured by DCM or HCM GWAS. These regions were located on:

- Region 924 – chromosome 5 (chr5:178,595,253–179,794,710; GRCh37):

  Although univariately subthreshold for both DCM ($P = 2.43 \times 10^{-8}$) and HCM ($P = 6.05 \times 10^{-14}$), this region showed a significant **negative local genetic correlation** ($\rho = -0.52$, $P = 2.07 \times 10^{-4}$). The lead variant in DCM GWAS was **rs4701067** ($P = 0.03$, $\beta = 0.04$), and in HCM GWAS, **rs7733548** ($P = 0.001$, $\beta = -0.09$), both mapping to *ADAMTS2*, a gene associated with dermatosparaxis-type Ehlers–Danlos syndrome (OMIM).
  Region 1277 – Chromosome 8 (chr8:32,454,963–33,982,537; GRCh37): This region showed moderate univariate association in DCM ($P = 5.90 \times 10^{-6}$) and strong association in HCM ($P = 3.42 \times 10^{-11}$), with a robust **inverse local correlation** ($\rho = -0.61$, $P = 4.54 \times 10^{-4}$). The lead DCM variant was **rs62510527** ($P = 0.0001$, $\beta = -0.067$; near **POFUT3**), and the top HCM variant was **rs17665441** ($P = 0.7$, $\beta = -0.005$; near **NRG1**), a gene previously associated with schizophrenia susceptibility (OMIM).)

- Region 1948 – chromosome 13 (chr13:109,813,577–110,995,432 (GRCh37)): This region was also identified in both CC-MTAG and HCM-MTAG analyses and prioritized to the COL4A1 locus, this region demonstrated significant heritability in both traits (DCM $P = 1.63 \times 10^{-11}$; HCM $P = 1.97 \times 10^{-9}$), with a notable opposing effect direction ($\rho = -0.64$, $P = 1.10 \times 10^{-5}$).

Together, these findings underscore the power of local genetic correlation analyses to uncover biologically relevant loci beyond conventional GWAS significance thresholds, particularly those with antagonistic effects across disease subtypes.

Genomic structural equation modeling

After having computed $r_{g,global}$ between DCM and HCM, we then aimed to re-compute $r_{g,global}$ accounting for the effect of other heritable traits. First, we aimed to account for blood pressure and body habitus traits, given that these extracardiac traits have been mentioned in literature as being risk factors for both DCM and HCM with concordant directionality (high blood pressure and higher body weight have been described as risk factors for DCM and HCM). To account for these heritable traits in our analysis, we used genomic Structural Equation Modeling, implemented in the *GenomicSEM* R-package. We first used the *ldsc()* function to compute pairwise $r_{g,global}$ values for all pairs of traits from DCM, HCM, systolic blood pressure[8] (SBP), diastolic blood pressure[8] (DBP), body-weight[8] and body-mass-index[9] (BMI). We then used the *usermodel()* function to fit a Structural Equation model using the following approach:

DCM ~ a1*SBP + a2*DBP + a3*BMI + a4*weight

HCM ~ b1*SBP + b2*DBP + b3*BMI + b4*weight

DCM ~~ r*HCM

SBP ~~ SBP

SBP ~~ DBP

SBP ~~ BMI

SBP ~~ weight

DBP ~~ BMI

DBP ~~ weight

DBP ~~ DBP

BMI ~~ BMI

BMI ~~ weight

weight ~~ weight

DCM ~~ DCM

HCM ~~ HCM

Essentially, this model regresses the four risk factors on DCM, and separately also on HCM, while allowing for covariance between each of the risk factors and between DCM and HCM. From the resulting model fit, we extracted the scaled covariance between DCM and HCM representing the $r_{g,global}$ between DCM and HCM conditional on the heritable components of SBP, DBP, weight and BMI.

In a similar fashion, we also computed the $r_{g,global}$ between DCM and HCM, conditional on cardiac endophenotypes from MRI - namely those most strongly associated with DCM and HCM including LVESVi, LV concentricity (LVconc) and global circumferential strain (Ecc)[3]. The input model was specified as follows:

DCM ~ a1*Ecc + a2*LVESVi + a3*LVconc

HCM ~ b1*Ecc + b2*LVESVi + b3*LVconc

DCM ~~ r*HCM

Ecc ~~ Ecc

Ecc ~~ LVESVi

Ecc ~~ LVconc

LVESVi ~~ LVconc

LVESVi ~~ LVESVi

LVconc ~~ LVconc

DCM ~~ DCM

HCM ~~ HCM

Case-case GWAS

To identify genetic variants that differentiate between DCM and HCM, we applied CC-GWAS, a summary-statistics-based method that estimates genetic divergence between cases of two disorders using case–control GWAS results. To construct CC-GWAS from our available GWAS summary statistics, we used the *CCGWAS* R-package (v0.1.0) [10] This method calculates allele frequency differences between DCM and HCM cases (A1 vs. B1) by leveraging the observed effects in DCM vs. controls (A1A0) and HCM vs. controls (B1B0). Central to this approach is the genetic distance measure $F_{ST,causal}$, defined as the average normalized squared difference in allele frequencies at causal SNPs across case-controls GWAS, which reflects the degree of genetic separation between the two phenotypes.

CC-GWAS estimates the case–case effect size ($\hat{\beta}_{A1B1}$) for each SNP using a weighted linear combination of the case–control GWAS effect sizes:

$$\hat{\beta}_{A1B1} = \omega_{A1A0} \cdot \hat{\beta}_{A1A0} + \omega_{B1B0} \cdot \hat{\beta}_{B1B0}$$

where ω_{A1A0} and ω_{B1B0} are trait-specific weights. Two weighting schemes are implemented:

(1) **CC-GWAS$_{OLS}$ weights**, optimized to minimize the expected squared error between the estimated and true A1B1 effect sizes, accounting for SNP heritabilities, disease prevalences, genetic correlation, sample sizes, and sample overlap; and (2) **CC-GWAS$_{exact}$ weights**, a conservative, sample-size-independent formulation based only on population prevalences:

$$\hat{\beta}_{A1B1}^{exact} = (1 - K_A)\widehat{\beta}_{A1A0} - (1 - K_B)\widehat{\beta}_{A1A0}\widehat{\beta}_{B1B0}$$

While the OLS weights provide higher power, they may be susceptible to type I error at so-called *stress test SNPs*—variants with significant and similarly directed effects in both case–control GWASs (nonzero A1A0 and B1B0) but no true case–case difference (A1B1 = 0). These variants can appear falsely significant due to random sampling variation. To address this, CC-GWAS applies a dual-threshold strategy: a

SNP is declared significant only if it passes genome-wide significance ($P < 5 \times 10^{-8}$) with OLS weights and a secondary threshold ($P < 10^{-4}$) with exact weights, thereby maximizing power while controlling type I error. This approach enhances sensitivity to opposite-direction genetic effects, while suppressing signals that are shared between diseases.

For this analysis, we used the processed summary statistics from DCM and HCM GWAS, along with the following input parameters: i) the assumed population prevalences (0.4% for DCM, 0.2% for HCM); ii) the case/control numbers in DCM GWAS and HCM GWAS, with some attenuation for potential missingness (see below); iii) the heritabilities from LDSC (14.2% for DCM, 18% for HCM); iv) the $r_{g,global}$ between DCM and HCM, and its intercept, from LDSC ($r_{g,globl}$=-0.56 and $error_{covariance}$=0.012)[6]; and v) the number of effectively independent causal variants for DCM (1200; see below)[10]. Naturally, CC-GWAS was restricted to genetic variants found in the processed DCM and HCM GWAS summary statistics; after additional automatic filtering by the *CCGWAS* package, 4987309 high-quality variants remained in the CC-GWAS analysis and resulting summary statistics. Genome-wide significance was defined as $P<5\times10-8$, and all hypothesis tests were two-sided.

To account for some degree of per-variant sample missingness, we adjusted the input case/control numbers - used as input to *CCGWAS*. Notably, because the meta-analytical case/control numbers in the summary statistics were based on the maximum sample size of contributing cohorts, the provided numbers reflect maximum values and are therefore broadly overestimated. This is relevant because *CCGWAS* computes expected effect sizes based on the case/control numbers. Indeed, *CCGWAS* raised warnings indicating that the expected effect sizes were not well-calibrated. Consistent with some overestimation of case/control numbers, we found that *CCGWAS* gave well-calibrated effect sizes when we assumed some attenuation of case/control numbers across all variants. For DCM GWAS, we attenuated case/control numbers to 90% of the maximum numbers, while for HCM GWAS we attenuated the numbers by 85%, when inputted to the *CCGWAS* software.

To estimate the number of independent causal variants, we used stratified fourth moments regression (https://github.com/lukejoconnor/SLD4M)[11]. This method computes from a GWAS the polygenicity, expressed as the effective number of independently associated causal variants (where the 'effective' clause accounts not only for the potential number of causal variants but also the relative effect size of causal variants). Assuming DCM to be more polygenic than HCM, we put forward the polygenicity statistic estimated from DCM GWAS (1223) to the *CCGWAS* software.

Locus definitions, variant annotation and gene prioritization

Locus definition, variant annotation, and gene prioritization were performed using a unified pipeline across all summary statistics, including DCM GWAS, DCM MTAG, HCM GWAS, HCM MTAG, CC-GWAS, and CC-MTAG. (**Code availability**)

Processing of summary statistics using FUMA

Each set of summary statistics was first processed using Functional Mapping and Annotation (FUMA)[12] v1.6.1 (https://fuma.ctglab.nl/). Among other analyses, FUMA applies Multi-marker Analysis of GenoMic Annotation (MAGMA; v.1.08) to perform an initial gene-based association analysis, by aggregating variant-level signals into gene-level statistics while accounting for linkage disequilibrium[13]. The MAGMA gene-level scores were also used by FUMA to test for tissue-specific enrichment of RNA expression profiles, based on transcriptomic profiles across dozens of tissues from the GTEx v8 dataset (GTEx/v8/gtex_v8_ts_general_avg_log2TPM)[14]. The MAGMA gene scores and tissue enrichment statistics were used as input for our gene prioritization pipeline, as described in detail below.

Fine-mapping and credible set formatting

Our gene nomination pipeline required finding credible sets that likely include the causal variants from the respective GWAS. To identify such credible sets, we performed fine-mapping using the SuSiER algorithm (v0.12.35)[15,16]. For each GWAS dataset, the SuSiER algorithm was run within separate LD blocks derived from UK Biobank European-ancestry reference data.[17] The minimum squared correlation was set to 0.5 (the default), unless the algorithm failed to converge, in which case we relaxed the threshold to 0.25. If SuSiE continued to fail in a region harboring genome-wide significant variants, we flagged the respective LD region and generated an artificial credible set using only the most significant variant in the region **(Supplementary Tables 24-36)**

Gene prioritization using FLAMES

To perform gene prioritization, we used the recently-described 'fine-mapped locus assessment model of effector genes' (FLAMES) approach (v1.1.1)[18]. FLAMES combines two main approaches to gene prioritization in a weighted framework to compute causal gene predictions that outperform prior methods. In particular, FLAMES first uses pre-fit machine learning models (based on XG-Boost) to link fine-mapped variants to likely effector genes based on various parameters including variant-to-gene distance, epigenomic context, and quantitative trait loci. Second, FLAMES uses the Polygenic Priority Score (PoPS[19]) method

to learn gene features associated with the trait based on functional networks; features consist of cell-type-specific gene expression, biological pathways and protein–protein interactions (PPIs).

We then applied the FLAMES framework to each of our GWAS datasets. To this end, for a given GWAS dataset, we first ran PoPS (v0.2),[19] using the MAGMA Z-scores as input and using the full feature matrix provided by the PoPS developers. We then annotated each credible set using the *annotate* module from FLAMES, which combines variant-to-gene mappings, MAGMA Z-scores, PoPS scores, and GTEx tissue enrichment data.

FLAMES then returned a ranked list of genes per locus in *FLAMES_scores.preds*, including raw and scaled FLAMES scores, XG-Boost scores, PoPS scores, and estimated precision.

Locus definition and consolidation across studies

For each credible set, we selected the top variant based on the highest posterior inclusion probability (PIP), or, in cases where fine-mapping failed, the variant with the lowest *P*-value. All index variants were then sorted by chromosome and genomic position. Index variants located within 1Mb of one another were merged into one locus, to define non-overlapping genomic loci. Each locus was assigned a unique identifier based on its genomic position, with consistent numbering maintained across all analyses (**Supplementary Table 2**).

Gene prioritization across studies

Despite applying a harmonized pipeline for gene prioritization across the various GWAS datasets, it was possible for the FLAMES algorithm to nominate different causal genes within the same locus in different GWAS datasets. To consolidate gene-level evidence within and across datasets, we therefore applied a scoring framework to prioritize effector genes at each locus. For each study, genes identified as top-ranked by either PoPS or FLAMES were assigned 0.5 points per method per study. Scores were then aggregated across all studies. For example, locus 12 (chr1:212,107,306–212,277,107) appeared in both DCM MTAG and CC MTAG. In DCM MTAG, *DTL* was prioritized by both PoPS and FLAMES (score = 1), while in CC MTAG, *BATF3* was prioritized by PoPS and *DTL* again by FLAMES. This resulted in cumulative rank scores of *DTL* = 1.5 and *BATF3* = 0.5. Accordingly, *DTL* was selected as the reported gene for this locus.

For each locus, the gene(s) with the highest total score were designated as lead candidates. In cases where multiple genes had equal scores, or where the difference between top-scoring genes was <1.0, all were retained as joint candidates. This strategy enabled the identification of both study-specific and consensus lead genes across DCM, HCM, and case–case analyses. While we acknowledge that the approach is to an extent arbitrary, we applied this approach to transparently indicate instances where gene prioritization produced potentially inconsistent results. Reassuringly, we found that a single effector gene was nominated in the vast majority of loci using this approach (**Extended Data Figure 3**). The final locus-level summary included genomic coordinates, contributing studies, top-ranked genes, prioritization scores, and selected lead gene(s).**(Supplementary Table 2).**

Functional enrichment analysis

We used the g:Profiler platform[20] (v. February 2025) to test for enrichment of gene sets from several predefined sources for genes curated from CC-GWAS and CC-MTAG. The g:Profiler algorithm uses one-sided Fisher's exact tests to test for enrichment of a prespecified list of genes across many gene sets, and subsequently adjusts one-sided P values for multiple testing while taking into account the correlation between gene sets (g:SCS method76). We used default settings with a multiple testing correction based on the Benjamini–Hochberg FDR and retained terms with adjusted P-values < 0.05. Gene set categories included Gene Ontology (GO: Biological Process, Molecular Function, and Cellular Component), KEGG, Reactome, WikiPathways, CORUM, Human Protein Atlas, Human Phenotype Ontology, transcription factor targets, and miRNA–target interactions.

First we tested all genes from CC-GWAS, then genes unique for CC-GWAS or CC-MTAG for loci that were not significant in other DCM and HCM GWAS and MTAG.

To quantify the strength of association for each term, we computed the odds ratio (OR) using a custom function based on contingency table parameters derived from term size, query size, and domain background size. Continuity correction was applied to avoid division by zero where needed.

To reduce redundancy in GO terms and annotate broader biological themes, we used REVIGO[21] to group enriched GO terms by semantic similarity. We parsed REVIGO output and linked each original term to a representative parent term, which was then used to group and annotate terms across GO:BP, GO:CC, and GO:MF domains.

A custom R workflow was developed to integrate enrichment results with REVIGO clusters, calculate ORs, and visualize results. We generated a volcano plot with odds ratio on the x-axis and $-\log_{10}$(adjusted P-value) on the y-axis. Select representative terms were labeled using the REVIGO group name.

This approach allowed us to highlight key functional pathways enriched among prioritized genes, including protein binding (GO:MF), sarcomeric and cytoskeletal structure (GO:CC), and cell junction organization and signal transduction (GO:BP). Since our prioritized genes may have been preselected towards genes with high cardiac expression (that is, through gene features learnt by PoPS), we performed a sensitivity analysis using nearest genes.

To generate a nearest-gene annotation for loci identified in the CC GWAS and CC MTAG analyses, we used the get_nearest_gene() function from the *gwasRtools* package (v0.1.0; available via GitHub: lcpilling/gwasRtools). For each lead SNP, the nearest protein-coding gene within ±500 kb (500,00 base pairs) was identified using coordinates aligned to human genome build GRCh37. This approach was applied separately to loci from CC GWAS and CC MTAG, producing two corresponding gene lists. These lists were used for pathway enrichment analysis alongside genes prioritized using the FLAMES/PoPS framework.

Pathway enrichment of nearest-gene sets (**Extended Data Figure 6b,d**) revealed broadly consistent biological pathways compared to functionally informed prioritization (**Extended Data Figure 6a,c**), including strong enrichment for muscle structure development, actin binding, cytoskeletal organization, and myofibril assembly. Notably, terms such as "actin binding", "cytoskeleton", and "myofibril" remained significant under both strategies, suggesting that core cardiomyocyte structural processes are recurrently implicated across methods.

However, enrichment significance was generally reduced when using nearest-gene annotation, and several key terms observed with FLAMES/PoPS—such as sarcomere organization or transcriptional regulation—were absent. This may reflect the limited precision of proximity-based gene assignment, especially in regions with multiple genes or regulatory elements acting at a distance.

Taken together, these results support the robustness of the key functional pathways implicated in our study while highlighting the additional specificity provided by functionally informed gene prioritization.

Cell type enrichment methods

Using the snRNA-seq data obtained from Reichart et al., 2022 (ref.[22]), we performed cell type enrichment analyses. The dataset consisted of samples from several anatomical locations (including several locations across the left and right ventricle) from 61 cardiomyopathy patients - of which 52 with DCM - and 18 non-failing controls. We focused on the 18 non-failing donors, and generated cell type-specific and cell state-specific annotations for enrichment testing using stratified linkage disequilibrium score regression within the *sc-linker* framework[23].

First, we defined cell types from cell type and cell state annotations provided with the publicly-available dataset. We removed variants flagged as 'native' or 'low-QC'. Nuclei with cell state 'PC1', 'PC2' or 'PC3' were then collapsed into 'Pericytes'. Nuclei with cell state 'SMC1.1', 'SMC1.2', or 'SMC2' were collapsed into 'VSMC'. Nuclei with cell state 'EC7' were assigned 'Endocardial'. Nuclei with cell state 'Meso' were assigned 'Epicardial'. Nuclei
with cell state 'EC8' were assigned 'Lymphatic endothelial'. Nuclei with cell state 'EC1.0', 'EC2.0', 'EC5.0', or 'EC6.0' were assigned 'Cardiac endothelial'. For remaining nuclei (those with cell states not mentioned above) the cell type annotations provided with the original dataset were retained. This approach left 11 distinct cardiac cell types.

To test for enrichment of cell type specific gene programs in our GWAS/MTAG datasets, we created cell-type specific gene programs. To this end, we performed 'pseudo-bulk' aggregation by summing gene counts across nuclei for each donor/tissue region combination, by cell type. We only retained a given donor/tissue region combination if they had at least 50 nuclei of that cell type. Lowly expressed genes identified with the filterByExpr() function in edgeR were removed. We normalized the pseudo-bulk expression with DESeq2 and fit the differential expression model ~0+cell_type+donor_tissue using limma-voom. Notably, we included a covariate for the donor/tissue region combination because each

559 donor/tissue region will be represented across most cell types. We then extracted
560 contrasts comparing gene expression in each focal cell type to all other cell types.
561 Cell type-specific gene programs were subsequently computed by ranking and
562 scoring genes based on their enrichment statistics, as described in previous work[23].
563 Notably, however, we adapted the algorithm to set all genes with negative
564 enrichment scores (ie, those depleted within the focal cell type as compared to the
565 other cell types) to 0; this was applied to avoid cell type enrichments driven by genes
566 that were in fact enriched in other cell types.
567
568 Using the cell type-specific gene expression profiles, we then performed heritability
569 enrichment analyses using the *sc-linker* pipeline (https://github.com/kkdey/GSSG)[23].
570 To this end, we used the epigenomic variant-to-gene mapping data for heart and
571 fetal heart, which were provided with the software, as input. We used the CC-GWAS
572 dataset as GWAS input. We then used the provided scripts to apply stratified LD-
573 score regression to compute heritability enrichment statistics for the cell type-specific
574 gene programs[24]. As recommended[24], we report test statistics and corresponding
575 one-sided *P*-values from the tau 'coefficient' - which is conditional on all other
576 annotations included in the model including the 'baseline LD' annotations. To
577 account for the 11 cell types tested, we applied a Bonferroni significance cutoff by
578 setting significance at 0.05/11=0.0.0045.
579

580 ## Partitioned heritability of CC loci using LDSC

581 To evaluate the contribution of loci identified through case–case GWAS (CC GWAS
582 and CC MTAG) to overall SNP-based heritability of DCM and HCM, we performed
583 partitioned heritability analysis using LDSC[25]. First, we selected genome-wide
584 significant variants ($P < 5 \times 10^{-8}$) from both CC GWAS and CC MTAG, then defined
585 500 kb flanking windows upstream and downstream of each lead SNP. These
586 regions were merged using *bedtools merge*[26] to create a non-redundant set of
587 genomic intervals (LD regions) comprising all CC-significant loci.

588 These merged regions were used to generate binary annotation files according to
589 LDSC documentation[25]. To estimate partial heritability, we applied LDSC *--h2* with
590 both baselineLD v2.2 annotations provided by LDSC developers and the newly

591 defined CC locus annotations. Heritability estimates were calculated separately for
592 DCM and HCM GWAS summary statistics, and enrichment was quantified as the
593 proportion of heritability explained divided by the proportion of SNPs annotated in
594 each category.

595 CC-significant loci comprised only 2.7% of genome-wide SNPs but explained a large
596 fraction of heritability for both cardiomyopathies. In DCM, CC loci captured 34.6% of
597 total SNP-based heritability (Enrichment = 12.7-fold; P=6.88×10^{-18}). In HCM, the
598 same loci explained 53.4% of SNP-based heritability (Enrichment = 19.6-fold;
599 P=9.26×10^{-16}). (**Supplementary Table 9)**

600 Genetic correlation between the DCM–HCM shared meta-analysis and
601 cardiometabolic traits

602 We applied linkage disequilibrium score regression (LDSC)[6] to assess genetic
603 correlations (r_g)[27] between the shared-effects meta-analysis and a set of 65
604 quantitative traits.[8,9] Analyses were performed using the European ancestry LD
605 reference panel and the default ldsc.py --rg settings using tutorials from developers.

606 Among all tested traits, four phenotypes remained significantly correlated with the
607 DCM–HCM meta-analysis after Bonferroni correction (P < 0.05/65): diastolic and
608 systolic blood pressure (DBP: r_g=0.415, P = 5.00 × 10^{-5}, SBP: r_g = 0.375, P = 1.50 ×
609 10^{-4}), body mass index (BMI) (r_g = 0.407, P = 1.00 × 10^{-4}), body weight (r_g = 0.392,
610 P = 1.50 × 10^{-4}) and C reactive protein (CRP) (r_g = 0.268, P = 6.00 ×
611 10^{-4}).(**Supplementary Table 19)** Several other traits were nominally significant,
612 including C-reactive protein, urate, creatinine, and multiple red blood cell indices
613 (e.g., nucleated RBC percentage, reticulocyte counts), suggesting shared polygenic
614 mechanisms related to inflammation, renal function, and hematopoiesis.

615 Some traits showed negative correlations (e.g., sex hormone-binding globulin
616 (SHBG), HDL cholesterol, basophil count), particularly in the shared model, though
617 these did not reach Bonferroni significance.

618 If compared to DCM and HCM, shared meta-analysis showed stronger and more
619 consistent genetic correlations with cardiometabolic risk traits. DCM and HCM
620 GWAS, when assessed separately, showed fewer Bonferroni-significant correlations.

While many traits trended in the same direction, effect sizes were attenuated and *P*-values were less robust, underscoring the increased power of the shared meta-analysis to detect shared polygenic architecture.

Despite the strong negative correlation, certain extracardiac risk factors - including hypertension and obesity - have been linked to both DCM and HCM[1,3,28,29]. Indeed, when accounting for the genetic components of blood pressure and body habitus (**Methods**)[30], the inverse genetic correlation between DCM and HCM became nominally stronger ($r_{g,global}$=-0.63, SE=0.07, P=9.4×10^{-16}). In contrast, when conditioned on the LV endophenotypes related to contractility and chamber size, the genetic correlation was substantially weakened, but not abolished ($r_{g,global}$=-0.26, SE=0.07, P=9.4×10^{-2}). These findings indicate that the genetic pathways intrinsic to myocardial function/structure may be largely inverse between DCM and HCM, while certain extracardiac pathways may be concordant.

Genetic correlations between cardiomyopathy GWAS and other cardiovascular diseases

We performed pairwise genetic correlation analyses using LDSC to evaluate shared polygenic architecture between cardiomyopathy GWAS/MTAG results and other cardiovascular traits[31], including coronary artery disease (CAD)[32], atrial fibrillation (AF)[33], and subtypes of heart failure (HF)[34].

The shared-effects meta-analysis correlations

The shared-effects meta-analysis of DCM and HCM demonstrated positive genetic correlation with all tested traits: HF types[34], AF[33], CAD[32]. The strongest correlations that also reached Bonferroni corrected statistical significance level was observed with non-ischemic heart failure (niHF)[34] (r_g = 0.69, P = 3.05 × 10^{-19}), all-cause HF[34] (r_g = 0.677, P = 7.87 × 10^{-34}). (**Supplementary Figure 9, Supplementary Table 19,20)**

Case–case analyses (CC GWAS/MTAG) genetic correlations

In contrast, the cardiomyopathy spectrum derived from CC GWAS and CC MTAG showed little or no genetic correlation with any of the tested cardiovascular traits (AF, CAD, HF)[32,33] (**Supplementary Figure 9, Supplementary Table 20)**

This lack of correlation suggests that the genetic signals captured by CC-based analyses represent distinct axes of trait differentiation, rather than shared susceptibility loci contributing broadly to cardiovascular disease. Notably, while CC MTAG and CC GWAS were highly correlated with each other (r_g = 1.09, P < 0.01), they remained largely orthogonal to external cardiac disease traits. (**Supplementary Figure 9)**

Replication of CC-GWAS

Source for replication datasets

HCM cases were recruited from the Sarcomeric Human Cardiomyopathy Registry (SHaRe). The registry's structure and initial findings have been detailed by Ho et al. (2018)[35]. HCM was diagnosed by each SHaRe site and is defined as unexplained left ventricular (LV) hypertrophy with a maximal LV wall thickness exceeding 15 mm, or over 13 mm in family members with HCM (or an equivalent LV wall thickness z score in pediatric patients).

Whole genome sequencing samples were collected from DCM patients across multiple cohorts and studies (GO-DCM (N=565), Bratislava (N=15), RBH Biobank (N=596), SMARTER-DCM (N=29), TRED-HF2 (N=22), MATCH and MATCH2 (N=179), the Heart Hive (N=109) and MitoDCM (N=10)) which have been described in detail elsewhere. Briefly, the GO-DCM study was a whole genome sequencing initiative that aimed to recruit 2000 patients with DCM, collecting blood samples, from 2020 to 2027 across England (London, Leeds, Oxford, Leicester, Liverpool, Southampton and Manchester) and Scotland (Glasgow) (clinicaltrials.gov/study/NCT03843255). Additional whole genome sequencing samples were obtained from Bratislava, approved by the Ethics Committee of Národný ústav srdcových a cievnych chorôb, a.s. The RBH Biobank is a biobank of patients recruited from heart, lung and critical care departments of the Royal Brompton & Harefield NHS Foundation Trust. Patients were invited to give biological samples including whole blood, serum and plasma (IRAS ID 264059.; www.hra.nhs.uk/planning-and-improving-research/application-summaries/research-summaries/royal-brompton-harefield-cardiovascular-research-centre-biobank.). The SMARTER study aimed to further the genetic understanding of cardiomyopathy and patients were invited to give samples for whole genome sequencing (IRAS ID 313058; www.hra.nhs.uk/planning-and-improving-research/application-summaries/research-summaries/the-smarter-cm-study). The TRED-HF2 study aimed to recruit recovered DCM patients between 2023 and 2026 to determine the therapies required to maintain heart failure remission (clinicaltrials.gov/study/NCT06091475) . The MATCH and MATCH2 studies recruited DCM patients to establish the relationship between heart failure and type 2 diabetes (IRAS ID 228222; www.hra.nhs.uk/planning-and-improving-research/application-

summaries/research-summaries/myocardial-tissue-characteristics-and-glycaemic-status. IRAS ID 273547; www.hra.nhs.uk/planning-and-improving-research/application-summaries/research-summaries/myocardial-tissue-characteristics-and-glycaemic-status-2). The Heart Hive is an online portal for individuals with cardiomyopathy to actively engage with researchers and research studies. Participants were invited to provide saliva samples for whole genome sequencing, and indicate studies they would like to engage in (IRAS ID 246395; www.hra.nhs.uk/planning-and-improving-research/application-summaries/research-summaries/the-heart-hive). The MitoDCM study was a double blind randomised controlled trial of mitoquinol mesylate on patients with DCM to assess the effect of reducing oxidative stress on the heart (clinicaltrials.gov/study/NCT05410873)

Sequencing reads were aligned to the hg38 reference genome and variants were called using the Illumina DRAGEN pipeline (v3.10.12). Individual gVCF files were joint-called in Hail (v0.2.128). Genotypes were set to missing if genotype quality (GQ) < 20, depth (DP) < 10, or allele balance (AB) in heterozygotes < 0.2. Analyses were restricted to individuals of non-Finnish European (NFE) ancestry. Ancestry assignment was performed by projecting study samples onto the gnomAD v4.1 principal component (PC) space and clustering with NFE reference individuals. Variant- and sample-level quality control was carried out in Hail (v0.2.128). Sample QC was performed first, excluding individuals with call rate < 98.5%, mean depth < 10, or mean genotype quality < 20. Genome-wide SNVs were also used to compute the heterozygous/homozygous variant ratio and the transition/transversion (Ti/Tv) ratio. Samples deviating by more than six median absolute deviations (MAD) from the median het/hom ratio, or with Ti/Tv ratios outside the expected range of 1.8–2.2, were excluded. Additional exclusions were applied for relatedness (pi-hat > 0.125) or sex discordance. After sample filtering, variant QC was restricted to biallelic autosomal SNPs. Variants were excluded if call rate < 0.98, Hardy–Weinberg equilibrium $p < 1 \times 10^{-10}$, or minor allele frequency < 0.01. Following QC, approximately 7.9 million SNPs across 1,158 HCM and 1,525 DCM cases were retained for analysis.

Replication CC-GWAS

The case–case GWAS (HCM vs DCM) was performed in Hail using an additive genetic model with SNP dosage. HCM cases were treated as “cases” and DCM cases as “controls.” Covariates included sex and the top 20 principal components. Association statistics were reported as log-odds ratios per effect allele with corresponding standard errors and p-values.

Validation of novel loci

Novel lead SNPs were defined from the discovery case–case GWAS and MTAG analyses (see manuscript **Methods**). For each locus, the discovery beta and standard error were taken from the analysis in which the SNP was originally reported. When a lead SNP appeared in both the case–case GWAS and MTAG, the MTAG estimate was used. The corresponding SNPs were extracted from our individual-level case–case GWAS (with positions lifted over from GRCh37 to GRCh38). Alleles were harmonized so that effect estimates correspond to the same effect allele across datasets.

Concordance between discovery and validation was assessed using several complementary approaches:

1. Direction concordance
    - The proportion of SNPs with the same direction of effect was tested against the null expectation of 50% using a binomial test.
2. One-sided look-ups
    - For each lead SNP, the one-sided p-value was evaluated in the validation dataset in the discovery-predicted direction of effect, testing whether the SNP showed enrichment of association beyond chance.
3. Effect-size concordance
    - Discovery and validation effect estimates were compared directly using correlation and linear regression (validation ~ discovery).
    - Both raw betas (log-odds scale) and standardized Z-scores (β/SE) were considered to account for differences in effect size scaling.

Seventeen novel loci with 18 corresponding lead SNPs were identified in the discovery analysis (locus 99, mapped to *NFATC3*, was tagged by two distinct variants:

rs8059305 in CC-GWAS and rs12599178 in CC-MTAG), Of these, 17 lead SNPs were present in the validation cohort, corresponding to 16 novel loci. For the 17 lead SNPs, we first assessed concordance in the direction of effect between discovery and validation. Sixteen of the 17 SNPs (94.1%) shared the same direction of effect, significantly greater than the 50% expected under the null (binomial test, $p = 0.000137$).

We next performed one-sided look-ups for the 17 lead SNPs to test whether the validation GWAS showed enrichment of association in the discovery-predicted direction. Under the null hypothesis of 5% replication by chance, fewer than one SNP would be expected to reach nominal significance. In contrast, 11 of the 17 SNPs (64.7%) did so in the validation dataset, representing a highly significant enrichment (binomial test, $p = 4.6 \times 10^{-11}$).

To further evaluate concordance, we compared effect sizes between discovery and validation. Effect directions were highly correlated ($r = 0.83$, $p = 3.3e\text{-}5$), although effect estimates were generally larger in the validation dataset due to differences in scale (**Supplementary Figure 7** and **Supplementary Figure 8**). Forest plots of individual loci illustrate the consistency in direction of effect across discovery and validation.

Validation of CASQ2 variant

The enrichment of the CASQ2 lead SNP was assessed separately in individuals in the HCM and DCM cohorts that pass QC (described above) relative to non-Finnish European (NFE) population controls from gnomAD (v4.1). Enrichment was evaluated using Fisher's exact tests (one sided p-value) under three genetic models: (i) an additive model, testing enrichment of the effect allele in cases versus controls; (ii) a dominant model, testing enrichment of effect-allele carriers (heterozygous + homozygous) versus non-carriers; and (iii) a recessive model, testing enrichment of homozygous effect-allele carriers versus all other genotypes.

The discovery GWAS identified a novel locus associated with increased risk of both HCM and DCM, with the lead variant corresponding to a missense substitution in CASQ2 (p.Thr66Ala). We validated this signal by comparing allele and genotype counts in HCM and DCM cases with population reference data from non-Finnish European individuals in gnomAD (v4.1).

Across 1,158 HCM and 1,525 DCM cases, we observed enrichment of the effect allele relative to gnomAD controls under multiple inheritance models. In an additive model, the effect allele was significantly enriched in both HCM (OR = 1.14, one-sided p = 0.002) and DCM (OR = 1.07, one-sided p = 0.037). Interestingly, however, the strength of association differed under dominant and recessive models. In a dominant model, carriers of the effect allele (heterozygotes and homozygotes) were significantly enriched in HCM cases compared with gnomAD (OR = 1.20, one-sided p = 0.001), whereas DCM cases were not significantly enriched. By contrast, under a recessive model, homozygous carriers were significantly enriched in DCM relative to gnomAD (OR = 1.19, one-sided p = 0.027), but not in HCM.

Together, these results suggest a dosage-dependent relationship: heterozygous carriers of the CASQ2 variant appear more likely to develop HCM, whereas homozygous carriers are enriched among DCM cases.

Validation of case-case PGS

Polygenic scores (PGS) were calculated for 1,158 HCM cases and 1,525 DCM cases that passed sample and variant quality control and were included in the validation GWAS, together with 7,296 population control individuals from the UK Biobank. Three PGS were derived for each individual: an HCM PGS, a DCM PGS, and a case–case (CC) PGS. Each score was generated by summing the number of effect alleles carried, weighted by per-allele effect sizes estimated using SBayesRC from the respective MTAG summary statistics. The first 20 genetic principal components (PCs) were calculated for each individual using gnomAD loadings. To minimise confounding by population structure, ancestry adjustment was performed in two stages using UK Biobank controls only. First, a linear model was fitted with the PGS as the outcome and the first 20 PCs as predictors to model ancestry-related differences in the PGS mean. The squared residuals from this model were then regressed on the same PCs to capture ancestry-related differences in variance. Both models were applied to all samples to predict the expected mean and variance of the PGS given ancestry, and each raw PGS was adjusted by subtracting the predicted mean and dividing by the predicted standard deviation. The resulting ancestry-

adjusted scores were then standardised using the mean and standard deviation of the control group, yielding z-scored values with mean 0 and variance 1 in controls. To validate whether each PGS could discriminate DCM cases from HCM cases, we fitted logistic regression models with DCM versus HCM status as the outcome. All models were adjusted for the first 20 ancestry PCs; sex was included as a covariate except in univariate models. From these models, we derived performance metrics including: (i) the log-odds ratio per standard deviation increase in PGS, (ii) the area under the receiver operating characteristic curve (AUC, univariate model), (iii) the area under the precision–recall curve (AUPRC, univariate model), and (iv) the improvement in Nagelkerke's pseudo-R^2.

We assessed whether case–case GWAS data could be leveraged to position individuals along the polygenic spectrum of cardiomyopathy using genome-wide genetic data. To this end, we constructed polygenic scores (PGS) from MTAG summary statistics for DCM (PGS-DCM), HCM (PGS-HCM), and the case–case analysis (PGS-CC), and tested their performance in an independent validation cohort comprising 1,158 HCM cases, 1,525 DCM cases, and 7,296 controls from the UK Biobank. PGS-CC provided the strongest discrimination between DCM and HCM, with an odds ratio of 3.12 per standard deviation (95% CI 2.83–3.44; $p = 3.1 \times 10^{-132}$), and the highest predictive performance (AUC = 0.85, AUPRC = 0.84) (Figure 5).

Drugability

To assess the therapeutic potential of prioritized genes, we performed a comprehensive druggability annotation by integrating tractability profiles from the Open Targets Platform (queried April 2025)[36] with quantitative predictions from DrugnomeAI[37] (**Methods**). Among the 146 prioritized genes across 113 loci identified across all summary statistics (DCM, HCM, CC GWAS/MTAG) and the shared-effects meta-analysis, 12 (8.2%) were classified as “Very High tractability” due to existing approved drugs, and an additional 6 (4.1%) had “High tractability” based on late-stage clinical development. Another 49 genes (33.6%) showed moderate to low–moderate tractability, supported by structural or mechanistic features, and represent promising targets for preclinical investigation. A large proportion of genes (69, or 47.3%) had only minimal supportive evidence, and 10 genes (6.8%) lacked any tractability annotation—potentially reflecting unexplored biology rather than true undruggability. (**Supplementary Table 13)** In addition to categorical annotations, we incorporated DrugnomeAI, a machine learning framework that predicts druggability using 324 gene-level features across 15 data sources, including protein–protein interaction networks, expression profiles, and functional annotations. For each gene, we extracted Tier 1, Tclin, and combined Tier1+Tclin probability scores from DrugnomeAI. The Tclin score reflects the likelihood that a gene encodes a target of a drug with established clinical evidence—namely, compounds that have entered human clinical trials. The Tier 1 score estimates the probability that a gene encodes a target of an FDA-approved therapeutic agent. These probabilities are derived using a semi-supervised learning framework trained on the features of known drug targets, integrating 324 gene-level features across 15 biological and pharmacological data domains (e.g., protein interaction networks, gene expression, structural data, and functional annotations). Higher scores in either category suggest greater alignment with attributes of clinically validated targets and thus higher potential for successful pharmacologic intervention. Quantitative scores from DrugnomeAI highly correlated with therapeutic profiles from OpenTargets, with median scores ranging from 0.975 in the approved drug category to 0.027 among genes with no tractability annotation. This concordance across different scoring systems supports the robustness of our prioritization framework and identifies a set of gene targets with strong translational potential.

To identify the most translationally promising targets, we first focused on genes with either existing pharmacological agents or high druggability likelihood (Tclin+Tier1 probability > 0.4, DrugnomeAI). In total, 35 prioritized genes met these criteria, of which 18 have been previously targeted in drug development efforts. Notably, drugs for 12 of these genes have reached the market.(**Extended Data Figure** 9**a,b**).

While many of the known drugs targeting prioritized genes were developed for cancer (e.g., regorafenib for *RAF1*[38], afatinib for *ERBB4*[39]), a subset also demonstrates direct relevance to cardiovascular and neuromuscular diseases.(**Supplementary table 11**, **Extended Data Figure 7a,b**) For instance, *PDE3A* is the target of milrinone, a phosphodiesterase inhibitor approved for acute heart failure, where it enhances myocardial contractility and reduces afterload.[40] *ADM* (Adrenomedullin) is another cardiovascular-relevant target; a non-neutralizing antibody (adrecizumab) has advanced to phase II trials in heart failure and sepsis,[41] and circulating levels of its prohormone (MR-proADM) are used as a biomarker of hemodynamic stress[42,43]. Additionally, some drugs, that has initially been used for non-cardiac reasons, might also hold promise to be applied in cardiovascular treatment. For example, *PGR* (Progesterone receptor) is primarily known for its role in reproductive physiology, however genetic variation near the *PGR* locus—particularly in the intergenic region between *PGR* and *TRPC6*—has been associated with hypertensive disorders of pregnancy (HDPs), including preeclampsia. *TRPC6* encodes a calcium-permeable channel involved in renal function and blood pressure regulation. Nearby, *ARHGAP42*, which modulates vascular tone, has shown reduced expression in preeclamptic placentas. Together, these findings suggest that the broader *PGR* region may contribute to vascular regulation and highlight it as a potential target for therapeutic exploration in cardiovascular and hypertensive conditions. [44] *GNRHR*, which emerged as a prioritized gene in our CC-GWAS analysis, encodes the gonadotropin-releasing hormone receptor and is the molecular target of abarelix, a GnRH antagonist developed for prostate cancer. While there is currently no direct evidence linking *GNRHR* itself to cardiovascular disease, related hormonal pathways may be relevant. Notably, *GNRH1*, which encodes the ligand for *GNRHR*, has been shown in Mendelian randomization studies to be positively associated with increased risk of ischemic heart disease (IHD). This suggests that pharmacologic modulation of the GnRH axis—while originally intended for oncologic

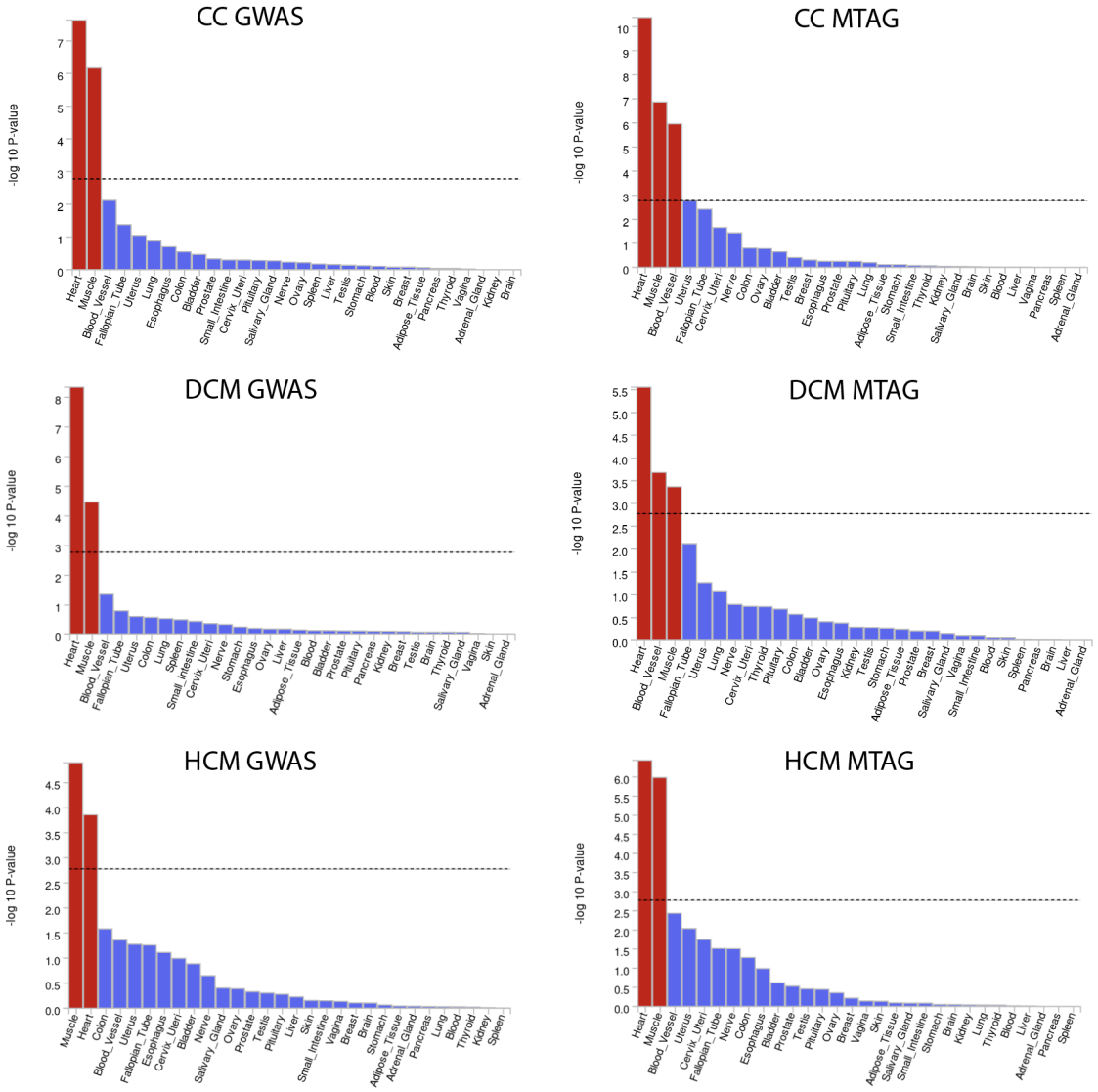
indications—could have broader systemic effects, including potential relevance in cardiovascular contexts. [45]

Notably, *RPL22* exhibits low tractability and druggability based on both structural and clinical evidence. However, it has been listed among the targets of ataluren, a drug approved for Duchenne muscular dystrophy. Rather than acting selectively on *RPL22*, ataluren is believed to exert a broad mechanism of action by modulating the translational machinery. It affects multiple ribosomal components, including *RPL22*, with a total of 78 annotated protein targets. This suggests that *RPL22* may not be the primary pharmacologic target, and its inclusion likely reflects the pleiotropic effects of ataluren on ribosomal function. (**Supplementary Table 13**)

Among the prioritized genes without clinically developed drugs, several nonetheless demonstrated high tractability based on the availability of chemical probes (**Extended Data Figure 9d-f**). A chemical probe is a highly potent, selective, and cell-permeable compound that binds to a target protein and modulates its function in a predictable and reversible way. It serves as a tool to study the biology of that target. If a high-quality chemical probe exists for a gene product, it's strong evidence that the protein is ligandable (i.e., it can bind a small molecule), which supports the tractability of that target for drug development.

Notably, targets such as *PLK2*, *MAP3K7*, and *KAT2B* were supported by multiple probes, including high-quality entries as defined by established scoring frameworks (**Methods**). This provide valuable opportunities for early-phase preclinical investigation, offering routes for target validation, mechanistic dissection, and pharmacological modulation in the absence of approved compounds. For instance, *MAP3K7 (TAK1),* a serine/threonine kinase central to stress and inflammatory signaling, represents a compelling druggable candidate in cardiomyopathy. It is supported by two validated chemical probes and has been shown to regulate key signaling pathways downstream of TGF-β, IL-1, and TNF-α. Recent evidence highlights its cardioprotective role in restraining inflammasome activation and pyroptosis under pressure overload, suggesting potential therapeutic relevance in hypertrophy and heart failure. Pathogenic variants in MAP3K7 have been implicated in congenital syndromes with structural cardiac manifestations. (**Supplementary Table 13**)[46]

Supplementary Figures



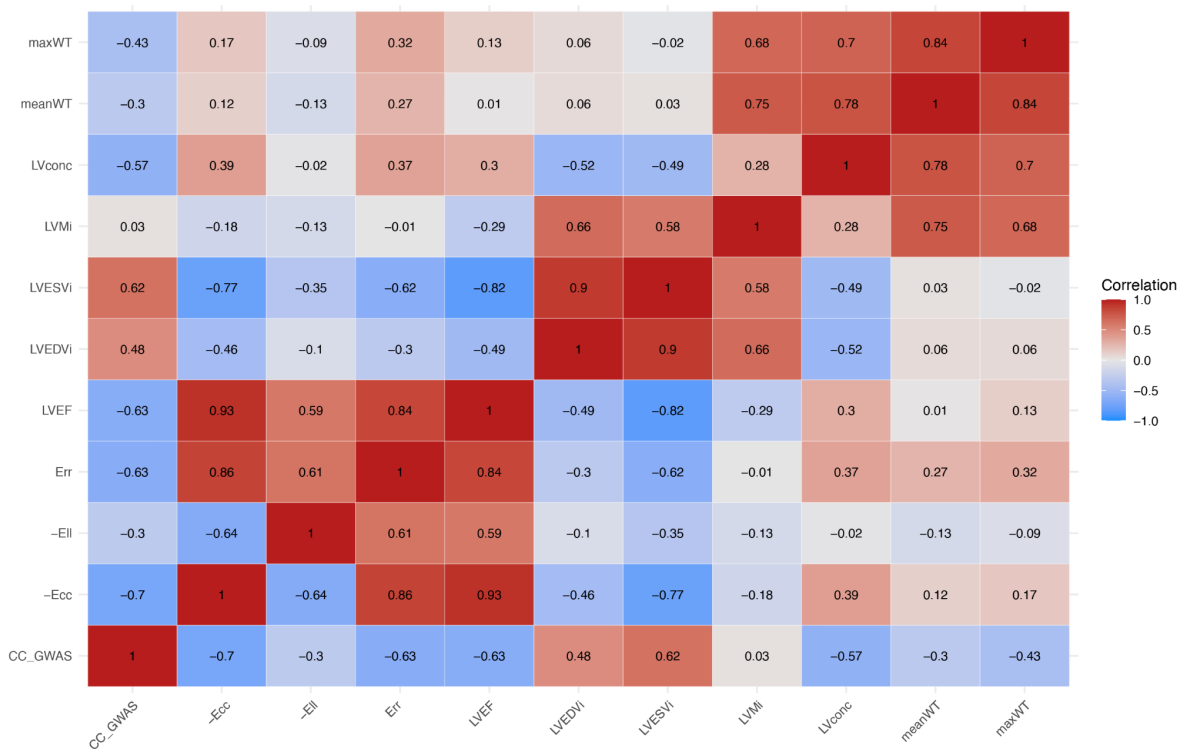
Supplementary Figure 1: Tissue enrichment of heritability for cardiomyopathy traits from bulk RNA sequencing data in GTEx v8.

Bar plots represent the $-\log_{10}(P)$ values from tissue-specific enrichment analyses, with tissues from GTEx v8 shown on the x-axis. Each panel corresponds to a different GWAS or MTAG result: CC GWAS and CC MTAG (top row), DCM GWAS and DCM MTAG (middle row), and HCM GWAS and HCM MTAG (bottom row).

Tissues surpassing the Bonferroni-corrected significance threshold are shown in red; the horizontal dashed line marks the significance cutoff. Tissues are ordered by their significance within each panel.

Enrichment *P*-values were obtained using a stratified LD score regression framework; they are unadjusted and can be interpreted as one-sided.

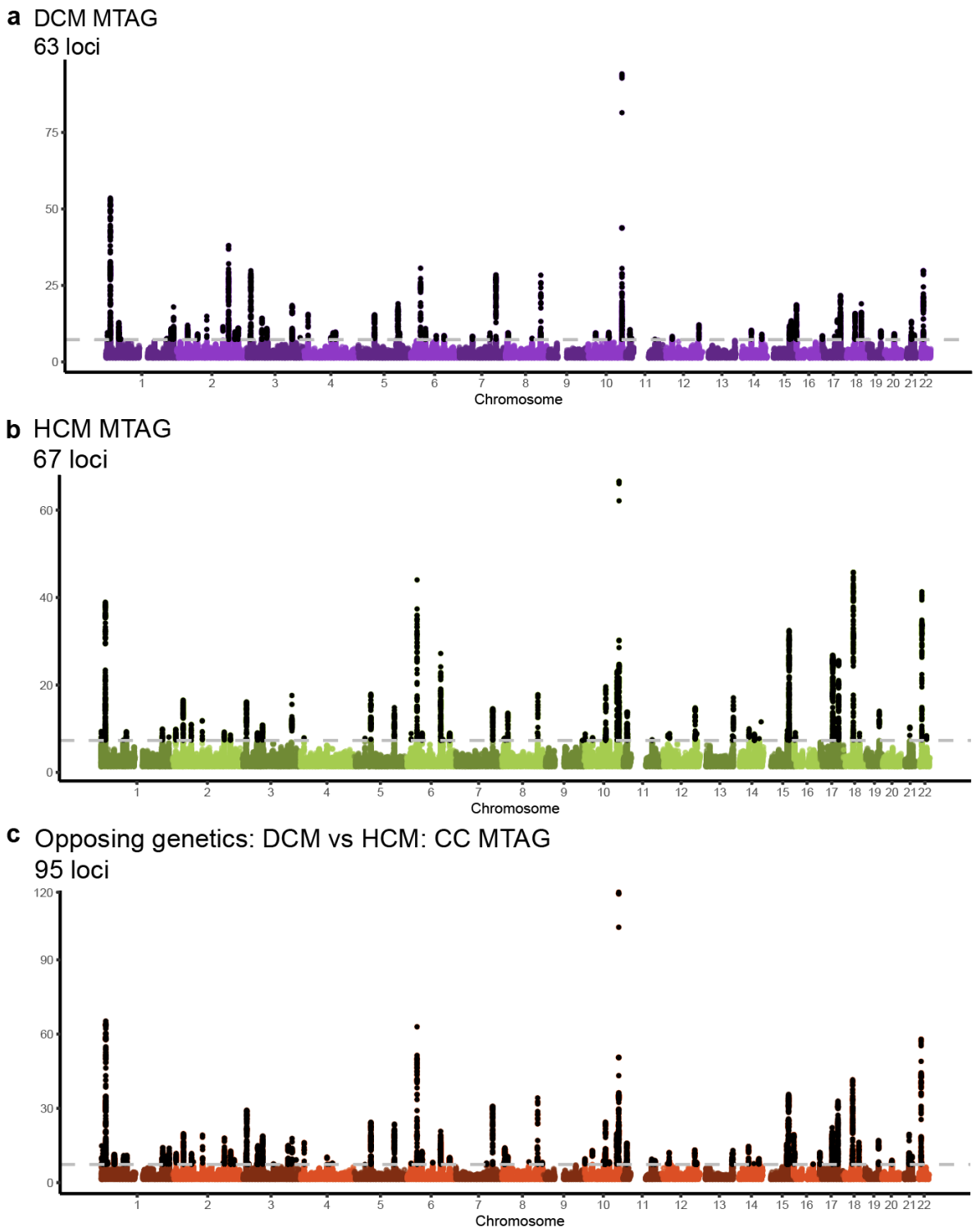
Abbreviations: GWAS, genome-wide association study; MTAG, multi-trait analysis of GWAS; DCM, dilated cardiomyopathy; HCM, hypertrophic cardiomyopathy; CC, case-case analysis



Supplementary Figure 2: Matrix of genetic correlations between cardiomyopathy spectrum (from CC GWAS) and left ventricular traits from cardiac MRI

A heatmap of bivariate genetic correlations estimated from GWAS data, showing CC GWAS and relevant cardiac MRI traits. The color represents the level of genetic correlation, with red and blue representing positive and negative correlation, respectively.

Note: CC-GWAS, case-case genome-wide association study; Ecc, global circumferential strain; Ell, global longitudinal strain; Err, global radial strain; LVEF, left ventricular ejection fraction; LVEDVi, left ventricular end-diastolic volume indexed to body-mass index; LVESVi, left ventricular end-systolic volume indexed to body-mass index; LVMi, left ventricular mass indexed to body-mass index; LVconc, left ventricular concentricity; maxWT, maximum wall thickness; meanWT, mean wall thickness. Since Ecc and Ell are always negative values, -Ecc and -Ell are plotted to facilitate interpretation of effect direction.



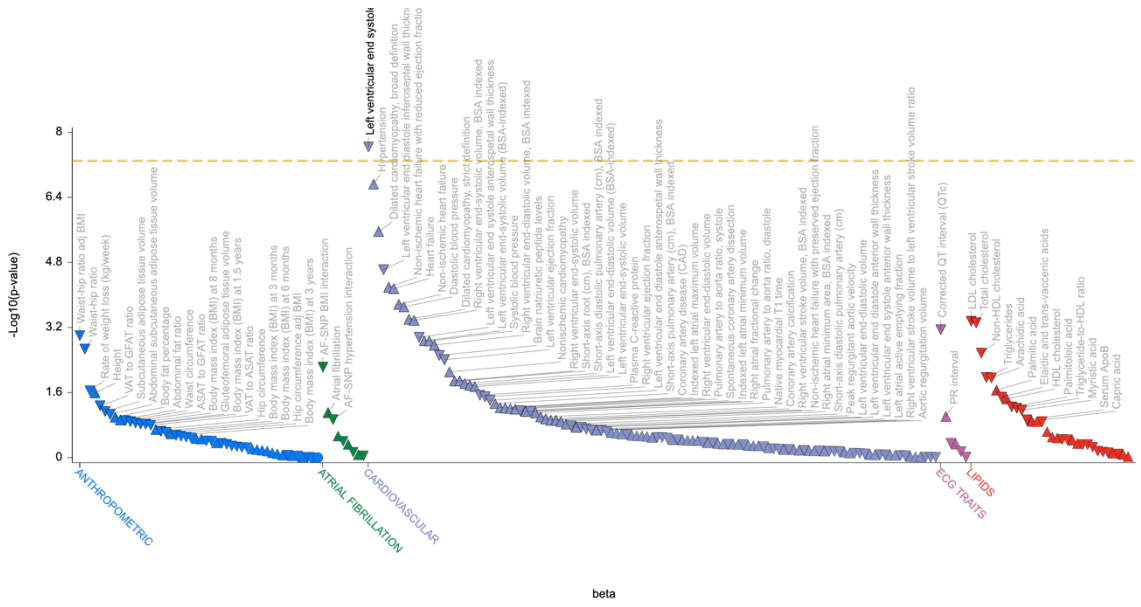
Supplementary Figure 3: Manhattan plots for DCM MTAG, HCM MTAG and CC MTAG

All panels are Manhattan plots where each dot represents a genetic variant, with genomic positions on the x-axis and -log10 of the association *P*-value on the y-axis. Panels **a** and **b** show results for published DCM MTAG (63 loci) and HCM MTAG (67

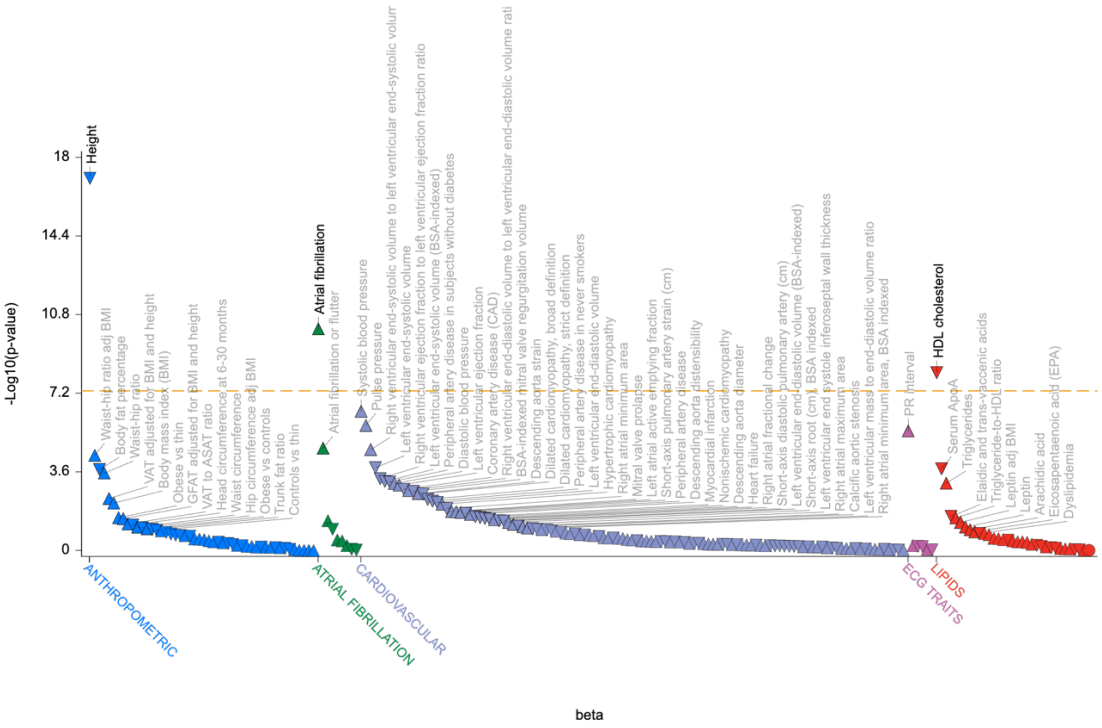
loci), respectively. Panel **c** shows results for a case-case MTAG (included CC GWAS and MRI traits MTAG), where DCM and HCM are statistically modeled as opposites on a singular genetic spectrum, yielding 95 significant loci.

Supplementary Figures 4a-s. PheWAS for prioritized gene from 17 novel CC GWAS/CCMTAG (below)

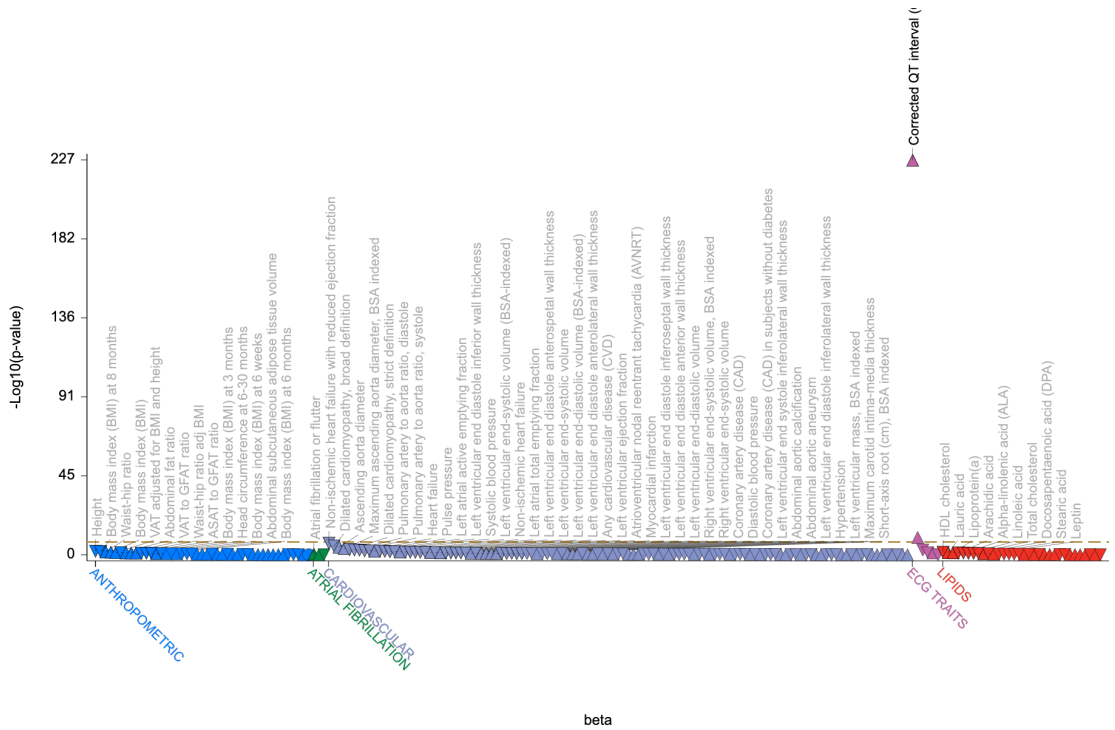
Panels below show broader results from PheWAS for lead variants of 17 novel CC GWAS/CCMTAG loci using data from the Cardiovascular Disease Knowledge Portal, specifically showing results from cardiovascular-related traits. The x-axis represents different traits grouped by trait clusters, while the y-axis represents the -log10 P-value for the association between the selected variant and the respective traits. The Variant Page summarizes variant impact predictions and genetic associations. The variant ID is in the format Chromosome:Position (hg19 genome build):Reference allele:Alternate allele.



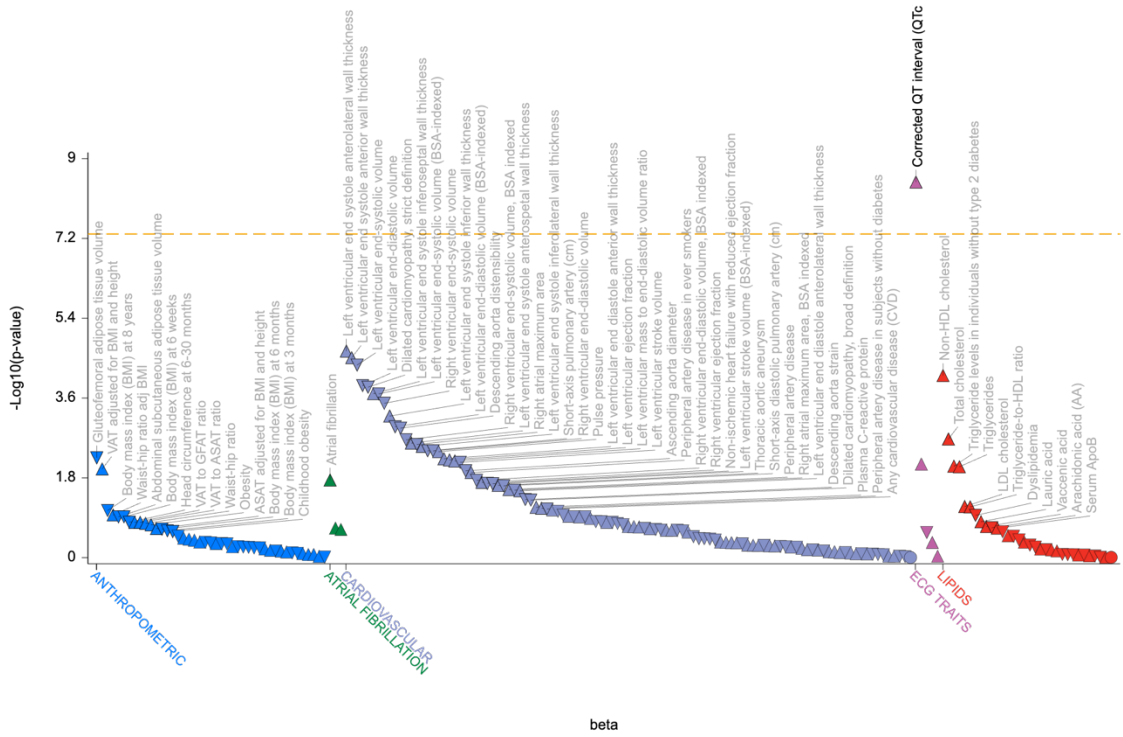
Supplementary Figure 4a: 1:3199217:C:T / rs16823802 PheWAS associations (locus 2, *PRDM16*)



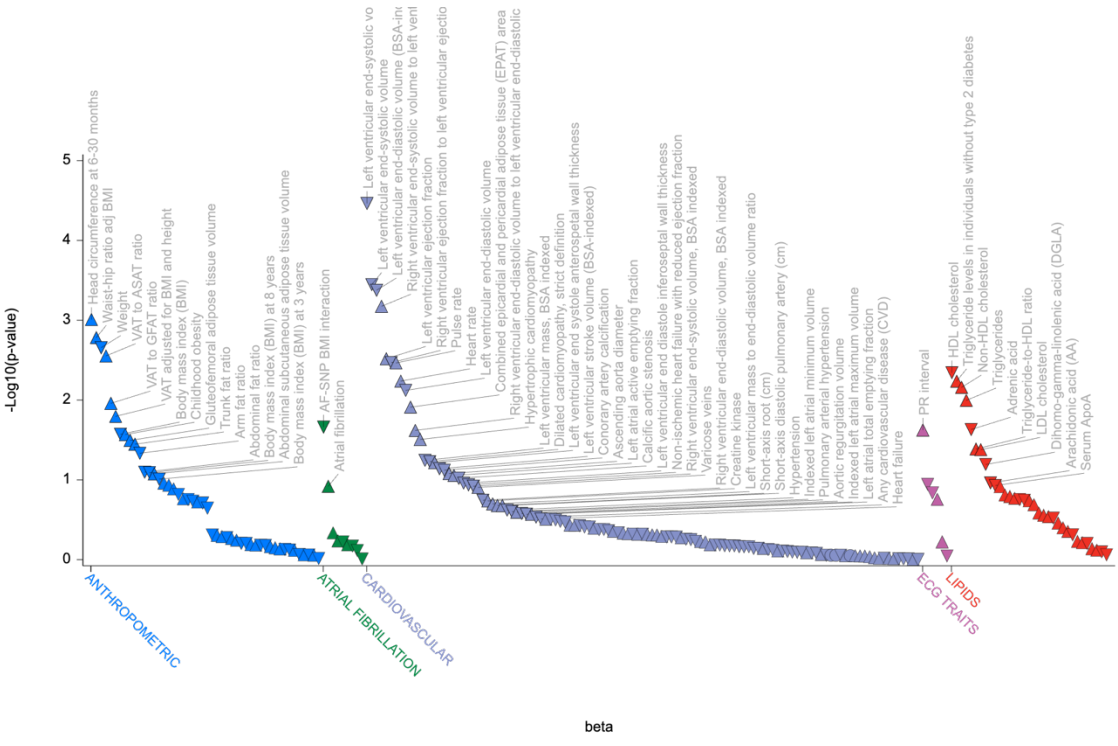
Supplementary Figure 4b: 1:22248881:G:A / rs10799719 PheWAS associations (locus 5, *HSPG2)*



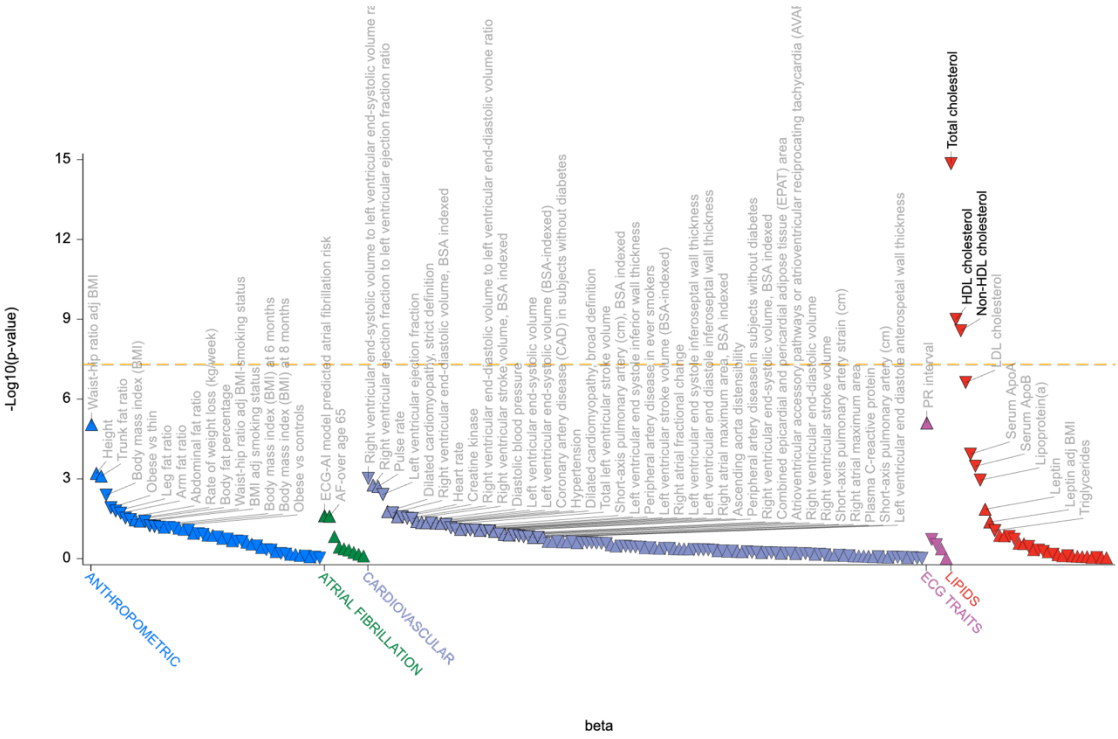
Supplementary Figure 4c: 1:162023184:T:C / rs60129000 PheWAS associations (locus 11, *NOS1AP)*



Supplementary Figure 4d: 2:174888351:G:T / rs35717017 PheWAS associations (locus 22, *SP3*)



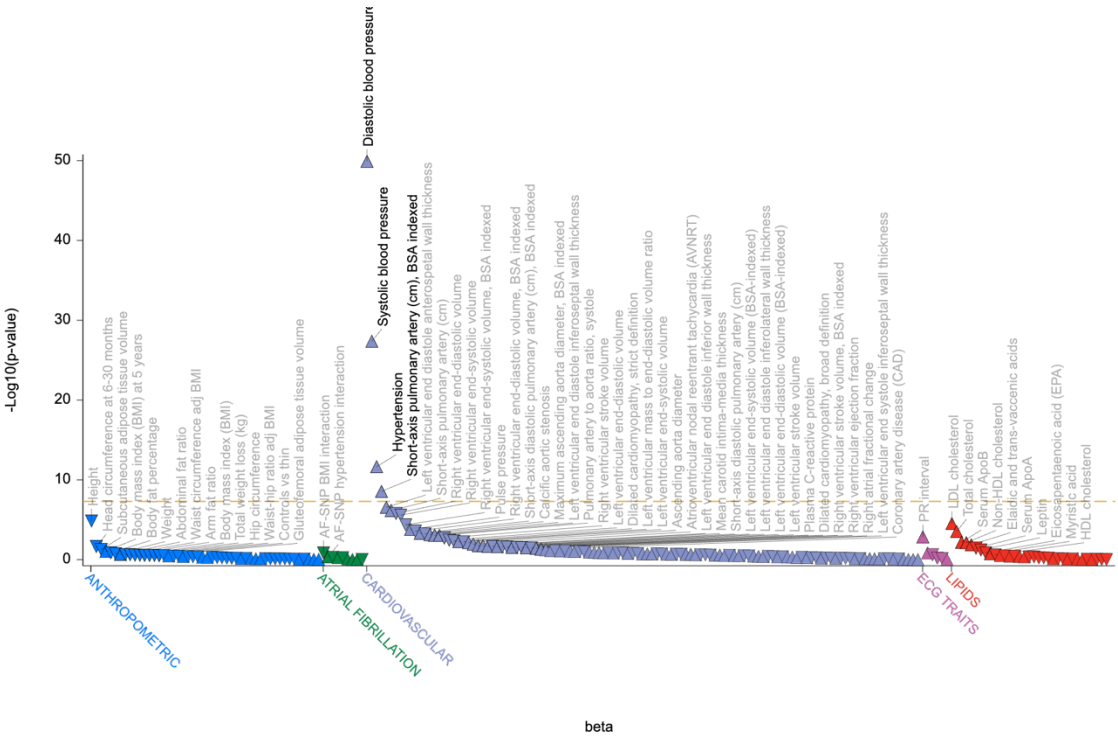
Supplementary Figure 4e: 3:20003611:T:G / rs4241539 PheWAS associations (locus 29, *KAT2B*)



1028

1029 Supplementary Figure 4f: 3:50306249:T:C / rs3806708 PheWAS associations (locus

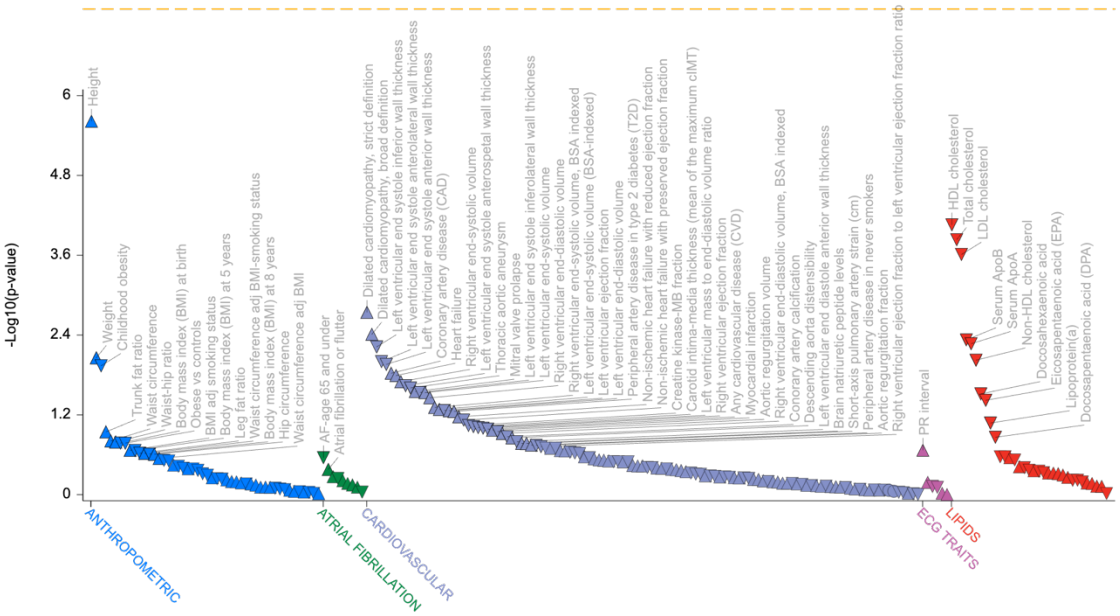
1030 30, *CACNA2D2*)



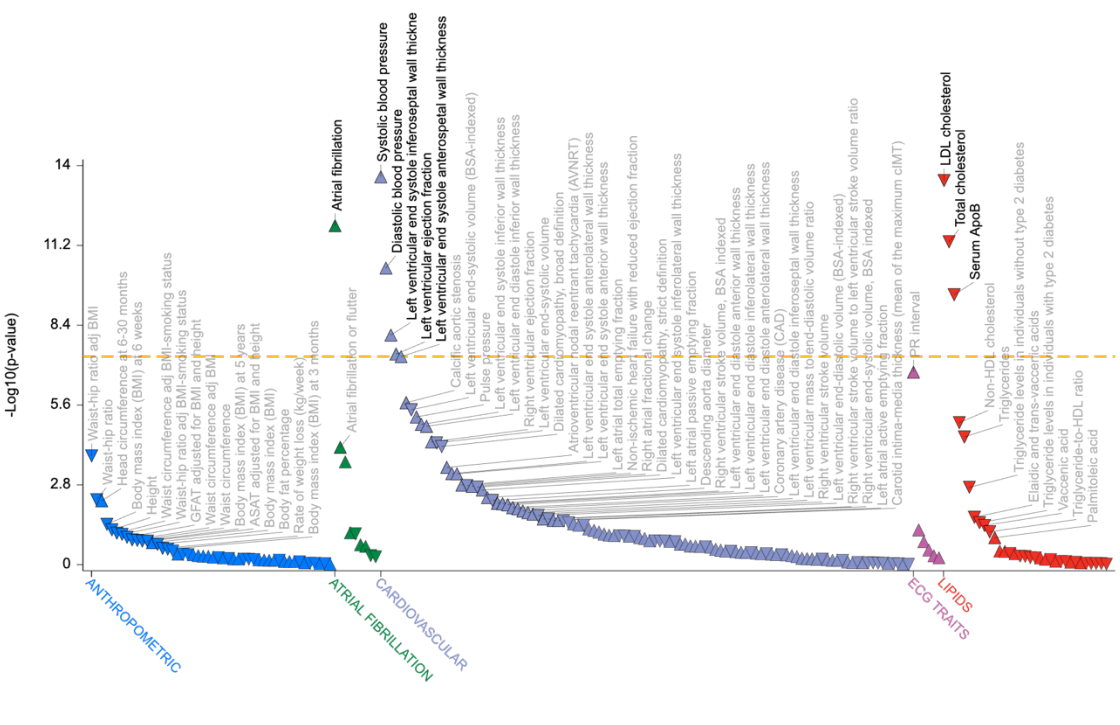
1031

1032 Supplementary Figure 4g: 3:169177924:C:T / rs9850919 PheWAS associations

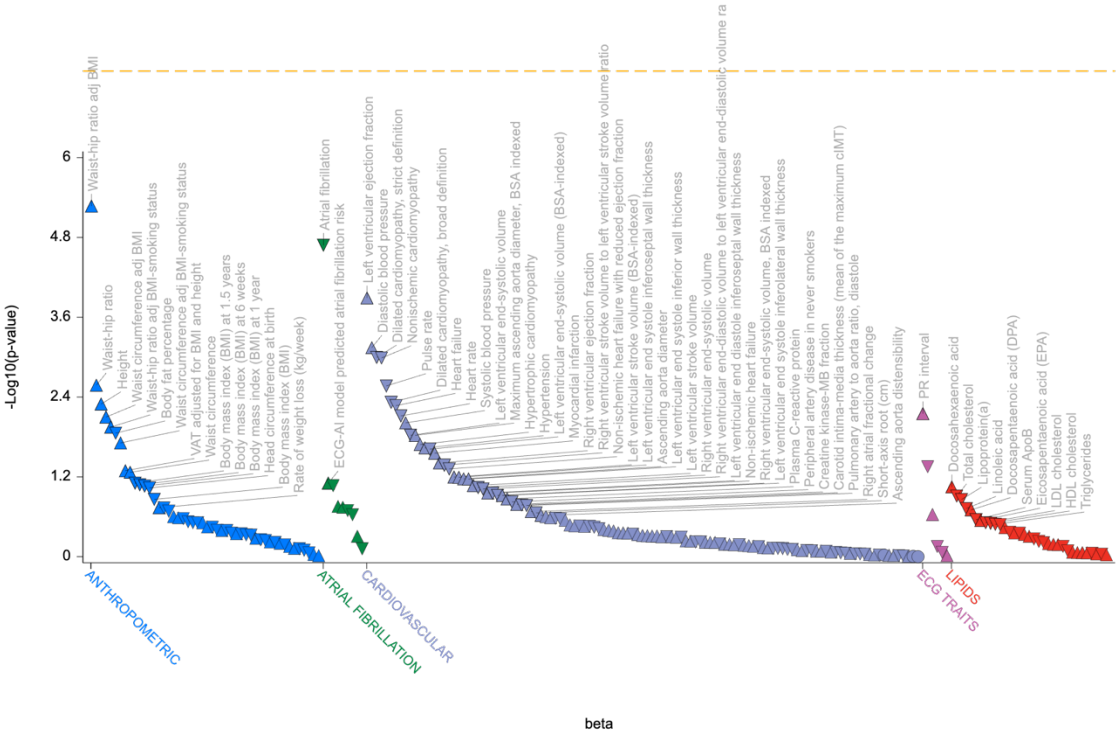
1033 (locus 35, *SEC62/MECOM*)



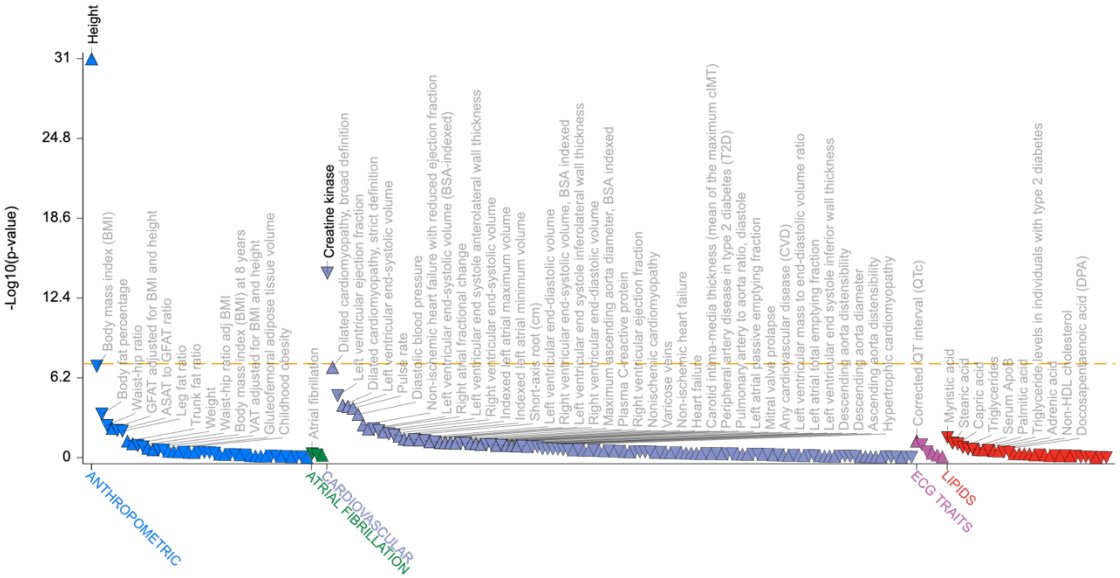
Supplementary Figure 4h: 4:68556786:A:G / rs9998837 PheWAS associations (locus 39, *GNRHR/UBA6*)



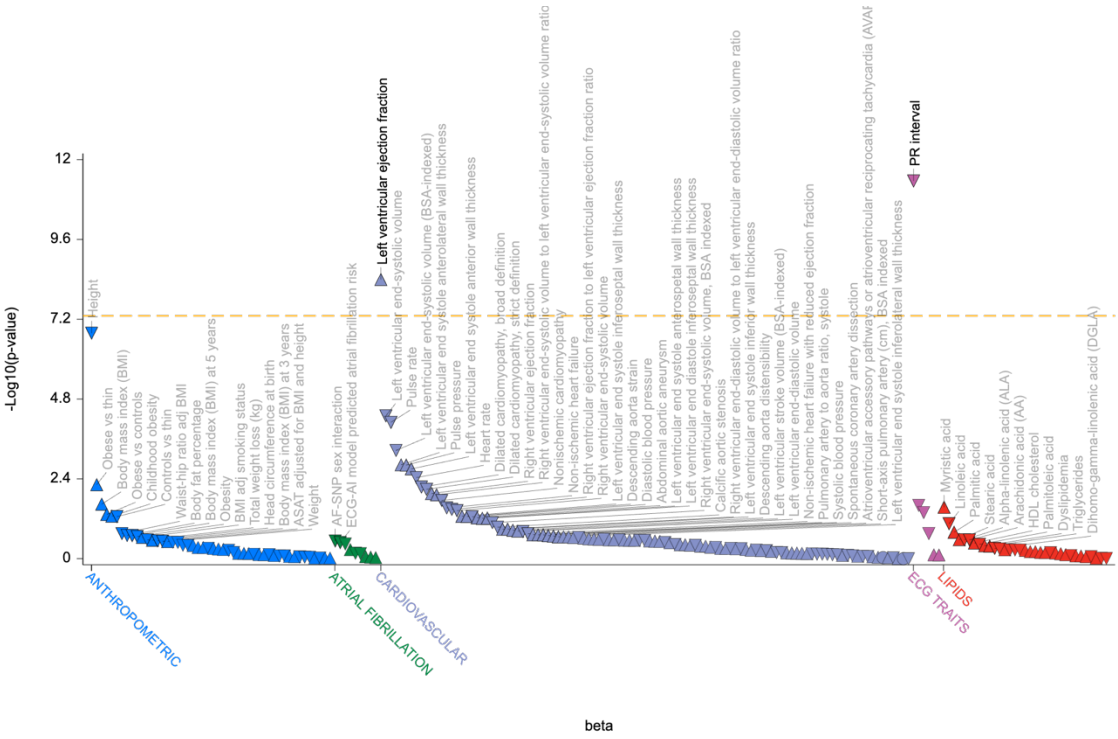
Supplementary Figure 4i: 8:141740868:A:C / rs6994744 PheWAS associations (locus 65, *PTP4A3/PTK2*)



Supplementary Figure 4j: 9:111865232:C:A / rs7028081 PheWAS associations (locus 66, *TMEM245*)



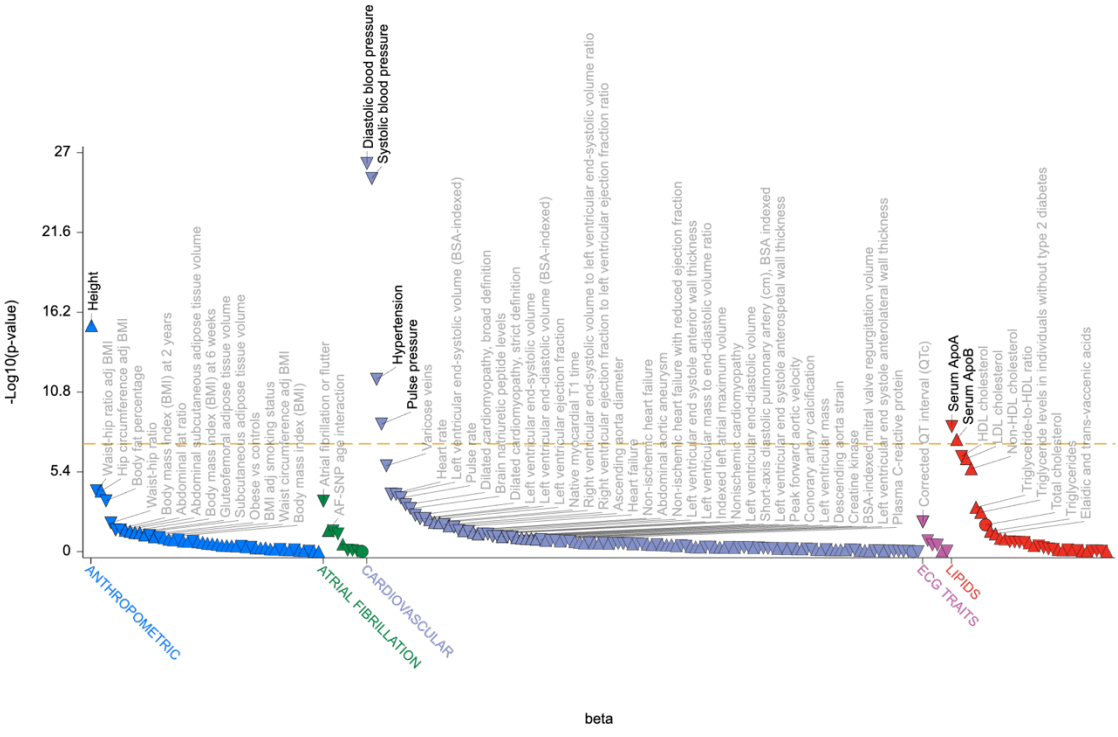
Supplementary Figure 4k: 10:69929058:T:C / rs7911060 PheWAS associations (locus 70, *MYPN*)



1052

1053 Supplementary Figure 4l: 10:88450058:C:A / rs12251655 PheWAS associations

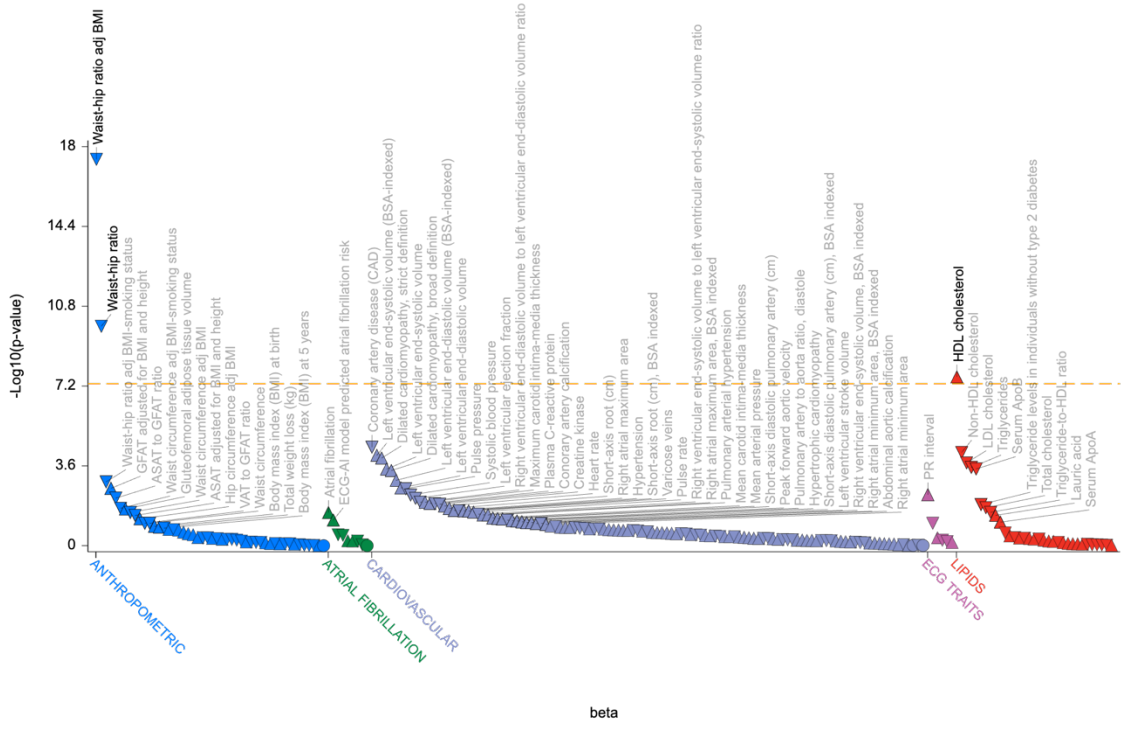
1054 (locus 72, *LDB3*)



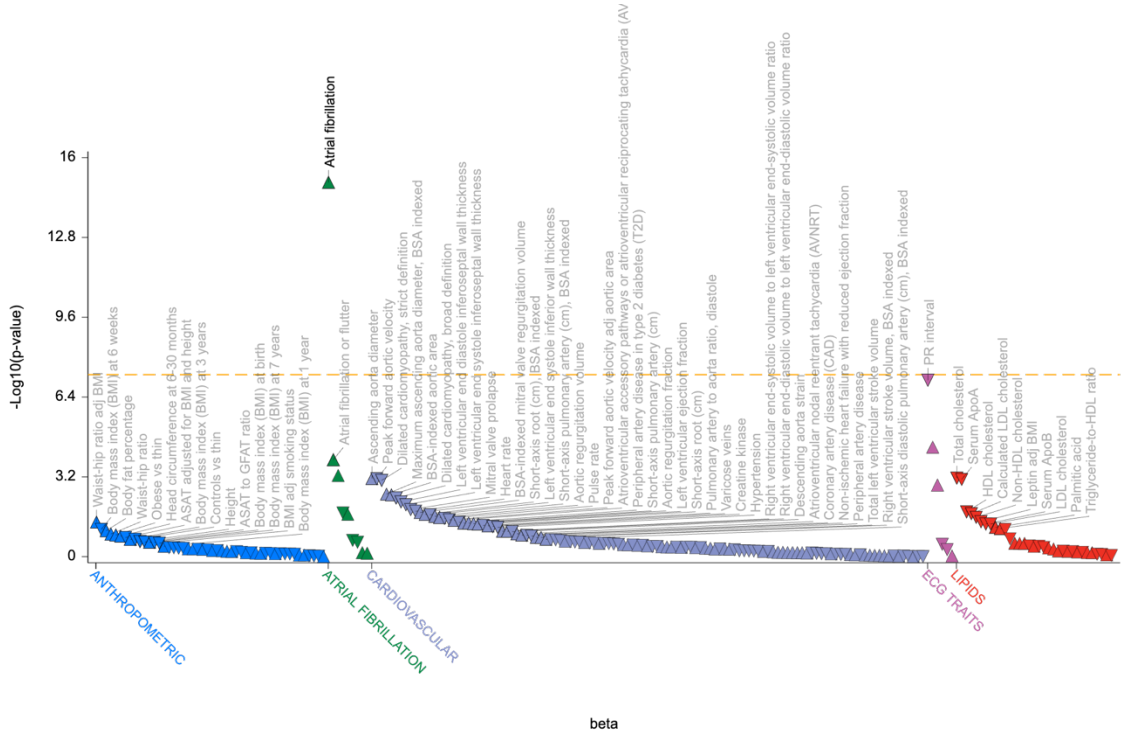
1055

1056 Supplementary Figure 4m: 11:10238974:A:G / rs1822293 PheWAS associations

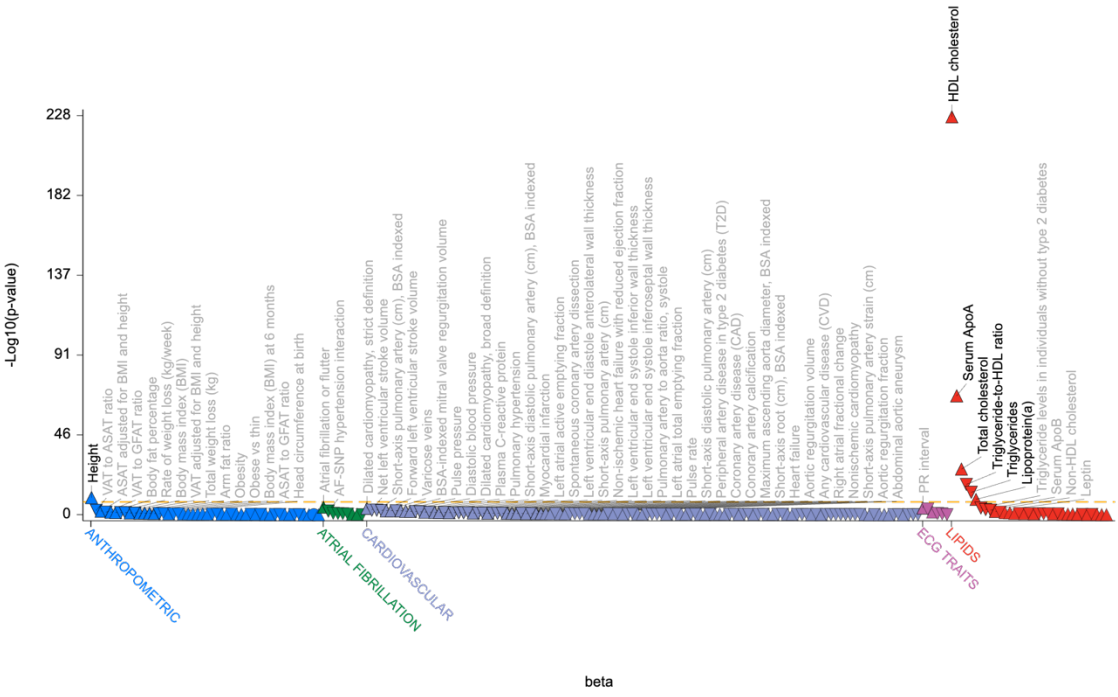
1057 (locus 76, *ADM*)



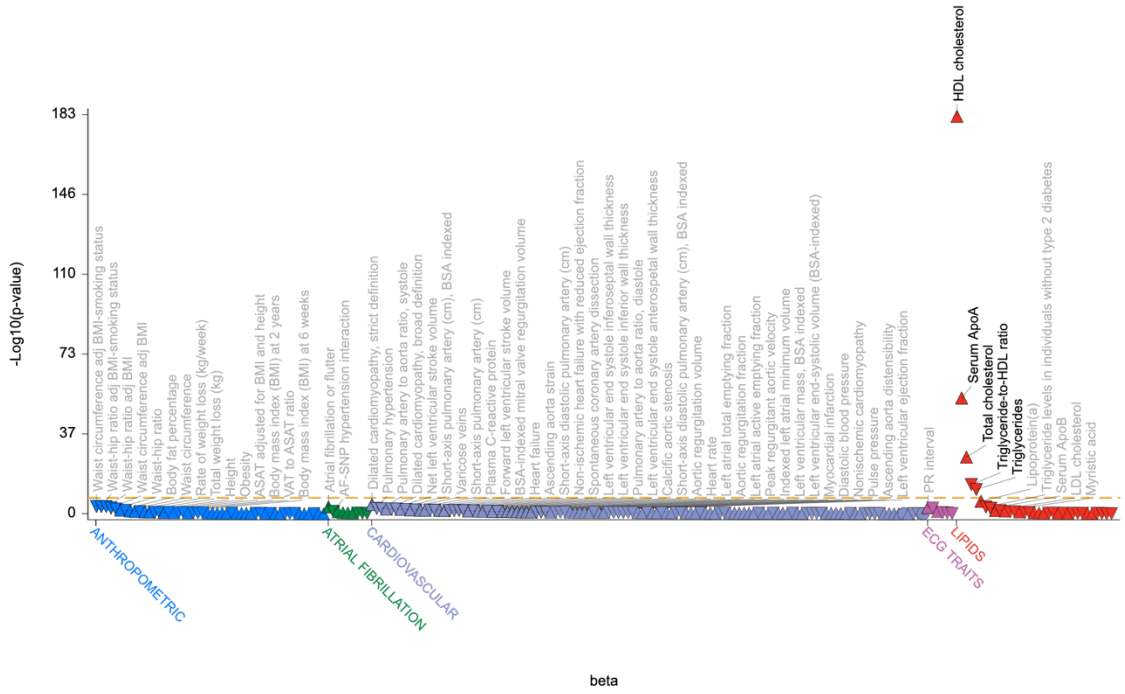
Supplementary Figure 4n: 11:111787962:G:T / rs10891299 PheWAS associations (locus 80, *CRYAB*)



Supplementary Figure 4o: 12:32980161:G:T / rs2045172 PheWAS associations (locus 83, *PKP2*)



Supplementary Figure 4p: 16:68036666:A:C / rs8059305 PheWAS associations (locus 99, *NFATC3*)

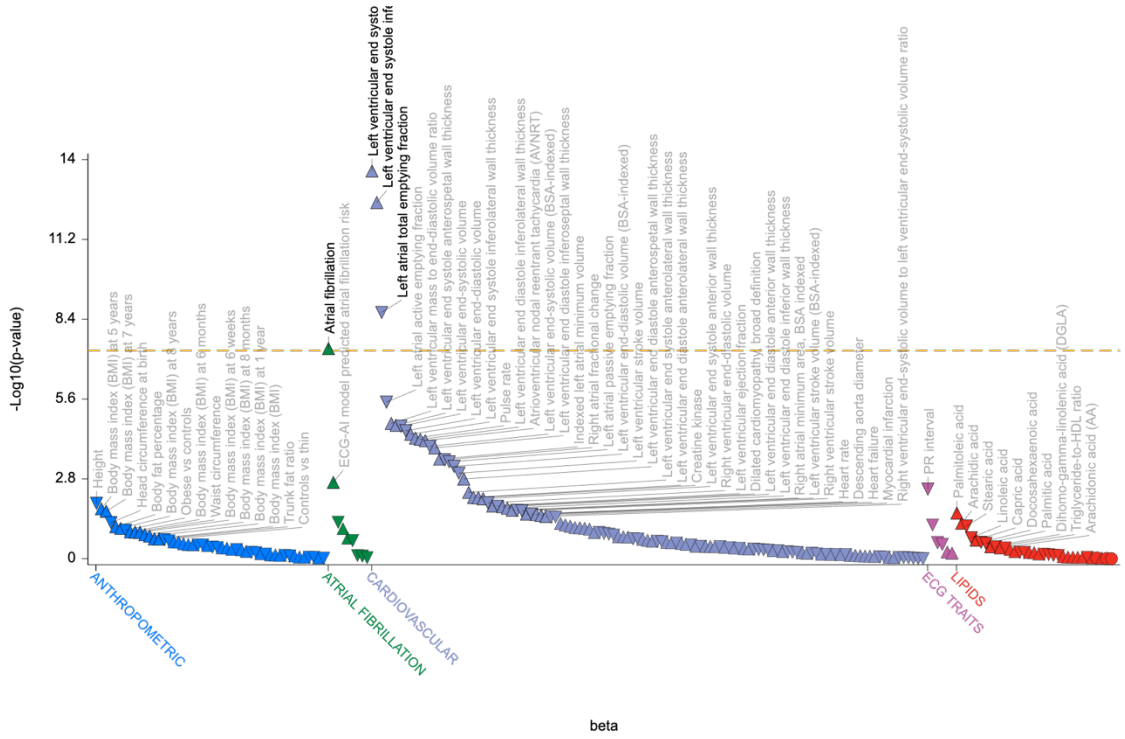


1069

1070 Supplementary Figure 4r: 16:68128104:A:G / rs12599178 PheWAS associations

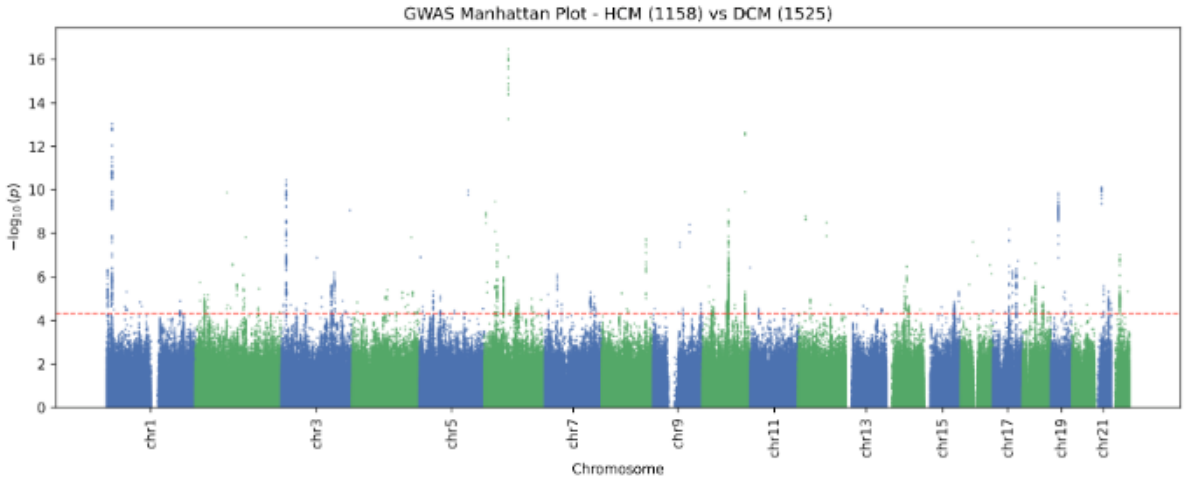
1071 (locus 99, *NFATC3*)

1072



1073

Supplementary Figure 4s: 22:26162902:A:G / rs4820654 PheWAS associations (locus 112, *MYO18B*)



Supplementary Figure 5: Manhattan plot of the replication case-case GWAS

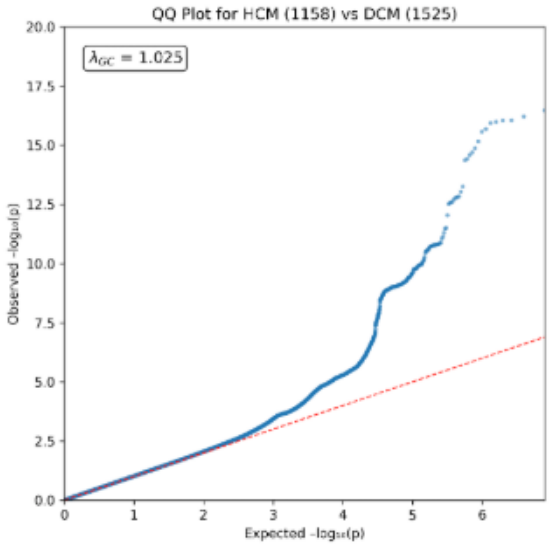
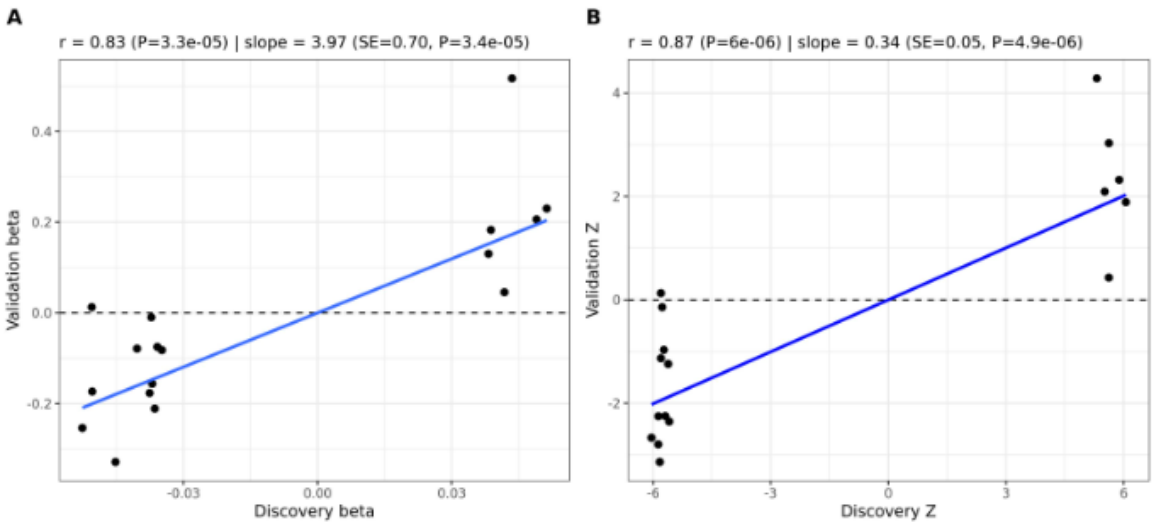


Figure 2: QQ plot of the case–case GWAS (HCM vs DCM).

Supplementary figure 6: QQ plot of the case–case GWAS (HCM vs DCM).

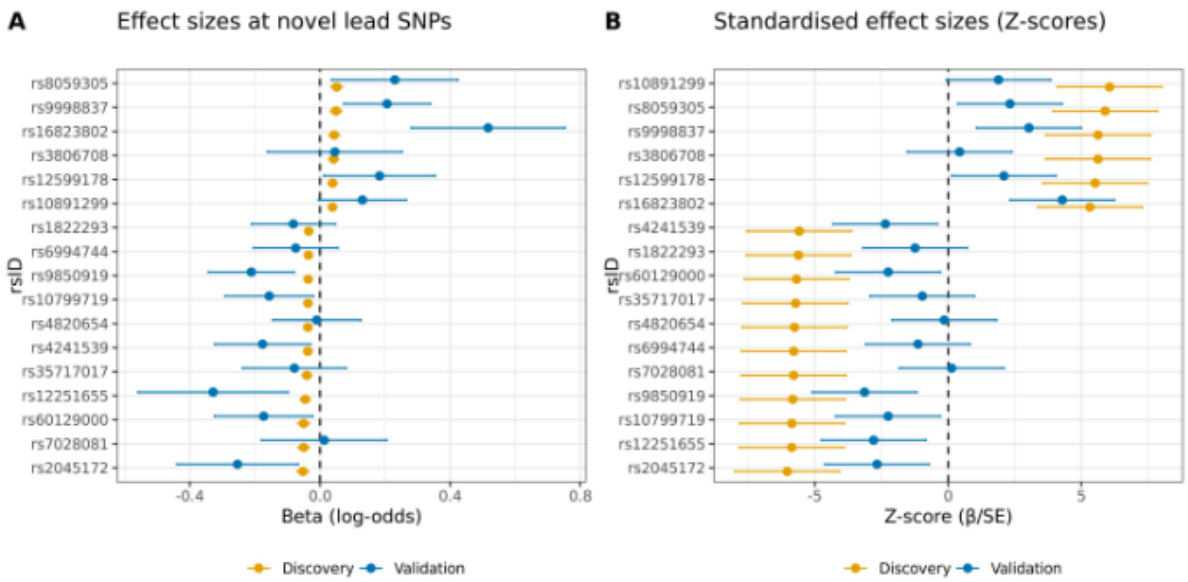


1086

1087 Supplementary figure 7: Scatterplots showing concordance of novel SNP
1088 effect estimates between discovery and validation.

1089 A) correlation of raw betas (log-odds ratios per allele). B) correlation of standardized
1090 effect sizes (Z-scores, β/SE). The solid line shows the fitted regression through the
1091 origin.

1092



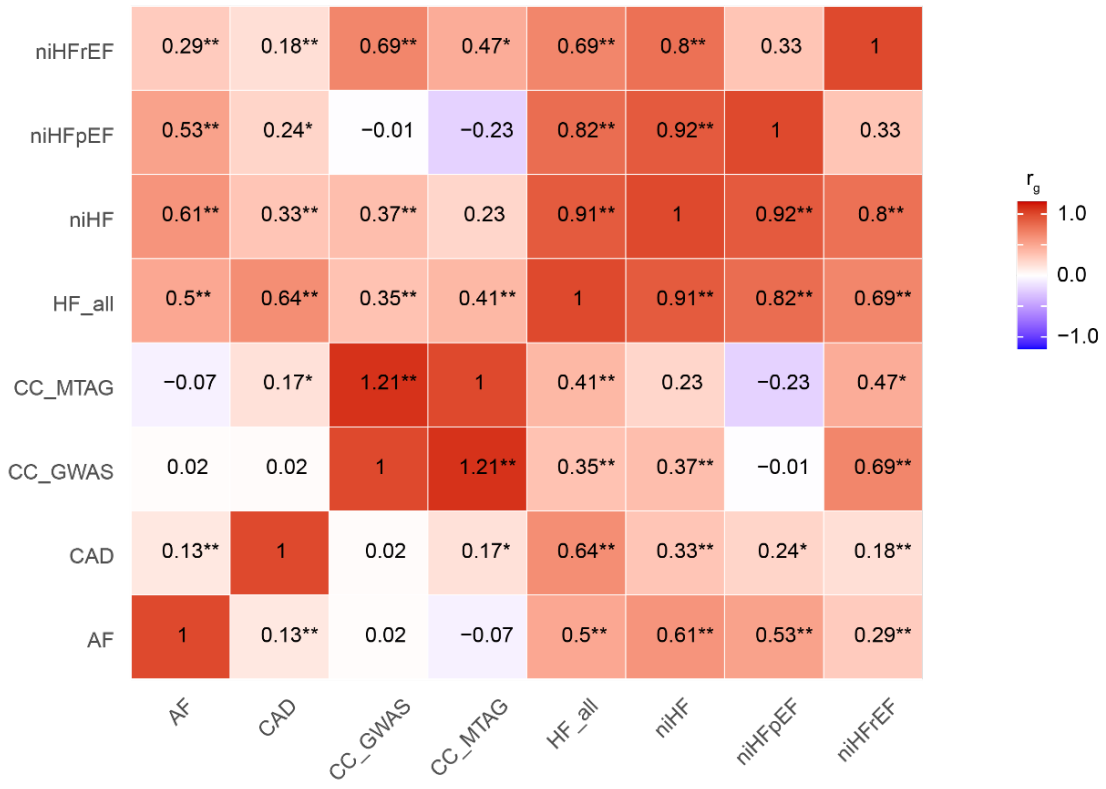
1093

1094 Supplementary Figure 8: Forest plots showing effect estimates for novel
1095 lead SNPs in discovery and validation.

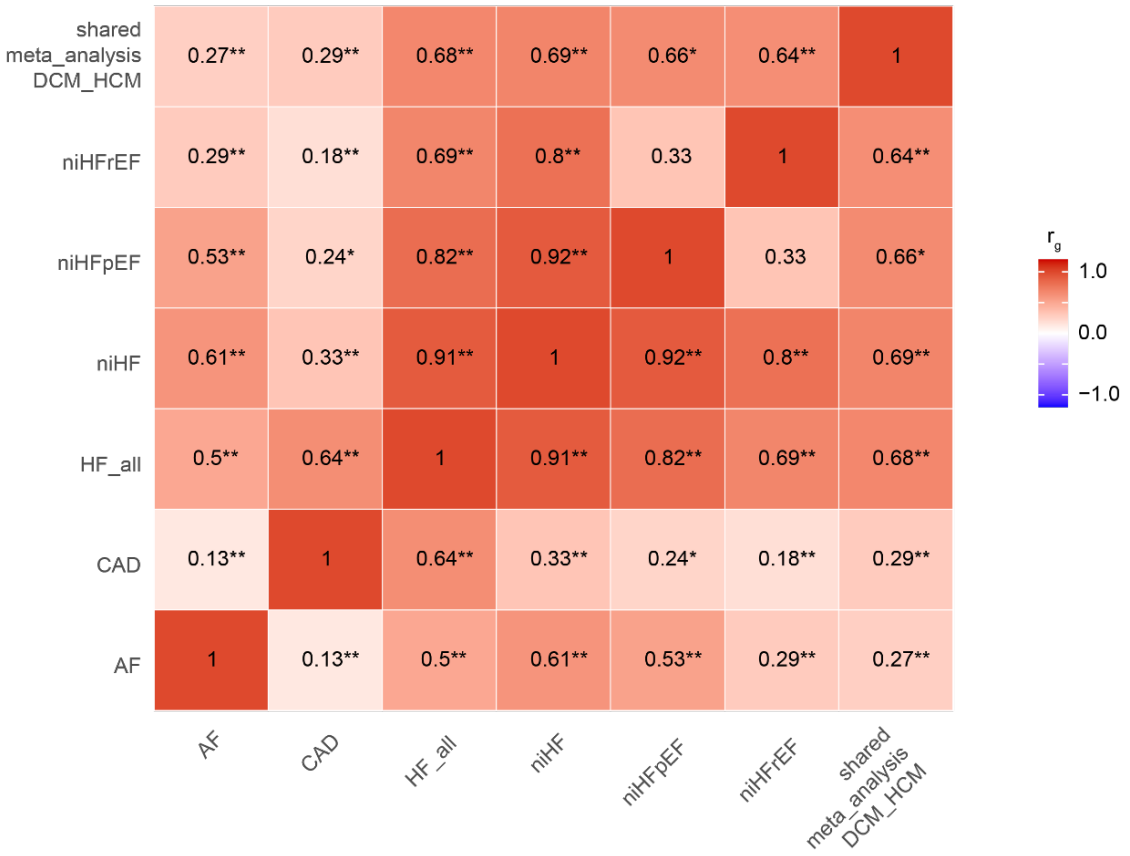
1096 A) raw betas (log-odds ratios per allele) with 95% confidence intervals. B)
1097 standardized effect sizes (Z-scores, β/SE) with 95% confidence intervals. SNPs are

ordered by discovery effect size. Discovery and validation estimates are shown side by side for each SNP.

a

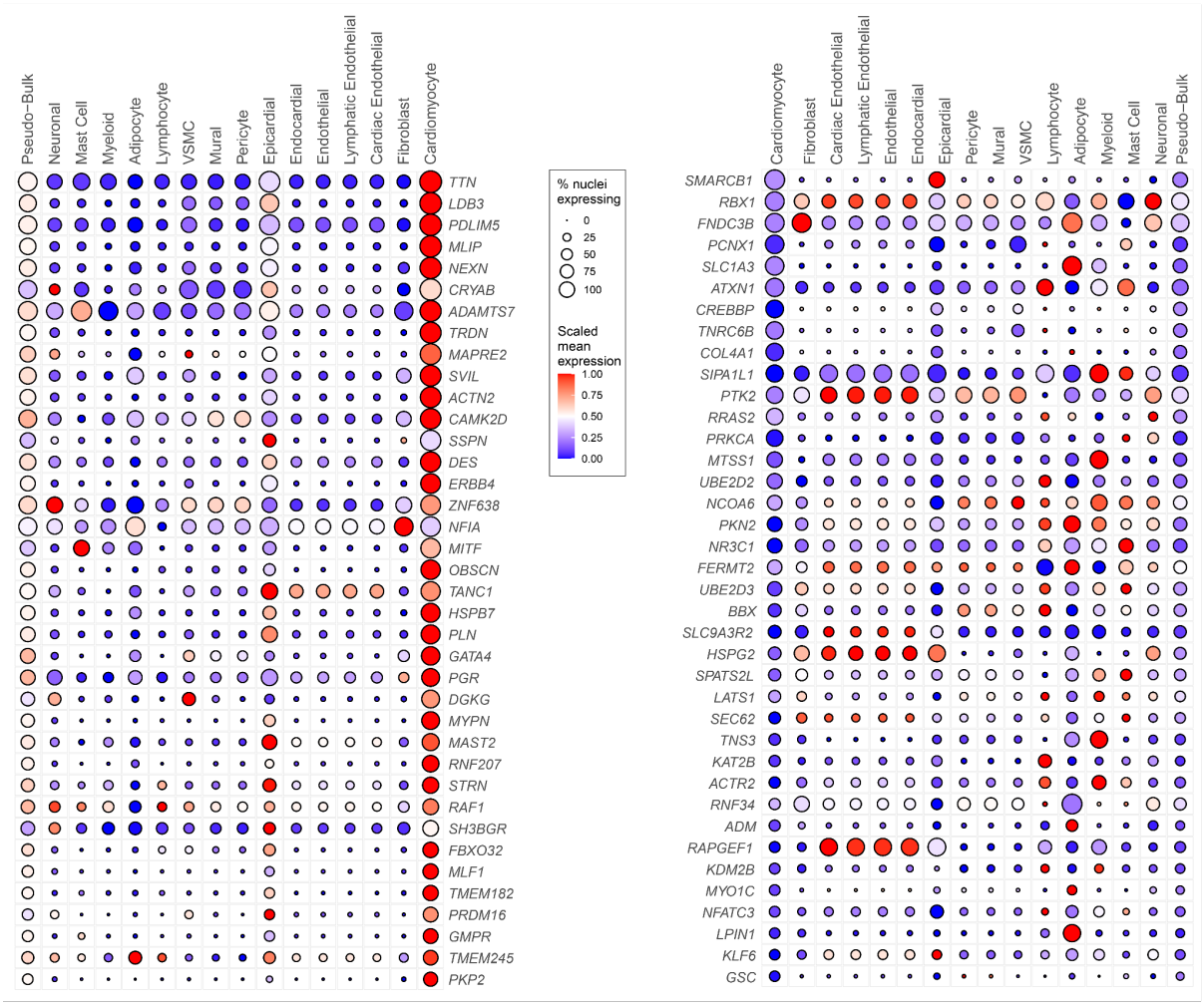

niHFrEF 0.29** 0.18** 0.69** 0.47* 0.69** 0.8** 0.33 1
niHFpEF 0.53** 0.24* −0.01 −0.23 0.82** 0.92** 1 0.33
niHF 0.61** 0.33** 0.37** 0.23 0.91** 1 0.92** 0.8**
HF_all 0.5** 0.64** 0.35** 0.41** 1 0.91** 0.82** 0.69**
CC_MTAG −0.07 0.17* 1.21** 1 0.41** 0.23 −0.23 0.47*
CC_GWAS 0.02 0.02 1 1.21** 0.35** 0.37** −0.01 0.69**
CAD 0.13** 1 0.02 0.17* 0.64** 0.33** 0.24* 0.18**
AF 1 0.13** 0.02 −0.07 0.5** 0.61** 0.53** 0.29**
AF CAD CC_GWAS CC_MTAG HF_all niHF niHFpEF niHFrEF
r_g
1.0
0.0
−1.0

b


shared meta_analysis DCM_HCM 0.27** 0.29** 0.68** 0.69** 0.66* 0.64** 1
niHFrEF 0.29** 0.18** 0.69** 0.8** 0.33 1 0.64**
niHFpEF 0.53** 0.24* 0.82** 0.92** 1 0.33 0.66*
niHF 0.61** 0.33** 0.91** 1 0.92** 0.8** 0.69**
HF_all 0.5** 0.64** 1 0.91** 0.82** 0.69** 0.68**
CAD 0.13** 1 0.64** 0.33** 0.24* 0.18** 0.29**
AF 1 0.13** 0.5** 0.61** 0.53** 0.29** 0.27**
AF CAD HF_all niHF niHFpEF niHFrEF shared meta_analysis DCM_HCM
r_g
1.0
0.0
−1.0

Supplementary Figure 9: Pairwise genetic correlations between cardiomyopathy analyses and related cardiovascular phenotypes.

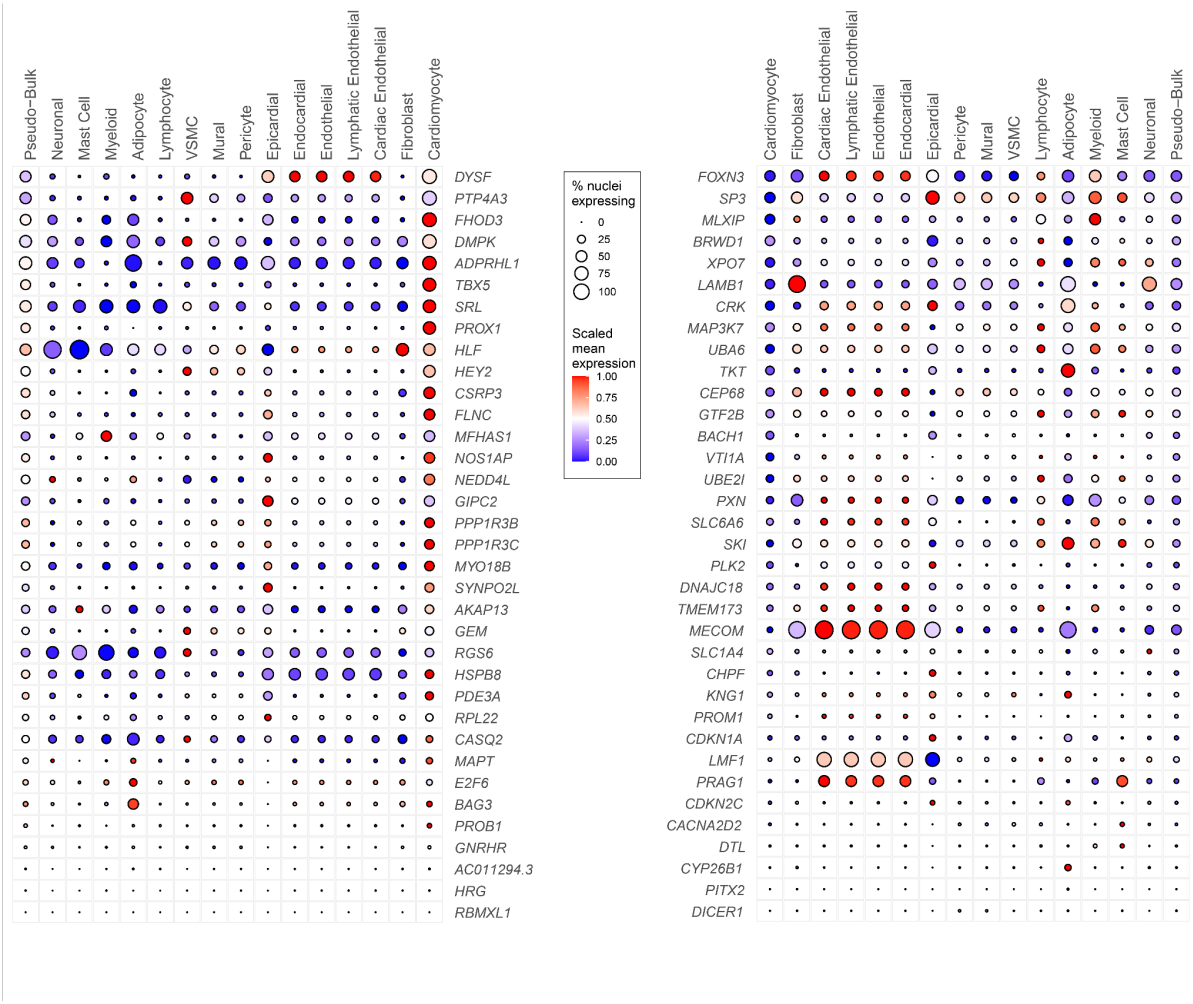
a, Genetic correlations (r_g) between case–case analyses (CC GWAS and CC MTAG), non-ischemic heart failure subtypes, coronary artery disease (CAD), and atrial fibrillation (AF).
b, Genetic correlations between the shared-effects DCM–HCM meta-analysis and the same cardiovascular phenotypes.
Pairwise correlations were estimated using linkage disequilibrium score regression (LDSC). Cell color reflects the magnitude and direction of (r_g (red = positive correlation, blue = negative correlation), and values are annotated in each cell. Asterisks indicate significance ($P < 0.05$: *, $P < 0.002$: ** (Bonferroni correction).



Supplementary Figure 10a: Cell-type-specific expression of the top prioritized genes for cardiomyopathies

Bubble heatmap showing gene expression data for prioritized genes across major cardiac cell types, based on integration of three published single-nucleus and single-cell RNA sequencing (sn/scRNA-seq) datasets of human left ventricles (LV) from non-failing control donors (maximum n = 18; **Supplementary Table 13**). The y-axis displays a shortlist of 76 highly prioritized genes from CC GWAS and MTAG analyses (with an additional 70 genes shown in **Supplementary Figure 10b**), while the x-axis shows 16 harmonized LV cell types. For each gene–cell type pair, dot size indicates the percentage of nuclei or cells expressing the gene at nonzero levels, and dot color represents the scaled, relative normalized expression within that cell type (compared to all other cell types). Expression values were aggregated after data harmonization and scaling. Note: VSMC, vascular smooth muscle cell; Pericyte, mural cell associated with microvasculature; Epicardial, epicardial-derived cells;

Endocardial, inner lining endothelial cells; Capillary Endothelial, capillary-associated endothelial cells; Lymphatic Endothelial, lymphatic vessel endothelial cells; Cardiomyocyte, cardiac muscle cells; Scaled mean expression, relative gene expression per cell type; % nuclei expressing, proportion of cells/nuclei with detectable gene expression; Padj, multiple-testing-adjusted two-sided P-value from DESeq2 differential expression



Supplementary Figure 10b: Cell-type-specific expression and DE of the top prioritized genes for cardiomyopathies

References

1. Jurgens, S. J. *et al.* Genome-wide association study reveals mechanisms underlying dilated cardiomyopathy and myocardial resilience. *Nat Genet* **56**, 2636–2645 (2024).
2. Turley, P. *et al.* Multi-trait analysis of genome-wide association summary statistics using MTAG. *Nat Genet* **50**, 229–237 (2018).
3. Tadros, R. *et al.* Large-scale genome-wide association analyses identify novel genetic loci and mechanisms in hypertrophic cardiomyopathy. *Nat Genet* **57**, 530–538 (2025).
4. Tadros, R. *et al.* Shared genetic pathways contribute to risk of hypertrophic and dilated cardiomyopathies with opposite directions of effect. *Nat Genet* **53**, 128–134 (2021).
5. Zheng, S. L. *et al.* Evaluation of polygenic score for hypertrophic cardiomyopathy in the general population and across clinical settings. *Nat. Genet.* **in press**, (2025).
6. Bulik-Sullivan, B. K. *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Genet.* **47**, 291–295 (2015).
7. Tadros, R. *et al.* Large scale genome-wide association analyses identify novel genetic loci and mechanisms in hypertrophic cardiomyopathy. *medRxiv* (2023) doi:10.1101/2023.01.28.23285147.
8. Jurgens, S. J. *et al.* Adjusting for common variant polygenic scores improves yield in rare variant association analyses. *Nat Genet* **55**, 544–548 (2023).
9. Pulit, S. L. *et al.* Meta-analysis of genome-wide association studies for body fat distribution in 694 649 individuals of European ancestry. *Hum Mol Genet* **28**,

166–174 (2019).

10. Peyrot, W. J. & Price, A. L. Identifying loci with different allele frequencies among cases of eight psychiatric disorders using CC-GWAS. *Nat. Genet.* **53**, 445–454 (2021).
11. O'Connor, L. J. *et al.* Extreme Polygenicity of Complex Traits Is Explained by Negative Selection. *Am J Hum Genet* **105**, 456–476 (2019).
12. Watanabe, K., Taskesen, E., van Bochoven, A. & Posthuma, D. Functional mapping and annotation of genetic associations with FUMA. *Nat Commun* **8**, 1826 (2017).
13. de Leeuw, C. A., Mooij, J. M., Heskes, T. & Posthuma, D. MAGMA: generalized gene-set analysis of GWAS data. *PLoS Comput Biol* **11**, e1004219 (2015).
14. GTEx Consortium. The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science* **369**, 1318–1330 (2020).
15. Wang, G., Sarkar, A., Carbonetto, P. & Stephens, M. A simple new approach to variable selection in regression, with application to genetic fine mapping. *J R Stat Soc Series B Stat Methodol* **82**, 1273–1300 (2020).
16. Zou, Y., Carbonetto, P., Wang, G. & Stephens, M. Fine-mapping from summary data with the 'Sum of Single Effects' model. *PLoS Genet* **18**, e1010299 (2022).
17. Karczewski, K. J. *et al.* Pan-UK Biobank GWAS improves discovery, analysis of genetic architecture, and resolution into ancestry-enriched effects. *bioRxiv* (2024) doi:10.1101/2024.03.13.24303864.
18. Schipper, M. *et al.* Prioritizing effector genes at trait-associated loci using multimodal evidence. *Nat Genet* **57**, 323–333 (2025).
19. Weeks, E. M. *et al.* Leveraging polygenic enrichments of gene features to predict genes underlying complex traits and diseases. *Nat Genet* **55**, 1267–1276

(2023).

20. Kolberg, L. *et al.* g:Profiler-interoperable web service for functional enrichment analysis and gene identifier mapping (2023 update). *Nucleic Acids Res* **51**, W207–W212 (2023).
21. Supek, F., Bošnjak, M., Škunca, N. & Šmuc, T. REVIGO summarizes and visualizes long lists of gene ontology terms. *PLoS One* **6**, e21800 (2011).
22. Reichart, D. *et al.* Pathogenic variants damage cell composition and single cell transcription in cardiomyopathies. *Science* **377**, eabo1984 (2022).
23. Jagadeesh, K. A. *et al.* Identifying disease-critical cell types and cellular processes by integrating single-cell RNA-sequencing and human genetics. *Nat Genet* **54**, 1479–1492 (2022).
24. Finucane, H. K. *et al.* Heritability enrichment of specifically expressed genes identifies disease-relevant tissues and cell types. *Nat Genet* **50**, 621–629 (2018).
25. Finucane, H. K. *et al.* Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nat Genet* **47**, 1228–1235 (2015).
26. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
27. Bulik-Sullivan, B. *et al.* An atlas of genetic correlations across human diseases and traits. *Nat Genet* **47**, 1236–1241 (2015).
28. Park, J.-B. *et al.* Obesity and metabolic health status are determinants for the clinical expression of hypertrophic cardiomyopathy. *Eur J Prev Cardiol* **27**, 1849–1857 (2020).
29. Harper, A. R. *et al.* Common genetic variants and modifiable risk factors underpin hypertrophic cardiomyopathy susceptibility and expressivity. *Nat Genet*

53, 135–142 (2021).

30. Grotzinger, A. D. *et al.* Genomic structural equation modelling provides insights into the multivariate genetic architecture of complex traits. *Nat Hum Behav* **3**, 513–525 (2019).
31. Jurgens, S. J. *et al.* Analysis of rare genetic variation underlying cardiometabolic diseases and traits among 200,000 individuals in the UK Biobank. *Nat Genet* **54**, 240–250 (2022).
32. Aragam, K. G. *et al.* Discovery and systematic characterization of risk variants and genes for coronary artery disease in over a million participants. *Nat Genet* **54**, 1803–1815 (2022).
33. Roselli, C. *et al.* Meta-analysis of genome-wide associations and polygenic risk prediction for atrial fibrillation in more than 180,000 cases. *Nat Genet* **57**, 539–547 (2025).
34. Henry, A. *et al.* Genome-wide association study meta-analysis provides insights into the etiology of heart failure and its subtypes. *Nat Genet* **57**, 815–828 (2025).
35. Ho, C. Y. *et al.* Genotype and Lifetime Burden of Disease in Hypertrophic Cardiomyopathy. *Circulation* (2018) doi:10.1161/CIRCULATIONAHA.117.033200.
36. Buniello, A. *et al.* Open Targets Platform: facilitating therapeutic hypotheses building in drug discovery. *Nucleic Acids Res* **53**, D1467–D1475 (2025).
37. Raies, A. *et al.* DrugnomeAI is an ensemble machine-learning framework for predicting druggability of candidate drug targets. *Commun Biol* **5**, 1291 (2022).
38. Strumberg, D. & Schultheis, B. Regorafenib for cancer. *Expert Opin Investig Drugs* **21**, 879–889 (2012).
39. Jiang, Y. *et al.* Afatinib for the Treatment of NSCLC with Uncommon EGFR

Mutations: A Narrative Review. *Curr Oncol* **30**, 5337–5349 (2023).

40. Polidovitch, N. *et al.* Phosphodiesterase type 3A (PDE3A), but not type 3B (PDE3B), contributes to the adverse cardiac remodeling induced by pressure overload. *J Mol Cell Cardiol* **132**, 60–70 (2019).
41. Voors, A. A. *et al.* Adrenomedullin in heart failure: pathophysiology and therapeutic application. *Eur J Heart Fail* **21**, 163–171 (2019).
42. Gaggin, H. K. & Januzzi, J. L., Jr. Biomarkers and diagnostics in heart failure. *Biochim Biophys Acta* **1832**, 2442–2450 (2013).
43. Pandhi, P. *et al.* Clinical value of pre-discharge bio-adrenomedullin as a marker of residual congestion and high risk of heart failure hospital readmission. *Eur J Heart Fail* **22**, 683–691 (2020).
44. Honigberg, M. C. *et al.* Polygenic prediction of preeclampsia and gestational hypertension. *Nat Med* **29**, 1540–1549 (2023).
45. Schooling, C. M. & Ng, J. C. M. Reproduction and longevity: A Mendelian randomization study of gonadotropin-releasing hormone and ischemic heart disease. *SSM Popul Health* **8**, 100411 (2019).
46. Li, X. *et al.* TAK1 Activation by NLRP3 Deficiency Confers Cardioprotection Against Pressure Overload-Induced Cardiomyocyte Pyroptosis and Hypertrophy. *JACC Basic Transl Sci* **8**, 1555–1573 (2023).