Common- and rare-variant genetic architecture of heart failure across the allele frequency spectrum

David S.M. Lee, MD, PhD[1], John S. DePaolo, MD, PhD[2], Krishna G. Aragam, MD[3,4], Kiran Biddinger[4], Mitchell Conery, BA[5], Ozan Dilitikas, MD[6], Lily Hoffman-Andrews, MS, CGC[7], Renae L. Judy, MS[2], Atlas Khan, PhD[8], Iftikhar Kulo, MD[6], Mugan J. Puckelwartz, PhD[9], Nosheen Reza, MD[6], Benjamin A. Satterfield, MD, PhD[6], Pankhuri Singhal, PhD[10], Regeneron Genetics Center[11], Zoltan P. Arany, MD, PhD[7,12], Thomas P. Cappola, MD, ScM[7], Eric Carruth, PhD[13], Sharlene M. Day, MD[7,12], Ron Do, PhD[14,15,16], Christopher M. Haggarty, PhD[13], Jacob Joseph, MD[17,18], Elizabeth McNally, MD, PhD[19], Girish Nadkarni, MD, MPH[20], Anjali T. Owens, MD[12], Daniel J. Rader, MD[10,21], Marylyn D. Ritchie, PhD[10,22], Yan Sun, PhD, MS[23], Benjamin F. Voight, PhD[10,24,25,26], Michael G. Levin, MD[7,26]*, Scott M. Damrauer, MD[2,26]*

[1]Department of Medicine, Northwestern University Feinberg School of Medicine, Chicago, IL

[2]Department of Surgery, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA

[3]Center for Genomic Medicine, Massachusetts General Hospital, Harvard Medical School, Boston, MA

[4]Program in Medical and Population Genetics and Cardiovascular Disease Initiative, Broad Institute of MIT and Harvard, Cambridge, MA

[5]Genomics and Computational Biology Graduate Group, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA

[6]Department of Cardiovascular Medicine, Mayo Clinic, Rochester, MN

[7]Division of Cardiovascular Medicine, Department of Medicine, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA

[8]Division of Nephrology, Department of Medicine, Columbia University Vagelos College of Physicians and Surgeons, New York, NY

[9]Department of Pharmacology, Center for Genetic Medicine, Northwestern University Feinberg School of Medicine, Chicago, IL

[10]Department of Genetics, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA

[11]Regeneron Genetics Center, Regeneron Pharmaceuticals, Tarrytown, NY

[12]Cardiovascular Institute, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA

[13]Department of Translational Data Science and Informatics, Geisinger, Danville, PA

[14]The Charles Bronfman Institute for Personalized Medicine, Mount Sinai Icahn School of Medicine, New York, NY

[15]Biome Phenomics Center, Mount Sinai Icahn School of Medicine, New York, NY

[16]Department of Genetics and Genomic Sciences, Mount Sinai Icahn School of Medicine, New York, NY

[17]Massachusetts Veterans Epidemiology Research and Information Center, VA Boston Healthcare System, Boston, MA
[18]Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA
[19]Center for Genetic Medicine, Bluhm Cardiovascular Institute, Northwestern University Feinberg School of Medicine, Chicago, IL
[20]Division of Nephrology, Department of Medicine, Mount Sinai Icahn School of Medicine, New York, NY
[21]Division of Translational Medicine and Human Genetics, Department of Medicine, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA
[22]Institute for Biomedical Informatics, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA
[23]Deparment of Biomedical Informatics, Emory University School of Medicine, Atlanta, GA
[24]Institute for Translational Medicine and Therapeutics, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA
[25]Department of Systems Pharmacology and Translational Therapeutics, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA
[26]Corporal Michael J. Crescenz VA Medical Center, Philadelphia, PA
*These authors contributed equally to this work.

Abstract

Heart failure (HF) is a complex trait, influenced by environmental and genetic factors, that affects over 30 million individuals worldwide. Historically, the genetics of HF have been studied in Mendelian forms of disease, linking rare genetic variants to familial cardiomyopathies. More recently, genome-wide association studies (GWAS) have successfully identified common genetic variants associated with risk of HF. However, the relative importance of genetic variants across the allele-frequency spectrum remains incompletely characterized. Here, we report the results of common- and rare-variant association studies of all-cause heart failure, applying recently developed methods to quantify the heritability of HF attributable to different classes of genetic variation. We combine GWAS data across multi-ancestry populations including 207,346 individuals from HF and 2,151,210 without, identifying 176 risk loci at genome-wide significance ($p < 5x10^{-8}$). Signals at newly identified common-variant loci include coding variants in Mendelian cardiomyopathy genes (*MYBPC3*, *BAG3*), as well as regulators of lipoprotein (*LPL*) and glucose metabolism (*GIPR*, *GLP1R*), and are enriched in cardiac, muscle, nerve, and vascular tissues, as well as myocyte and adipocyte cell types. Gene burden studies in the Penn Medicine BioBank uncover exome-wide significant ($p < 3.03x10^{-6}$) associations for HF and rare predicted loss-of-function variants in *TTN* and *MYBPC3.* Although the total burden heritability of

rare coding variants in HF (3.5%, 95% CI 0.94-6.0%) is comparable to common variant heritability (4.3%, 95% CI 3.9-4.7%), burden heritability is highly concentrated in a small set of Mendelian cardiomyopathy genes (30.4% explained by *TTN*, *LMNA*, *FLNC*, and *MYBPC3)*, while common-variant heritability is more diffusely spread throughout the genome. Finally, we demonstrate that common-variant background, in the form of a polygenic risk score (PRS), significantly modifies the risk of HF among carriers of pathogenic truncating variants in the Mendelian cardiomyopathy gene TTN. These findings suggest a significant polygenic component to HF exists that is not captured by current clinical genetic testing.

Introduction

Heart failure (HF) is a heterogeneous clinical syndrome affecting >30 million individuals worldwide and is the leading cause of unplanned hospital admissions in the United States for individuals >65 years old. HF arises from diverse etiologic causes, including chronic ischemic heart disease, myocardial infarction, diabetes, hypertension, obesity, arrhythmias, infiltrative and inflammatory disorders, and exposure to drugs or environmental toxins[1]. Although there has been progress in mapping genes involved in Mendelian forms of cardiomyopathy, the genetic architecture of common HF remains less well understood. Unlike for other common cardiometabolic traits, genetic loci identified in association with HF have remained comparatively limited despite increasing sample sizes, partly due to the heterogeneity of HF and its diverse etiologies[2–6]. In a subset of individuals, pathogenic variants affecting a core set of cardiomyopathy genes can be identified as the putative cause of HF using clinical panel-based genetic testing[7,8]. As with other common diseases and traits[4,6,9], common less penetrant variants influencing these core genes might also contribute to HF risk; yet, the catalog of such variants and their cumulative contribution to HF has not been fully enumerated. More broadly, few studies have integrated effects across common and rare variants to quantify the extent they implicate similar genes and pathways in HF.

Here, we conduct a comprehensive analysis of genetic variants across the allele frequency spectrum to examine shared and distinct pathways implicated by common and rare genetic variants in HF. Specifically, we perform a multi-population GWAS meta-analysis to identify new loci where common genetic variants are associated with all-cause HF and, in parallel, examine the contribution of rare coding variants in a large hospital-based biobank. We leverage this data to estimate the heritability of HF, and use the recently developed Abstract Mediation Model and Burden Heritability Regression methods[10,11] to quantify enrichment of both common and rare variants in a core set of known cardiomyopathy genes and other functional annotations. Finally, we construct a multi-population PRS to evaluate how common variants act collectively to modify the penetrance of loss-of-function truncating variants in *TTN*, the most common Mendelian risk factor for cardiomyopathy. Overall, our findings highlight the significant polygenic component of HF risk within the general population that is currently not captured by clinical genetic testing, and suggest that polygenic background may be important even among carriers of highly penetrant loss-of-function variants in known monogenic cardiomyopathy genes.

Results

GWAS of All-Cause Heart Failure

We conducted a multi-population GWAS meta-analysis of all-cause HF using an inverse-variance weighted fixed effects model, assembled from participants of large consortia and medical biobanks (Figure 1). Comprised of 207,346 individuals with and 2,151,210 individuals without HF, this analysis represents an approximate doubling in the effective sample size of the previously largest published HF GWAS (Figure 1)[3]. We combined individuals across four population groups with genetic similarity to 1000 Genomes Project reference superpopulation groupings[12]: a European group (EUR, 81.1% of effective sample size), an East Asian group (EAS, 6.5%), an African group including admixed African Americans (AFR, 9.7%), and an Admixed American group (AMR, 2.6%). The analysis identified 9,990 variants at 176 loci where genetic associations reached genome-wide significance ($P<5x10^{-8}$ - Figure 2, Suppl. Fig. 1 and Suppl. Table 1), representing an approximately 3.7-fold increase in the number of genome-wide significant loci compared to prior[3].

Of the 47 lead variants reported in the most recent HF GWAS[3], 46 variants (or perfect proxies) were associated with HF with concordant direction of effects (two-tailed binomial $P = 6.82 \times 10^{-13}$). 41/46 (89%) were significant at the genome-wide threshold, while 5/46 (11%) (rs1712355, rs12325072, rs138178446, rs150966198, rs6944634) remained significant at a conservative Bonferroni-corrected replication threshold ($p < 0.05/46$) with concordant directions of effect. Of the 176 loci identified in our study (Fig. 2A), 146 were located >500 kb from lead variants identified in the previous largest HF GWAS (where multiple signals were present at some loci)[3]. Approximately one-third of new lead variants in our study are associated with several known HF risk factors in the NHGRI/EBI GWAS Catalog, including BMI, atrial fibrillation, systolic and diastolic blood pressure, coronary artery disease, lipid metabolism, diabetes, and kidney function (Suppl. Table 2). Several newly discovered HF lead variants have also been associated with lifestyle factors with known contributions to HF risk, including alcohol consumption (rs988748, rs28712821)[13,14], smoking (rs1435428, rs34517439)[15,16], and physical activity (rs34517439)[17], highlighting the importance of behavioral phenotypes in HF risk.

Fine-mapping of risk loci

We performed fine-mapping of the 176 known and novel loci reaching genome-wide significance in our multi-population meta-analysis to identify candidate causal variants.

Compared to fine-mapping of the EUR population alone, the multi-population analysis improved fine-mapping resolution, decreasing the median size of the 99% credible sets by 35.1%, from 55.5 to 36 fine-mapped variants per locus (Suppl. Fig. 4). At 45 loci (25%), the fine-mapped credible set contained 10 or fewer variants. We explored protein coding consequences for fine-mapped variants within 99% credible sets across all 176 lead loci (n=14,628) using Ensembl VEP. Approximately one-third of our lead loci (61/176) contained a fine-mapped variant affecting a protein-coding sequence, with 42 loci harboring variants predicted to produce either missense mutations or protein truncations (Suppl. Table 3). Among these variants, four were predicted to be high-confidence protein truncations and were located in novel loci: in *SLC23A1* (rs11242462), *LPL* (rs328), *OR4P4* (rs76160133), and the microtubule-associated *MAPT* (rs754512) gene, respectively. *SLC23A1* encodes a sodium-dependent L-ascorbic acid transporter associated with circulating plasma vitamin C levels in humans[18], while *OR4P4* encodes an olfactory receptor gene at the 11q11 locus which has previously been linked to obesity[19]. Other variants uncovered by our approach include a missense variant in *MYBPC3* (rs11570076, tagged by the novel HF locus at rs2904129) located within a key domain for regulating the cMyBP-C interaction with S2 domain of myosin, and the previously described *BAG3* missense variant rs2234962 (tagged by rs17617337) that is enriched among patients with dilated cardiomyopathy[20].

Gene-prioritization

To nominate genes implicated by GWAS loci, we used an ensemble approach combining nearest-gene, coding consequence, TWAS, and MAGMA (see Methods) to prioritize 577 genes, with 73 genes prioritized by at least 3 methods (Suppl. Table 4). Among the subset of highly prioritized genes at novel loci were the glucose-dependent insulinotropic polypeptide receptor (*GIPR*) and glucagon-like peptide-1 receptor (*GLP1R*) genes. Of note, neither of these genes have previously been associated with HF in humans, although cardioselective deletion of *GIPR* has previously been shown to improve survival and reduce adverse remodeling following myocardial ischemic injury in mice[21]. Given the role of these genes in glucose homeostasis and obesity, we performed multi-trait colocalization using GWAS summary statistics to evaluate the genetic evidence supporting shared or independent signals for HF, diabetes[5], hemoglobin A1c[22], and body mass index[23]. At the *GIPR* locus there was 99.4% posterior probability of a shared genetic variant influencing both HF and BMI. In contrast, at the *GLP1R* locus we did not find evidence of colocalization, indicating the possibility that the HF, diabetes, and obesity signals at this locus may be due to separate mechanisms.

Tissue, cell, and pathway-enrichment

Using the full set of genes nominated at GWAS loci (n = 577), we queried public databases to identify relevant tissues and biological pathways implicated by loci from our GWAS meta-analysis. As expected, heart and blood vessels were among the tissues containing significantly up-regulated expression of GWAS-prioritized genes (Fig. 2B) while hallmark gene sets for myogenesis and hypoxia were also enriched after correction for multiple testing (Suppl. Fig 5). Among phenotype-associated gene sets published in the NHGRI/EBI GWAS catalog, GWAS-prioritized genes overlapped most significantly with gene sets associated with coronary artery disease, systolic blood pressure, body mass index, and atrial fibrillation, reflecting the multifactorial nature of HF (Suppl. Fig. 6). Consistent with an emerging role for metabolic processes in HF, when matched against single cell expression data from 13 cell types across 8 different tissues published by GTEx[24], GWAS-prioritized genes were most significantly enriched in myocyte and adipocyte cells in the heart (Fig. 2C). Indeed cardiac adipose tissue is thought to contribute to cardiac pathology through a myriad of mechanisms[25], and this enrichment in cardiac adipocytes suggests part of their contribution to HF is mediated by gene expression changes resulting from germline genetic variation. Notably, prior GWAS had only reported enrichment of GWAS genes in cardiac myocytes[3].

Rare Variant Genetic Architecture of All-cause Heart Failure

To assess the rare variant architecture of HF, we conducted an exome-collapsing gene-based association test using linked electronic health record data and whole exome sequencing for 43,731 participants from the Penn Medicine Biobank (PMBB). We performed gene-based group tests for rare (MAF < 0.1% in gnomAD exomes and in PMBB) predicted loss-of-function (pLoF) variants using the mixed-model framework SAIGE-GENE[26] (Figure 3A) for 5,769 participants with and 23,593 participants without HF from the EUR population and 2,235 participants with and 7,982 participants without HF from the AFR population.

Under this framework, we replicated the previously-reported association between rare pLoF variants in *TTN* and HF at exome-wide significance in EUR (P-value 1.8×10^{-24}, Suppl. Table 5)[7]. This association was nominally significant in AFR (P-value 5.6×10^{-3}, Suppl. Table 6). We also uncovered an exome-wide significant signal for rare pLoF variants in *MYBPC3* in EUR (P-value 4.6×10^{-8}) but did not observe this association among AFR participants (P-value 0.58).

Previously, pLoF variants in *TET2* had been associated with HF in an exome-collapsing study of HF containing participants from two large clinical trials[7], however we found no association between *TET2* and all-cause HF in our analysis (P-value 0.98 in EUR, 0.85 in AFR).

Rare-variant associations in Mendelian cardiomyopathy genes

Taking a targeted approach, we assessed the prevalence of pLoF variants in a set of 65 expert-curated (NIH ClinGen) genes associated with Mendelian forms of dilated or hypertrophic cardiomyopathy (the two most prevalent forms of inherited cardiomyopathy[27]). We grouped these genes by levels of evidence supporting their role as a monogenic cause of cardiomyopathy as determined by expert panels[28,29]. This resulted in 19 definitive evidence genes, 6 strong or moderate evidence genes, 32 limited or disputed evidence genes, and 8 genes without evidence for involvement in human cardiomyopathy despite experimental evidence in the literature (Suppl. Table 7). Using these groupings, we observed that the prevalence of HF among cardiomyopathy pLoF carriers declined with decreasing levels of evidence of involvement in cardiomyopathy (Fig. 3B), and the yield of finding a cardiomyopathy pLoF carrier among participants with HF approximately doubled when all potential cardiomyopathy genes are considered compared to just definitive cardiomyopathy genes (Fig. 3C – see Suppl. Note).

We next asked whether pLoF burden in each cardiomyopathy gene was associated with increased odds of all-cause HF. For this secondary analysis, we used a strict significance threshold of $P<1.29\times10^{-3}$ as determined by permutation to control Type 1 Error at 5% (P-value < 1.53×10^{-3} for AFR), and a more permissive suggestive significance threshold ($P < 0.05$). Among the 19 genes classified as definitive cardiomyopathy genes, 4 were significantly associated with all-cause HF at the strict threshold (*TTN*, *MYBPC3*, *LMNA*, *FLNC*), while an additional 2 (*BAG3, RBM20*) were associated at the permissive threshold (Suppl. Fig. 8, Suppl. Tables 8 and 9). For genes with moderate to strong evidence supporting their association with cardiomyopathy, 2 of 6 demonstrated associations with increased odds of all-cause HF (*JPH2*, *DSP*) in EUR pLoF variant carriers at the permissive threshold. For genes with limited or disputed evidence of association with any cardiomyopathy, 4 of 30 were significantly associated with all-cause HF at the permissive threshold (*CALR3*, *ANKRD1*, *MYPN*, *EYA4*), and for genes deemed to have no evidence of an association with cardiomyopathies only pLoF variation in *LRRC10* in AFR was associated with a diagnosis of all-cause HF at the permissive threshold. Because missense variants may produce distinct consequences compared to pLoF variants[30], we additionally

repeated testing associations with carrier status for a predicted damaging missense variant and all-cause HF status. Compared to carrier status for pLoF variants, predicted damaging missense variants did not increase risk of HF in most genes (see Suppl. Note).

Heritability of Heart Failure based on Common and Rare Variants

To investigate the contribution of common genetic variants to the heritability of all-cause HF, we use LD-score regression with corresponding population-specific LD scores from Pan-UKB[31]. Common variant heritability (AF 5-50%) for all-cause HF ranged from 3-10% across the 4 populations included in our analysis (AFR 8.6%, 95% CI 6.1-11%, AMR 10%, 95% CI 6.0-14%, EAS 3.4%, 95% CI 1.9-4.9%, EUR 4.3%, 95% CI 3.9-4.7%) on the liability scale, using population specific prevalence data for heart failure from NHANES based on self-reported race/ethnicity[32] (Suppl. Fig. 7). Overall the range of these estimates was consistent with previous GWAS meta analyses of all-cause HF, which have estimated common variant heritability around 8.8%[2].

We assessed the contribution of rare coding genetic variation to the heritability of all-cause HF using whole exome sequencing in the PMBB with Burden Heritability Regression (BHR)[11]. Our estimate of the total burden heritability, mediated by rare (AF < 0.001) damaging missense and pLoF variants, was comparable to that of common variant heritability (3.5%, 95% CI 0.94-6.0%) in EUR. As expected, burden heritability was concentrated in pLoF variants, where biological impacts are the most severe, and in the rarest variants (Figure 4A), consistent with the impact of negative selection. In contrast, heritability for missense variants was low - even at the rarest allele frequencies (Suppl. Fig. 12). As seen in other common complex disease phenotypes[11], pLoF burden heritability in all-cause HF was enriched in more constrained genes (as defined by the gnomAD LOEUF quintile – Figure 4B), and was overall lower than common variant heritability (Figure 4A). These findings support a polygenic model of HF whereby common variants affecting many peripheral genes explain a similar proportion of heritability as rare coding variants affecting a comparatively limited set of core cardiomyopathy genes.

Next, we asked what proportion of both heritability estimated from rare variant burden and common variants could be attributed to the set of 65 genes implicated in monogenic cardiomyopathies. BHR revealed significant enrichment in burden heritability across all cardiomyopathy genes, but this enrichment was primarily concentrated in genes with “definitive”

evidence for a causal role in cardiomyopathy syndromes (60-fold enrichment - Figure 4B). We did not observe enrichment in burden heritability for the set of genes with only limited evidence supporting their role in Mendelian cardiomyopathies in the EUR population. We found similar patterns of enrichment in rare variant heritability within the AFR population, although this analysis was comparatively limited due to reduced sample size and power (Suppl. Fig. 11).

We estimated heritability mediated through common variants in our set of curated cardiomyopathy genes using the Abstract Mediation Model (AMM)[10]. While rare variant heritability was 27-fold enriched across all monogenic cardiomyopathy genes, this gene set was enriched only 3.4-fold for common variants, accounting for only a small proportion of overall common-variant heritability (Figure 5A). This difference was even more striking for definitive monogenic cardiomyopathy genes, where rare variant heritability was 60-fold enriched compared to 8.1-fold for common variants. The smaller enrichment of heritability across known cardiomyopathy genes for common variants, coupled with the comparable common- and burden heritability estimates for HF, is consistent with common variant HF heritability being significantly more polygenic than rare coding variant HF heritability. Indeed, the fraction of common variant heritability mediated through the 509 protein-coding genes linked to GWAS-significant loci in our study accounted for approximately 37.5% of the total common variant heritability estimate (Fig. 5A). In contrast, the 18 definitive cardiomyopathy genes explained approximately 34.2% of rare variant burden heritability in the EUR population, with 30.4% of overall rare variant burden heritability concentrated in just 4 definitive cardiomyopathy genes (*TTN*, *LMNA*, *FLNC*, and *MYBPC3;* Figure 5B, Suppl. Table 11).

Modification of HF risk for loss-of-function *TTN* carriers by polygenic risk factors

In some cases, the etiology of HF may be assessed clinically using panel-based testing for rare pathogenic variants in known cardiomyopathy genes. However, the role of testing for common variants is unclear. While individually common variants tend to have smaller effects than rare deleterious variants, in aggregate they explain a comparable portion of HF heritability, suggesting they may play an important role in predicting HF risk. To evaluate how common variants may clinically contribute to HF risk, we generated a multi-population polygenic risk score (PRS) for HF using summary statistics from the GWAS meta-analysis (excluding PMBB participants), using PRS-CSx[33]. The PRS was normally distributed in the held-out PMBB population, and the mean PRS was significantly greater among individuals with HF than those

without (Suppl. Fig. 14A). The PRS was predictive of prevalent HF in PMBB (AUC 0.72, 95% CI 0.72-0.73), and significantly improved upon a naïve baseline model (Brier skill score 0.065, 95% CI 0.063-0.068). Prevalence of HF ranged from 8.1% (95% CI 7.3-9.0) among the lowest PRS decile to 16.2% (95% CI 15.1-17.3) among the highest PRS decile (Suppl. Fig. 14B).

We then asked whether polygenic risk might modify the penetrance of high PSI (>90% proportion spliced into cardiac-expressed transcripts) truncating variants in *TTN* (TTNtv), a "definitive"-evidence cardiomyopathy gene which accounts for the majority of the rare variant burden heritability of HF[27]. This analysis demonstrated a significant interaction between PRS and TTNtv status (interaction OR 1.32; 95% CI 1.06-1.64; p = 0.01), indicating that the risk of HF for TTNtv carriers depends significantly on their individual polygenic background. Most of this signal was driven by the relationship between PRS and *TTN*tv status in the EUR population (P-value 0.001 - Suppl. Fig. 15, Suppl. Table 14). Among all participants, a PRS two standard deviations above the mean conferred an approximately 7-fold increased odds of HF for TTNtv carriers, compared with an approximately 2-fold increase in risk for TTNtv carriers at average polygenic risk. These results imply that HF risk assessment for individuals carrying loss-of-function mutations in cardiomyopathy genes may improve if polygenic background is also considered.

Discussion

We present a comprehensive analysis of genetic contributions to heart failure across the allele frequency spectrum. Our study highlights the shared and unique genetic architectures of common and rare variants in HF, finding significant common and rare variant heritability, which is concentrated in different functional annotations. We characterize the contribution of known cardiomyopathy genes to HF risk across both common and rare variants, and explore the interaction between common and rare genetic variation. Overall, these findings have implications for future genetic association studies of HF, and for HF risk-stratification.

In our multi-population genome-wide meta-analysis, we report 176 loci associated with HF, substantially increasing the number of HF risk-loci compared to prior studies[2,3,34]. Among the newly discovered loci are several associated with lifestyle factors, including smoking and alcohol consumption. Metabolism and nutrition are also broad processes implicated by our novel HF loci, with fine-mapped variants implicating the obesity-related gene *OR4P4*, new HF

associations with loci near the insulin signaling and glucose homeostasis genes *GIPR* and *GLP1R*, and the enrichment of expression of GWAS-prioritized genes in cardiac adipocytes. Mendelian randomization has previously implicated *GLP1R* signaling in HF[35], and clinical trials for the dual GIP/GLP-1 receptor agonist tirzepatide as a potential therapy for select patients with heart failure are ongoing[36]. Although medications targeting these mechanisms have potent effects on diabetes, glycemic traits, and obesity, our colocalization findings indicate that the HF effects at these loci may be separate from glycemic traits, raising the possibility of direct HF-specific benefits. Although we report a significant increase in the number of HF-associated loci in our study, genes prioritized from our ensemble approach explain 16-37% of common-variant heritability, indicating that more common-variant HF loci remain undiscovered.

In parallel, we investigate rare coding variant contributions to HF in a large hospital-based biobank. Our exome-wide association study identified exome-wide associations for the known cardiomyopathy genes *TTN* and *MYBPC3*, and nominal associations for several other cardiomyopathy genes including *LMNA* and *FLNC*. In addition to these definitive cardiomyopathy genes, we observed that pLoF variants in the limited evidence gene *LRRC10* are associated with HF in AFR individuals. Although this association was not significant after accounting for multiple testing, it is consistent with prior studies showing loss-of-function mutations in *LRRC10* can cause heart failure and dilated cardiomyopathy in zebrafish and mice[37,38], and homozygous mutations in *LRRC10* have been implicated as a cause of pediatric dilated cardiomyopathy in humans[39]. Finally, in contrast to common variants, rare variants in known cardiomyopathy genes explain a large fraction of burden heritability in HF. Despite this, the overall fraction of burden heritability explained by known cardiomyopathy genes is less than half, and we anticipate that larger cohorts will lead to the discovery of additional genes whose loss-of-function is associated with HF.

Across both our common- and rare-variant analyses, we find that heritability of HF is enriched within a core set of cardiomyopathy genes. This finding provides empirical support for the current expert pathogenicity categorization of these genes, and supports their inclusion in cardiomyopathy testing panels[28,29]. Importantly, although current clinical cardiovascular testing panels focus on damaging coding variants in these core genes, we find that common variants at these same loci also contribute to risk of HF. In contrast to rare variants, for which most of the heritability is accounted for by these core genes, we find that common variant heritability is substantially more polygenic, implicates a far broader range of risk factors, and explains a

comparable fraction of HF heritability as rare coding variants. In combination, these results are consistent with an polygenic model of inheritance,[40] where a set of core genes and pathways are influenced by a much broader set of interconnected regulatory networks. Supporting this model for HF, we find significant common-variant heritability in non-coding functional annotations, while rare-variant heritability is concentrated in low-frequency, loss-of-function variation in constrained genes. We also observe that common variation and burden heritability explain a similar fraction of overall heritability in HF. These results suggest that at a population level, common genetic variation plays a significant role in risk of HF comparable to the known contribution of rare pathogenic variants in familial cardiomyopathy genes.

Finally, we observe an interaction between common and rare genetic variation, which may have implications for HF risk stratification. Previously it had been shown that an individual's polygenic background can modify the penetrance of pathogenic variants across several monogenic conditions[41,42]. Here, we observe a significant interaction between polygenic background and TTNtv carrier status, suggesting that risk stratification for rare pLoF variant carriers in monogenic cardiomyopathies may be improved if additional information is available regarding polygenic background. A similar interaction has previously been reported for TTNtv and a PRS for atrial fibrillation, highlighting the diverse manifestations of cardiovascular disease among TTNtv carriers[43]. Future studies are warranted to investigate the use of polygenic background to adjudicate cardiomyopathy genes with less certain levels of evidence and variants of unknown significance, as well as the role of PRS testing more broadly when assessing HF etiology and risk.

Limitations

The results of our study should be interpreted within the context of their limitations. First, the current analysis focused on an all-cause heart failure phenotype, which may limit discovery of genetic variants that discordantly influence different forms of heart failure (eg. heart failure with reduced vs. preserved ejection fraction, dilated vs. hypertrophic vs. arrhythmogenic cardiomyopathy, ischemic vs. nonischemic cardiomyopathy). For example, we and others have previously observed HF risk loci with opposite directions of effect on HF endophenotypes and subtypes[3,42]. By studying an all-cause HF phenotype, the effects of these variants may be biased toward the null. Future large multi-population studies of HF subtypes and endophenotypes are necessary to identify these associations. Similarly, known cardiomyopathy genes increase HF risk by diverse mechanisms beyond pLoF. While we considered both pLoF

and predicted damaging missense variants, domain and gene-specific expertise is likely necessary for accurate pathogenicity evaluation. Next, although the prevalence of HF is similar across sexes, whether specific loci affect HF through sex-specific mechanisms requires future study[44]. Finally, most published GWAS consist primarily of participants from European populations, which remains a limitation across much of human genetics research. We hope that future studies continue to expand on the diversity of participants included from historically underrepresented populations. The inclusion of participants from the VA Million Veterans Program (MVP) significantly expanded the genetic diversity of our study compared to previous publications (Fig. 1). Additional promising future research directions include examining both common and rare variant contributions across subtypes of HF and identifying the genetic pathways responsible for governing an individual's prognosis and response to specific therapies.

In summary, these analyses highlight similarities and differences in the genetic architecture of HF across the allele-frequency spectrum, which is characterized by a highly concentrated component of heritability in rare deleterious variants in known cardiomyopathy genes, and a large polygenic component of heritability that is more diffusely spread across GWAS loci and enriched for non–coding variation.

Methods

Genome-wide meta analysis of heart failure

GWAS summary statistics for HF were obtained from non-overlapping analyses of 8 separate cohorts/consortia (HERMES, VA MVP, FinnGen, Mount Sinai BioMe, Penn Medicine Biobank, eMERGE, Geisinger DiscovEHR, and the Global Biobank Meta-analysis Initiative). All-cause HF was defined using cohort-specific definitions (pheCodes or ICD9/10 codes documented within the electronic health record for all studies except HERMES, which additionally included expert adjudication among some cohorts).

Analyses were performed separately by cohort and population, adjusted for age, sex, and population structure. Fixed-effects meta-analysis was performed using METAL[45] using the inverse-variance weighted (standard error) method within and across populations to generate population-specific and multi-population meta-analysis summary statistics. Fine-grained

population estimation from summary statistics was performed using *bigsnpr*[46]. Independent significant genomic risk loci were defined by grouping genome-wide significant variants ($p < 5x10^{-8}$) into 1MB clusters using the *genome_cluster* function of the *tidygenomics* package in R[47]. To minimize spurious associations, loci were required to have at least 2 genome-wide significant associations. Lead variants at each locus were defined by the variant with the lowest p-value, present in at least 2 studies.

Gene Prioritization and Identification of variants within the 99% credible set

Credible sets were generated surrounding the index variant at each 1MB locus identified in the multi-population meta-analysis, centered on the index variant (+/-500kb). Credible sets were constructed by first estimating approximate Bayes factors as described in Graham et. al, which assumes a single causal variant per locus. [48]Posterior probabilities were then generated by dividing the Bayes factor for each variant at a locus by the sum of Bayes factors at each locus[49]. Finally, variants at each locus were ranked in descending order of posterior probability and included until the cumulative posterior probability at each locus attained 99%. Possible protein-coding consequences were annotated for each variant in this 99% credible set using Ensembl VEP[50]. TWAS was performed using S-MultiXcan[51] using GTEx tissue expression data while additional gene candidates were identified from GWAS loci using MAGMA[52] following the standard protocol, with statistical significance determined using a Bonferroni correction for multiple testing (0.05/# of genes in each TWAS or MAGMA analysis). Gene set enrichment analysis was performed using FUMA[53] for tissue-specific enrichments. To identify cell-type specific enrichment, processed single-cell expression data was downloaded from GTEx v8[24]. Cell-type specific expression profiles were constructed by taking median expression values across samples for each annotated cell-type, and then enrichment for cell-type specific expression was determined by assessing the number of prioritized GWAS genes that were expressed by each cell-type using the hypergeometric test.

Multi-trait colocalization

Multi-trait colocalization was performed using HyPrColoc to evaluate the presence of a shared causal variant influencing HF and diabetes/obesity related traits[54]. This Bayesian algorithm is designed to evaluate for the presence of a causal variant shared among several traits, under the single causal variant assumption. GWAS summary statistics for diabetes[5], hemoglobin A1c[22], and body mass index[23] were queried at the *GLP1R* and *GIPR* loci to identify variants +/-500kb of the lead variant for HF at each locus. Because the diabetes, A1c, and BMI GWAS were

performed among European populations, the European population GWAS of heart failure was used to minimize differences in colocalization due to linkage disequilibrium. Evidence for colocalization was determined based on the default variant specific regional and alignment priors, with colocalization determined with posterior probability >= 0.25.

Curation of Cardiomyopathy Genes

Cardiomyopathy genes for dilated and hypertrophic cardiomyopathies were selected from two previous publications from the NIH ClinGen working group and classified according to the highest level of evidence supporting each gene across DCM or HCM collectively. We focused this analysis on genes from Ingles et al.[28] for hypertrophic cardiomyopathy, and Jordan et al. for dilated cardiomyopathy[29]. Except for *FHL1*, *ALPK3*, and *PLN*, genes involved in syndromic HCM were not included in our set of 65 cardiomyopathy genes.

Rare variant association analysis

Sample preparation, exome sequencing, and variant QC of participants in the Penn Medicine Biobank have been described previously[55]. We further processed variants to remove variants flagged during joint variant calling based on GATK's hard-filtering thresholds[56]. When >10% of variant alternate calls across PMBB participants at a particular site were flagged for QC these variant sites were removed from the analysis. Otherwise, flagged variants composing less than 10% of all alternate calls at a variant site were changed to missing.

Variant annotation was performed using Ensembl VEP. Variants were removed from the analysis if they were missing in >10% of sequenced participants in the PMBB. Damaging missense variants for each gene were identified using publicly available REVEL scores[57]. For predicted loss-of-function (pLoF) variants (essential splice site disrupting, stop gain, or frameshift variants) annotated by Ensembl VEP, we further restricted analysis to the subset of variants passing all LOFTEE filters to produce a filtered high quality set of pLoF variants[58]. Rare variant association studies were performed using SAIGE-GENE for individuals of European or African populations separately, adjusting for age, sex, and the first 10 genetic principal components. P-values were derived from the sequence-kernel association test (SKATO) implemented in SAIGE-GENE, and adjusted for case-control imbalance using Firth's logistic regression. The exome-wide analyses amounted to 16,517 association tests in EUR participants, and 14,232 association tests in AFR participants, corresponding to Bonferroni-adjusted significance thresholds of 3.02×10^{-6} and 3.51×10^{-6} respectively. Unless otherwise

noted, all other rare variant studies in the present study relied on pLoF variant annotations produced from this analysis.

To determine the odds of having HF among pLoF carriers in the subset of 65 curated cardiomyopathy genes, the burden of all LOFTEE-annotated high quality loss of function variants were tested for an association with HF in the Penn Medicine Biobank using logistic regression controlling for age, age^2, sex, and the first 10 genetic principal components. Odds ratios were corrected by Firth's adjusted logistic regression. This procedure was repeated for predicted damaging missense variants for each gene, defined as having a REVEL score >0.5. To determine the empirical threshold for statistical significance while controlling type 1 Error at 5%, we permuted HF vs. non-HF labels while maintaining the same total number of HF cases versus controls and determined the P-value threshold representing the lowest 5th percentile generated from 1000 permutations. We conducted this procedure for pLoF and missense REVEL >0.5 variants in European and African populations separately.

Estimation of common variant heritability and enrichment

Common variant heritability was estimated using LD score regression (LDSC)[59,60] for each population separately in the HF GWAS meta-analysis using reference panels published by the Pan-UK Biobank.

To determine heritability enrichment and fraction of common variant heritability distributed across different functional categories we used stratified LD-score regression[61] and precomputed baseline 2.2 model LD scores available online here. Since pre-computed LD scores were not available for African populations, we generated custom LD scores using bed files corresponding to the baseline 2.2 model (available here) to estimate heritability enrichment across functional categories using summary statistics derived from the African GWAS fraction.

To estimate the enrichment and fraction of common variant heritability mediated through specific gene sets, we used the Abstract Mediation Model with default parameters[10]. Briefly, for analysis of cardiomyopathy genes, we separated cardiomyopathy genes into sets based on their level of supporting evidence as determined by expert review, and assessed each set separating for common variant heritability enrichment and the fraction of overall common variant heritability explained. To determine the enrichment of common variant heritability explained by all genome-wide significant loci, we used the set of nearest mapped protein coding genes for

each loci. Of 176 independent genome-wide significant loci, we were able to identify a nearest protein-coding gene (as annotated by Ensembl VEP) for 509 genes prioritized through our ensemble approach from 157 loci that could be used in our estimate of common variant heritability mediated through various GWAS-significant gene sets. To calculate the fraction of common variant heritability mediated through these genes we multiplied their estimated enrichment in heritability as determined by AMM by the fraction of genes in each gene set respectively divided by the total number of genes included in the AMM model.

Estimation of rare variant heritability and enrichment

To estimate rare variant burden heritability, we used Burden Heritability Regression (BHR)[11]. Unless otherwise noted, results are reported for singleton variants in the PMBB. Only variants annotated within protein coding exons on autosomes were used for this analysis. To estimate rare variant heritability, we performed BHR using sets of singleton variants grouped by functional consequence (synonymous, missense, or pLoF) for European and African American populations separately in the PMBB using the default parameters. To estimate rare variant heritability enrichment across cardiomyopathy genes, we supplied BHR with separate annotation lists of cardiomyopathy gene IDs in addition to the baseline model. For the fraction of heritability explained by cardiomyopathy genes, we supplied BHR with separate sets of cardiomyopathy genes, split by level of evidence supporting their role in monogenic cardiomyopathy syndromes as determined by expert review. In this analysis, some genes with low variant counts within the PMBB were estimated to contribute small but negative fractions to fraction of burden heritability and these were not considered in the overall estimate for the fraction of burden heritability explained by significant genes analysis.

Interaction between common and rare variants to determine HF risk in PMBB

Carriers of predicted loss-of-function variants in TTN affecting high-percent spliced in (PSI) exons were identified in the PMBB, and an interaction effect between hi-PSI TTNtv carrier status and HF polygenic risk was tested for among EUR and AFR participants of PMBB. The multi-population PRS was constructed using the “auto” setting of PRS-CSx[33], using population-specific LD reference panels from the UK Biobank. Classification performance for prevalent HF was assessed in PMBB using repeated cross validation (3 repeats of 10 folds), using the area under the receiver operator characteristic curve as a measure of discrimination, and using the Brier skill score as an integrated measure of discrimination and calibration in comparison to a naïve/no-skill baseline model[62]. TTNtv carrier status was included in a logistic model with the

multi-population PRS derived from the multi-population GWAS using HF (defined as 2 instances of phecode 428) as an outcome, including an interaction term between the PRS and TTNtv carrier status. All PRS models were adjusted for age, sex, and 5 genetic principal components, with PRS scaled/centered within each target population (EUR or AFR) prior to modeling.

Data Availability

Summary statistics from the genome-wide association study meta-analysis and rare-variant gene-burden association study of all-cause heart failure will be made available upon publication. Polygenic risk scores for heart failure will be deposited in the Polygenic Score Catalog.

Acknowledgements

We thank the participants of the Penn Medicine Biobank and other cohorts/consortia contributing to this analysis. The PMBB is supported by Perelman School of Medicine at University of Pennsylvania, a gift from the Smilow family, and the National Center for Advancing Translational Sciences of the National Institutes of Health under CTSA award number UL1TR001878. M.G.L. is supported by the Institute for Translational Medicine and Therapeutics of the Perelman School of Medicine at the University of Pennsylvania, the NIH/NHLBI National Research Service Award postdoctoral fellowship (T32HL007843), and the Measey Foundation.

Main Figures

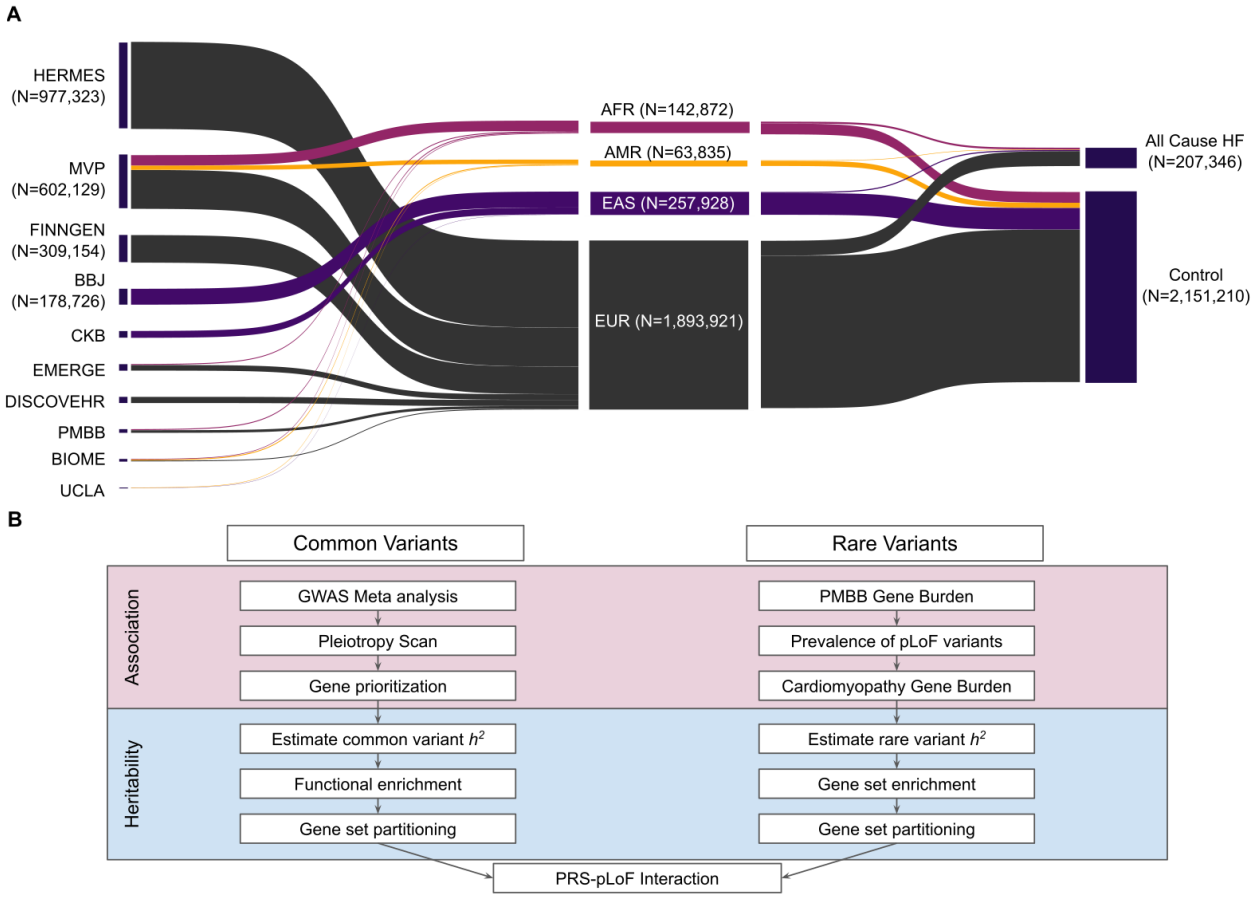

Figure 1: (A) Study population used in GWAS meta-analysis. We conducted a meta-analysis of heart failure across 10 cohorts comprising a combined count of N=207,346 participants with HF and 2,151,210 participants without HF. Summary statistics were collected across 10 cohorts and meta-analyzed using METAL. (B) Overview of study design, which encompasses GWAS meta-analysis and estimate of common variant h^2 in parallel with an analysis of rare coding variant associations with heart failure and rare variant h^2 in the Penn Medicine BioBank (PMBB).

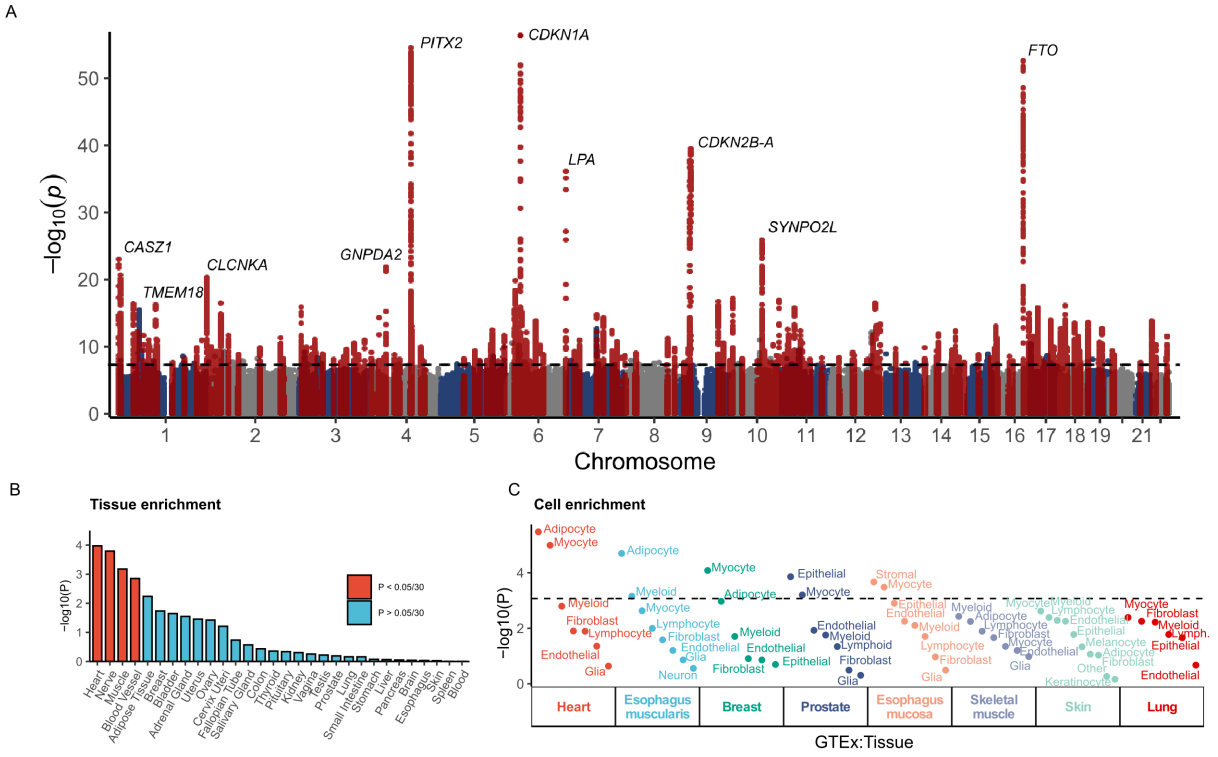

Figure 2: (A) Manhattan plot for common variant association in multi population meta-analysis (207,346 HF, 2,151,210 non-HF). Each point represents a variant, organized by chromosomal position along the x-axis. Lead variants are annotated with the closest mapped gene and the surrounding +/- 500kb is highlighted in red. (B) Enrichment for tissue-specific differentially increased gene expression of GWAS prioritized genes. P-values determined by hypergeometric test. Enrichment p-values above the Bonferroni-corrected significance threshold (0.05/30) are shaded red. (C) Enrichment for expression of GWAS prioritized genes in GTEx multi-tissue single cell sequencing data. P-values determined by hypergeometric test (see Methods). Dashed black line represents the Bonferroni-corrected significance threshold (0.05/59).

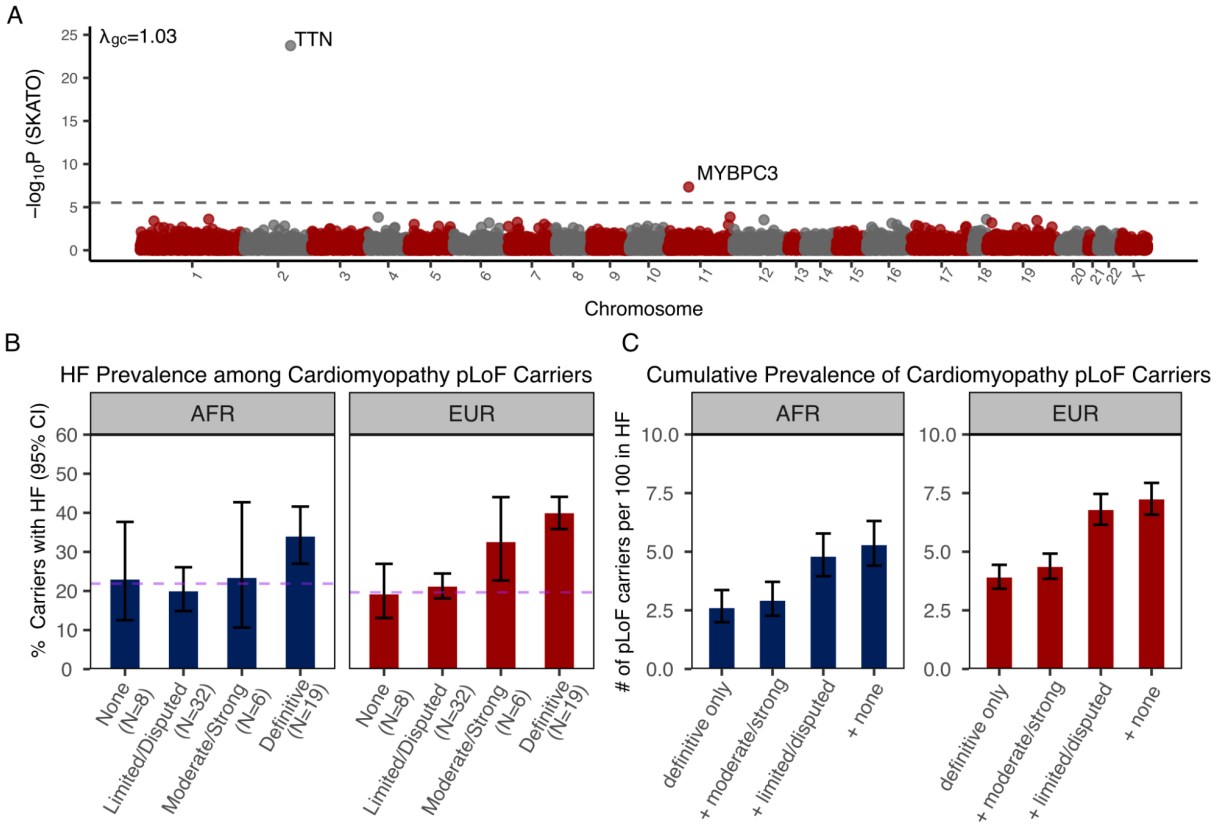

Figure 3: (A) Gene burden results for EUR population. Predicted loss-of-function (pLoF) variants were aggregated across individuals within the PMBB and the burden of pLoF variants was tested for an association with heart failure as determined by a combination of ICD codes. Dashed line represents the Bonferroni-corrected significance threshold. *P*-values derived from burden testing implemented in SAIGE-GENE(+). (B) Prevalence of HF among rare (MAF < 0.001) pLoF variant carriers across genes implicated in monogenic cardiomyopathy syndromes. Cardiomyopathy genes were categorized by their level of evidence for a role in monogenic cardiomyopathies as determined by expert review. Purple dashed line represents overall prevalence of HF among PMBB participants for each population respectively. (C) Cumulative prevalence of rare (MAF < 0.001) pLoF variant carriers in cardiomyopathy genes among PMBB participants, with addition of cardiomyopathy gene sets of decreasing evidence from left to right.

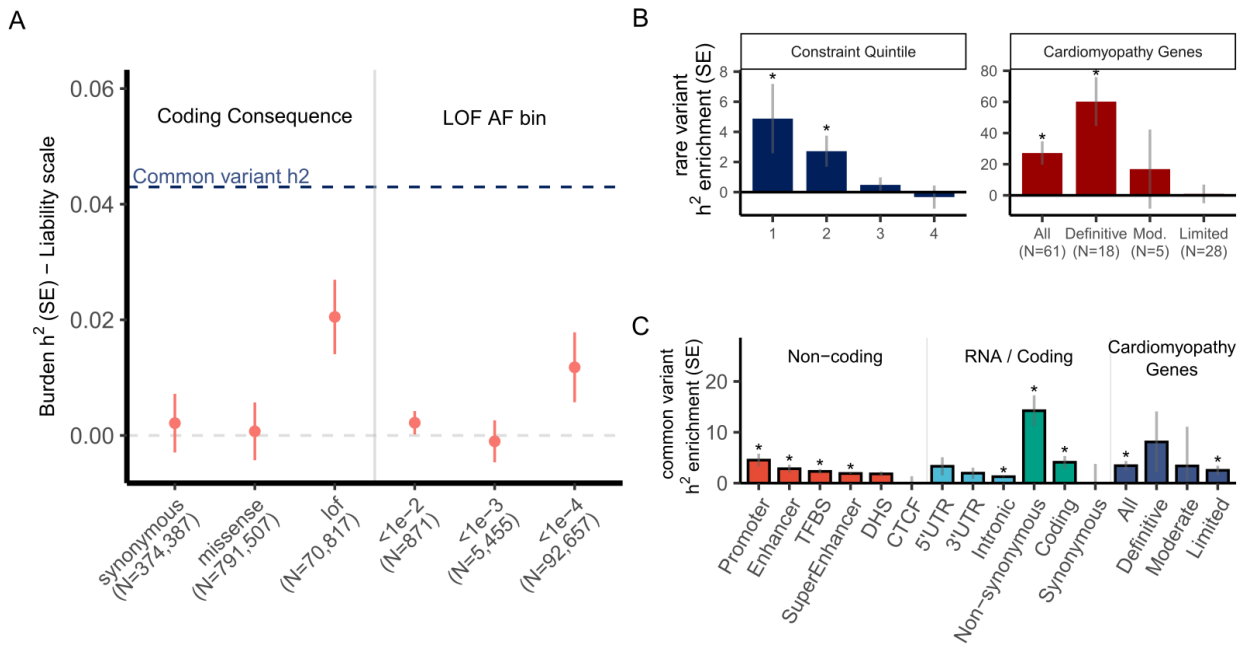

Figure 4: Estimates of rare coding variant burden heritability and comparison of heritability enrichment across several categories for common and rare variants in the PMBB and European all-cause HF meta-analysis. (A) BHR heritability estimate for rare coding variants split by (i) coding consequence and (ii) AF bins in the EUR population. For coding consequence, burden heritability estimates were determined using singletons. The common variant heritability estimate from the EUR GWAS population is represented by the blue dashed line. Burden heritability is significantly concentrated in predicted loss-of-function (LOF) variants and increases with decreasing allele frequency bin. (B) Enrichment of rare variant heritability across (i) constraint quintiles or (ii) expert-reviewed genes involved in cardiomyopathies. (C) Enrichment in common variant heritability across genomic annotations and in genes implicated in cardiomyopathies as determined by stratified LDSC for Non-coding and RNA/Coding annotations and AMM for cardiomyopathy genes.

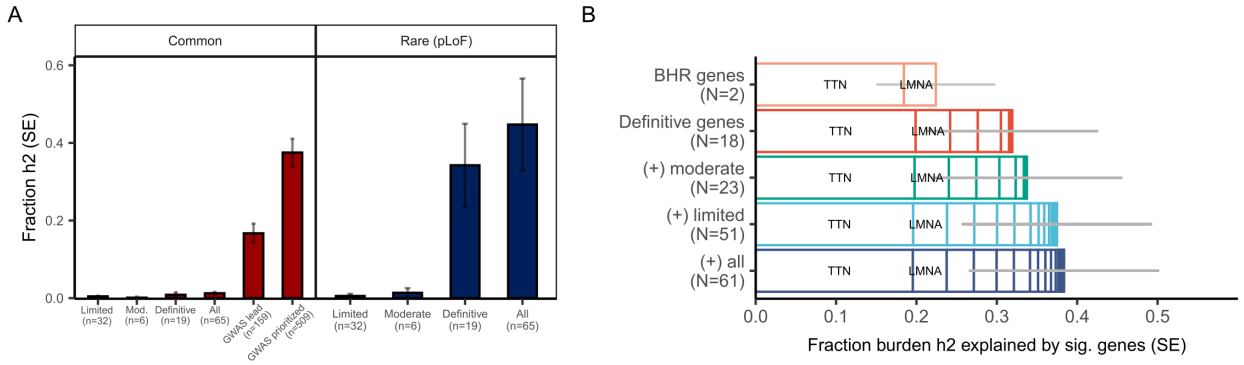

Figure 5: (A) Fraction of common (red) or rare (blue) variant heritability explained by known cardiomyopathy genes in HF for the MVP-Meta EUR population and PMBB respectively. Fraction of common variant heritability was estimated using AMM for all cardiomyopathy gene sets, all protein coding genes nearest to GWAS significant loci (n=159), and all protein coding genes implicated by GWAS significant loci (n=509). (B) Effect of adding known cardiomyopathy genes to the estimate of fraction burden heritability explained by each gene set. BHR genes refers to the set of genes identified as having an exome-wide significant association by BHR (without specifying significant genes in the model). See Suppl. Table 11 for full enrichment statistics.

References

1. Ziaeian, B. & Fonarow, G. C. Epidemiology and aetiology of heart failure. *Nat. Rev. Cardiol.* **13**, 368–378 (2016).
2. Shah, S. *et al.* Genome-wide association and Mendelian randomisation analysis provide insights into the pathogenesis of heart failure. *Nat. Commun.* **11**, 163 (2020).
3. Levin, M. G. *et al.* Genome-wide association and multi-trait analyses characterize the common genetic architecture of heart failure. *Nat. Commun.* **13**, 6914 (2022).
4. 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).
5. Mahajan, A. *et al.* Multi-ancestry genetic study of type 2 diabetes highlights the power of diverse populations for discovery and translation. *Nat. Genet.* 1–13 (2022) doi:10.1038/s41588-022-01058-3.
6. Graham, S. E. *et al.* The power of genetic diversity in genome-wide association studies of lipids. *Nature* **600**, 675–679 (2021).
7. Povysil, G. *et al.* Assessing the Role of Rare Genetic Variation in Patients With Heart Failure. *JAMA Cardiol.* **6**, 379–386 (2021).
8. Dellefave-Castillo, L. M. *et al.* Assessment of the Diagnostic Yield of Combined Cardiomyopathy and Arrhythmia Genetic Testing. *JAMA Cardiol.* **7**, 966–974 (2022).
9. Yengo, L. *et al.* A saturated map of common genetic variants associated with human height. *Nature* **610**, 704–712 (2022).
10. Weiner, D. J., Gazal, S., Robinson, E. B. & O'Connor, L. J. Partitioning gene-mediated disease heritability without eQTLs. *Am. J. Hum. Genet.* **109**, 405–416 (2022).
11. Weiner, D. J. *et al.* Polygenic architecture of rare coding variation across 400,000 exomes. 2022.07.06.22277335 Preprint at https://doi.org/10.1101/2022.07.06.22277335 (2022).

12. Auton, A. *et al.* A global reference for human genetic variation. *Nature* **526**, 68–74 (2015).
13. Clarke, T.-K. *et al.* Genome-wide association study of alcohol consumption and genetic overlap with other health-related traits in UK Biobank (N=112 117). *Mol. Psychiatry* **22**, 1376–1384 (2017).
14. Evangelou, E. *et al.* New alcohol-related genes suggest shared genetic mechanisms with neuropsychiatric disorders. *Nat. Hum. Behav.* **3**, 950–961 (2019).
15. Liu, M. *et al.* Association studies of up to 1.2 million individuals yield new insights into the genetic etiology of tobacco and alcohol use. *Nat. Genet.* **51**, 237–244 (2019).
16. Saunders, G. R. B. *et al.* Genetic diversity fuels gene discovery for tobacco and alcohol use. *Nature* **612**, 720–724 (2022).
17. Hanscombe, K. B. *et al.* The genetic case for cardiorespiratory fitness as a clinical vital sign and the routine prescription of physical activity in healthcare. *Genome Med.* **13**, 180 (2021).
18. Timpson, N. J. *et al.* Genetic variation at the SLC23A1 locus is associated with circulating concentrations of l-ascorbic acid (vitamin C): evidence from 5 independent studies with >15,000 participants. *Am. J. Clin. Nutr.* **92**, 375–382 (2010).
19. Jarick, I. *et al.* Novel common copy number variation for early onset extreme obesity on chromosome 11q11 identified by a genome-wide analysis. *Hum. Mol. Genet.* **20**, 840–852 (2011).
20. Villard, E. *et al.* A genome-wide association study identifies two loci associated with heart failure due to dilated cardiomyopathy. *Eur. Heart J.* **32**, 1065–1076 (2011).
21. Ussher, J. R. *et al.* Inactivation of the Glucose-Dependent Insulinotropic Polypeptide Receptor Improves Outcomes following Experimental Myocardial Infarction. *Cell Metab.* **27**, 450-460.e6 (2018).
22. Chen, J. *et al.* The trans-ancestral genomic architecture of glycemic traits. *Nat. Genet.* **53**, 840–860 (2021).

23. Yengo, L. *et al.* Meta-analysis of genome-wide association studies for height and body mass index in ~700000 individuals of European ancestry. *Hum. Mol. Genet.* **27**, 3641–3649 (2018).
24. Eraslan, G. *et al.* Single-nucleus cross-tissue molecular reference maps toward understanding disease gene function. *Science* **376**, eabl4290 (2022).
25. Oikonomou, E. K. & Antoniades, C. The role of adipose tissue in cardiovascular health and disease. *Nat. Rev. Cardiol.* **16**, 83–99 (2019).
26. Zhou, W. *et al.* SAIGE-GENE+ improves the efficiency and accuracy of set-based rare variant association tests. *Nat. Genet.* **54**, 1466–1469 (2022).
27. McKenna, W. J. & Judge, D. P. Epidemiology of the inherited cardiomyopathies. *Nat. Rev. Cardiol.* **18**, 22–36 (2021).
28. Ingles, J. *et al.* Evaluating the Clinical Validity of Hypertrophic Cardiomyopathy Genes. *Circ. Genomic Precis. Med.* **12**, e002460 (2019).
29. Jordan, E. *et al.* Evidence-Based Assessment of Genes in Dilated Cardiomyopathy. *Circulation* **144**, 7–19 (2021).
30. Park, J. *et al.* A genome-first approach to rare variants in hypertrophic cardiomyopathy genes MYBPC3 and MYH7 in a medical biobank. *Hum. Mol. Genet.* **31**, 827–837 (2022).
31. Pan UKBB | Pan UKBB. https://pan.ukbb.broadinstitute.org/downloads/index.html.
32. Rethy, L. *et al.* Trends in the Prevalence of Self-reported Heart Failure by Race/Ethnicity and Age From 2001 to 2016. *JAMA Cardiol.* **5**, 1425–1429 (2020).
33. Ruan, Y. *et al.* Improving polygenic prediction in ancestrally diverse populations. *Nat. Genet.* **54**, 573–580 (2022).
34. Joseph, J. *et al.* Genetic architecture of heart failure with preserved versus reduced ejection fraction. *Nat. Commun.* **13**, 7753 (2022).
35. Daghlas, I. *et al.* Genetic Evidence for Repurposing of GLP1R (Glucagon-Like Peptide-1 Receptor) Agonists to Prevent Heart Failure. *J. Am. Heart Assoc.* **10**, e020331 (2021).

36. Eli Lilly and Company. *A Randomized, Double-Blind, Placebo-Controlled, Phase 3 Study Comparing the Efficacy and Safety of Tirzepatide Versus Placebo in Patients With Heart Failure With Preserved Ejection Fraction and Obesity (SUMMIT)*. https://clinicaltrials.gov/ct2/show/NCT04847557 (2023).
37. Kim, K.-H. *et al.* Lrrc10 is required for early heart development and function in zebrafish. *Dev. Biol.* **308**, 494–506 (2007).
38. Brody, M. J. *et al.* Ablation of the Cardiac-Specific Gene Leucine-Rich Repeat Containing 10 (Lrrc10) Results in Dilated Cardiomyopathy. *PLOS ONE* **7**, e51621 (2012).
39. Woon, M. T. *et al.* Pediatric Dilated Cardiomyopathy-Associated LRRC10 (Leucine-Rich Repeat–Containing 10) Variant Reveals LRRC10 as an Auxiliary Subunit of Cardiac L-Type Ca2+ Channels. *J. Am. Heart Assoc.* **7**, e006428.
40. Boyle, J. J. *et al.* Activating Transcription Factor 1 Directs Mhem Atheroprotective Macrophages Through Coordinated Iron Handling and Foam Cell Protection. *Circ. Res.* **110**, 20–33 (2012).
41. Fahed, A. C. *et al.* Polygenic background modifies penetrance of monogenic variants for tier 1 genomic conditions. *Nat. Commun.* **11**, 3635 (2020).
42. 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).
43. Choi, S. H. *et al.* Monogenic and Polygenic Contributions to Atrial Fibrillation Risk. *Circ. Res.* **126**, 200–209 (2020).
44. Lam, C. S. P. *et al.* Sex differences in heart failure. *Eur. Heart J.* **40**, 3859–3868c (2019).
45. Willer, C. J., Li, Y. & Abecasis, G. R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190–2191 (2010).
46. Privé, F., Aschard, H., Ziyatdinov, A. & Blum, M. G. B. Efficient analysis of large-scale genome-wide data with two R packages: bigstatsr and bigsnpr. *Bioinformatics* **34**, 2781–2787 (2018).

47. Tidy Verbs for Dealing with Genomic Data Frames. https://const-ae.github.io/tidygenomics/.
48. Graham, S. E. *et al.* The power of genetic diversity in genome-wide association studies of lipids. *Nature* **600**, 675–679 (2021).
49. Mägi, R. *et al.* Trans-ethnic meta-regression of genome-wide association studies accounting for ancestry increases power for discovery and improves fine-mapping resolution. *Hum. Mol. Genet.* **26**, 3639–3650 (2017).
50. McLaren, W. *et al.* The Ensembl Variant Effect Predictor. *Genome Biol.* **17**, 122 (2016).
51. Barbeira, A. N. *et al.* Integrating predicted transcriptome from multiple tissues improves association detection. *PLOS Genet.* **15**, e1007889 (2019).
52. Leeuw, C. A. de, Mooij, J. M., Heskes, T. & Posthuma, D. MAGMA: Generalized Gene-Set Analysis of GWAS Data. *PLOS Comput. Biol.* **11**, e1004219 (2015).
53. Watanabe, K., Taskesen, E., van Bochoven, A. & Posthuma, D. Functional mapping and annotation of genetic associations with FUMA. *Nat. Commun.* **8**, 1826 (2017).
54. Foley, C. N. *et al.* A fast and efficient colocalization algorithm for identifying shared genetic risk factors across multiple traits. *Nat. Commun.* **12**, 764 (2021).
55. Verma, A. *et al.* The Penn Medicine BioBank: Towards a Genomics-Enabled Learning Healthcare System to Accelerate Precision Medicine in a Diverse Population. *J. Pers. Med.* **12**, 1974 (2022).
56. Hard-filtering germline short variants. *GATK* https://gatk.broadinstitute.org/hc/en-us/articles/360035890471-Hard-filtering-germline-short-variants.
57. Ioannidis, N. M. *et al.* REVEL: An Ensemble Method for Predicting the Pathogenicity of Rare Missense Variants. *Am. J. Hum. Genet.* **99**, 877–885 (2016).
58. Karczewski, K. J. *et al.* The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* **581**, 434–443 (2020).
59. Bulik-Sullivan, B. K. *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Genet.* **47**, 291–295 (2015).

60. Bulik-Sullivan, B. *et al.* An atlas of genetic correlations across human diseases and traits. *Nat. Genet.* **47**, 1236–1241 (2015).

61. Finucane, H. K. *et al.* Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nat. Genet.* **47**, 1228–1235 (2015).

62. Huang, C. *et al.* Performance Metrics for the Comparative Analysis of Clinical Risk Prediction Models Employing Machine Learning. *Circ. Cardiovasc. Qual. Outcomes* **14**, e007526 (2021).