Large scale genome-wide association analyses identify novel genetic loci and mechanisms in hypertrophic cardiomyopathy

Rafik Tadros[1,2,3*], Sean L Zheng[4,5,6*], Christopher Grace[7,8], Paloma Jordà[1,2], Catherine Francis[4,6], Sean J Jurgens[3,9], Kate L Thomson[7,10], Andrew R Harper[7,8], Elizabeth Ormondroyd[7,8], Dominique M West[7,8], Xiao Xu[5], Pantazis I Theotokis[4,5,6], Rachel J Buchan[4,5,6], Kathryn A McGurk[4,5], Francesco Mazzarotto[4,11], Beatrice Boschi[12], Elisabetta Pelo[12], Michael Lee[4], Michela Noseda[4], Amanda Varnava[4,13], Alexa MC Vermeer[3,14,15], Roddy Walsh[3], Ahmad S Amin[3,15,16], Marjon A van Slegtenhorst[17], Nicole Roslin[18], Lisa J Strug[19,20,21], Erika Salvi[22], Chiara Lanzani[23,24], Antonio de Marvao[4,5], Hypergenes InterOmics Collaborators, Jason D Roberts[25], Maxime Tremblay-Gravel[1,2], Genevieve Giraldeau[1,2], Julia Cadrin-Tourigny[1,2], Philippe L L'Allier[1,2], Patrick Garceau[1,2], Mario Talajic[1,2], Yigal M Pinto[3,15,16], Harry Rakowski[26], Antonis Pantazis[6], John Baksi[4,6], Brian P Halliday[4,6], Sanjay K Prasad[4,6], Paul JR Barton[4,5,6], Declan P O'Regan[5], Stuart A Cook[5,27,28], Rudolf A de Boer[29], Imke Christiaans[30], Michelle Michels[15,29], Christopher M Kramer[31], Carolyn Y Ho[32], Stefan Neubauer[33], HCMR Investigators, Paul M Matthews[34], Arthur A Wilde[3,15,16,35], Jean-Claude Tardif[1,2], Iacopo Olivotto[36], Arnon Adler[37,38], Anuj Goel[7,8*], James S Ware[4,5,6,39*], Connie R Bezzina[3,15*], Hugh Watkins[7,8*]

[*] These authors contributed equally to this work

Correspondence to:

rafik.tadros@umontreal.ca

j.ware@imperial.ac.uk

c.r.bezzina@amsterdamumc.nl

hugh.watkins@rdm.ox.ac.uk

Authors affiliations

[1]Cardiovascular Genetics Centre, Montreal Heart Institute, Montreal, QC, Canada, [2]Faculty of Medicine, Université de Montréal, Montreal, QC, Canada, [3]Department of Experimental Cardiology, Amsterdam Cardiovascular Sciences, University of Amsterdam, Amsterdam UMC, Amsterdam, the Netherlands, [4]National Heart & Lung Institute, Imperial College London, London, UK, [5]MRC London Institute of Medical Sciences, Imperial College London, London, UK, [6]Royal Brompton & Harefield Hospitals, Guy's and St. Thomas' NHS Foundation Trust, London, UK, [7]Radcliffe Department of Medicine, University of Oxford, Division of Cardiovascular Medicine, John Radcliffe Hospital, Oxford,

UK, [8]Wellcome Centre for Human Genetics, University of Oxford, Oxford, UK, [9]Cardiovascular Disease Initiative, Broad Institute of MIT and Harvard, Cambridge, MA, USA, [10]Oxford Genetics Laboratories, Churchill Hospital, Oxford, UK, [11]Department of Molecular and Translational Medicine, University of Brescia, Brescia, Italy, [12]Genetics Unit, Careggi University Hospital, Florence, Italy, [13]Imperial College Healthcare NHS Trust, Imperial College London, London, UK, [14]Department of Clinical Genetics, University of Amsterdam, Amsterdam UMC, Amsterdam, the Netherlands, [15]European Reference Network for Rare and Low Prevalence Complex Diseases of the Heart, (ERN GUARD-HEART; https://guardheart.ern-net.eu), [16]Department of Clinical Cardiology, Amsterdam Cardiovascular Sciences, University of Amsterdam, Amsterdam UMC, Amsterdam, the Netherlands, [17]Department of Clinical Genetics, Erasmus Medical Center, University Medical Center Rotterdam, Rotterdam, the Netherlands, [18]The Centre for Applied Genomics, Genetics and Genome Biology, The Hospital for Sick Children, Toronto, ON, Canada, [19]Departments of Statistical Sciences and Computer Science, Data Sciences Institute, University of Toronto, Toronto, ON, Canada, [20]The Centre for Applied Genomics, The Hospital for Sick Children, Toronto, ON, Canada, [21]Ontario Regional Centre, Canadian Statistical Sciences Institute, University of Toronto, Toronto, ON, Canada, [22]Neuroalgology Unit, Fondazione IRCCS Istituto Neurologico Carlo Besta, Milan, Italy, [23]Genomics of Renal Diseases and Hypertension Unit, Nephrology Operative Unit, IRCCS San Raffaele Hospital, Milan, Italy, [24]Chair of Nephrology, Vita-Salute San Raffaele University, Milan, Italy, [25]Section of Cardiac Electrophysiology, Division of Cardiology, Department of Medicine, Western University, London, ON, Canada, [26]Peter Munk Cardiac Centre, Toronto, ON, Canada, [27]National Heart Centre Singapore, Singapore, [28]Duke-National University of Singapore Medical School, Singapore, [29]Department of Cardiology, Thoraxcenter, Erasmus University Medical Center, Rotterdam, the Netherlands, [30]Department of Genetics, University of Groningen, University Medical Center Groningen, Groningen, the Netherlands, [31]Department of Medicine, Cardiovascular Division, University of Virginia Health, Charlottesville, VA, USA, [32]Cardiovascular Division, Brigham and Women's Hospital, Boston, MA, USA, [33]Division of Cardiovascular Medicine, Radcliffe Department of Medicine, NIHR Oxford Health Biomedical Research Centre, University of Oxford, Oxford, UK, [34]Department of Brain Sciences and UK Dementia Research Institute, Imperial College London, London, UK, [35]ECGen, Cardiogenetics Focus Group of EHRA, France, [36]Department of Experimental and Clinical Medicine, Meyer Children Hospital, University of Florence, Florence, Italy, [37]Division of Cardiology, Peter Munk Cardiac Centre, University Health Network,

Toronto, ON, Canada, [38]Department of Medicine, University of Toronto, Toronto, ON, Canada, [39]Program in Medical & Population Genetics, Broad Institute of MIT and Harvard, Cambridge, MA, USA

Hypergenes InterOmics Collaborators: Daniele Cusi, Paolo Manunta, Lorena Citterio, Nicola Glorioso

HCMR Investigators: Theodore Abraham, Lisa Anderson, Evan Appelbaum, Camillo Autore, Colin Berry, Elena Biagini, William Bradlow, Chiara Bucciarelli-Ducci, Amedeo Chiribiri, Lubna Choudhury, Andrew Crean, Dana Dawson, Milind Desai, Patrice Desvigne-Nickens, Eleanor Elstein, Andrew Flett, Matthias Friedrich, Nancy Geller, Stephen Heitner, Adam Helms, Daniel Jacoby, Dong-Yun Kim, Han Kim, Bette Kim, Eric Larose, Masliza Mahmod, Heiko Mahrholdt, Martin Maron, Gerry McCann, Saidi Mohiddin, Francois-Pierre Mongeon, Sherif Nagueh, David Newby, Anjali Owens, Sven Plein, Ornella Rimoldi, Michael Salerno, Jeanette Schulz-Menger, Mark Sherrid, Albert van Rossum, Jonathan Weinsaft, James White, Eric Williamson.

Hypertrophic cardiomyopathy (HCM) is an important cause of morbidity and mortality with both monogenic and polygenic components. We here report results from the largest HCM genome-wide association study (GWAS) and multi-trait analysis (MTAG) including 5,900 HCM cases, 68,359 controls, and 36,083 UK Biobank (UKB) participants with cardiac magnetic resonance (CMR) imaging. We identified a total of 70 loci (50 novel) associated with HCM, and 62 loci (32 novel) associated with relevant left ventricular (LV) structural or functional traits. Amongst the common variant HCM loci, we identify a novel HCM disease gene, *SVIL*, which encodes the actin-binding protein supervillin, showing that rare truncating *SVIL* variants cause HCM. Mendelian randomization analyses support a causal role of increased LV contractility in both obstructive and non-obstructive forms of HCM, suggesting common disease mechanisms and anticipating shared response to therapy. Taken together, the findings significantly increase our understanding of the genetic basis and molecular mechanisms of HCM, with potential implications for disease management.

HCM is a disease of cardiac muscle characterized by thickening of the LV wall with an increased risk of arrhythmia, heart failure, stroke and sudden death. Previously viewed as a Mendelian disease with rare pathogenic variants in cardiac sarcomere genes identified in ~35% of cases (HCM_{SARC+}), HCM is now known to have complex and diverse genetic architectures.[1] Prior studies have established that common genetic variants underlie a large portion of disease heritability in HCM not caused by rare pathogenic variants (HCM_{SARC-}) and partly explain the variable expressivity in HCM patients carrying pathogenic variants (HCM_{SARC+}), but such studies had limited power to identify a large number of significant loci.[2,3]

We report a new meta-analysis of 7 case-control HCM GWAS datasets, including 3 not previously published, comprising a total of 5,900 HCM cases, 68,359 controls and 9,492,702 variants with a minor allele frequency (MAF)>1% (**Supplementary Table 1**; Study flowchart in **Figure 1**). Using the conventional genome-wide significance threshold ($P< 5x10^{-8}$), 34 loci were significantly associated with HCM, of which 15 were novel (**Table 1**). We then performed 2 stratified analyses in HCM_{SARC+} (1,776 cases) and HCM_{SARC-} (3,860 cases), and identified an additional 1 locus and 4 loci, respectively (**Table 1**; **Supplementary Table 2**; **Supplementary Figure 1**). Using conditional analysis[4], we identified additional suggestive and independent associations with HCM, HCM_{SARC+}, and HCM_{SARC-} with a false discovery rate (FDR) <1% (**Supplementary Table 3**). A locus on chromosome 11 which includes

31 *MYBPC3,* a well-established disease gene, is associated with HCM and HCM_{SARC+}, but not HCM_{SARC-},
32 implying that this association is tagging known founder pathogenic variants in *MYBPC3*.[2,3] We
33 estimated the heritability of HCM attributable to common genetic variation (h^2_{SNP}) in the all-comer
34 analysis to be 0.17±0.02 using LD score regression (LDSC)[5], and, as expected, found higher estimates
35 (0.25±0.02) using genome-based restricted maximum likelihood (GREML)[6], with higher h^2_{SNP} in
36 HCM_{SARC-} (0.29±0.02) compared to HCM_{SARC+} (0.16±0.04) (**Supplementary Table 4**).

37 Rare variants in sarcomere genes that cause HCM and dilated cardiomyopathy (DCM) are known to
38 have opposing effects on contractility[7] and we previously demonstrated that HCM and DCM GWAS
39 loci similarly overlap, with opposite direction of effect.[3] We leveraged such opposing genomic effects
40 in HCM and DCM to identify additional loci involved in HCM. Bayesian pairwise analysis (GWAS-PW[8])
41 including the present HCM GWAS meta-analysis and a published DCM GWAS[9] identified four genomic
42 regions where the same variant was deemed causal for both diseases with a posterior probability >0.9
43 (**Supplementary Table 5**). In all 4 genomic regions, opposing directional effects were observed in HCM
44 and DCM. The top mapped genes at these loci using OpenTargets[10] were *HSPB7, BAG3*, *CCT8* and *SVIL*.
45 The former 3 loci were associated with HCM at $P<5x10^{-8}$ while the locus mapped to *SVIL* did not reach
46 GWAS significance ($P=4.4x10^{-6}$ in HCM; $P=2.9x10^{-5}$ in DCM; **Figure 2A-B**) and required further evidence
47 to support implication in HCM. *SVIL* encodes supervillin, a large, multi-domain actin and myosin
48 binding protein with multiple muscle and non-muscle isoforms, of which the muscle isoform has
49 known roles in myofibril assembly and Z-disk attachment.[11] SVIL is highly expressed in cardiac,
50 skeletal, and smooth muscle myocytes in the Genotype Tissue Expression (GTEx) v9 single-nuclei RNA
51 sequencing (snRNA-seq) dataset[12], and *SVIL* morpholino knockdown in zebrafish produces cardiac
52 abnormalities.[13] In humans, loss of function (LoF) *SVIL* variants have been associated with smaller
53 descending aortic diameter[14] and homozygous LoF *SVIL* variants have been shown to cause a skeletal
54 myopathy with mild cardiac features (left ventricular hypertrophy).[15] To provide further evidence
55 linking *SVIL* to human HCM and to explore the association of *SVIL* LoF variants with HCM, we
56 performed rare variant burden analysis including 1,845 clinically-diagnosed unrelated HCM cases and
57 37,481 controls. We demonstrate a 10.5-fold (95% CI: 4.1-26.8; $P:2.3x10^{-7}$) excess burden of *SVIL* LoF
58 variants in HCM cases (**Figure 2C-D**; List of annotated variants in **Supplementary Table 6a**). Notably,
59 the excess burden is even greater at 15.3-fold (95% CI: 5.7-41.3; $P:7x10^{-7}$) when restricting the
60 analysis to high confidence LoF variants affecting the predominant SVIL transcript in LV

61 (ENST00000375400) (**Supplementary Table 6b**). In one family, the *SVIL* LoF variant (p.(Gln255*)) was
62 carried by two cousins with HCM (parents deceased), providing some evidence of co-segregation.
63 Taken together, these data support *SVIL* as a novel HCM disease gene.

64 To further maximize locus discovery in HCM, we performed a multi-trait analysis of GWAS (MTAG[16];
65 **Figure 3**). We first completed a GWAS of 10 cardiomyopathy-relevant LV traits in 36,083 participants
66 of the UKB without cardiomyopathy and with available CMR, with machine learning assessment[17] of
67 LV volumes, wall thickness (mean and maximal) and myocardial strain (**Supplementary Table 7**;
68 **Supplementary Figures 2-11**). We discovered 62 loci associated with LV traits (32 novel), of which 30
69 showed association with HCM with nominal significance (P<0.001) and 13 were mapped to genes
70 associated with Mendelian heart disease (**Supplementary Table 8**). LDSC analyses[18] demonstrated
71 high genetic correlations (rg) between LV traits within 3 clusters (contractility, volume and mass) and
72 with HCM (**Figure 3**, **Supplementary Table 9**). Leveraging such correlations, we then performed MTAG
73 with HCM and 3 LV traits including the most correlated trait with HCM from each cluster, namely
74 global circumferential strain (contractility cluster; rg -0.62), LV end-systolic volume (volume cluster; rg
75 -0.48), and the ratio of LV mass to end-diastolic volume (mass cluster; rg 0.63). MTAG resulted in a
76 significant increase in mean χ^2 equivalent to ~29% increase in effective sample size of the HCM GWAS
77 (from 21,725 to 28,106), with an estimated upper bound of the false discovery rate (maxFDR)[16] of
78 0.027. MTAG resulted in a substantial step up in loci discovered, identifying a total of 68 loci
79 associated with HCM at $P<5x10^{-8}$, including 48 that have not been previously published (13 novel ones
80 also identified in the single-trait HCM meta-analysis, and 35 were additionally novel by MTAG) (**Figure**
81 **4**, **Supplementary Table 10**). Two of the 34 loci reaching genome-wide significance in the HCM GWAS
82 were not significant in MTAG (loci mapped to *TRDN*/*HEY2* and *CHPF*). The total number of loci
83 identified in GWAS or MTAG is therefore 70, of which 50 have not been previously published. Notably,
84 the locus mapped to *SVIL* which was uncovered from the GWAS-PW analysis reached genome-wide
85 significance in MTAG ($P=1.1x10^{-8}$). Although it was not possible to test for replication for the 35 novel
86 MTAG loci, a prior study strongly supports the robustness of the HCM-LV traits MTAG approach.[3]

87 MAGMA[19] gene-set analysis identified multiple significant gene sets linked to muscle, contractility and
88 sarcomeric function (**Supplementary Table 11**) and tissue expression analysis pointed to cardiac
89 tissue (LV and atrial appendage, AA), and to a lower degree, other tissues with smooth muscle
90 content, including arterial tissues (**Supplementary Table 12**). Within cardiac tissue, we further

91 explored the contribution of specific cell types in HCM by leveraging available snRNA-seq data from
92 donor human hearts.[20] Using sc-linker[21], we identified significant enrichment of heritability in
93 cardiomyocyte and adipocyte cell types (cardiomyocyte: FDR-adjusted $P=1.8x10^{-6}$; adipocyte: FDR-
94 adjusted $P=3.0x10^{-3}$) and state gene programs (**Supplementary Figure 12**).

95 Prioritization of potential causal genes in HCM MTAG loci was performed using OpenTargets variant
96 to gene (V2G) mapping[10] (**Supplementary Table 13**) and FUMA[22] (**Supplementary Table 14**). Of all
97 prioritized genes, 26 were selected based on concordance in both OpenTargets (top 3 genes per
98 locus) and FUMA, as well as LV specific expression in bulk RNAseq data (GTEx v8) and expression in
99 cardiomyocytes using publicly available snRNA-seq data from a recent study[23] (**Supplementary Figure**
100 **13** and **Supplementary Tables 13-14**). Of those, 7 are known Mendelian cardiomyopathy genes (*PLN,*
101 *FLNC, FHOD3* and *ALPK3* were previously reported[2,3], while *ACTN2, TTN* and *NEXN* are in novel
102 common variant HCM loci). Among the other 19 predominantly LV-expressed genes, 5 are mapped to
103 previously published known HCM loci, while 14 are in novel loci and include genes involved in
104 cardiomyocyte energetics and metabolism (*RNF207*[24], *MLIP*[25]), myocyte differentiation and
105 transcriptional regulation (*MITF*[26], *PROX1*[27], *TMEM182*[28]), myofibril assembly (*SVIL*[11]), and calcium
106 handling and contractility (*PDE3A*[29], *SRL*[30]). Last, a transcriptome-wide association study (TWAS) with
107 S-MultiXcan[31] using the MTAG summary statistics with cardiac tissues (LV and AA) from GTEx V8
108 identified 127 genes significantly associated with HCM at $P<3.7x10^{-6}$ (**Supplementary Table 15**). Of
109 those, 50 were not mapped to MTAG loci using either FUMA or OpenTargets, including *HHATL*
110 ($P=1x10^{-11}$), a gene of uncertain function prioritized based on dominant LV expression, and whose
111 depletion in zebrafish may lead to cardiac hypertrophy.[32]

112 Rare sarcomeric variants associated with HCM have been shown to result in increased contractility,
113 and cardiac myosin-inhibitors attenuate the development of sarcomeric HCM in animal models.[33]
114 Prior data from GWAS and Mendelian randomization (MR) also support a causal association of
115 increased LV contractility with HCM, extending beyond rare sarcomeric variants.[3] Pharmacologic
116 modulation of LV contractility using myosin inhibitors has recently been approved in the treatment of
117 HCM associated with LV obstruction (oHCM)[34,35], but remains of uncertain utility in non-obstructive
118 HCM (nHCM) which represents a significant proportion of the HCM patient population (both HCM_{SARC-}
119 and HCM_{SARC+}) and where no specific therapy currently exists. To further dissect the specific
120 implication of LV contractility in nHCM and oHCM, we performed two-sample MR, testing the causal

121 association of LV contractility as exposure, with HCM, nHCM and oHCM as outcomes. LV contractility
122 was assessed with CMR using a volumetric method (LV ejection fraction, LVEF), and tridimensional
123 tissue deformation methods (i.e. global LV strain in the longitudinal ($strain^{long}$), circumferential
124 ($strain^{circ}$) and radial ($strain^{rad}$) directions). Results from the primary MR inverse variance weighted
125 (IVW) analysis are shown in **Figure 5A** and sensitivity analyses results appear in **Supplementary Table**
126 **16** and **Supplementary Figures 14-15**. Although significant heterogeneity in the exposure–outcome
127 effects are limitations, MR findings support a causal association between increased LV contractility
128 and increased risk for both nHCM and oHCM, with a substantial risk increase of 12-fold and 29-fold
129 per standard deviation increase in $strain^{circ}$, respectively (**Figure 5A**). Altogether, these data suggest
130 that increased contractility is involved in both oHCM and nHCM development, and thus myosin
131 inhibitors currently approved for symptom control in oHCM may also be of clinical benefit in nHCM.
132 Last, we also performed MR analyses exploring whether increased systolic (SBP) and diastolic (DBP)
133 blood pressure, and pulse pressure (PP=SBP-DBP) are causally associated with nHCM and oHCM. As
134 for LV contractility, the causal association of SBP and DBP with HCM^2 extended to both oHCM and
135 nHCM subgroups (**Figure 5B**, **Supplementary Table 16** and **Supplementary Figure 16**), suggesting that
136 lowering blood pressure may be a therapeutic target to mitigate disease progression for both nHCM
137 and oHCM.

138 In conclusion, the large number of new susceptibility loci arising from this work support new
139 inferences regarding disease mechanisms in HCM. With the identification of the role of *SVIL*, we have
140 uncovered further evidence that a subset of genes underlies both monogenic and polygenic forms of
141 the condition. However, this shared genetic architecture does not extend to the core sarcomere
142 genes which cause monogenic HCM; instead, the common variant loci implicate processes outside the
143 myofilament, thereby widening our biological understanding and pointing to the importance of
144 downstream remodeling pathways. These insights have therapeutic implications. The shared
145 mechanistic pathways between obstructive and non-obstructive forms of HCM suggest that the new
146 class of myosin inhibitors may be effective in both settings, while the further exploration of newly
147 implicated loci and pathways may in the future yield new treatment targets.

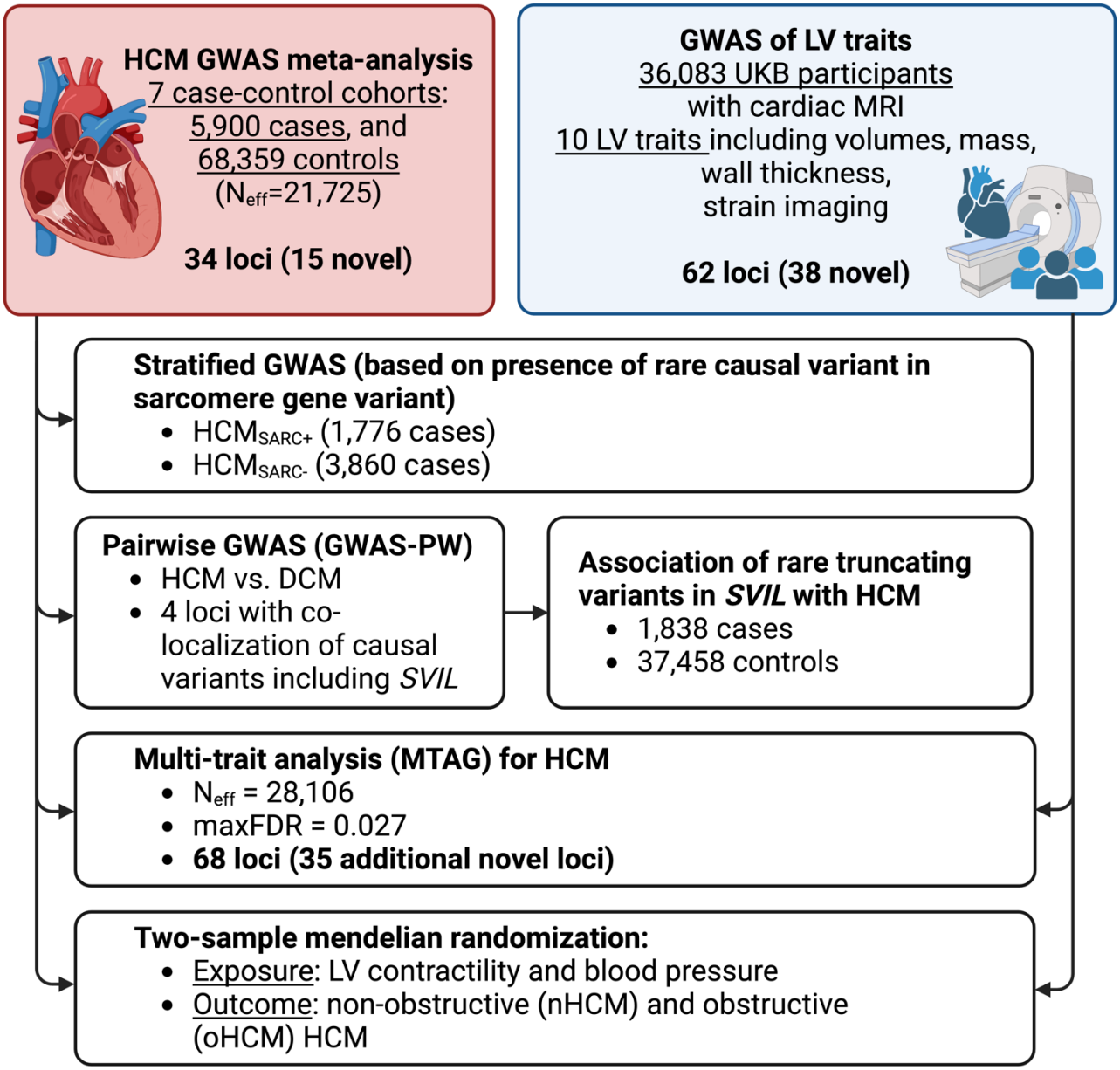
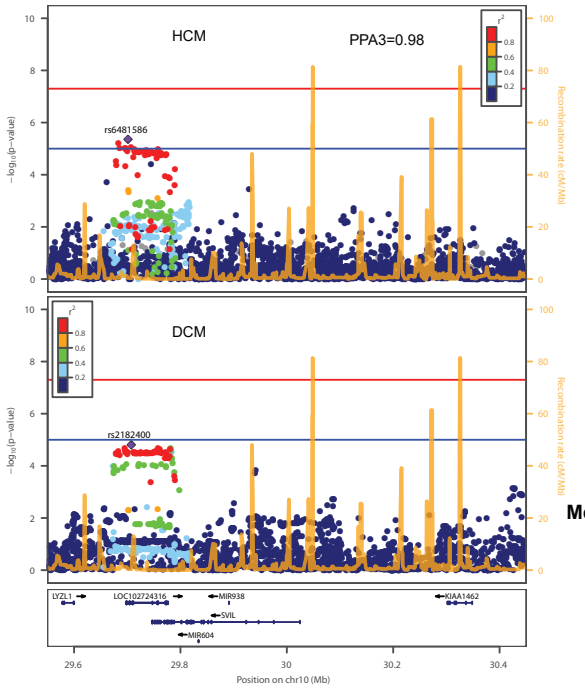


Figure 1: Study Flowchart. Abbreviations: DCM, dilated cardiomyopathy; HCM, hypertrophic cardiomyopathy; LV, left ventricular; maxFDR, upper bound of the estimated false discovery rate computed using MTAG; MRI, magnetic resonance imaging; N_{eff}, effective sample size (see methods); UKB, UK Biobank.

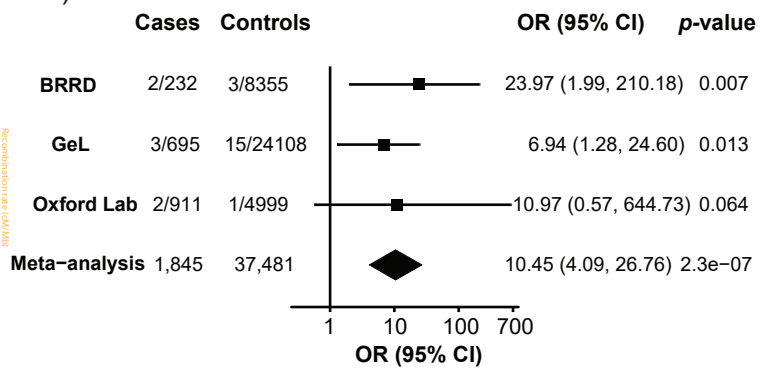
a)



b)

		HCM		DCM	
SNP	**EA/NEA**	**Effect**	***p*-value**	**Effect**	***p*-value**
rs6481586	G/C	0.12	4.40E-06	-0.16	2.86E-05

c)



d)

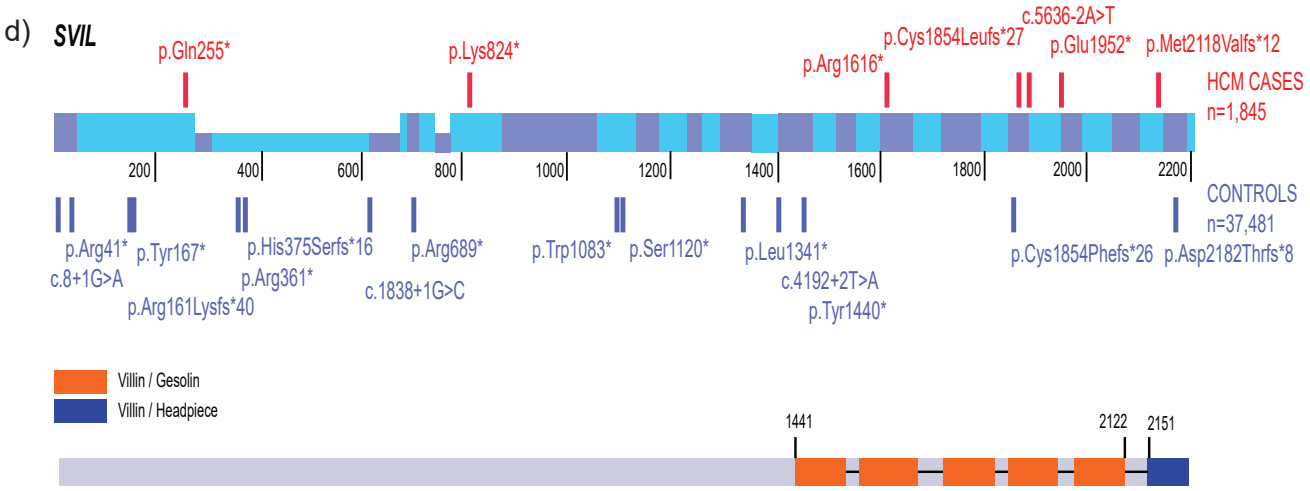


Figure 2: GWAS and rare variant association analyses identify *SVIL* as a novel HCM gene. A) GWAS in HCM and DCM[9] identify a subthreshold locus near *SVIL*. GWAS-PW analysis identifies this locus as sharing the same causal variant (model 3) in both DCM and HCM (posterior probability of model 3, PPA3, 0.98). **B**) Summary statistics of the lead HCM variant (rs6481586) showing effect and non-effect alleles (EA/NEA) and opposite directions of effect (regression coefficient) in HCM and DCM. **C**) Forest plot showing excess of rare loss of function (LoF) variants in *SVIL* in HCM vs. controls in the Rare Disease Bioresource (BRRD), Genomics England (GeL) and Oxford laboratory. **D**) Schematic of the rare LoF *SVIL* variants in HCM cases (top, total N=1,845) and controls (bottom, total N=37,481) along the linear structure of SVIL. The coordinates reflect the codon numbers, and the coloured bars are the exons. The height of the exons reflects expression in cardiac isoforms and is not to scale. Detailed variant annotation appears in Supplementary Table 6.

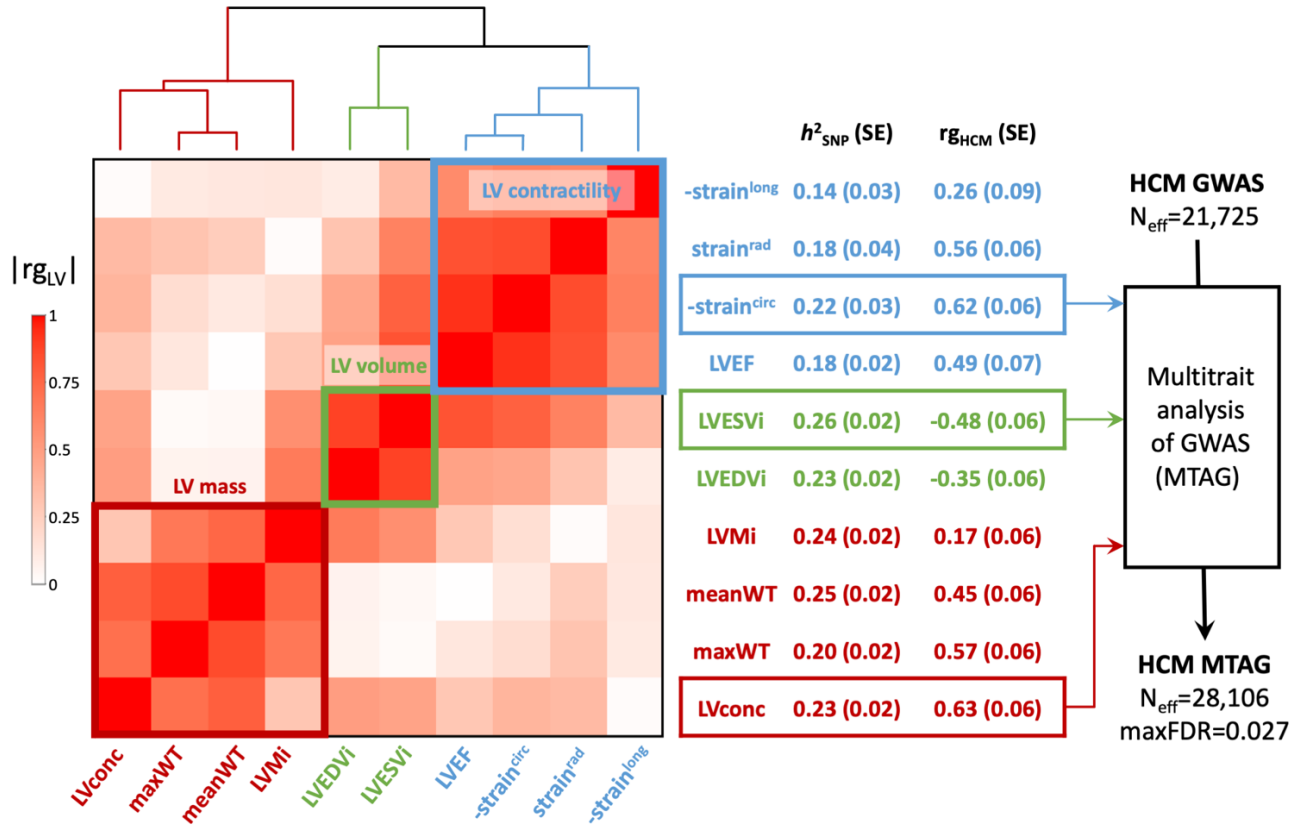


Figure 3: LV traits and HCM genetic correlations and use of MTAG to empower locus discovery. Pairwise genetic correlation between left ventricular (LV) traits shown in heatmap as absolute values ($|rg_{LV}|$) ranging from 0 (white) to 1 (red). LV traits are sorted based on $|rg_{LV}|$ along the x and y axes using Euclidean distance and complete hierarchical clustering into 3 clusters: LV contractility (blue), volume (green) and mass (dark red). See dendrogram on top. The table in the middle shows the individual LV trait common variant heritability (h^2_{SNP}) and genetic correlation with HCM (rg_{HCM}), with corresponding standard errors (SE). The trait with the strongest correlation (based on rg_{HCM}) in each of the 3 clusters was carried forward for multi-trait analysis of GWAS (MTAG) to empower locus discovery in HCM. MTAG resulted in an increase of the effective sample size (N_{eff}, based on number of cases and controls and increase in mean χ^2 statistic) from 21,816 to 28,224, with an estimated upper bound of the false discovery rate (maxFDR) of 0.027. Other abbreviations: LVconc, LV concentricity index (LVM/LVEDV); LVEDVi, LV end-diastolic volume indexed for body surface area; LVEF, LV ejection fraction; LVESVi, LV end-systolic volume indexed for body surface area; LVMi, LV mass indexed for body surface area; maxWT, maximal LV wall thickness; meanWT, mean LV wall thickness; $strain^{circ}$, global LV circumferential strain; $strain^{long}$, global LV longitudinal strain; $strain^{rad}$, global LV radial strain. Note: Since $strain^{circ}$ and $strain^{long}$ are negative values where increasingly negative values reflect increased contractility, we show $-strain^{circ}$ and $-strain^{long}$ to facilitate interpretation rg_{HCM} sign. Full rg_{LV} and rg_{HCM} results are shown in Supplementary Table 9.

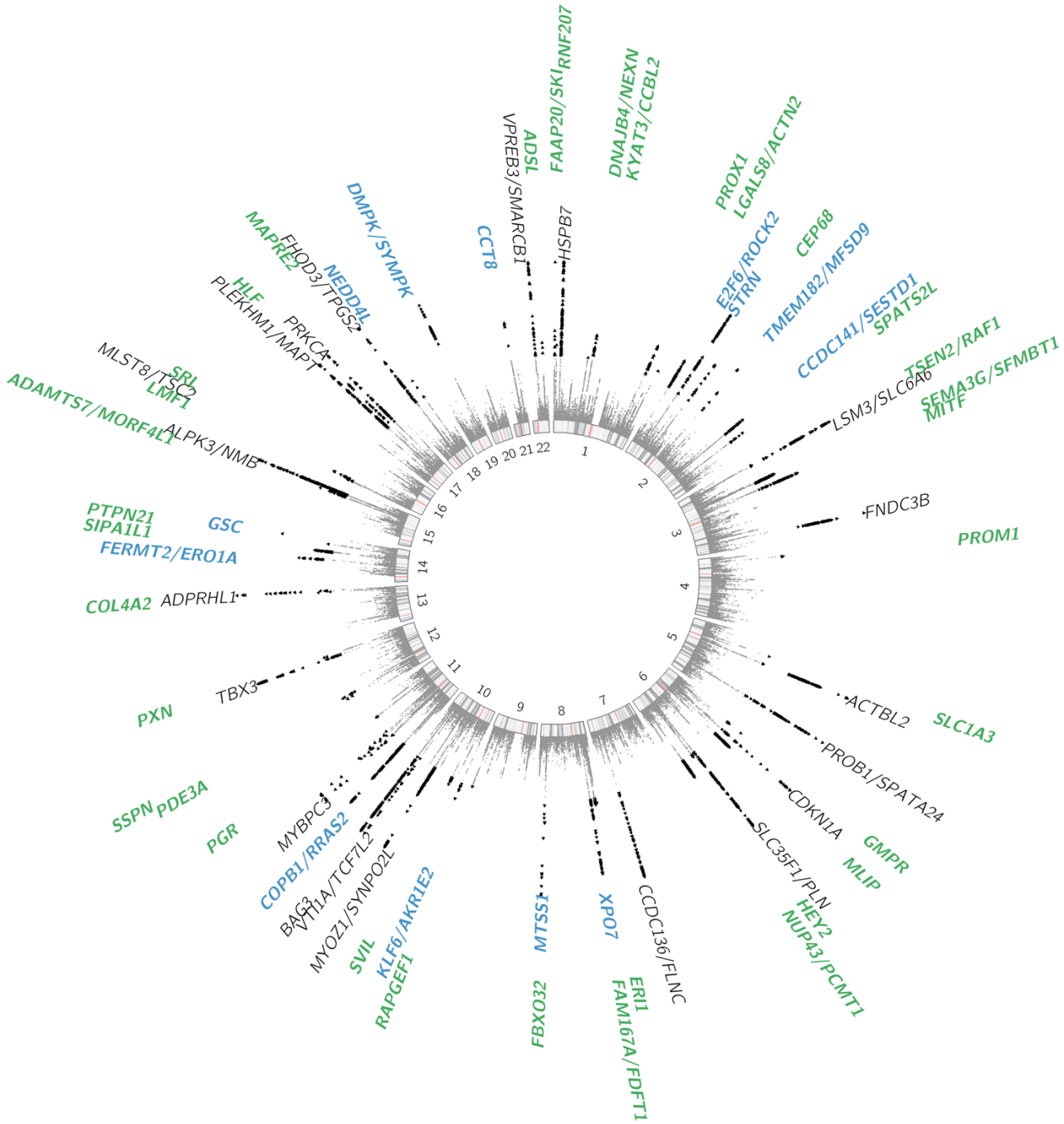


Figure 4: Circular Manhattan plot of HCM summary statistics from MTAG analysis. Previously published loci are identified in black (N=20), novel loci discovered by single trait all-comer GWAS meta-analysis are identified in blue (N=13) and additional novel loci from MTAG are identified in green (N=35). Two other loci reaching GWAS significance threshold in the single trait HCM GWAS meta-analysis but not reaching significance in MTAG are not shown (mapped to *TRDN*/*HEY2* and *CHPF*, see Table 1). Results with $P<1x10^{-15}$ are assigned $P=1x10^{-15}$. Variants with $P<5x10^{-8}$ are shown as black triangles. Locus naming was performed primarily by OpenTargets gene prioritisation considering FUMA and prior gene association with Mendelian HCM. See Supplementary Table 10 for loci details.

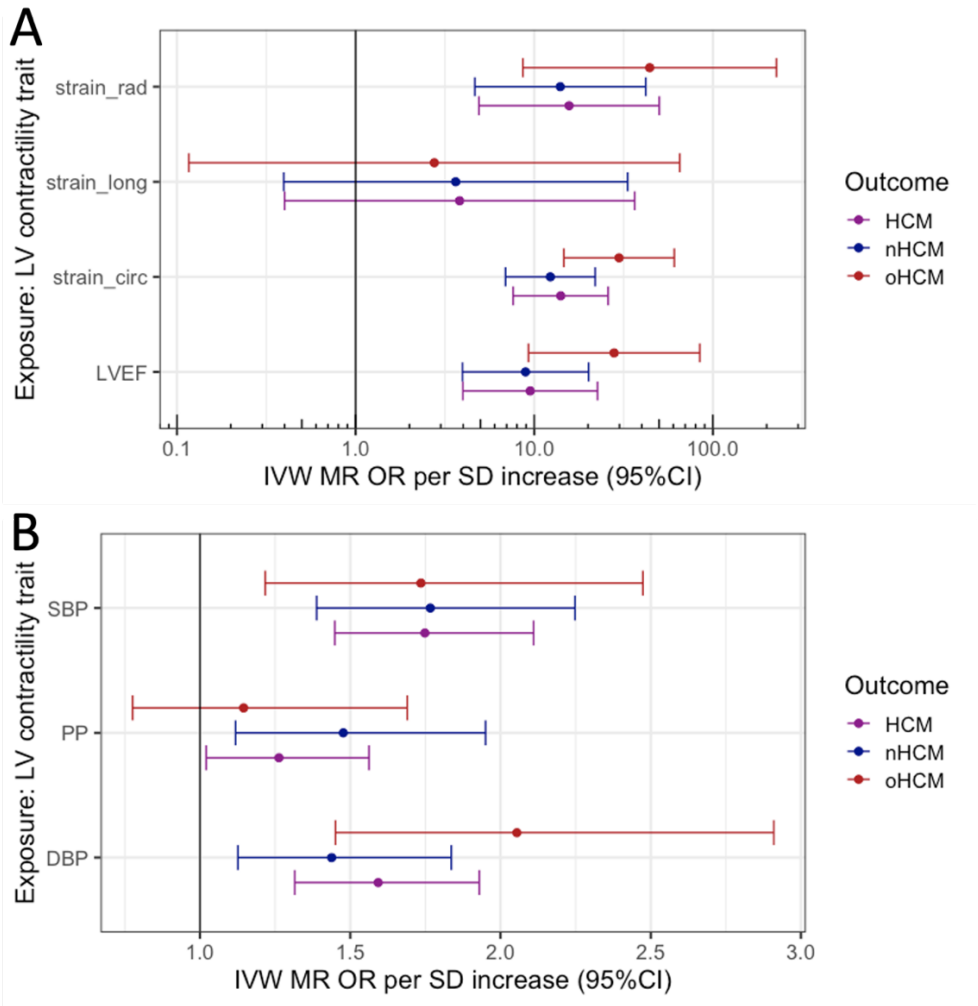


Figure 5: Mendelian randomization (MR) analysis of LV contractility and blood pressure on risk of obstructive (oHCM) and non-obstructive (nHCM) hypertrophic cardiomyopathy (HCM). Odds ratio (OR) represented are those inferred from the inverse variance weighted (IVW) two-sample MR per standard deviation increase (SD). The error bars represent the 95% confidence interval of the OR. **A)** MR suggests causal association of LV contractility (exposure) with HCM, oHCM and nHCM (outcomes), where increased contractility increases disease risk. Genetic instruments for LV contractility were selected from the present GWAS of left ventricular ejection fraction (LVEF), and strain in the radial (stain_rad), longitudinal (stain_long) and circumferential (strain_circ) directions in 36,083 participants of the UKB without cardiomyopathy and with available CMR. To facilitate interpretation of effect directions, OR for strain_circ and strain_long reflect those of increased contractility (more negative strain_circ and strain_long values). The outcome HCM GWAS included 5,927 HCM cases vs. 68,359 controls. Of those, 964 cases and 27,163 controls were included in the oHCM GWAS, and 2,491 cases and 27,109 were included in the nHCM GWAS. Note a logarithmic scale in the x-axis. **B)** MR suggests causal associations of systolic (SBP) and diastolic (DBP) blood pressure with HCM, nHCM and oHCM. Genetic instruments for SBP, DBP and pulse pressure (PP = SBP-DBP) were selected from a published GWAS including up to 801,644 individuals.[36] See Supplementary Table 16 for full MR results.

Table 1: Lead variants from the HCM meta-analysis.

Lead SNP	GRCh37	EA/NEA	EAF	OR (95% CI)	P-value	Locus name	GWS in HCM_{SARC+}	GWS in HCM_{SARC-}
a) Genome wide significant loci from all HCM meta-analysis								
rs2234962	10:121429633	C/T	0.21	1.45 (1.38 - 1.52)	1.39E-49	*BAG3*	•	•
rs2644262	18:34223566	C/T	0.29	1.38 (1.32 - 1.45)	1.79E-43	*FHOD3/TPGS2*	•	•
rs78310129	11:56793878	T/C	0.01	3.53 (2.92 - 4.27)	9.79E-39	*MYBPC3*	•	
rs1048302	1:16340879	T/G	0.33	1.28 (1.23 - 1.34)	8.47E-30	*HSPB7*		•
rs2070458	22:24159307	A/T	0.22	1.30 (1.24 - 1.37)	5.93E-25	*VPREB3/SMARCB1*		•
rs3176326	6:36647289	A/G	0.21	1.30 (1.24 - 1.37)	3.18E-24	*CDKN1A*		•
rs12212795	6:118654308	C/G	0.05	1.51 (1.39 - 1.65)	4.76E-22	*SLC35F1/PLN*		•
rs4577128	17:64308473	C/T	0.57	1.23 (1.18 - 1.29)	3.26E-21	*PRKCA*		•
rs393838	17:43705756	C/G	0.23	1.26 (1.20 - 1.32)	5.02E-21	*CRHR1/MAPT*		•
rs8033459	15:85253258	T/C	0.46	1.20 (1.15 - 1.25)	7.04E-18	*ALPK3/NMB*	•	•
rs11196085	10:114505037	C/T	0.28	1.22 (1.16 - 1.28)	1.85E-17	*VTI1A/TCF7L2*		•
rs7301677	12:115381147	C/T	0.74	1.22 (1.16 - 1.29)	7.01E-16	*TBX3*		•
rs2177843	10:75409877	T/C	0.16	1.26 (1.19 - 1.34)	2.80E-15	*MYOZ1/SYNPO2L*		•
rs41306688	13:114078558	C/A	0.03	1.60 (1.42 - 1.80)	3.04E-15	*ADPRHL1*		•
rs2191445	5:57011469	T/A	0.80	1.23 (1.17 - 1.30)	8.22E-14	*ACTBL2*		•
rs4894803	3:171800256	G/A	0.41	1.18 (1.13 - 1.24)	2.19E-13	*FNDC3B*		•
rs13061705	3:14291129	C/T	0.69	1.19 (1.13 - 1.25)	5.67E-13	*SLC6A6/LSM3*		•
rs13021775	**2:37059557**	**C/G**	**0.50**	**1.17 (1.12 - 1.23)**	**5.98E-13**	***STRN***		•
rs8006225	**14:95219657**	**G/T**	**0.83**	**1.22 (1.15 - 1.30)**	**2.64E-11**	***GSC***		•
rs10052399	5:138668504	T/C	0.27	1.18 (1.12 - 1.24)	3.99E-11	*SPATA24*		
rs66520020	7:128438284	T/C	0.16	1.21 (1.14 - 1.28)	5.87E-11	*CCDC136/FLNC*		
rs12460541	**19:46312077**	**G/A**	**0.66**	**1.16 (1.11 - 1.21)**	**6.01E-11**	***DMPK/SYMPK***		
rs7461129	**8:125861374**	**T/C**	**0.31**	**1.16 (1.11 - 1.21)**	**8.19E-11**	***MTSS1***		
rs56005624	**2:179774634**	**G/T**	**0.14**	**1.21 (1.14 - 1.28)**	**8.31E-11**	***CCDC141/SESTD1***		•
rs7824244	**8:21802432**	**A/G**	**0.14**	**1.22 (1.14 - 1.29)**	**2.39E-10**	***XPO7***	•	
rs12270374	**11:14375079**	**C/T**	**0.36**	**1.14 (1.09 - 1.20)**	**6.85E-10**	***RRAS2/COPB1***		
rs62222424	**21:30530131**	**G/A**	**0.93**	**1.32 (1.20 - 1.44)**	**1.21E-09**	***CCT8***		
rs11687178	**2:11584197**	**C/A**	**0.65**	**1.14 (1.09 - 1.19)**	**7.70E-09**	***E2F6/ROCK2***		
rs9320939	**6:123818871**	**A/G**	**0.49**	**1.13 (1.08 - 1.18)**	**1.04E-08**	***TRDN/HEY2***		•
rs2540277	**2:103426177**	**C/T**	**0.94**	**1.32 (1.19 - 1.45)**	**2.31E-08**	***TMEM182/MFSD9***		
rs6566955	**18:55922789**	**G/A**	**0.31**	**1.14 (1.08 - 1.19)**	**2.93E-08**	***NEDD4L***		
rs13004994	**2:220406239**	**T/G**	**0.46**	**1.13 (1.08 - 1.18)**	**3.02E-08**	***CHPF***		
rs2645210	**10:4098453**	**A/G**	**0.19**	**1.16 (1.10 - 1.23)**	**3.94E-08**	***KLF6/AKR1E2***		
rs113907726	**14:53316867**	**G/T**	**0.19**	**1.16 (1.10 - 1.22)**	**4.10E-08**	***FERMT2/ERO1A***		
b) Additional loci discovered in HCM_{SARC+} or HCM_{SARC-}								
rs9311485	3:52987645	T/G	0.25	1.13 (1.08 - 1.19)	1.86E-07	*ITIH3/SFMBT1*		•
rs77963625	12:46446897	C/T	0.03	1.38 (1.22 - 1.57)	2.97E-07	*SCAF11*		•
rs846111	1:6279370	G/C	0.73	1.14 (1.08 - 1.20)	6.32E-07	*RNF207*		•
rs58747679	12:26348304	T/C	0.71	1.12 (1.07 - 1.18)	1.30E-06	*SSPN*		•
rs112787369	14:68252852	T/A	0.04	1.21 (1.08 - 1.35)	6.04E-04	*ZYVE26*	•	

All reported summary statistics refer to the all HCM case-control meta-analysis results, including for loci identified only in the sarcomere-positive and -negative stratified analyses (HCM_{SARC+} and HCM_{SARC-}). Table sorted increasing order of the all-comer p-value. Novel loci are shown in **bold**. Locus naming was performed primarily by OpenTargets[10] gene mapping, also considering FUMA[22] mapping and prior rare variant associations with HCM.[37] **Abbreviations**: EA/NEA, effect and non-effect alleles; EAF, effect allele frequency; GRCh37, genomic coordinates using the Genome Reference Consortium Human Build 37; GWS, genome-wide significant ($P \leq 5x10^{-8}$); OR (95% CI), odds ratio with 95% confidence interval.

Methods

GWAS of hypertrophic cardiomyopathy

The HCM GWAS included HCM cases and controls from 7 strata: the Hypertrophic Cardiomyopathy Registry (HCMR), a Canadian HCM cohort, a Netherlands HCM cohort, the Genomics England 100K Genome Project (GEL), the Royal Brompton HCM cohort, an Italian HCM cohort and the BioResource for Rare Disease (BRRD) project. Quality control (QC) and association analyses were performed per strata, followed by a meta-analysis. The 7 strata are described in the **Supplementary Note** and in **Supplementary Table 1**. Cases consisted of unrelated patients diagnosed with HCM in presence of unexplained left ventricular (LV) hypertrophy defined as a LV wall thickness (LVWT) >15mm, or >13mm and either presence of family history of HCM or a pathogenic or likely pathogenic genetic variant causing HCM. HCM cases underwent gene panel sequencing as per clinical indications. Variants identified within 8 core sarcomere genes (*MYBPC3*, *MYH7*, *TNNI3*, *TNNT2*, *MYL2*, *MYL3*, *ACTC1* and *TPM1*) were centrally assessed at the Oxford laboratory using the American College of Medical Genetics and Genomics (ACMG) guidelines.[38] HCM cases were dichotomised into sarcomere-positive and sarcomere-negative groups using a classification framework previously reported in Neubauer et al.[39] In addition to the primary all-comer GWAS analyses including all cases with HCM (total of 5,900 cases and 68,359 controls), analyses stratified for sarcomere status in cases and randomly allocated controls were performed, including a total of 1,776 cases vs. 29,414 controls in the sarcomere-positive analysis (HCM_{SARC+}) and 3,860 cases vs. 38,942 controls in the sarcomere-negative analysis (HCM_{SARC-}).

Meta-analyses for the all-comer HCM GWAS was performed on betas and standard errors using GWAMA.[40] We kept variants where meta-analysis came from 2 or more studies and also had a sample size >5,000. Genomic inflation was estimated from the median χ^2 distribution and using HapMap3 European ancestry LD scores using LD Score Regression.[5] All variants were mapped to Genome Reference Consortium Human Build 37 (GRCh37) extrapolated using the 1000 Genome phase 3 genetic maps. A genome wide significant locus was assigned where two variants had a meta-analysis $P<5x10^{-8}$ and were 0.5 cM distance apart. A similar approach was implemented for the HCM_{SARC+} and HCM_{SARC-} stratified analyses which comprised 5 and 7 strata, respectively (the GEL and BRRD strata did not include enough sarcomere-positive HCM cases). Variants were retained where meta-analysis came from 2 or more studies and had sample size >5,000 for sarcomere-negative and >2,500 for

31 sarcomere-positive. The final dataset included 9,492,702 (all comer), 7,614,734 (HCM_{SARC+}) and
32 9,226,079 (HCM_{SARC-}) variants after filtering. The results of the all-comer HCM GWAS meta-analysis
33 and stratified analyses are presented in **Table 1**, **Supplementary Figure 1** and **Supplementary Table 2**.

34 A false discovery rate (FDR) 1% P value cut-off was derived from the all-comer, HCM_{SARC+} and HCM_{SARC-}
35 summary statistics using Simes method (Stata 10.1) and the corresponding P-values were $8.5x10^{-6}$,
36 $1.6x10^{-6}$ and $7.8x10^{-6}$ respectively. Using the 1% FDR P value thresholds, we then performed a
37 stepwise model selection to identify 1% FDR independently associated variants using GCTA.[4] The
38 analysis was performed chromosome wise using default window of 10Mbp, 0.9 collinearity and UKB
39 reference panel containing 60K unrelated European ancestry participants. The results of this
40 conditional analysis are presented in **Supplementary Table 3**.

41 <u>HCM heritability attributable to common variants</u>

42 We estimated the heritability of HCM attributable to common genetic variation (h^2_{SNP}) in the all-
43 comer HCM, as well as HCM_{SARC+} and HCM_{SARC} using LD score regression (LDSC)[5] and genome-based
44 restricted maximum likelihood (GREML)[6]. For LDSC, HapMap3 SNPs were selected from the summary
45 statistics corresponding to HCM, HCM_{SARC+} and HCM_{SARC} meta-analyses. The h^2_{SNP} was computed on
46 the liability scale assuming a disease prevalence of 0.002.[41] Since LDSC tends to underestimate h^2_{SNP},
47 we also estimated h^2_{SNP} using GREML, as previously performed.[2,3] We first computed h^2_{SNP} for HCM,
48 HCM_{SARC+} and HCM_{SARC}, using GREML for each of the largest 3 strata (HCMR, the Canadian HCM cohort
49 and the Netherlands HCM cohort), followed by fixed-effects and random-effects meta-analyses
50 combining all 3 strata. To exclude the contribution of rare founder HCM causing variants, we excluded
51 the *MYBPC3* locus for the Canadian and Netherlands strata and the *TNNT2* locus for the Canadian
52 stratum.[3] The results of h^2_{SNP} analyses are presented in **Supplementary Table 4**.

53 <u>Locus colocalization in dilated (DCM) and hypertrophic cardiomyopathy (HCM)</u>

54 We explored colocalization of HCM and DCM loci using GWAS-PW.[8] The genome was split into 1,754
55 approximately independent regions and the all-comer HCM meta-analysis results were analysed with
56 those of a publicly available DCM GWAS[9] using a Bayesian approach. GWAS-PW fits each locus into
57 one of the 4 models where model 1 is association in only the first trait, model 2 is association in only
58 the second trait, model 3 when the two traits co-localize and model 4 when the genetic signals are
59 independent in the two traits. We considered a locus to show colocalization when either trait

60 harbours a genetic signal with $P<1x10^{-5}$ and the GWAS-PW analysis demonstrates a posterior
61 probability of association for model 3 (PPA3) greater than 0.8. Results of GWAS-PW are presented in
62 **Supplementary Table 5** and **Figure 2** (panels **A** and **B**, for the *SVIL* locus).

63 Association of rare *SVIL* loss of function (LoF) variants with HCM

64 We assessed the association of LoF variants in *SVIL* with HCM in 3 cohorts (BRRD[42], GEL[43] and the
65 Oxford laboratory) followed by a meta-analysis. For BRRD, HCM cases were probands within the bio-
66 resource project HCM. Controls were all remaining individuals within the BRRD projects except for
67 those within the GEL and GEL2 projects (the Genomics England pilot data), since there is overlap of
68 individuals with the GEL analysis in these two projects. For GEL, HCM cases were probands with a
69 primary disease of HCM. Controls were probands without any primary or secondary cardiovascular
70 disease and without any primary and secondary congenital myopathy, since *SVIL* has previously been
71 associated with myopathy.[15] For the Oxford laboratory, cases were clinically diagnosed with HCM and
72 referred for diagnostic panel testing. The control group for the Oxford analysis consisted of 5,000
73 individuals randomly selected from the UK Biobank (UKB), which were all white British and unrelated.
74 They had normal LV volume and function and no clinical diagnosis of cardiomyopathy (HCM or DCM).
75 Genetic variants were identified using next generation sequencing (whole-genome sequencing for
76 BRRD and GEL, panel/exome sequencing for Oxford cases and UKB controls) and annotated using the
77 Ensembl variant effect predictor (VEP).[44] LoF variants in *SVIL* were defined as those with the following
78 VEP terms: stop lost, stop gained, splice donor variant, splice acceptor variant and frameshift variant.
79 Only variants with a $MAF<10^{-4}$ in the non-Finish European ancestral group of gnomAD v2.1.1[45] were
80 selected. Only LoF variants present in the Matched Annotation from NCBI and EMBL-EBI (MANE) /
81 canonical transcript (NM_021738.3; ENST00000355867.9) were retained for the analysis. The
82 proportion of cases and controls with *SVIL* LoF variants were compared using the Fisher Exact test for
83 each of the 3 case-control datasets, followed by a fixed-effect model meta-analysis. We also
84 performed a secondary analysis where association of *SVIL* LoF variants with HCM was restricted to
85 variants that cause LoF in the primary LV transcript (ENST00000375400), and excluding variants
86 expected to escape nonsense-mediated decay. The results of *SVIL* LoF variant association with HCM
87 are shown in **Figure 2C**, and the list of *SVIL* LoF variants identified in cases and controls is shown in
88 **Figure 2D** with annotation in **Supplementary Table 6a**. Results of the secondary analysis restricted to
89 high confidence LoF variants are shown in **Supplementary Table 6b**.

90 GWAS of cardiac magnetic resonance-derived left ventricular traits

91 *UK Biobank (UKB) study population*. The UKB is an open-access population cohort resource that has
92 recruited half a million participants in its initial recruitment phase, from 2006-2010. At the time of
93 analysis, CMR imaging data was available from 39,559 individuals in the imaging substudy. The UKB
94 CMR acquisition protocol has been described previously.[46] In brief, images were acquired according to
95 a basic cardiac imaging protocol using clinical 1.5 Tesla wide bore scanners (MAGNETOM Aera, Syngo
96 Platform VD13A, Siemens Healthcare, Erlangen, Germany) in three separate imaging centers.
97 Extensive clinical and questionnaire data and genotypes are available for these individuals. Clinical
98 data were obtained at the time of the imaging visit. These included sex (31), age (21003), weight
99 (21002), height (50), SBP (4080), DBP (4079), self-reported non-cancer illness code (20002), and ICD10
100 codes (41270). The mean age at the time of CMR was 63 ± 8 (range 45-80), and 46% of participants
101 were male. Cohort anthropometrics, demographics and comorbidities are reported in **Supplementary**
102 **Table 7**. Exclusion criteria for the UKB imaging substudy included childhood disease, pregnancy and
103 contraindications to MRI scanning. For the current analysis, we also excluded, by ICD-10 code and/or
104 self-reported diagnoses, any subjects with heart failure, cardiomyopathy, a previous myocardial
105 infarction, or structural heart disease. After imaging quality control and exclusions for comorbidities
106 or genotype quality control, we had a maximum cohort size of 36,083 individuals. The UKB received
107 National Research Ethics Approval (REC reference 11/NW/0382). The present study was conducted
108 under terms of UKB access approval 18545.

109 *LV trait phenotyping*. Description of CMR image analysis has previously been published[3] and is
110 detailed in the **Supplementary Note**. We included ten LV phenotypes for GWAS analyses: end-
111 diastolic volume (LVEDV), end-systolic volume (LVESV), ejection fraction (LVEF), mass (LVM),
112 concentricity index (LVconc = LVM/LVEDV), mean wall thickness (meanWT), maximum wall thickness
113 (maxWT) as well as global peak strain in radial ($strain^{rad}$), longitudinal ($strain^{long}$) and circumferential
114 ($strain^{circ}$) directions. The means and standard deviations of all ten LV phenotypes, overall and
115 stratified by sex, are shown in **Supplementary Table 7**.

116 *LV trait genome-wide association analyses.* A description of genotyping, imputation and QC appears in
117 the **Supplementary Note**. The GWAS model for LVEF, LVconc, meanWT, maxWT, $strain^{rad}$, $strain^{long}$
118 and $strain^{circ}$ included age, sex, mean arterial pressure (MAP), body surface area (BSA, derived from
119 the Mosteller formula) and the first eight genotypic principal components as covariates. LVEDV,

LVESV and LVM were indexed to body surface area for the analysis, as commonly performed in clinical practice. For indexed values (LVEDVi, LVESVi, LVMi), the GWAS model did not include BSA as a covariate, but all other covariates were the same as for non-indexed phenotypes. BOLT-LMM (v2.3.2)[47] was used to construct mixed models for association with around 9.5 million directly genotyped and imputed SNPs. A high-quality set of directly genotyped model SNPs was selected to account for random effects in the genetic association analyses. These were selected by MAF (>0.001), and LD-pruned (r^2 <0.8) to create an optimum SNP set size of around 500,000. The model was then applied to the > 9.8 million imputed SNPs passing quality control and filtering. Results of the LV traits GWAS are shown in **Supplementary Table 8** and **Supplementary Figures 2-11**.

Locus definition and annotation. Genomic loci associated with all LV traits were annotated jointly. Specifically, summary statistics were combined and a P value corresponding to the minimal P value (minP) across all 10 summary statistics. The minP summary statics was then used to define loci using FUMA v1.4.2[22] using a maximum lead SNP P-value of $5x10^{-8}$, maximum GWAS P-value of 0.05 and r^2 threshold for independent significant SNPs of 0.05 (using the European 1000 Genomes Project dataset), and merging LD blocks within 250kb. Loci were then mapped to genes using positional mapping (<10kb), eQTL mapping using GTEx v8 restricted to atrial appendage, left ventricle and skeletal muscle tissues, and chromatin interaction mapping using left and right ventricles. See FUMA tutorial for detailed methods. Genes mapped using FUMA were further prioritized by querying the Clinical Genomes Resource (ClinGen)[48] for genes linked to Mendelian heart disease with moderate, strong or definitive evidence, and using a recent review of overlapping GWAS and Mendelian cardiomyopathy genes.[37] In addition to FUMA locus to gene mapping, we also report closest gene and top gene mapped using OpenTargets.[10] Annotated LV trait loci are shown in **Supplementary Table 8**.

<u>Genetic correlations between HCM and LV traits</u>

Pairwise genetic correlations for HCM and the 10 LV traits were assessed using LD score regression (LDSC, v.1.0.1).[18] The analysis was restricted to well-imputed non-ambiguous HapMap3 SNPs, excluding SNPs with MAF<0.01 and those with low sample size, using default parameters. We then assessed genetic correlations for each of the 55 pairs (HCM and 10 LV traits) using precomputed LD scores from the European 1000 Genomes Project dataset. We did not constrain the single-trait and cross-trait LD score regression intercepts. The results of the genetic correlation analyses are shown in **Figure 3** and **Supplementary Table 9**.

150 Multitrait analysis of GWAS (MTAG)

151 We performed multi-trait analysis of GWAS summary statistics using MTAG (v.1.0.8)[16] to increase
152 power for discovery of genetic loci associated with HCM. MTAG jointly analyzes multiple sets of GWAS
153 summary statistics of genetically correlated traits to enhance statistical power. Due to high
154 computation needs to calculate the maximum false discovery rate (maxFDR) with MTAG, we limited
155 the number of GWAS summary statistics to 4 (HCM + 3 LV traits). The 3 LV traits to include were
156 selected as follows. First, we performed hierarchical clustering of the 10 LV traits using the absolute
157 value of the pairwise genetic correlations, Euclidean distance and the complete method, predefining
158 the number of clusters to 3. This resulted in clustering of LV traits into a LV contractility cluster (LVEF,
159 $strain^{rad}$, $strain^{long}$ and $strain^{circ}$), a LV volume cluster (LVEDVi, LVESVi) and a LV mass cluster (LVMi,
160 LVconc, meanWT, maxWT) (**Figure 3**). We then selected the trait with the highest genetic correlation
161 with HCM for each cluster ($strain^{circ}$, LVESVi and LVconc) to include in MTAG together with HCM. Only
162 SNPs included in all meta-analyses (that is HCM and LV traits) were used in MTAG. The
163 coded/noncoded alleles were aligned for all 4 studies before MTAG, and multi-allelic SNPs were
164 removed. All summary statistics refer to the positive strand of GRCh37 and, as such,
165 ambiguous/palindromic SNPs (having alleles A/T or C/G) were not excluded. Regression coefficients
166 (beta) and their standard errors were used as inputs for MTAG. The maxFDR was calculated as
167 suggested by the MTAG developers.[16] MaxFDR calculates the type I error in the analyzed dataset for
168 the worst-case scenario. We estimated the gain in statistical power by the increment in the effective
169 sample size (N_{eff}). The N_{eff} for the HCM GWAS was calculated using the following formula.[16,49]

170 $$N_{eff(GWAS)} = \frac{4}{Ncases^{-1} + Ncontrols^{-1}}$$

171 The N_{eff} for the HCM MTAG was estimated by means of the fold-increase in mean χ^2, using the
172 following formula.[16]

173 $$N_{eff(MTAG)} = N_{eff(GWAS)} \times \left({\chi^2}_{MTAG,mean} - 1 \Big/ {\chi^2}_{GWAS,mean} - 1 \right)$$

174 The MTAG N_{eff} corresponds to the approximate the sample size needed to achieve the same mean χ^2
175 value in a standard GWAS. The results of HCM MTAG are presented in **Figure 4** and **Supplementary**
176 **Table 10**.

177

178 Genome-wide annotation

179 Genome-wide analyses following MTAG were performed using MAGMA v.1.08, as implemented in
180 FUMA[22], including gene-set and tissue expression analyses. We used Gene Ontology (GO) gene sets
181 from the Molecular Signatures Database (MSigDB, v.6.2) for the gene-set analysis and the Genotype-
182 Tissue Expression project (GTEx, v.8) for the tissue specificity analysis. The results of MAGMA analyses
183 are shown in **Supplementary Table 11** (gene-set analyses) and **Supplementary Table 12** (tissue
184 specificity analyses).

185 Cardiac cell type heritability enrichment analysis

186 Gene programs derived from single nuclei RNA sequencing (snRNA-seq) were used to investigate
187 heritability enrichment in cardiac cell types and states using the sc-linker framework.[21] This approach
188 uses snRNA-seq data to generate gene programs that characterize individual cell types and states.
189 These programs are then linked to genomic regions and the SNPs that regulate them by incorporating
190 Roadmap Enhancer-Gene Linking[50,51] and Activity-by-Contact models[52,53]. Finally, the disease
191 informativeness of resulting SNP annotations is tested using stratified LD score regression (S-LDSC)[54]
192 conditional on broad sets of annotations from the baseline-LD model.[55,56] Cell type and state-specific
193 gene programs were generated from snRNA-seq data of ventricular tissue from 12 control subjects,
194 with cell type and state annotations made as part of a larger study of ~880,000 nuclei (samples from
195 61 DCM and 12 control subjects.[20] Cell states that may not represent true biological states (for
196 example, technical doublets) were excluded from analysis. Results of sc-linker cardiac cell type
197 heritability enrichment analysis are shown in **Supplementary Figure 12**.

198 Locus to gene annotation

199 A genome wide significant HCM MTAG loci was assigned where two variants had a MTAG $P<5x10^{-8}$
200 and were 0.5 cM distance apart, as performed for the HCM GWAS. Prioritization of potential causal
201 genes in HCM MTAG loci was performed using OpenTargets variant to gene (V2G) mapping[10] and
202 FUMA[22]. The lead SNP at each independent locus was used as input for OpenTargets V2G using the
203 release of October 12th, 2022. Locus to gene mapping with FUMA v1.3.7 was performed based on 1)
204 position (within 100kb), 2) eQTL associations in disease-relevant tissues (GTEx V8 left ventricle, atrial
205 appendage and skeletal muscle) and 3) chromatin interactions in cardiac tissue (left ventricle and
206 right ventricle, FDR $P<10^{-6}$).

207 We further annotated genes mapped using OpenTargets and/or FUMA with their implication in
208 mendelian cardiomyopathy. Specifically, we queried the Clinical Genome Resource (ClinGen[48,57]) for
209 genes associated with any cardiomyopathy phenotype with a level of evidence of moderate, strong or
210 definitive and included genes with robust recent data supporting an association with Mendelian
211 cardiomyopathy.[37]

212 We also prioritized genes based on RNA expression data from bulk tissue RNAseq data in the GTEx[58]
213 v8 dataset accessible at the GTEx portal and snRNA-seq data from Chaffin et al[23] accessible through
214 the Broad Institute single cell portal (singlecell.broadinstitute.org). Using the GTEx v8 data, we
215 assessed specificity of LV expression by computing the ratio of median LV transcripts per million
216 (TPM) to the median TPM in other tissues excluding atrial appendage and skeletal muscle and
217 averaging tissue within types (e.g., all arterial tissues, all brain tissues, etc.). High and Mid LV
218 expression specificity were empirically defined as >10-fold and >1.5-fold LV to other tissues median
219 TPM ratios, respectively. Using snRNA-seq data from Chaffin et al[23], we report the expression in the
220 cardiomyocyte_1 cell type using scaled mean expression (relative to each gene's expression across all
221 cell types) and percentage of cells expressing. High and Mid expression in cardiomyocytes were
222 empirically defined as percentage expressing cells ≥80% and 40-80%, respectively. Prioritized genes
223 were defined as genes mapped using both OpenTargets (top 3 genes) and FUMA, AND had either 1)
224 High LV specific expression, OR 2) High cardiomyocyte expression, OR 3) both Mid LV specific
225 expression and Mid cardiomyocyte expression.

226 Gene mapping data including ClinGen cardiomyopathy genes at HCM loci, LV expression specificity
227 and cardiomyocyte expression are shown in **Supplementary Table 13** (OpenTargets genes) and
228 **Supplementary Table 14** (FUMA genes). Prioritized genes are illustrated in **Supplementary Figure 13**.

229 <u>Transcriptome-wide association study (TWAS)</u>

230 We used MetaXcan to test the association between genetically predicted-gene expression and HCM
231 using summary results from MTAG analysis.[31,59] Biologically-informed MASHR-based prediction
232 models of gene expression for left ventricle (LV) and atrial appendage (AA) tissue from GTEx v8[60] were
233 analysed individually with S-PrediXcan[59], and then analysed together using S-MultiXcan.[31] GWAS
234 MTAG summary statistics were harmonised and imputed to match GTEx v8 reference variants present
235 in the prediction model. To account for multiple testing, TWAS significance was adjusted for the total

236 number of genes present in S-MultiXcan analysis (13,558 genes, $P=3.7x10^{-6}$). TWAS results are shown
237 in **Supplementary Table 15**.

238 Two-sample Mendelian randomization (MR)

239 We assessed whether increased contractility and blood pressure are causally linked to increased risk
240 of HCM globally and its obstructive (oHCM) and non-obstructive (nHCM) forms using two-sample MR.
241 LV contractility and blood pressure parameters were used as exposure variables, and HCM, oHCM and
242 nHCM as outcomes. Analyses were performed using the TwoSampleMR (MRbase) package[61] (v.0.5.6)
243 in R (v.4.2.0). Four exposure variables corresponding to measures of LV contractility were used
244 separately: LVEF as a volumetric marker of contractility, and global strain ($strain^{circ}$, $strain^{rad}$ and
245 $strain^{long}$) as contractility markers based on myocardial tissue deformation. Instrument SNPs for
246 contractility were selected based on the LV trait GWAS presented here using a P value threshold of
247 $<5x10^{-8}$. Only independent SNPs (using $r^2<0.01$ in the European 1000 Genomes population) were
248 included. Instrument SNPs for the blood pressure analysis were selected with a similar approach using
249 a published blood pressure GWAS.[36] The outcome summary statistics were those of the single-trait
250 HCM case-control meta-analysis (5,927 cases and 68,359 controls). We also performed a GWAS meta-
251 analysis including data from HCMR and the Canadian HCM cohort (**Supplementary Table 1**) for nHCM
252 (2,491 cases and 27,109 controls) and oHCM (964 cases and 27,163 controls) to use as outcomes. For
253 these stratified analyses, oHCM was defined as HCM in presence of a LV outflow tract gradient
254 ≥30mmHg at rest or during Valsalva/exercise at any time point. All other HCM cases were considered
255 nHCM. Notably, nHCM and oHCM show high genetic correlation (rg=0.87 with standard error,
256 SE=0.13; $P=4x10^{-11}$), suggesting a substantially shared genetic basis.

257 Insertions/deletions and palindromic SNPs with intermediate allele frequencies (MAF>0.42) were
258 excluded, and other SNPs in the same locus were included only if $P<5x10^{-8}$. An inverse variance
259 weighted MR model was used as a primary analysis. We used three additional methods as sensitivity
260 analyses: weighted median, weighted mode and MR Egger. Cochran's Q statistics were calculated to
261 investigate heterogeneity between SNP causal effects using IVW. Evidence of directional pleiotropy
262 was also assessed using the MR Egger intercept. Mean F-statistics were calculated to assess the
263 strength of the genetic instruments used. Leave-one-out analyses were also performed to ensure the
264 SNP causal effects are not driven by a particular SNP. The summary results of MR analyses are shown
265 in **Figure 5** and **Supplementary Table 16**, with effect plots shown in **Supplementary Figures 14**

266 (contractility) and **Supplementary Figure 16** (blood pressure), and leave-one-out analyses for the
267 contractility MR in **Supplementary Figure 15**. The MR effects are shown per unit change (% for
268 contractility; mmHg for blood pressure) in **Supplementary Table 16** and **Supplementary Figures 14-**
269 **16**, and per SD change in **Figure 5**. OR per SD increase are calculated as follows $OR = e^{\beta_{MR} \times SD}$. SDs
270 are reported in **Supplementary Table 16** and correspond to those in the current UKB CMR dataset (for
271 contractility) and those reported by Evangelou et al[36] in the UK Biobank (for blood pressure).

272 **Data availability**

273 Data from the Genome Aggregation Database (gnomAD, v.2.1.1) are available at
274 https://gnomad.broadinstitute.org. Data from the UKB can be requested from the UKB Access
275 Management System (https://bbams.ndph.ox.ac.uk). Data from the GTEx consortium are available at
276 the GTEx portal (https://gtexportal.org). Published snRNA-seq data are available at the Broad Single
277 Cell Portal (singlecell.broadinstitute.org) and at the Cellxgene tool website
278 (https://cellxgene.cziscience.com/collections/e75342a8-0f3b-4ec5-8ee1-245a23e0f7cb/private).
279 Other datasets generated during and/or analyzed during the current study can be made available
280 upon reasonable request to the corresponding authors. Individual level data sharing is subject to
281 restrictions imposed by patient consent and local ethics review boards. Summary statistics of GWAS
282 and MTAG will be made available in the GWAS catalog upon publication following peer-review, and
283 interactive Manhattan and regional plots will be made available at www.well.ox.ac.uk/hcm.

284 **Acknowledgements**

285 This work was supported by funding from the British Heart Foundation (BHF, RG/18/9/33887,
286 RE/18/4/34215, FS/15/81/31817); the National Heart, Lung, and Blood Institute (NIH grant
287 U01HL117006-01A1); the Wellcome Trust (201543/B/16/Z, 107469/Z/15/Z, 200990/A/16/Z); the
288 Wellcome Trust core awards (090532/Z/09/Z, 203141/Z/16/Z); the National Institute for Health
289 Research (NIHR) Oxford Biomedical Research Centre; the NIHR Imperial College Biomedical Research
290 Centre; NIHR Royal Brompton Cardiovascular Biomedical Research Unit; Sir Jules Thorn Charitable
291 Trust [21JTA]; the Medical Research Council (MRC, UK); the Dutch Heart Foundation (CVON 2018-30
292 PREDICT2); the Horstingstuit Foundation; the Montreal Heart Institute Foundation; the Philippa and
293 Marvin Carsley Cardiology Chair; the Fonds de la Recherche du Québec-Santé (254616, 265449); the
294 Canadian Institutes for Health Research (CIHR, 428321). For the purpose of open access, the authors
295 have applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from
296 this submission. The views expressed in this work are those of the authors and not necessarily those
297 of the funders.

298 R.T. holds the Canada Research Chair in translational cardiovascular genetics. S.L.Z. received support
299 from BHF Centre of Research Excellence Clinical Research Fellowship (RE/18/4/34215). C.F. received
300 funding from a BHF Clinical Research Training Fellowship (FS/15/81/31817). A.R.H. received support
301 from the MRC Doctoral Training Partnership. X.X. is currently a post-doc scientist funded by MRC-

302 London Institute of Medical Sciences. R.W. received support from an Amsterdam Cardiovascular
303 Sciences fellowship. M.T. receives support from Monat Foundation. S.J.J. was supported by a Junior
304 Clinical Scientist Fellowship (03-007-2022-0035) from the Dutch Heart Foundation, and by an
305 Amsterdam UMC Doctoral Fellowship. Y.M.P. receives support from the Dutch Heart Foundation
306 (CVON PRIME). B.P.H. is funded by the BHF Intermediate Fellowship (FS/ICRF/21/26019). D.P.O'R. is
307 supported by the MRC (MC_UP_1605/13) and BHF (RG/19/6/34387). R.A.d.B. is supported by the
308 Dutch Heart Foundation (2020B005), by the Leducq Foundation (Cure-PLaN), and by the European
309 Research Council (ERC CoG 818715). I.C. receives support from the Dutch Heart Foundation (CVON
310 2015-12 eDETECT). P.M.M. received funding from the Edmond J. Safra Foundation and Lily Safra and
311 an NIHR Senior Investigator Award, the UK Dementia Research Institute, which receives its funding
312 from UK DRI Ltd., funded by the MRC, Alzheimer's Society, and Alzheimer's Research UK and the
313 Imperial College British Heart Foundation Centre of Excellence. J.-C.T. holds the Canada Research
314 Chair in personalized medicine and the University of Montreal endowed research chair in
315 atherosclerosis, and he is also the principal investigator of the Montreal Heart Institute André and
316 France Desmarais hospital cohort funded by the Montreal Heart Institute Foundation. A.G. has
317 received support from the BHF, European Commission [LSHM-CT- 2007-037273, HEALTH-F2-2013-
318 601456] and TriPartite Immunometabolism Consortium [TrIC]- NovoNordisk Foundation
319 [NNF15CC0018486], BHF-DZHK (SP/19/2/344612). C.R.B. received support from EJP-RD (LQTS-NEXT,
320 ZonMW project 40-46300-98-19009) and the Leducq Foundation (project 17CVD02). HW is member
321 of the Oxford BHF Centre for Research Excellence (RE/13/1/30181). HW and JW are supported by
322 CureHeart, the British Heart Foundation's Big Beat Challenge award (BBC/F/21/220106).

323 This research has been conducted in part using the UK Biobank Resource under application numbers
324 18545 and 47602. This research was made possible through access to the data and findings generated
325 by the 100,000 Genomes Project. The 100,000 Genomes Project is managed by Genomics England
326 Limited (a wholly owned company of the Department of Health and Social Care), and funded by the
327 National Institute for Health Research and NHS England. The Wellcome Trust, Cancer Research UK and
328 the Medical Research Council have also funded research infrastructure. The 100,000 Genomes Project
329 uses data provided by patients and collected by the National Health Service as part of their care and
330 support. The Genotype-Tissue Expression (GTEx) Project was supported by the Common Fund of the
331 Office of the Director of the National Institutes of Health, and by NCI, NHGRI, NHLBI, NIDA, NIMH, and

NINDS. The data used for the analyses described in this manuscript were obtained from the GTEx Portal.

Competing Interests

R.T. has received research support from Bristol Myers Squibb. A.R.H. is a current employee and stockholder of AstraZeneca. D.P.O'R. has received grants and consultancy fees from Bayer. R.d.B. has received research grants and/or fees from AstraZeneca, Abbott, Boehringer Ingelheim, Cardior Pharmaceuticals GmbH, Ionis Pharmaceuticals, Inc., Novo Nordisk, and Roche; and also has speaker engagements with Abbott, AstraZeneca, Bayer, Bristol Myers Squibb, Novartis, and Roche. P.G. receives research funds from Abbott Cardiovascular and Medtronics. C.M.K. received research grants from Cytokinetics and Bristol Myers Squibb. P.M.M. has received consultancy fees from Roche, Biogen, Nodthera and Sangamo Pharmaceuticals and has received research or educational funds from Biogen, Novartis, Merck and Bristol Myers Squibb. J.-C.T. has received research grants from Amarin, AstraZeneca, Ceapro, DalCor, Esperion, Ionis, Novartis, Pfizer and RegenXBio; honoraria from AstraZeneca, DalCor, HLS Therapeutics, Pendopharm and Pfizer; holds minor equity interest in DalCor; and is an author of a patent on pharmacogenomics-guided CETP inhibition. J.S.W. has received research support or consultancy fees from Myokardia, Bristol-Myers Squibb, Pfizer, and Foresite Labs. C.R.B. has consulted for Illumina. H.W. has consulted for Cytokinetics, BridgeBio and BioMarin.

References

1 Watkins, H. Time to Think Differently About Sarcomere-Negative Hypertrophic Cardiomyopathy. *Circulation* **143**, 2415-2417, doi:10.1161/CIRCULATIONAHA.121.053527 (2021).

2 Harper, A. R. *et al.* Common genetic variants and modifiable risk factors underpin hypertrophic cardiomyopathy susceptibility and expressivity. *Nat Genet* **53**, 135-142, doi:10.1038/s41588-020-00764-0 (2021).

3 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, doi:10.1038/s41588-020-00762-2 (2021).

4 Yang, J. *et al.* Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. *Nat Genet* **44**, 369-375, S361-363, doi:10.1038/ng.2213 (2012).

5 Bulik-Sullivan, B. K. *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat Genet* **47**, 291-295, doi:10.1038/ng.3211 (2015).

6 Lee, S. H., Wray, N. R., Goddard, M. E. & Visscher, P. M. Estimating missing heritability for disease from genome-wide association studies. *Am J Hum Genet* **88**, 294-305, doi:10.1016/j.ajhg.2011.02.002 (2011).

7 Robinson, P., Griffiths, P. J., Watkins, H. & Redwood, C. S. Dilated and hypertrophic cardiomyopathy mutations in troponin and alpha-tropomyosin have opposing effects on the calcium affinity of cardiac thin filaments. *Circ Res* **101**, 1266-1273, doi:10.1161/CIRCRESAHA.107.156380 (2007).

8 Pickrell, J. K. *et al.* Detection and interpretation of shared genetic influences on 42 human traits. *Nat Genet* **48**, 709-717, doi:10.1038/ng.3570 (2016).

9 Aragam, K. G. *et al.* Phenotypic Refinement of Heart Failure in a National Biobank Facilitates Genetic Discovery. *Circulation*, doi:10.1161/CIRCULATIONAHA.118.035774 (2018).

10 Mountjoy, E. *et al.* An open approach to systematically prioritize causal variants and genes at all published human GWAS trait-associated loci. *Nat Genet* **53**, 1527-1533, doi:10.1038/s41588-021-00945-5 (2021).

11 Lee, M. A. *et al.* Archvillin anchors in the Z-line of skeletal muscle via the nebulin C-terminus. *Biochem Biophys Res Commun* **374**, 320-324, doi:10.1016/j.bbrc.2008.07.036 (2008).

12 Eraslan, G. *et al.* Single-nucleus cross-tissue molecular reference maps toward understanding disease gene function. *Science* **376**, eabl4290, doi:10.1126/science.abl4290 (2022).

13 Deo, R. C. *et al.* Prioritizing causal disease genes using unbiased genomic features. *Genome Biol* **15**, 534, doi:10.1186/s13059-014-0534-8 (2014).

14 Pirruccello, J. P. *et al.* Deep learning enables genetic analysis of the human thoracic aorta. *Nat Genet* **54**, 40-51, doi:10.1038/s41588-021-00962-4 (2022).

15 Hedberg-Oldfors, C. *et al.* Loss of supervillin causes myopathy with myofibrillar disorganization and autophagic vacuoles. *Brain* **143**, 2406-2420, doi:10.1093/brain/awaa206 (2020).

16 Turley, P. *et al.* Multi-trait analysis of genome-wide association summary statistics using MTAG. *Nat Genet* **50**, 229-237, doi:10.1038/s41588-017-0009-4 (2018).

17 Bai, W. *et al.* A population-based phenome-wide association study of cardiac and aortic structure and function. *Nat Med* **26**, 1654-1662, doi:10.1038/s41591-020-1009-y (2020).

18 Bulik-Sullivan, B. *et al.* An atlas of genetic correlations across human diseases and traits. *Nat Genet* **47**, 1236-1241, doi:10.1038/ng.3406 (2015).

19 de Leeuw, C. A., Mooij, J. M., Heskes, T. & Posthuma, D. MAGMA: generalized gene-set analysis of GWAS data. *PLoS Comput Biol* **11**, e1004219, doi:10.1371/journal.pcbi.1004219 (2015).

20 Reichart, D. *et al.* Pathogenic variants damage cell composition and single cell transcription in cardiomyopathies. *Science* **377**, eabo1984, doi:10.1126/science.abo1984 (2022).

21 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, doi:10.1038/s41588-022-01187-9 (2022).

22 Watanabe, K., Taskesen, E., van Bochoven, A. & Posthuma, D. Functional mapping and annotation of genetic associations with FUMA. *Nat Commun* **8**, 1826, doi:10.1038/s41467-017-01261-5 (2017).

23 Chaffin, M. *et al.* Single-nucleus profiling of human dilated and hypertrophic cardiomyopathy. *Nature* **608**, 174-180, doi:10.1038/s41586-022-04817-8 (2022).

24 Mizushima, W. *et al.* The novel heart-specific RING finger protein 207 is involved in energy metabolism in cardiomyocytes. *J Mol Cell Cardiol* **100**, 43-53, doi:10.1016/j.yjmcc.2016.09.013 (2016).

25 Cattin, M. E. *et al.* Deletion of MLIP (muscle-enriched A-type lamin-interacting protein) leads to cardiac hyperactivation of Akt/mammalian target of rapamycin (mTOR) and impaired cardiac adaptation. *J Biol Chem* **290**, 26699-26714, doi:10.1074/jbc.M115.678433 (2015).

26 Tshori, S. *et al.* Transcription factor MITF regulates cardiac growth and hypertrophy. *J Clin Invest* **116**, 2673-2681, doi:10.1172/JCI27643 (2006).

27 Risebro, C. A. *et al.* Prox1 maintains muscle structure and growth in the developing heart. *Development* **136**, 495-505, doi:10.1242/dev.030007 (2009).

28 Luo, W. *et al.* TMEM182 interacts with integrin beta 1 and regulates myoblast differentiation and muscle regeneration. *J Cachexia Sarcopenia Muscle* **12**, 1704-1723, doi:10.1002/jcsm.12767 (2021).

29 Beca, S. *et al.* Phosphodiesterase type 3A regulates basal myocardial contractility through interacting with sarcoplasmic reticulum calcium ATPase type 2a signaling complexes in mouse heart. *Circ Res* **112**, 289-297, doi:10.1161/CIRCRESAHA.111.300003 (2013).

30 Yoshida, M. *et al.* Impaired Ca2+ store functions in skeletal and cardiac muscle cells from sarcalumenin-deficient mice. *J Biol Chem* **280**, 3500-3506, doi:10.1074/jbc.M406618200 (2005).

31 Barbeira, A. N. *et al.* Integrating predicted transcriptome from multiple tissues improves association detection. *PLoS Genet* **15**, e1007889, doi:10.1371/journal.pgen.1007889 (2019).

32 Shi, X. *et al.* Zebrafish hhatla is involved in cardiac hypertrophy. *J Cell Physiol* **236**, 3700-3709, doi:10.1002/jcp.30106 (2021).

33 Green, E. M. *et al.* A small-molecule inhibitor of sarcomere contractility suppresses hypertrophic cardiomyopathy in mice. *Science* **351**, 617-621, doi:10.1126/science.aad3456 (2016).

34 Olivotto, I. *et al.* Mavacamten for treatment of symptomatic obstructive hypertrophic cardiomyopathy (EXPLORER-HCM): a randomised, double-blind, placebo-controlled, phase 3 trial. *Lancet* **396**, 759-769, doi:10.1016/S0140-6736(20)31792-X (2020).

35 Desai, M. Y. *et al.* Myosin Inhibition in Patients With Obstructive Hypertrophic Cardiomyopathy Referred for Septal Reduction Therapy. *J Am Coll Cardiol* **80**, 95-108, doi:10.1016/j.jacc.2022.04.048 (2022).

36 Evangelou, E. *et al.* Genetic analysis of over 1 million people identifies 535 new loci associated with blood pressure traits. *Nat Genet* **50**, 1412-1425, doi:10.1038/s41588-018-0205-x (2018).

37 Walsh, R., Offerhaus, J. A., Tadros, R. & Bezzina, C. R. Minor hypertrophic cardiomyopathy genes, major insights into the genetics of cardiomyopathies. *Nat Rev Cardiol* **19**, 151-167, doi:10.1038/s41569-021-00608-2 (2022).

38 Richards, S. *et al.* Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med* **17**, 405-424, doi:10.1038/gim.2015.30 (2015).

39 Neubauer, S. *et al.* Distinct Subgroups in Hypertrophic Cardiomyopathy in the NHLBI HCM Registry. *J Am Coll Cardiol* **74**, 2333-2345, doi:10.1016/j.jacc.2019.08.1057 (2019).

40 Magi, R. & Morris, A. P. GWAMA: software for genome-wide association meta-analysis. *BMC Bioinformatics* **11**, 288, doi:10.1186/1471-2105-11-288 (2010).

41 Semsarian, C., Ingles, J., Maron, M. S. & Maron, B. J. New perspectives on the prevalence of hypertrophic cardiomyopathy. *J Am Coll Cardiol* **65**, 1249-1254, doi:10.1016/j.jacc.2015.01.019 (2015).

42 Turro, E. *et al.* Whole-genome sequencing of patients with rare diseases in a national health system. *Nature* **583**, 96-102, doi:10.1038/s41586-020-2434-2 (2020).

43 Genomics England: The National Genomics Research and Healthcare Knowledgebase v5. (2019). <doi:10.6084/m9.figshare.4530893.v5>.

44 McLaren, W. *et al.* Deriving the consequences of genomic variants with the Ensembl API and SNP Effect Predictor. *Bioinformatics* **26**, 2069-2070, doi:10.1093/bioinformatics/btq330 (2010).

45 Karczewski, K. J. *et al.* The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* **581**, 434-443, doi:10.1038/s41586-020-2308-7 (2020).

46 Petersen, S. E. *et al.* UK Biobank's cardiovascular magnetic resonance protocol. *J Cardiovasc Magn Reson* **18**, 8, doi:10.1186/s12968-016-0227-4 (2016).

47 Loh, P. R. *et al.* Efficient Bayesian mixed-model analysis increases association power in large cohorts. *Nat Genet* **47**, 284-290, doi:10.1038/ng.3190 (2015).

48 Rehm, H. L. *et al.* ClinGen--the Clinical Genome Resource. *N Engl J Med* **372**, 2235-2242, doi:10.1056/NEJMsr1406261 (2015).

49 Willer, C. J., Li, Y. & Abecasis, G. R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190-2191, doi:10.1093/bioinformatics/btq340 (2010).

50 Ernst, J. *et al.* Mapping and analysis of chromatin state dynamics in nine human cell types. *Nature* **473**, 43-49, doi:10.1038/nature09906 (2011).

51 Roadmap Epigenomics, C. *et al.* Integrative analysis of 111 reference human epigenomes. *Nature* **518**, 317-330, doi:10.1038/nature14248 (2015).

52 Fulco, C. P. *et al.* Activity-by-contact model of enhancer-promoter regulation from thousands of CRISPR perturbations. *Nat Genet* **51**, 1664-1669, doi:10.1038/s41588-019-0538-0 (2019).

53 Nasser, J. *et al.* Genome-wide enhancer maps link risk variants to disease genes. *Nature* **593**, 238-243, doi:10.1038/s41586-021-03446-x (2021).

54 Finucane, H. K. *et al.* Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nat Genet* **47**, 1228-1235, doi:10.1038/ng.3404 (2015).

55 Gazal, S. *et al.* Linkage disequilibrium-dependent architecture of human complex traits shows action of negative selection. *Nat Genet* **49**, 1421-1427, doi:10.1038/ng.3954 (2017).

56 Gazal, S., Marquez-Luna, C., Finucane, H. K. & Price, A. L. Reconciling S-LDSC and LDAK functional enrichment estimates. *Nat Genet* **51**, 1202-1204, doi:10.1038/s41588-019-0464-1 (2019).

57 Ingles, J. *et al.* Evaluating the Clinical Validity of Hypertrophic Cardiomyopathy Genes. *Circ Genom Precis Med* **12**, e002460, doi:10.1161/CIRCGEN.119.002460 (2019).

58 The Genotype-Tissue Expression (GTEx) project. *Nat Genet* **45**, 580-585, doi:10.1038/ng.2653 (2013).

59 Barbeira, A. N. *et al.* Exploring the phenotypic consequences of tissue specific gene expression variation inferred from GWAS summary statistics. *Nat Commun* **9**, 1825, doi:10.1038/s41467-018-03621-1 (2018).

60 Barbeira, A. N. *et al.* Exploiting the GTEx resources to decipher the mechanisms at GWAS loci. *Genome Biol* **22**, 49, doi:10.1186/s13059-020-02252-4 (2021).

61 Hemani, G. *et al.* The MR-Base platform supports systematic causal inference across the human phenome. *Elife* **7**, doi:10.7554/eLife.34408 (2018).