

1 **Common genetic variants, and modifiable risk factors, underpin hypertrophic**
2 **cardiomyopathy susceptibility and expressivity**

3

4 Andrew R. Harper^{1,2}, Anuj Goel^{1,2}, Christopher Grace^{1,2}, Kate L. Thomson^{1,2,3}, Steffen E.
5 Petersen⁴, Xiao Xu⁵, Adam Waring², Elizabeth Ormondroyd^{1,2}, Christopher M. Kramer⁶,
6 Carolyn Y. Ho⁷, Stefan Neubauer¹, HCMR Investigators, Rafik Tadros⁸, James S. Ware^{5*},
7 Connie R. Bezzina^{9*}, Martin Farrall^{1,2*}, and Hugh Watkins^{1,2,10*}

8

9 ¹Radcliffe Department of Medicine, University of Oxford, Division of Cardiovascular
10 Medicine, John Radcliffe Hospital, Oxford, UK.

11 ²Wellcome Centre for Human Genetics, Roosevelt Drive, Oxford, UK.

12 ³Oxford Medical Genetics Laboratories, Churchill Hospital, Oxford, UK.

13 ⁴William Harvey Research Institute, Barts and The London School of Medicine and
14 Dentistry, Queen Mary University of London, London, UK.

15 ⁵National Heart & Lung Institute, Imperial College London, London, UK.

16 ⁶University of Virginia Health System, Charlottesville, VA, USA.

17 ⁷Cardiovascular Division, Brigham and Women's Hospital, Boston, MA, USA.

18 ⁸Cardiovascular Genetics Centre, Montreal Heart Institute, Montreal, QC, Canada.

19 ⁹Amsterdam UMC, AMC Heart Center, Amsterdam, The Netherlands.

20 ¹⁰NIHR Oxford Biomedical Research Centre, Oxford University Hospitals NHS Foundation
21 Trust, John Radcliffe Hospital, Oxford, UK.

22 *These authors contributed equally to this work.

23 e-mail: hugh.watkins@rdm.ox.ac.uk

24 **Hypertrophic cardiomyopathy (HCM) is a common, serious, genetic heart disorder.**
25 **Rare pathogenic variants in sarcomere genes cause HCM, but with unexplained**
26 **phenotypic heterogeneity. Moreover, most patients do not carry such variants. We**
27 **report a genome-wide association study of 2,780 cases and 47,486 controls that**
28 **identified 12 genome-wide significant susceptibility loci for HCM. SNP-heritability**
29 **indicated a strong polygenic influence, especially for sarcomere-negative HCM (64% of**
30 **cases; $h^2_g = 0.34 \pm 0.02$). A genetic risk score showed substantial influence on odds of**
31 **HCM in a validation study, halving odds in the lowest quintile and doubling in the**
32 **highest quintile, and also influenced phenotypic severity in sarcomere variant carriers.**
33 **Mendelian randomization identified diastolic blood pressure (DBP) as a key modifiable**
34 **risk factor for sarcomere-negative HCM, with 1 standard deviation increase in DBP**
35 **increasing HCM risk four-fold. Common variants and modifiable risk factors have**
36 **important roles in HCM that we suggest will be clinically actionable.**

37

38 HCM is common, affecting at least 1 in 500 individuals, and presents substantial unmet
39 medical need¹. It is a leading cause of sudden death, embolic stroke and heart failure in early
40 and mid-adult life. Sarcomeric HCM, caused by mutations in myofilament genes, is inherited
41 as an autosomal dominant disorder. However, as commonly seen in adult onset heterozygous
42 disorders, HCM is characterized by reduced penetrance and variable expressivity, providing
43 challenges for diagnosis and prognosis^{2,3}. In the more common sarcomere-negative setting,
44 cases are often isolated, but clustering in nuclear families is still frequent, requiring clinical
45 surveillance in families^{4,5}. To investigate the contribution of common genetic variants to
46 HCM risk, we performed two independent multi-ancestry case-control genome-wide

47 association studies (GWAS) of unrelated HCM patients recruited to the Hypertrophic
48 Cardiomyopathy Registry (HCMR, 2,541 unselected cases vs. 40,283 UK Biobank controls)
49 and the BioResource for Rare Disease (BRRD, 239 sarcomere-negative cases vs. 7,203
50 controls) (Fig. 1, Supplementary Table 1 and Supplementary Note). SNP-heritability (h^2_g)
51 estimates calculated using GREML-LDMS indicated that a significant proportion of HCM
52 risk was attributable to the additive effects of common (minor allele frequency (MAF) >
53 0.01) SNPs (HCMR $h^2_g = 0.35 \pm 0.01$; BRRD $h^2_g = 0.68 \pm 0.16$).

54 We performed fixed-effects inverse-variance meta-analysis of the HCMR and BRRD
55 GWAS datasets for 8,590,397 single nucleotide polymorphisms (SNPs) across a total of
56 2,780 HCM cases and 47,486 age and sex-matched controls. All-comer analysis (i.e.
57 inclusive of sarcomere-positive and sarcomere-negative HCM cases) identified 13
58 independent genome-wide significant variants in 12 loci ($P < 5 \times 10^{-8}$) using a stepwise
59 model selection procedure with Genome-wide Complex Trait Analysis (GCTA) and
60 confirmed with conditional analysis (Table 1, Supplementary Table 2 and Online Methods).
61 We identified an additional 16 independent variants at a 5% FDR significance threshold ($P <$
62 1.82×10^{-6}) (Supplementary Table 3). We replicated 11 of the 13 genome-wide significant
63 variants and 4 of the 16 FDR variants in a smaller, independent HCM meta-analysis ($n =$
64 1,643 cases and 6,628 controls; Table 1, Supplementary Table 3 and Online Methods).
65 Additionally, we obtained similar discovery findings with a European-only analysis
66 (Supplementary Table 2 and Supplementary Note).

67 The *FHOD3* locus was found to harbor two independent genome-wide significant
68 variants, rs4799426 and rs118060942, in linkage equilibrium ($r^2 = 0.01$). Sentinel SNPs in the
69 HCM susceptibility loci conferred relatively large susceptibility effect sizes (median OR =
70 1.25, range 1.18–2.16) across a range of effect allele frequencies (range 0.012–0.83)
71 (Supplementary Table 4). Tissue enrichment tests, performed in FUMA using gene-level data

72 (Supplementary Table 5) and tissue expression data from GTEx (v8.0) showed enrichment in
73 left ventricular myocardium ($\beta = 0.04 \pm 0.01$; $P = 7.46 \times 10^{-6}$), skeletal muscle ($\beta =$
74 0.03 ± 0.01 ; $P = 1.13 \times 10^{-5}$), and atrial appendage ($\beta = 0.04 \pm 0.01$; $P = 1.18 \times 10^{-5}$)
75 (Supplementary Table 6 and Supplementary Note)^{6,7}. Functional GWAS supported these
76 findings and revealed cell types where sentinel SNPs were most enriched (Supplementary
77 Tables 7 and 8).

78 We dichotomized HCM cases in HCMR into sarcomere-positive (34.3%) and
79 sarcomere-negative (64.3%) groups using a published framework (Supplementary Tables 9
80 and 10)⁸. The GREML heritability estimate for sarcomere-negative HCM exceeded that of
81 sarcomere-positive HCM ($h^2_g = 0.34 \pm 0.02$ vs. 0.16 ± 0.04) (Supplementary Table 11). This
82 supports the hypothesis that where there is familial aggregation that is not explained by co-
83 segregation with a rare variant, as in sarcomere-negative HCM, a greater role for common
84 variants may be expected. This applies in particular to the BRRD samples, which were
85 enriched for positive family history despite negative gene-panel testing, where heritability
86 was indeed greatest. A meta-analysis of sarcomere-negative HCM (1,874 HCM cases vs.
87 27,344 controls) identified 10 independent genome-wide significant variants in 9 loci and a
88 further 15 independent variants in 13 loci below a 5% FDR threshold ($P < 1.56 \times 10^{-6}$)
89 (Supplementary Table 12). Three loci (*FHOD3*, *TBX3* and *PLN*) harbored a secondary
90 independent variant following conditional association analysis (Supplementary Table 2).
91 Sarcomere-positive HCM GWAS analysis (871 HCMR cases vs. 20,142 UKBB controls)
92 yielded 7 independent genome-wide significant variants and a further 11 independent variants
93 below a 5% FDR threshold ($P < 1.50 \times 10^{-6}$) from 12 loci. This includes 7 variants in the
94 peri-centromeric region of chromosome 11 neighboring *MYBPC3*, a prominent cause of
95 monogenic HCM (Supplementary Tables 2 and 13). Haplotype analysis of individual-level
96 sequence data demonstrated long-range linkage disequilibrium and potential spurious

97 associations between frequently observed rare pathogenic variants in *MYBPC3*
98 (NM_000256.3), specifically p.R502W and p.Trp792ValfsTer41, and common imputed
99 variants in the chr11:44,976,681-57,917,265 genomic interval⁹⁻¹¹. Modelling the impact of
100 both rare and common variants with multiple logistic regression confirmed that HCM risk
101 could be entirely attributed to the rare variants (Supplementary Note). Common variants in
102 chr11:44,000,000-58,000,000 were masked from subsequent analyses, leaving 2 independent
103 variants of genome-wide significance and 9 below a 5% FDR threshold. Excluding
104 chr11:44,976,681-57,917,265 had a trivial effect on the heritability estimate.

105 Bivariate GREML analysis revealed a strong positive genetic correlation between
106 sarcomere-positive and sarcomere-negative HCM ($r_g = 1.00 \pm 0.12$). Pairwise GWAS
107 comparison revealed overlapping signals between sarcomere-negative and sarcomere-positive
108 loci for 59% of regions ($n = 22/37$) (Supplementary Table 14). Most of the sarcomere-
109 negative GWAS loci were not reproduced at the genome-wide significance level in the
110 sarcomere-positive GWAS, which could potentially be explained by a relative lack of power
111 (Supplementary Table 15). Four SNPs (rs2312403, rs35469308, rs12299450 and rs2758215)
112 showed association only in the sarcomere-positive GWAS, and may represent modifier loci.

113 All loci were novel, apart from *FHOD3*, which has been previously reported in a
114 HCM GWAS¹². Previous candidate gene studies have reported rare variant associations in
115 different forms of cardiomyopathy for *BAG3* and *FHOD3*, and common variant associations
116 with dilated cardiomyopathy (DCM) have been reported for *BAG3* and *HSPB7* loci¹³⁻¹⁷. At
117 these loci shared by HCM and DCM, the direction of effect is opposite, with the HCM risk
118 allele being previously shown to decrease risk of DCM. The involvement of *BAG3*, *HSPB7*
119 and *FHOD3* points to the importance of homeostatic pathways for sarcomeric structural
120 integrity during mechanical stress (Supplementary Note). While some of the other loci also
121 encode known cardiomyopathy genes (*PLN*, *TTN*), the major HCM and DCM myofilament

122 loci are not represented, consistent with the cardiomyopathy-causing changes in these genes
123 altering protein structure rather than expression level. In the remaining loci, some early clues
124 implicate specific genes and mechanisms: a deleterious missense variant implicates
125 *ADPRHL1*, important for Z-disc and actin dynamics, and a *cis*-eQTL implicates *SLC6A6*,
126 which encodes a taurine transporter known to be responsible for cardiomyopathy in dogs
127 (Supplementary Note).

128 After excluding rs28768976 for extreme pleiotropy (Supplementary Note) and
129 rs78310129 due to long-range linkage disequilibrium with pathogenic *MYBPC3* variants
130 (Supplementary Note), 27 SNPs demonstrating independent associations with HCM at a 5%
131 FDR threshold in the all-comer HCM meta-analysis were aggregated into a scaled (i.e. per-
132 standard deviation effects) weighted genetic risk score (GRS) (Table 2 and Supplementary
133 Tables 16 and 17). The GRS predicts odds of HCM in a validation meta-analysis of three
134 independent HCM cohorts comprising 1,769 cases and 39,828 controls (OR = 1.73 per SD
135 (95%CI 1.63-1.83)) (Fig. 2). Using the largest replication cohort, we conducted a sensitivity
136 analysis and confirmed a 5% FDR threshold as representative of alternate SNP significance
137 thresholds (Supplementary Table 16).

138 Stratification of the HCMR cases by their average genetic ancestry, as determined by
139 principal components analysis, demonstrates similar effect sizes across all ancestry groups
140 (Supplementary Table 18). Using the central 60% of the population as the reference group,
141 there is a protective effect for individuals in the lowest quintile (OR = 0.53 (95%CI 0.45-
142 0.63)) and a greater than a two-fold increased odds of HCM for individuals in the highest
143 quintile (OR = 2.30 (95%CI 2.02–2.62)). In alignment with h^2_g estimates, the GRS
144 demonstrated larger effects in the sarcomere-negative subgroup (Fig. 2 and Supplementary
145 Note). Nevertheless, in young individuals carrying a pathogenic sarcomere mutation, who

146 might typically have a ~50% chance of developing overt cardiomyopathy in adulthood, a
147 halving or doubling of average risk of developing the cardiomyopathy is likely to be
148 clinically meaningful.

149 To determine whether the common susceptibility variants also influence disease
150 severity in sarcomere-positive HCM, i.e. through a modifier effect, we assessed the impact of
151 the GRS on LV hypertrophy in groups of cases with similar mutational mechanisms. A 1 SD
152 unit increase in GRS conferred a 0.71 ± 0.35 mm increase in maximum left ventricular wall
153 thickness ($P = 0.048$) in carriers of *MYBPC3* truncating variants ($n = 232$) and a 0.73 ± 0.36
154 mm increase ($P = 0.037$) in carriers of *MYH7* missense variants ($n = 186$) (Fig. 3). Allelic
155 heterogeneity currently limits single variant expressivity estimates; the most frequently
156 observed pathogenic variant in HCM (*MYBPC3*, p.R502W) is associated with a larger GRS
157 effect size (1.61 ± 0.80 mm increase per 1 SD unit increase in GRS) but is currently modestly
158 powered ($n = 48$).

159 Observational studies show that hypertension, obesity and type 2 diabetes are more
160 prevalent in individuals with HCM, but these could be secondary to reduced exercise^{8,18,19}.
161 We performed two-sample Mendelian randomization (2SMR) to leverage large-scale GWAS
162 for these heritable traits^{20–23}. We inferred causal relationships with HCM for hypertension
163 and obesity, but not diabetes (Fig. 4 and Supplementary Table 19). Most notably, diastolic
164 blood pressure appears to be a substantial risk factor for the development of sarcomere-
165 negative HCM (Fig. 4 and Supplementary Table 19). A 1 SD unit increase in diastolic blood
166 pressure (11.3 mmHg) confers a four-fold increased risk of HCM (OR = 3.93 (95%CI 2.86–
167 5.41); $P = 3.74 \times 10^{-16}$), more than double the risk typically observed for other diseases
168 associated with diastolic blood pressure (Fig. 4 and Supplementary Table 20)^{24–28}. The strong
169 association with hypertension raises the possibility that sarcomere-negative HCM may

170 represent, in part, an exaggerated response to hypertension in genetically susceptible
171 individuals. The association specifically with diastolic blood pressure likely reflects that this
172 is the dominant form of hypertension in young and mid-adult life^{29,30}.

173 The individual loci identified in this study hold great potential for driving new
174 insights into cardiomyopathy pathogenesis. Many of the association signals have already
175 been replicated; others will need further study to guard against false positive findings.
176 Collectively, our findings highlight the important influence of common variants on the risk of
177 developing HCM. The polygenic contribution is weaker in individuals with pathogenic
178 sarcomeric variants, but a common variant GRS may still be particularly useful here because
179 the high prior risk means that the modest (e.g. four-fold) changes in individual-specific
180 penetrance, which will apply to 40% of individuals, will have a large absolute effect on
181 outcome. Additionally, it appears that common variants explain part of the variable
182 expressivity of pathogenic sarcomeric variants. The clinical utility of a GRS now needs study
183 in adequately powered longitudinal surveys of HCM disease progression, especially in
184 sarcomere-positive individuals who were limited in number in the current study. In
185 individuals lacking cardinal pathogenic mutations in sarcomeric genes, we suggest that
186 extremes of the polygenic risk distribution (e.g. the top 1% of the population), combined with
187 causal risk factors, drive individual susceptibility. Managing sarcomere-negative HCM
188 patients and their relatives may be greatly facilitated by awareness of the strong influence of
189 polygenic risk and of diastolic blood pressure as a major modifiable risk factor.

190

191 **HCMR Investigators**

192

193 Paul Kolm¹¹, Raymond Kwong¹², Sarahfaye F. Dolman¹¹, Patrice Desvigne-Nickens¹³, John
194 P. Dimarco¹⁴, Nancy Geller¹³, Dong-Yun Kim¹³, Cheng Zhang¹¹, William Weintraub¹¹,
195 Theodore Abraham¹⁵, Lisa Anderson¹⁶, Evan Appelbaum¹⁷, Camillo Autore¹⁸, Colin Berry¹⁹,

196 Elena Biagini²⁰, William Bradlow²¹, Chiara Bucciarelli-Ducci²², Amedeo Chiribiri²³, Lubna
197 Choudhury²⁴, Andrew Crean²⁵, Dana Dawson²⁶, Milind Y. Desai²⁷, Eleanor Elstein²⁸,
198 Andrew Flett²⁹, Matthias Friedrich³⁰, Stephen Heitner³¹, Adam Helms³², Daniel L. Jacoby³³,
199 Han Kim³⁴, Bette Kim³⁵, Eric Larose³⁶, Masliza Mahmud³⁷, Heiko Mahrholdt³⁸, Martin
200 Maron³⁹, Gerry McCann⁴⁰, Michelle Michels⁴¹, Saidi Mohiddin⁴², Sherif Nagueh⁴³, David
201 Newby⁴⁴, Iacopo Olivetto⁴⁵, Anjali Owens⁴⁶, F. Pierre-Mongeon⁴⁷, Sanjay Prasad⁴⁸, Ornella
202 Rimoldi⁴⁹, Michael Salerno⁵⁰, Jeanette Schulz-Menger⁵¹, Mark Sherrid⁵², Peter Swoboda⁵³,
203 Albert van Rossum⁵⁴, Jonathan Weinsaft⁵⁵, James White⁵⁶, and Eric Williamson⁵⁷
204
205 ¹¹MedStar Heart and Vascular Institute, Washington, DC, USA.
206 ¹²Cardiovascular Division, Department of Medicine and Department of Radiology, Brigham
207 and Women's Hospital, Boston, MA, USA.
208 ¹³National Heart, Lung, and Blood Institute, Bethesda, MD, USA.
209 ¹⁴Cardiovascular Division, University of Virginia Health System, Charlottesville, VA, USA.
210 ¹⁵Hypertrophic Cardiomyopathy Center of Excellence, Johns Hopkins University, Baltimore,
211 MD, USA.
212 ¹⁶St Georges University Hospitals NHS Trust, London, UK.
213 ¹⁷Departments of Medicine, Cardiovascular Division and Radiology, Beth Israel Deaconess
214 Medical Center, Harvard Medical School, Boston, MA, USA.
215 ¹⁸Division of Cardiology, Department of Clinical and Molecular Medicine, St. Andrea
216 Hospital, Sapienza University, Rome, Italy.
217 ¹⁹British Heart Foundation Glasgow Cardiovascular Research Centre, Institute of
218 Cardiovascular and Medical Sciences, University of Glasgow, Glasgow, UK.
219 ²⁰Cardio-Thoraco-Vascular Department, University Hospital of Bologna, Policlinico S.
220 Orsola-Malpighi, Bologna, Italy.
221 ²¹Department of Cardiology, New Queen Elizabeth Hospital Birmingham, Birmingham, UK.
222 ²²Bristol Heart Institute, Bristol National Institute of Health Research (NIHR) Biomedical
223 Research Centre, University Hospitals Bristol NHS Trust and University of Bristol, Bristol,
224 UK.
225 ²³Cardiovascular Division, King's College London British Heart Foundation Centre of
226 Excellence, The Rayne Institute, St. Thomas' Hospital Campus, London, UK.
227 ²⁴Division of Cardiology, Department of Medicine, Bluhm Cardiovascular Institute,
228 Northwestern University Feinberg School of Medicine, Chicago, IL, USA.
229 ²⁵Division of Cardiology, Peter Munk Cardiac Center, University Health Network, University
230 of Toronto, Toronto, ON, Canada.
231 ²⁶Aberdeen Cardiovascular and Diabetes Centre, University of Aberdeen, Aberdeen, UK.
232 ²⁷Department of Cardiovascular Medicine, Center for Radiation Heart Disease, Heart and
233 Vascular Institute, Cleveland Clinic, Cleveland, OH, USA.
234 ²⁸Division of Cardiology, Department of Medicine, Royal Victoria Hospital, McGill
235 University Health Centre, Montreal, QC, Canada.
236 ²⁹Department of Cardiology, University Hospital Southampton NHS Foundation Trust,
237 Southampton, UK.
238 ³⁰Department of Medicine, Heidelberg University, Heidelberg, Germany.
239 ³¹Oregon Health and Sciences University (OHSU), Division of Cardiovascular Medicine,
240 Knight Cardiovascular Institute, Portland, OR, USA.
241 ³²Department of Internal Medicine, University of Michigan, Ann Arbor, MI, USA.
242 ³³Section of Cardiovascular Medicine, Department of Internal Medicine, Yale School of
243 Medicine, New Haven, CT, USA.
244 ³⁴Duke Cardiovascular Magnetic Resonance Center and Division of
245 Cardiology, Duke University Medical Center, Durham, NC, USA.

246 ³⁵Mount Sinai West, Icahn School of Medicine at Mount Sinai, New York, NY, USA.
247 ³⁶Quebec Heart and Lung Institute, Laval University, QC, Canada.
248 ³⁷Division of Cardiovascular Medicine, Radcliffe Department of Medicine, University of
249 Oxford, Oxford, UK.
250 ³⁸Department of Cardiology, Robert-Bosch-Krankenhaus, Stuttgart, Germany.
251 ³⁹Hypertrophic Cardiomyopathy Center and Research Institute, Tufts Medical Center,
252 Boston, MA, USA.
253 ⁴⁰Department of Cardiovascular Sciences, University of Leicester, Leicester, UK.
254 ⁴¹Erasmus University, Rotterdam, The Netherlands.
255 ⁴²Barts Heart Center, The Cardiovascular Magnetic Resonance Imaging Unit, St
256 Bartholomew's Hospital, London, UK.
257 ⁴³Methodist DeBakey Heart and Vascular Center, Houston, TX, USA.
258 ⁴⁴Centre for Cardiovascular Science, University of Edinburgh, Edinburgh, UK.
259 ⁴⁵Cardiomyopathy Unit and Genetic Unit, Careggi University Hospital, Florence, Italy.
260 ⁴⁶Center for Inherited Cardiovascular Disease, Division of Cardiovascular Medicine,
261 Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA.
262 ⁴⁷Montréal Heart Institute, Montréal, QC, Canada.
263 ⁴⁸National Heart and Lung Institute, Imperial College London, and Royal Brompton Hospital,
264 London, UK.
265 ⁴⁹Vita Salute University and San Raffaele Hospital, Milan, Italy.
266 ⁵⁰Department of Medicine, University of Virginia, Charlottesville, VA, USA.
267 ⁵¹Charité, Medical Faculty of the Humboldt University, Experimental and Clinical Research
268 Center, Berlin, and Helios Clinics, Cardiology, Berlin, Germany.
269 ⁵²Hypertrophic Cardiomyopathy Program, Leon Charney Division of Cardiology,
270 Department of Medicine, New York University School of Medicine, New York, NY, USA.
271 ⁵³Department of Cardiovascular Imaging Science, Leeds Institute of Cardiovascular and
272 Metabolic Medicine, University of Leeds, Leeds, UK.
273 ⁵⁴Department of Cardiology, Amsterdam UMC, Amsterdam, The Netherlands.
274 ⁵⁵Departments of Medicine and Radiology, Weill Cornell Medical College, New York, NY,
275 USA.
276 ⁵⁶Calgary Foothills Medical Centre, University of Calgary, Calgary, AB, Canada.
277 ⁵⁷Department of Radiology, Mayo Clinic, Rochester, MN, USA.

278

279 **Acknowledgements**

280 This work was supported by funding from the British Heart Foundation (BHF), the Medical
281 Research Council (MRC), the National Heart, Lung, and Blood Institute (NIH grant
282 U01HL117006-01A1), the Wellcome Trust (201543/B/16/Z), Wellcome Trust core awards
283 (090532/Z/09/Z, 203141/Z/16/Z) and the National Institute for Health Research (NIHR)
284 Oxford Biomedical Research Centre. A.R.H. has received support from the Medical Research
285 Council Doctoral Training Partnership. A.G. has received support from the BHF, European

286 Commission [LSHM-CT- 2007-037273, HEALTH-F2-2013-601456] and TriPartite
287 Immunometabolism Consortium [TrIC]- NovoNordisk Foundation [NNF15CC0018486].
288 S.E.P. acknowledges support from the NIHR Barts Biomedical Research Centre. A.W. has
289 received support from the Wellcome Trust. S.N., M.F. and H.W. are members of the Oxford
290 BHF Centre for Research Excellence (RE/13/1/30181). We are grateful for access to the
291 high-performance Oxford Biomedical Research Computing (BMRC) facility, a joint
292 development between the Wellcome Centre for Human Genetics and the Big Data Institute
293 supported by Health Data Research UK and the NIHR Oxford Biomedical Research Centre.
294 The views expressed are those of the author(s) and not necessarily those of the NHS, the
295 NIHR or the Department of Health or the Department of Health and Social Care.

296 We thank the NIHR BioResource volunteers for their participation, and gratefully
297 acknowledge NIHR BioResource centres, NHS Trusts and staff for their contribution. We
298 thank the National Institute for Health Research and NHS Blood and Transplant. This
299 research was made possible through access to the data and findings generated by the 100,000
300 Genomes Project, which is managed by Genomics England Limited (a wholly owned
301 company of the Department of Health and Social Care) and funded by the National Institute
302 for Health Research and NHS England with research infrastructure funding from the
303 Wellcome Trust, Cancer Research UK and the Medical Research Council. The 100,000
304 Genomes Project uses data provided by patients and collected by the National Health Service
305 as part of their care and support.

306 We acknowledge the contribution of the Oxford Medical Genetics Laboratories.

307

308 **Author Contributions**

309 A.R.H., M.F., and H.W. conceived and designed the study. A.R.H., A.G., C.G., K.L.T.,
310 S.E.P., A.W., E.O., C.M.K., S.N., and C.Y.H. acquired, analyzed, and interpreted the data.

311 X.X. and R.T. provided assistance with replication. A.R.H., M.F., and H.W. wrote the
312 manuscript. J.S.W., C.R.B., and R.T. critically revised the manuscript for important
313 intellectual content.

314

315 **Competing Interests**

316 As of April 2020, A.R.H. is an employee of AstraZeneca.

317

318 **References**

319

- 320 1. Yotti, R., Seidman, C. E. & Seidman, J. G. Advances in the genetic basis and
321 pathogenesis of sarcomere cardiomyopathies. *Annu. Rev. Genomics Hum. Genet.* **20**,
322 129–153 (2019).
- 323 2. Harper, A. R., Parikh, V. N., Goldfeder, R. L., Caleshu, C. & Ashley, E. A. Delivering
324 clinical grade sequencing and genetic test interpretation for cardiovascular medicine.
325 *Circ. Cardiovasc. Genet.* **10**, e001221 (2017).
- 326 3. Watkins, H., Ashrafian, H. & Redwood, C. Inherited cardiomyopathies. *N. Engl. J.*
327 *Med.* **364**, 1643–1656 (2011).
- 328 4. Thomson, K. L. *et al.* Analysis of 51 proposed hypertrophic cardiomyopathy genes
329 from genome sequencing data in sarcomere negative cases has negligible diagnostic
330 yield. *Genet. Med.* **21**, 1576–1584 (2019).
- 331 5. Ingles, J. *et al.* Nonfamilial hypertrophic cardiomyopathy: prevalence, natural history
332 and clinical implication. *Circ. Cardiovasc. Genet.* **10**, e001620 (2017).
- 333 6. Watanabe, K., Taskesen, E., Van Bochoven, A. & Posthuma, D. Functional mapping
334 and annotation of genetic associations with FUMA. *Nat. Commun.* **8**, 1826 (2017).
- 335 7. Aguet, F. *et al.* The GTEx Consortium atlas of genetic regulatory effects across human
336 tissues. *bioRxiv* (2019) doi:10.1101/787903.
- 337 8. Neubauer, S. *et al.* Distinct subgroups in hypertrophic cardiomyopathy in the NHLBI
338 HCM Registry. *J. Am. Coll. Cardiol.* **74**, 2333–2345 (2019).
- 339 9. Wray, N. R., Purcell, S. M. & Visscher, P. M. Synthetic associations created by rare
340 variants do not explain most GWAS results. *PLoS Biol.* **9**, e1000579 (2011).
- 341 10. Dickson, S. P., Wang, K., Krantz, I., Hakonarson, H. & Goldstein, D. B. Rare variants
342 create synthetic genome-wide associations. *PLoS Biol.* **8**, e1000294 (2010).
- 343 11. Orozco, G., Barrett, J. C. & Zeggini, E. Synthetic associations in the context of
344 genome-wide association scan signals. *Hum. Mol. Genet.* **19**, R137–R144 (2010).
- 345 12. Wooten, E. C. *et al.* Formin homology 2 domain containing 3 variants associated with
346 hypertrophic cardiomyopathy. *Circ. Cardiovasc. Genet.* **6**, 10–18 (2013).
- 347 13. Ochoa, J. P. *et al.* Formin Homology 2 Domain Containing 3 (FHOD3) is a genetic
348 basis for hypertrophic cardiomyopathy. *J. Am. Coll. Cardiol.* **72**, 2457–2467 (2018).
- 349 14. Esslinger, U. *et al.* Exome-wide association study reveals novel susceptibility genes to
350 sporadic dilated cardiomyopathy. *PLoS One* **12**, e0172995 (2017).
- 351 15. Selcen, D. *et al.* Mutation in BAG3 causes severe dominant childhood muscular

- dystrophy. *Ann. Neurol.* **65**, 83–89 (2009).
- 353 16. Odgerel, Z. *et al.* Inheritance patterns and phenotypic features of myofibrillar
354 myopathy associated with a BAG3 mutation. *Neuromuscul. Disord.* **20**, 438–442
355 (2010).
- 356 17. Villard, E. *et al.* A genome-wide association study identifies two loci associated with
357 heart failure due to dilated cardiomyopathy. *Eur. Heart J.* **32**, 1065–1076 (2011).
- 358 18. Fumagalli, C. *et al.* Association of obesity with adverse long-term outcomes in
359 hypertrophic cardiomyopathy. *JAMA Cardiol.* **5**, 65–72 (2019).
- 360 19. Ho, C. Y. *et al.* Genotype and lifetime burden of disease in hypertrophic
361 cardiomyopathy. *Circulation* **138**, 1387–1398 (2018).
- 362 20. Mahajan, A. *et al.* Fine-mapping type 2 diabetes loci to single-variant resolution using
363 high-density imputation and islet-specific epigenome maps. *Nat. Genet.* **50**, 1505–
364 1513 (2018).
- 365 21. Evangelou, E. *et al.* Genetic analysis of over 1 million people identifies 535 new loci
366 associated with blood pressure traits. *Nat. Genet.* **50**, 1412–1425 (2018).
- 367 22. Pulit, S. L. *et al.* Meta-analysis of genome-wide association studies for body fat
368 distribution in 694 649 individuals of European ancestry. *Hum. Mol. Genet.* **28**, 166–
369 174 (2019).
- 370 23. Yengo, L. *et al.* Meta-analysis of genome-wide association studies for height and body
371 mass index in ~700 000 individuals of European ancestry. *Hum. Mol. Genet.* **27**,
372 3641–3649 (2018).
- 373 24. van der Harst, P. & Verweij, N. Identification of 64 novel genetic loci provides an
374 expanded view on the genetic architecture of coronary artery disease. *Circ. Res.* **122**,
375 433–443 (2018).
- 376 25. Roselli, C. *et al.* Multi-ethnic genome-wide association study for atrial fibrillation.
377 *Nat. Genet.* **50**, 1225–1233 (2018).
- 378 26. Wuttke, M. *et al.* A catalog of genetic loci associated with kidney function from
379 analyses of a million individuals. *Nat. Genet.* **51**, 957–972 (2019).
- 380 27. Malik, R. *et al.* Multiancestry genome-wide association study of 520,000 subjects
381 identifies 32 loci associated with stroke and stroke subtypes. *Nat. Genet.* **50**, 524–537
382 (2018).
- 383 28. Shah, S. *et al.* Genome-wide association and Mendelian randomisation analysis
384 provide insights into the pathogenesis of heart failure. *Nat. Commun.* **11**, 163 (2020).
- 385 29. Franklin, S. S. *et al.* Predictors of new-onset diastolic and systolic hypertension: the
386 Framingham Heart Study. *Circulation* **111**, 1121–1127 (2005).
- 387 30. Franklin, S. S., Jacobs, M. J., Wong, N. D., L'Italien, G. J. & Lapuerta, P.
388 Predominance of isolated systolic hypertension among middle-aged and elderly US
389 hypertensives: analysis based on National Health and Nutrition Examination Survey
390 (NHANES) III. *Hypertension* **37**, 869–874 (2001).

391

392

393 **Figure legends**

394

395 **Figure 1 | Study design for the HCM genome-wide association analysis.** Two independent
396 HCM genome-wide association studies (GWAS) were performed before fixed-effects
397 inverse-variance meta-analysis was conducted. Genetic risk scores were generated and
398 stratified by sarcomere variant status. Findings were validated using three independent
399 cohorts (GeL, Genomics England; RBH, Royal Brompton Hospital; the Netherlands cohort
400 (Amsterdam, Rotterdam and Groningen). Two-sample Mendelian randomization was
401 performed, stratified by sarcomere variant status, to provide insight into heritable risk factors
402 for HCM. SNP-heritability (h^2_g) estimates were compared between component GWAS
403 studies using GREML-LDMS and stratified by sarcomere variant status. Standard errors for
404 h^2_g estimates are presented in square brackets.

405

406 **Figure 2 | Validation of an HCM genetic risk score.** A genetic risk score was generated
407 from 27 SNPs with <5% FDR and weighted by the beta estimate from the multi-ancestry
408 meta-analysis joint model GCTA results. The genetic risk score was evaluated in all-comers,
409 sarcomere-positive and sarcomere-negative HCM cases, in three validation cohorts. **a**, A
410 quintile-based analysis demonstrates the protective effects of the genetic risk score in the
411 lowest 20% of the population, compared with the middle 60%. Similarly, the upper 20%
412 demonstrate increased susceptibility towards risk of developing HCM, when compared with
413 the middle 60%. Odds ratios (x -axis) reported with error bars denoting 95% confidence
414 intervals. **b**, To facilitate comparison between other genetic risk scores, a per standard
415 deviation estimate is reported. Odds ratios (x -axis) reported with error bars denoting 95%
416 confidence intervals. Validation cohorts: GeL (Genomics England HCM cases ($n = 435$) vs.

417 controls ($n = 36,500$); RBH (Royal Brompton Hospital HCM cases ($n = 359$) vs. controls (n
418 $= 1,211$)); and the Netherlands HCM cases ($n = 975$) vs. controls ($n = 2,117$)).

419

420 **Figure 3 | Relationship between standardized genetic risk score (GRS) and maximum**
421 **left ventricular (LV) wall thickness.** Linear regression was performed to assess the most
422 frequently observed HCM variant classes: truncating variants in *MYBPC3*, *MYH7* missense
423 variants, and most frequently observed pathogenic variant (*MYBPC3* p.R502W). **a**, Carriers
424 of pathogenic or likely pathogenic *MYH7* missense variants ($n = 186$, $\beta = 0.73 \pm 0.35$, $P =$
425 0.036). **b**, Carriers of *MYBPC3* truncating variants ($n = 232$, $\beta = 0.71 \pm 0.35$, $P = 0.048$). **c**,
426 Carriers of the most frequently observed pathogenic variant in HCM, *MYBPC3* p.R502W (n
427 $= 48$, $\beta = 1.61 \pm 0.80$, $P = 0.051$) evaluated in HCMR cases ($n = 36$) and participants from
428 the UK Biobank ($n = 12$). Linear regression line is denoted in blue, alongside 95%
429 confidence interval in gray. P values are uncorrected for multiple testing.

430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448

Figure 4 | Two-sample inverse variance weighted Mendelian randomization identifies modifiable risk factors for HCM. **a**, Effect of presumed risk phenotypes, based on prior observational evidence, on sarcomere-positive ($n = 871$) and sarcomere-negative ($n = 1,635$) HCM. Odds ratios represented per standard deviation for SBP, DBP, BMI and WHRadjBMI. Error bars represent 95% confidence intervals. As T2D is a binary phenotype, risk is represented as the per log odds unit of T2D. **b**, Relative impact of DBP on sarcomere-positive and sarcomere-negative HCM susceptibility, in relation to other established hypertension-associated phenotypes. Odds ratio measured per standard deviation of DBP (11.3 mmHg). Error bars represent 95% confidence intervals. Ischemic stroke reflects all TOAST subtypes. Summary of cases/controls included: heart failure (47,309 cases and 930,014 controls), atrial fibrillation (65,446 cases and 522,744 controls), cardioembolic stroke (9,006 cases and 454,450 controls), ischemic stroke (60,341 cases and 454,450 controls), T2D (74,124 cases and 824,006 controls), chronic kidney disease (64,164 cases and 625,219 controls), and coronary artery disease (122,733 cases and 424,528 controls). Abbreviations: BMI, body mass index; DBP, diastolic blood pressure; HCM, hypertrophic cardiomyopathy; SBP, systolic blood pressure; T2D, type 2 diabetes mellitus; WHRadjBMI: waist-hip ratio adjusted for BMI.

449 **ONLINE METHODS**

450

451 **GWAS in multi-ancestry HCMR cases vs. UKBB controls.** As described in Neubauer et
452 al.⁸, 2,755 incident HCM cases were recruited from 44 sites across 6 countries in North
453 America and Europe. Cases were 18 to 65 years of age with evidence of unexplained left
454 ventricular hypertrophy (wall thickness > 15 mm)³¹. All participants provided written
455 informed consent (South Central - Oxford A Research Ethics Committee approval:
456 14/SC/0190; clinicaltrials.gov identifier: NCT01915615). Genotyping was performed using
457 the Axiom™ Precision Medicine Research Array (Affymetrix/ThermoFisher). Following
458 quality control, 2,541 individuals, not closely related (i.e. > 3 degrees of relatedness), were
459 available for analysis. Gene panel sequence data, generated using a custom-designed TruSeq
460 kit (Illumina), were available on 2,636 HCMR cases, as previously reported^{8,32}. Variant
461 classification was performed for the 8 core sarcomere genes
462 (*MYBPC3*, *MYH7*, *TNNI3*, *TNNT2*, *MYL2*, *MYL3*, *ACTC1* and *TPM1*) using the American
463 College of Medical Genetics and Genomics (ACMG) guidelines³³. Cases were systematically
464 dichotomized into sarcomere-positive ($n = 871$) or sarcomere-negative ($n = 1,635$) groups
465 using a published, evidence-based framework (Supplementary Tables 9 and 10)^{8,34}. Details of
466 the rare variants used to partition cases are reported (Supplementary Tables 21 and 22).
467 Access to the UK Biobank (UKBB) genotypes was provided through application 11223 (UK
468 REC approval: 11/NW/0382). Genotyping was performed using the UK Biobank
469 Axiom® array (Affymetrix). Individuals who underwent genotyping using the UK BiLEVE
470 array, or had asked to be withdrawn from the UK Biobank, as of 16 October 2018, were
471 excluded. Individuals with an ICD10 code indicating HCM (I420 or I421), or other
472 phenotypes that may confound HCM analyses (Supplementary Table 23) in Hospital Episode
473 Statistics (HES) data or self-reported questionnaire fields, were excluded ($n = 15,901$).

474 Individuals in the UKBB exome sequencing subset ($n = 49,959$) that harbor variants of
475 uncertain significance (VUS), likely pathogenic or pathogenic variants in the core sarcomere
476 genes were excluded. Closely related individuals, within 3 degrees of relatedness, and gender
477 mismatches were excluded. Of the remaining 270,260 individuals, 40,283 were randomly
478 selected for subsequent analysis, sampled using a 20:1 allocation against HCMR cases ($n =$
479 2,541), with approximate age (per decade) and genotype-assigned sex matching.

480 The HCMR (Precision Medicine Research Array (Affymetrix)) and UKBB (UK
481 Biobank Axiom® array (Affymetrix)) cohorts were genotyped on partially overlapping
482 arrays. 174,974 single nucleotide polymorphisms (SNPs) (MAF > 0.01, genotype missing
483 rate 1%, Hardy-Weinberg equilibrium with mid- P correction of 1×10^{-9}) present in both the
484 HCMR and UKBB cohorts were extracted for subsequent analysis. The UKBB and HCMR
485 SNPs were aligned to the HRC reference panel, using HRC-1000G-check-bim.pl from
486 <https://www.well.ox.ac.uk/~wrayner/tools/>, before being merged.

487 Principal components analysis was then performed using FlashPCA2 on a subset of
488 SNPs in approximate linkage equilibrium ($r^2 < 0.05$), determined using the `–indep-pairwise`
489 function in PLINK (version 1.90b3). Ancestry was inferred by projecting principal
490 components, derived from the 1000 Genomes Project (Phase 3), onto HCMR/UKBB
491 genotypes. A multinomial logistic regression model, performed using the *nnet* CRAN
492 package in R (<https://CRAN.R-project.org/package=nnet>), classified ancestral groups as
493 specified by the International Genome Sample Resource
494 (<http://www.internationalgenome.org/category/population/>) (Supplementary Table 1).
495 The Michigan Imputation Server³⁵ (<https://imputationserver.sph.umich.edu/>) performed
496 haplotype phasing with Eagle³⁶, and imputation against the Haplotype Reference Consortium
497 (HRC.r1.1.2016 reference panel)³⁷, generating genotypes for 38,954,302 imputed variants.

498 Imputed variants with an INFO score > 0.3 and MAF > 0.01 were retained for subsequent
499 analysis.

500 An all-comer analysis (2,541 HCM cases vs. 40,283 controls) and separate
501 sarcomere-positive (871 vs. 20,142) and sarcomere-negative HCM analyses (1,635 vs.
502 20,141) were performed. The UKBB controls were randomly allocated to either the
503 sarcomere-positive or sarcomere-negative GWAS.

504 Analyses were performed with logistic regression to fit an additive case-control
505 association model, using the SNPTEST v2.5.4-beta3 newml function, adjusting for the first
506 ten ancestry-informative principal components. As HCM is a disease of a relatively low
507 prevalence (~1 in 500), statistical power was maximized by not adjusting for age or sex³⁸.
508 There was no evidence of extreme population stratification in genomic control analyses (all-
509 comers: original $\lambda = 1.191$, sarcomere-positive: $\lambda = 1.089$, sarcomere-negative: pre- $\lambda_{GC} =$
510 1.142). A genomic control adjustment was performed when λ exceeded 1.1.

511

512 **Multi-ancestry BioResource for Rare Disease case-control GWAS.** Details regarding the
513 BioResource for Rare Disease (BRRD) cohort, a pilot study of the Genomics England
514 100,000 Genomes Project (GeL), have been described elsewhere³⁹. All participants provided
515 written informed consent (East of England - Cambridge South REC approval: 13/EE/0325).
516 Briefly, 13,037 individuals from 20 rare disease areas underwent genome sequencing,
517 including 243 individuals diagnosed with sarcomere-negative HCM⁴. Individuals clinically
518 diagnosed with HCM, with diagnostic criteria as for HCMR, were recruited via inherited
519 cardiac condition (ICC) clinics within the UK (Oxford University Hospitals NHS Foundation
520 Trust, Royal Brompton & Harefield NHS Foundation Trust, Guy's and St Thomas' NHS
521 Foundation Trust and the Newcastle upon Tyne Hospitals NHS Foundation Trust).
522 Individuals recruited were aged 18 to 70 years old, or >70 years when there was a positive

523 family history, with an absence of likely pathogenic or pathogenic variants across 13 well-
524 established HCM genes (sarcomeric genes:
525 *MYBPC3*, *MYH7*, *TNNI3*, *TNNT2*, *MYL2*, *MYL3*, *ACTC1*, *TPM1*; other non-sarcomeric, but
526 robustly associated HCM genes: *CSRP3*, *PLN*; and phenocopy genes:
527 *PRKAG2*, *GLA*, *FHL1*)⁴.

528 Reference controls were recruited from the other BRRD rare disease participants, or
529 their family members. Individuals recruited via the GeL pilot study for the purposes of
530 investigating an ICC were excluded. Overall, 239 cases and 7,203 controls were available for
531 analysis, and high-quality variants were extracted from the respective genome sequencing
532 variant call format (VCF) files. High quality variants were defined as those that had: PASS
533 filter status; MAF > 1%, a depth of at least 10 informative reads per site (DP > 10); a
534 genotype quality score of at least 20 (GQ > 20); and a genotype missingness of no more than
535 10% (CR > 0.9). Multiallelic sites were split. Ancestrally informative principal components
536 were derived using FastPCA2 and 1000 Genomes Phase 3 data (Supplementary Table 1).
537 Association analysis was performed using SAIGE (v0.29.4.2) with the first three principal
538 components included as covariates⁴⁰. SAIGE step 1 was performed using 123,903 genotypes
539 following a linkage disequilibrium pruning procedure in PLINK (v1.9), with a 500-kb
540 window, a step size of 50 markers, and a pairwise r^2 threshold of 0.2^{41,42}. SAIGE step 2
541 analysis was performed using genotypes with a minor allele count > 5 and MAF > 0.01.
542 Summary genetic association statistics for 9,341,129 autosomal variants were then computed
543 using a mixed logistic regression model; a genomic control analysis showed little evidence of
544 over-dispersion ($\lambda = 1.049$). The BRRD GWAS was included in the all-comer and the
545 sarcomere-negative meta-analyses.

546

547 **HCM sarcomere carrier stratification.** Up to two-thirds of variants of uncertain
548 significance (VUS) in confirmed sarcomere genes are considered causal of HCM³⁴. In order
549 to contrast the common-variant genetic architecture of HCM patients carrying pathogenic
550 variants in sarcomeric genes with non-carriers, individuals were assigned sarcomere-positive
551 status if they harbored a variant classified as either VUS-indeterminate, VUS-favors
552 pathogenic, likely pathogenic or pathogenic in *ACTC1*, *MYH7*, *MYL2*, *MYL3*, *TNNT2*, *TNNI3*
553 and *TPM1*, or a VUS-favors pathogenic, likely pathogenic or pathogenic in *MYBPC3*
554 (Supplementary Tables 9 and 10)⁸.

555

556 **Heritability estimates.** SNP-heritability (h^2_g) was estimated using GREML (Genomic
557 relatedness matrix Restricted Maximum Likelihood)⁴³ for SNPs demonstrating an INFO
558 score > 0.3 and MAF > 0.01. LD scores were assigned to SNPs from 200-kb blocks across
559 the genome, before SNPs were stratified into quartiles based on SNP LD scores to generate
560 genomic relatedness matrices (GRM). The GRMs were subjected to REML analysis of case-
561 control status with the first ten ancestry-informative principal components as covariates. h^2_g
562 estimates were approximated on a liability scale; this represents binary traits (i.e. cases vs.
563 controls) on a continuous scale, and above a liability threshold an individual will be affected.
564 Representation of a binary trait on this classic multifactorial liability scale is dependent on
565 both the population prevalence of disease (0.2% based on population-based epidemiological
566 estimates) and the sample prevalence of disease. The prevalence of sarcomere-negative and
567 positive HCM was set as 0.0012 and 0.0008³⁴.

568

569 **Quality control.** EasyQC (v9.2)⁴⁴ was used for genotype quality control. The HRC reference
570 panel was used for mapping and allele frequencies. Variants were removed if they were
571 monomorphic, demonstrated a minor allele count < 6, were absent from the HRC reference

572 panel or duplicated, or when the observed allele frequency deviated by > 0.2 from the HRC
573 allele frequency.

574 Genomic inflation was assessed across all cohorts through calculation of the genomic
575 control, λ , and by evaluating the overall P -value distribution. Genomic control correction was
576 performed when $\lambda > 1.1$, by adjusting the standard error ($SE_{gc} = SE \times \sqrt{\lambda}$) and re-calculating
577 adjusted χ^2 statistics ($\chi_{gc}^2 = (\frac{\beta}{SE_{gc}})^2$) and associated P -values (under 1 degree of freedom).

578 The overall P -value distribution, generated from each component study and meta-analysis,
579 were plotted and assessed. Local false discovery rates (FDR) were computed by the qvalue R
580 package (<https://github.com/StoreyLab/qvalue>)⁴⁵. The FDR provides a frequentist equivalent
581 to the empirical Bayesian posterior probability that the null hypothesis is true, based on the
582 distribution of generated p -values. For genome-wide significance, an *a priori* alpha threshold
583 of 5×10^{-8} was set, and a FDR threshold of 5% was calculated for each study⁴⁶.

584

585 **Meta-analysis.** All-comer (2,780 HCM cases vs. 47,486 controls) and sarcomere-negative
586 (1,874 cases vs. 27,344 controls) fixed-effects inverse-variance meta-analysis analyses,
587 incorporating the HCMR vs. UKBB and BRRD vs. BRRD component GWAS, were
588 conducted using GWAMA⁴⁷. Effect sizes, standard errors, effect allele frequency estimates
589 and heterogeneity statistics, specifically Cochran's statistic (Q), were reported alongside
590 q -values and FDR values.

591

592 **Replication.** Replication of HCM loci was performed in a smaller, independent dataset
593 composed of three HCM case-control studies from the Netherlands (975 cases, 2,117
594 controls), Royal Brompton Hospital (359 cases, 1,211 controls) and Canada (313 cases, 3,300
595 controls). No cases recruited to HCMR or BRRD were present in these replication cohorts.

596 Meta-analysis of these three GWAS studies was performed using METAL ($\lambda = 1.074$).

597 Detailed methods regarding these replication cohorts are available in Tadros et al.⁴⁸.

598

599 **Conditional association analysis.** To identify genetic variants that confer independent
600 risk effects, conditional association analysis was performed using a stepwise model
601 selection procedure (`--cojo-slc`) using GCTA^{49,50}. Summary statistics were extracted from
602 meta-analyses results, or component GWAS studies, and linkage disequilibrium metrics
603 were based on 62,018 unrelated, European individuals randomly selected from the UKBB.
604 For each analysis, the *P*-value threshold was determined by a 5% FDR level.

605

606 **Assessment of pleiotropy.** All independently associated variants, confirmed in a conditional
607 analysis, were evaluated for association with other diseases or traits by cross-referencing
608 publicly available databases via OpenTargets⁵¹ (<https://genetics.opentargets.org/>), a web
609 resource that synthesises data from both the NHGRI-EBI GWAS catalogue
610 (<https://www.ebi.ac.uk/gwas/>) and previously published UK Biobank summary statistics
611 (Supplementary Table 24)⁴⁰. Associations between sentinel SNPs and their tagging SNPs (r^2
612 > 0.8) and gene expression, proteins, metabolites or epigenetics were evaluated using
613 PhenoScanner v2.0 (Supplementary Table 25)⁵².

614

615 **Genetic correlation.** To measure the genetic correlation between sarcomere-positive and
616 sarcomere-negative HCM, bivariate GREML analysis was performed. The prevalence of
617 sarcomere-negative and sarcomere-positive HCM were set at 0.0012 and 0.0008. The first ten
618 ancestry-informative principal components were included as covariates. To assess the shared
619 genetic effects between sarcomere-positive and sarcomere-negative HCM at individual loci,
620 pairwise GWAS (GWAS-PW) calculated the probability a variant located within locus

621 contributes to either one, both, or neither traits, or whether two separate signals within the
622 same region contribute to each trait independently. GWAS-PW was performed using
623 genomic regions 500 kb upstream and downstream of independent loci from the all-comer,
624 sarcomere-positive and sarcomere-negative GWAS⁵³.

625

626 **Long-range linkage disequilibrium and spurious association.** Independent, genome-wide
627 significant, variants on chromosome 11 (chr11:44,976,681-57,917,265) were identified due
628 to their close proximity to *MYBPC3*, an HCM gene known to contain pathogenic founder
629 variants, raising the possibility of long-range linkage disequilibrium and spurious association.
630 Haplotypes were constructed using genotyped and imputed common variants (MAF > 0.2), in
631 combination with rare *MYBPC3* variants (p.R502W, p.Trp792ValfsTer41 and c.1224-
632 52G>A) derived from gene panel data (HCM cases) or exome data (UKBB controls), in
633 PLINK (version 1.90b3). HCM cases were limited to those individuals in whom a pathogenic
634 or likely pathogenic variant in a core sarcomere gene had previously been identified ($n =$
635 851). Controls were drawn from the UK Biobank ($n = 19,851$). Haplotype structure was
636 evaluated using a maximum likelihood method in Haploview (version 4.2)⁵⁴. Given the
637 presence of numerous zero-value genotype counts, multiple logistic regression association
638 analysis was performed using the R *logistf* function. This method was used to model the
639 independent effects of a *rare pathogenic variant and a common* variant on HCM risk, while
640 allowing for linkage disequilibrium between the two variants.

641

642 **Genetic risk score.** After removing SNPs demonstrating extreme pleiotropy or extreme
643 ancestral bias (see Supplementary Note), a genetic risk score (GRS) that combines
644 independent (i.e. in linkage equilibrium) SNPs identified through the conditional analysis of
645 the multi-ancestry meta-analysis was first tested in a component GWAS (HCMR vs. UKBB)

646 then validated in three independent studies: Genomics England's 100,000 Genomes Project
647 (GeL) (REC: 14/EE/1112), the Royal Brompton Hospital's HCM case-control series, and a
648 Netherlands HCM case-control series (Supplementary Table 16). Individuals recruited to
649 both a discovery cohort and a validation cohort were identified and excluded from the
650 validation cohort (51 individuals from the Royal Brompton Hospital series and 24 individuals
651 from the Netherlands series). The cumulative genetic effect of the SNPs was calculated for
652 each individual using the allelic scoring function in PLINK. The relative weight assigned to
653 each SNP was the beta estimate from the multi-ancestry meta-analysis joint model COJO
654 results. Raw GRSs were plotted and evaluated, before standardizing the GRS distribution to a
655 mean of 0 and variance of 1. A logistic regression model was fitted with affection status as
656 the outcome variable and standardized GRS score as an explanatory variable, with covariates
657 including the first ten principal components, age and gender. Cases from each validation
658 dataset were dichotomized based on the presence of a rare causal variant in an established
659 sarcomere gene into sarcomere-positive and sarcomere-negative cases (see Supplementary
660 Note).

661

662 **Expressivity analysis.** The two predominant classes of pathogenic HCM variants are
663 *MYBPC3* truncating variants and *MYH7* missense variants, the mechanisms of which have
664 been previously demonstrated^{55,56}. The most frequently detected pathogenic variant, present
665 in ~2% of HCM cases, is *MYBPC3* p.R502W. Variants, classified using ACMG guidelines as
666 pathogenic or likely pathogenic in *MYBPC3* and *MYH7* were identified^{33,57}. Ensembl's
667 Variant Effect Predictor was used to define variant consequences. Truncating variants
668 included frameshift, stop gained and splice acceptor/donor variants.

669 For individuals in the HCMR cohort, maximum wall thickness measurements were
670 derived from cardiac magnetic resonance (CMR) imaging (either 1.5-T or 3.0-T) performed

671 using a standardized protocol, multichannel phased-array chest coils and electrocardiographic
672 gating, as previously reported⁸.

673 Variant carriers for *MYBPC3* p.R502W were identified in both the HCMR and UKBB
674 cohorts. In the UKBB cohort, heterozygous *MYBPC3* p.R502W carriers, who had also
675 undertaken UKBB based CMR imaging, were identified using array-based genotypes ($n =$
676 12). Exome sequence data was available for six of these individuals; in all cases the presence
677 of *MYBPC3* p.R502W was confirmed, supporting prior analysis indicating that UKBB-based
678 genotyping for *MYBPC3* p.R502W was satisfactory^{58,59}. Demographic and phenotypic details
679 were reviewed for *MYBPC3* p.R502W carriers, including ICD10 classifications and self-
680 reported co-morbidities for HCM codes (I421 and I422). *MYBPC3* p.R502W carriers were
681 age and sex matched 1:1 with a non-variant carrier from a sample of unrelated UKBB
682 participants who had undertaken exome sequencing and CMR imaging and demonstrated no
683 potentially disease-causing variant (i.e. no HCM associated variant of uncertain significance,
684 likely pathogenic variant or pathogenic variant) or a previous diagnosis, reported via HES
685 data, that might confound CMR analyses (Supplementary Table 23). Long and short axis
686 cine-tagged CMR data, generated by the UKBB as previously reported, were reviewed by an
687 investigator blinded to variant carrier status, to minimise bias, when reporting maximum left
688 ventricular wall thickness⁶⁰. A linear regression model was used to approximate the effect of
689 a standardised GRS (mean of 0 and variance of 1) against maximum left ventricular wall
690 thickness in mm.

691

692 **Mendelian randomization.** Mendelian randomization (MR) uses the random meiotic
693 segregation of alleles to assess whether an association between a risk factor and an outcome
694 is consistent with a causal effect. Two-sample Mendelian randomization (2SMR) leverages
695 data from large-scale genome-wide association studies to infer causal relationships between

696 two heritable traits. Observational data suggest several modifiable risk factors that may
697 influence the phenotypic variability observed in HCM^{18,19,61,62}. Genome-wide significant loci
698 for blood pressure (systolic blood pressure, diastolic blood pressure)²¹, body mass
699 index/waist-hip ratio²² and type 2 diabetes²⁰ were identified and summary statistics were
700 collated as instrumental variables. Effect alleles were harmonized between instrumental
701 variables and the HCM summary statistics. Analyses were performed using MR-base⁶³. The
702 corresponding SNPs were extracted from the sarcomere-positive and sarcomere-negative
703 summary statistics and MR estimates generated using fixed and random-effects inverse-
704 variance weighted (IVW) MR. Sensitivity analysis to test for horizontal pleiotropy was
705 performed using MR-Egger, and for robustness by unweighted and weighted median
706 regression. Betas and standard errors were compared, for all risk factors, between sarcomere-
707 positive and sarcomere-negative HCM. Additional 2SMR was performed to further evaluate
708 the relative effect of diastolic blood pressure on HCM, relative to other well-established risk
709 factor-disease relationships (Supplementary Table 20).

710

711 **Functional mapping and annotation.** Functional annotation of GWAS summary statistics
712 was undertaken using FUMA (v1.3.5e) (<https://fuma.ctglab.nl/snp2gene>)⁶⁴ to link genotypes,
713 eQTLs and chromatin interactions. Tissue enrichment was performed using MAGMA (Multi-
714 marker Analysis of GenoMic Annotation) (v1.07), a gene-level analysis tool provided by
715 FUMA, with tissue expression data from GTEx (version 8.0)
716 (<https://www.gtexportal.org/home>)⁷. eQTLs were evaluated in heart tissue (atrial appendage
717 and left ventricle). Chromatin interaction data was evaluated in left and right ventricular
718 tissue, derived from previously reported Hi-C data (GSE87112)⁶⁵.

719

720 **Functional GWAS.** Functional GWAS (fGWAS) is software that assesses enrichment of

721 functional sites (such as histone marks and methylation data) within GWAS summary
722 statistics⁶⁶. fGWAS then uses these enrichment parameters to fine map and re-weight GWAS
723 loci.

724 Using multi-ancestry HCM meta-analysis summary statistics, fGWAS was performed using
725 chromatin marks (enhancers, flanking/active TSS, active TSS, genetic enhancers, repressed
726 polycomb, bivalent enhancers, transcription at gene 5' and 3', flanking/bivalent
727 TSS/enhancers and bivalent/poised TSS) from the ChromHMM dataset of the Roadmap
728 Epigenomic project, for cardiac tissues (left ventricle (E095), fetal heart (E083), right
729 ventricle (E105), right atrium (E104))^{67,68}. Enrichment estimates were generated for each
730 tissue type, together with a list of loci below a 5% FDR significance threshold.**Data**

731 **availability**

732 We confirm that all relevant data are included in the paper and/ or its supplementary
733 information files. The datasets generated during this study are available from the
734 corresponding author upon reasonable request. The following institutional domain
735 (www.well.ox.ac.uk/hcm) will provide summary level statistics.

736

737 **Code availability**

738 Publicly available software tools were used to analyze these data. This includes: SAIGE
739 (<https://github.com/weizhouUMICH/SAIGE>); SNPTEST
740 (https://mathgen.stats.ox.ac.uk/genetics_software/snptest/snptest.html); GCTA
741 (<https://cnsgenomics.com/software/gcta/>); PLINK ([https://www.cog-
genomics.org/plink/1.9/data](https://www.cog-
genomics.org/plink/1.9/data)); BGENIX (<https://bitbucket.org/gavinband/bgen/wiki/bgenix>);
743 QCTOOL (https://www.well.ox.ac.uk/~gav/qctool_v2/);
744 GWAMA (<https://genomics.ut.ee/en/tools/gwama>); and MR-base (<http://www.mrbase.org/>).

745 **Methods-only references**

746

- 747 31. Gersh, B. J. *et al.* 2011 ACCF/AHA guideline for the diagnosis and treatment of
748 hypertrophic cardiomyopathy: a report of the American College of Cardiology
749 Foundation/American Heart Association Task Force on Practice Guidelines.
750 *Circulation* **124**, e783-831 (2011).
- 751 32. Harper, A. R. *et al.* A re-evaluation of the South Asian *MYBPC3*^{Δ25bp} intronic deletion
752 in hypertrophic cardiomyopathy. *Circ. Genomic Precis. Med.* **13**, e002783 (2020).
- 753 33. Richards, S. *et al.* Standards and guidelines for the interpretation of sequence variants :
754 a joint consensus recommendation of the American College of Medical Genetics and
755 Genomics and the Association for Molecular Pathology. *Genet. Med.* **17**, 405–424
756 (2015).
- 757 34. Walsh, R. *et al.* Reassessment of Mendelian gene pathogenicity using 7,855
758 cardiomyopathy cases and 60,706 reference samples. *Genet. Med.* **19**, 192–203 (2017).
- 759 35. Das, S. *et al.* Next-generation genotype imputation service and methods. *Nat. Genet.*
760 **48**, 1284–1287 (2016).
- 761 36. Loh, P. R., Palamara, P. F. & Price, A. L. Fast and accurate long-range phasing in a
762 UK Biobank cohort. *Nat. Genet.* **48**, 811–816 (2016).
- 763 37. McCarthy, S. *et al.* A reference panel of 64,976 haplotypes for genotype imputation.
764 *Nat. Genet.* **48**, 1279–1283 (2016).
- 765 38. Pirinen, M., Donnelly, P. & Spencer, C. C. A. Including known covariates can reduce
766 power to detect genetic effects in case-control studies. *Nat. Genet.* **44**, 848–851 (2012).
- 767 39. Turro, E. *et al.* Whole-genome sequencing of patients with rare diseases in a national
768 health system. *Nature* **583**, 96–102 (2020).
- 769 40. Zhou, W. *et al.* Efficiently controlling for case-control imbalance and sample
770 relatedness in large-scale genetic association studies. *Nat. Genet.* **50**, 1335–1341
771 (2018).
- 772 41. Chang, C. C. *et al.* Second-generation PLINK: rising to the challenge of larger and
773 richer datasets. *Gigascience* **4**, 7 (2015).
- 774 42. Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based
775 linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
- 776 43. Yang, J., Zeng, J., Goddard, M. E., Wray, N. R. & Visscher, P. M. Concepts,
777 estimation and interpretation of SNP-based heritability. *Nat. Genet.* **49**, 1304–1310
778 (2017).
- 779 44. Winkler, T. W. *et al.* Quality control and conduct of genome-wide association meta-
780 analyses. *Nat. Protoc.* **9**, 1192–1212 (2014).
- 781 45. Storey, J. D. & Tibshirani, R. Statistical significance for genomewide studies. *Proc.*
782 *Natl. Acad. Sci. USA* **100**, 9440–9445 (2003).
- 783 46. Nelson, C. P. *et al.* Association analyses based on false discovery rate implicate new
784 loci for coronary artery disease. *Nat. Genet.* **49**, 1385–1391 (2017).
- 785 47. Mägi, R. & Morris, A. P. GWAMA: software for genome-wide association meta-
786 analysis. *BMC Bioinformatics* **11**, 288 (2010).
- 787 48. Tadros, R. *et al.* Shared genetic pathways contribute to risk of hypertrophic and dilated
788 cardiomyopathies with opposite directions of effect. *Nat. Genet.* **[CE: add reference**
789 **and doi for NG-LE54093R1 Tadros]**
- 790 49. Yang, J., Lee, S. H., Goddard, M. E. & Visscher, P. M. GCTA: a tool for genome-wide
791 complex trait analysis. *Am. J. Hum. Genet.* **88**, 76–82 (2011).
- 792 50. Yang, J. *et al.* Conditional and joint multiple-SNP analysis of GWAS summary
793 statistics identifies additional variants influencing complex traits. *Nat. Genet.* **44**, 369–
794 375 (2012).

- 795 51. Carvalho-Silva, D. *et al.* Open Targets Platform: new developments and updates two
796 years on. *Nucleic Acids Res.* **47**, D1056–D1065 (2019).
- 797 52. Kamat, M. A. *et al.* PhenoScanner V2: an expanded tool for searching human
798 genotype-phenotype associations. *Bioinformatics* **35**, 4851–4853 (2019).
- 799 53. Pickrell, J. K. *et al.* Detection and interpretation of shared genetic influences on 42
800 human traits. *Nat. Genet.* **48**, 709–717 (2016).
- 801 54. Barrett, J. C., Fry, B., Maller, J. & Daly, M. J. Haploview: analysis and visualization
802 of LD and haplotype maps. *Bioinformatics* **21**, 263–265 (2005).
- 803 55. Toepfer, C. N. *et al.* Hypertrophic cardiomyopathy mutations in *MYBPC3* dysregulate
804 myosin. *Sci. Transl. Med.* **11**, eaat1199 (2019).
- 805 56. Toepfer, C. N. *et al.* Myosin sequestration regulates sarcomere function,
806 cardiomyocyte energetics, and metabolism, informing the pathogenesis of
807 hypertrophic cardiomyopathy. *Circulation* **141**, 828–842 (2020).
- 808 57. Kelly, M. A. *et al.* Adaptation and validation of the ACMG/AMP variant classification
809 framework for MYH7-associated inherited cardiomyopathies: recommendations by
810 ClinGen’s Inherited Cardiomyopathy Expert Panel. *Genet. Med.* **20**, 351–359 (2018).
- 811 58. Wright, C. F. *et al.* Assessing the pathogenicity, penetrance, and expressivity of
812 putative disease-causing variants in a population setting. *Am. J. Hum. Genet.* **104**,
813 275–286 (2019).
- 814 59. Van Hout, C. V. *et al.* Exome sequencing and characterization of coding variation in
815 49,960 individuals in the UK Biobank. *Nature* **586**, 749–756 (2020).
- 816 60. Petersen, S. E. *et al.* UK Biobank’s cardiovascular magnetic resonance protocol. *J.*
817 *Cardiovasc. Magn. Reson.* **18**, 8 (2016).
- 818 61. Karam, R., Lever, H. M. & Healy, B. P. Hypertensive hypertrophic cardiomyopathy or
819 hypertrophic cardiomyopathy with hypertension? A study of 78 patients. *J. Am. Coll.*
820 *Cardiol.* **13**, 580–584 (1989).
- 821 62. Wasserstrum, Y. *et al.* The impact of diabetes mellitus on the clinical phenotype of
822 hypertrophic cardiomyopathy. *Eur. Heart J.* **40**, 1671–1677 (2019).
- 823 63. Hemani, G. *et al.* The MR-Base platform supports systematic causal inference across
824 the human phenome. *Elife* **7**, e34408 (2018).
- 825 64. Watanabe, K., Taskesen, E., Van Bochoven, A. & Posthuma, D. Functional mapping
826 and annotation of genetic associations with FUMA. *Nat. Commun.* **8**, 1826 (2017).
- 827 65. Schmitt, A. D. *et al.* A compendium of chromatin contact maps reveals spatially active
828 regions in the human genome. *Cell Rep.* **17**, 2042–2059 (2016).
- 829 66. Pickrell, J. K. Joint analysis of functional genomic data and genome-wide association
830 studies of 18 human traits. *Am. J. Hum. Genet.* **94**, 559–573 (2014).
- 831 67. Ernst, J. & Kellis, M. ChromHMM: automating chromatin-state discovery and
832 characterization. *Nat. Methods* **3**, 215–216 (2012).
- 833 68. Roadmap Epigenomics Consortium *et al.* Integrative analysis of 111 reference human
834 epigenomes. *Nature* **518**, 317–330 (2015).
- 835
836
837
838

839
840

Tables

Chr	SNP	Position (GRCh37)	NEA/EA	Freq EA	Discovery			Replication			Locus name
					OR	95% CI	P	OR	95% CI	P	
Genome-wide significant ($P < 5 \times 10^{-8}$)											
1	rs1048302	16,340,879	G/T	0.33	1.32	1.24-1.40	2.54×10^{-17}	1.26	1.16-1.37	$1.06 \times 10^{-7*}$	<i>HSPB7</i>
3	rs13061705	14,291,129	T/C	0.69	1.25	1.16-1.34	9.18×10^{-9}	1.13	1.04-1.24	4.49×10^{-3}	<i>SLC6A6</i>
6	rs3176326	36,647,289	G/A	0.21	1.28	1.19-1.38	2.22×10^{-11}	1.27	1.15-1.40	$1.87 \times 10^{-6*}$	<i>CDKN1A</i>
6	rs12212795	118,654,308	G/C	0.05	1.48	1.31-1.67	2.51×10^{-10}	1.72	1.45-2.05	$8.19 \times 10^{-10*}$	<i>PLN</i>
10	rs72840788	121,415,685	G/A	0.21	1.52	1.42-1.64	5.06×10^{-29}	1.42	1.29-1.56	$4.90 \times 10^{-13*}$	<i>BAG3</i>
12	rs7301677	115,381,147	T/C	0.73	1.24	1.15-1.33	1.26×10^{-8}	1.19	1.08-1.31	$2.74 \times 10^{-4*}$	<i>TBX3</i>
13	rs41306688	114,078,558	A/C	0.03	1.82	1.53-2.17	1.08×10^{-11}	1.38	1.10-1.73	5.03×10^{-3}	<i>ADPRHL1</i>
15	rs8033459	85,253,258	C/T	0.47	1.21	1.14-1.29	3.41×10^{-9}	1.18	1.09-1.28	$5.49 \times 10^{-5*}$	<i>ALPK3</i>
17	rs28768976	43,688,317	A/G	0.23	1.29	1.20-1.39	4.12×10^{-12}	1.29	1.17-1.42	$2.11 \times 10^{-7*}$	<i>SPPL2C</i>
17	rs7210446	64,307,014	G/A	0.58	1.25	1.16-1.34	6.82×10^{-10}	1.25	1.15-1.35	$8.93 \times 10^{-8*}$	<i>PRKCA</i>
18	rs4799426	34,280,891	A/G	0.35	1.38	1.29-1.47	4.00×10^{-23}	1.44	1.32-1.57	$1.13 \times 10^{-16*}$	<i>FHOD3</i>
18	rs118060942	34,280,732	C/T	0.01	1.79	1.45-2.20	2.35×10^{-8}	2.70	1.80-4.04	$1.49 \times 10^{-6*}$	<i>FHOD3</i>
22	rs2070458	24,159,307	T/A	0.22	1.34	1.25-1.44	7.12×10^{-15}	1.25	1.12-1.38	$2.81 \times 10^{-5*}$	<i>MMP11</i>

841
842
843
844
845
846
847
848
849
850
851
852
853
854
855

Table 1 | Variants independently associated with hypertrophic cardiomyopathy beneath the genome-wide significant threshold. Fixed-effects inverse-variance meta-analysis was performed for two multi-ancestry genome-wide association studies (Hypertrophic Cardiomyopathy Registry (HCMR) cases ($n = 2,541$) vs. UK Biobank ($n = 40,283$)) and BRRD (BioResource for Rare Disease hypertrophic cardiomyopathy cases ($n = 239$) vs. controls ($n = 7,203$)) using 8,590,397 SNPs. Variants conferring independent risk effects were identified through a stepwise model selection procedure. Other than *FHOD3*, all loci are novel. Most loci do not overlap with the myofilament genes known to carry rare variants causal for hypertrophic cardiomyopathy, but instead highlight important homeostatic pathways. 10 of the 12 genome-wide significant variants replicate. Replication was performed in non-overlapping cases from the Netherlands, Canada and Royal Brompton discovery series ($n = 1,643$ cases, 6,628 controls) with significant loci (alpha threshold = 0.05/29 (total number of independent variants beneath 5% FDR threshold, i.e. $P < 1.72 \times 10^{-3}$) highlighted (*). No significant heterogeneity was detected between the discovery and replication studies ($P > 0.05$) (Supplementary Table 3). Abbreviations: CI, confidence interval; EA, effect allele; FDR, false discovery rate; Freq EA, effect allele frequency; GWAS, genome-wide association study; NEA, non-effect allele; OR, odds ratio; SNP, single nucleotide polymorphism. *P* values are uncorrected for multiple testing.

Chr	SNP	NEA/ EA	Beta	Discovery <i>P</i>	Replication <i>P</i>	Locus name
10	rs72840788	A/G	-0.421	5.06 x 10 ⁻²⁹	4.90 x 10 ^{-13*}	<i>BAG3</i>
18	rs4799426	G/A	-0.321	4.00 x 10 ⁻²³	1.13 x 10 ^{-16*}	<i>FHOD3</i>
1	rs1048302	G/T	0.277	2.54 x 10 ⁻¹⁷	1.06 x 10 ^{-7*}	<i>HSPB7</i>
22	rs2070458	T/A	0.293	7.12 x 10 ⁻¹⁵	2.81 x 10 ^{-5*}	<i>MMP11</i>
13	rs41306688	C/A	-0.601	1.08 x 10 ⁻¹¹	5.03 x 10 ⁻³	<i>ADPRHL1</i>
6	rs3176326	A/G	-0.247	2.22 x 10 ⁻¹¹	1.87 x 10 ^{-6*}	<i>CDKN1A</i>
6	rs12212795	C/G	-0.393	2.51 x 10 ⁻¹⁰	8.19 x 10 ^{-10*}	<i>PLN</i>
17	rs7210446	A/G	-0.22	6.82 x 10 ⁻¹⁰	8.93 x 10 ^{-8*}	<i>PRKCA</i>
15	rs8033459	T/C	-0.19	3.41 x 10 ⁻⁹	5.49 x 10 ^{-5*}	<i>ALPK3</i>
3	rs13061705	T/C	0.224	9.18 x 10 ⁻⁹	4.49 x 10 ⁻³	<i>SLC6A6</i>
12	rs7301677	T/C	0.214	1.26 x 10 ⁻⁸	2.74 x 10 ^{-4*}	<i>TBX3</i>
18	rs118060942	T/C	-0.799	2.35 x 10 ⁻⁸	1.49 x 10 ^{-6*}	<i>FHOD3</i>
6	rs9320939	A/G	-0.174	5.78 x 10 ⁻⁸	4.48 x 10 ⁻²	<i>TRDN</i>
5	rs10052399	C/T	0.206	6.21 x 10 ⁻⁸	4.42 x 10 ^{-6*}	<i>PROB1</i>
10	rs11196085	C/T	-0.19	7.30 x 10 ⁻⁸	1.99 x 10 ^{-6*}	<i>TCF7L2</i>
2	rs2003585	C/T	0.174	8.60 x 10 ⁻⁸	1.12 x 10 ⁻²	<i>STRN</i>
9	rs734638	G/C	-0.186	1.09 x 10 ⁻⁷	3.39 x 10 ⁻¹	<i>RAPGEF1, POMT1</i>
5	rs66761011	G/A	-0.349	1.45 x 10 ⁻⁷	5.18 x 10 ⁻¹	<i>AK098570</i>
15	rs1814880	C/T	0.179	1.59 x 10 ⁻⁷	4.84 x 10 ⁻³	<i>CHRN4</i>
12	rs1480036	C/T	-0.211	3.50 x 10 ⁻⁷	2.60 x 10 ⁻¹	<i>SSPN</i>
3	rs4894803	G/A	-0.179	3.51 x 10 ⁻⁷	9.91 x 10 ^{-6*}	<i>FNDC3B</i>
2	rs7556984	A/G	0.186	5.21 x 10 ⁻⁷	4.94 x 10 ⁻²	<i>E2F6, ROCK2</i>
2	rs62177303	T/C	0.175	7.00 x 10 ⁻⁷	7.79 x 10 ⁻³	<i>TTN</i>
19	rs117710064	T/C	-0.222	8.55 x 10 ⁻⁷	9.36 x 10 ⁻¹	<i>AZU1</i>
21	rs2832230	T/G	0.251	9.57 x 10 ⁻⁷	2.49 x 10 ⁻³	<i>MAP3K7CL</i>
8	rs7003871	T/C	0.169	1.14 x 10 ⁻⁶	4.21 x 10 ⁻²	<i>MTSS1</i>
11	rs1390519	G/A	-0.203	1.16 x 10 ⁻⁶	2.44 x 10 ⁻²	<i>CYP2R1</i>

857

858

859

860

861

Table 2 | Variants independently associated with hypertrophic cardiomyopathy beneath the 5% FDR threshold used in a genetic risk score instrument. Beta estimates derived from a fixed-effects inverse-variance meta-analysis that incorporated two multi-ancestry genome-wide association studies (Hypertrophic Cardiomyopathy Registry (HCMR) cases ($n = 2,541$) vs. UK Biobank ($n = 40,283$)) and BRRD (BioResource for Rare Disease hypertrophic cardiomyopathy cases ($n = 239$) vs. controls ($n = 7,203$)) using 8,590,397 SNPs. Variants conferring independent

862 risk effects beneath a 5% FDR threshold ($P < 1.82 \times 10^{-6}$) were identified through a stepwise model selection procedure. Abbreviations: Chr,
863 chromosome; EA, effect allele; FDR, false discovery rate; NEA, non-effect allele; SNP, single nucleotide polymorphism. *denotes independent
864 replication at Bonferroni-corrected significance level ($P < 1.72 \times 10^{-3}$), as described in Table 1. P values are uncorrected for multiple testing.

