

1 **Genetic determinants of interventricular septal anatomy and the risk of ventricular septal**
2 **defects and hypertrophic cardiomyopathy.**

3
4 Mengyao Yu PhD^{1,2*}, Andrew R. Harper MRCP DPhil^{3,4,5*}, Matthew Aguirre AB^{1,6}, Maureen
5 Pittman^{7,8}, Catherine Tcheandjieu DVM PhD^{1,2,9}, Dulguun Amgalan PhD^{10,11}, Christopher
6 Grace PhD^{3,4}, Anuj Goel MBBS MSc^{3,4}, Martin Farrall FRCPath^{3,4}, Ke Xiao MS¹², Jesse
7 Engreitz PhD^{10,11}, Katherine Pollard PhD^{7,8,13}, Hugh Watkins MD PhD^{3,4}, James R. Priest
8 MD^{1,2,13,14}

9
10 Affiliations:

- 11 1) Department of Pediatrics, Division of Pediatric Cardiology, Stanford University School
12 of Medicine, Stanford, California, USA
- 13 2) Stanford Cardiovascular Institute, Stanford University, Stanford, California, USA
- 14 3) Radcliffe Department of Medicine, University of Oxford, Division of Cardiovascular
15 Medicine, John Radcliffe Hospital, Oxford, UK.
- 16 4) Wellcome Centre for Human Genetics, Roosevelt Drive, Oxford, UK.
- 17 5) Centre for Genomics Research, Discovery Sciences, BioPharmaceuticals R&D,
18 AstraZeneca, Cambridge, UK
- 19 6) Department of Biomedical Data Science, Stanford Medical School
- 20 7) University of California, San Francisco, San Francisco, CA, USA
- 21 8) Gladstone Institute of Data Science & Biotechnology, San Francisco, CA, USA
- 22 9) Department of Medicine, Division of Cardiovascular Medicine, Stanford University
23 School of Medicine, Stanford, California, USA
- 24 10) Department of Genetics, Stanford University, Stanford, CA, USA
- 25 11) Basic Sciences and Engineering Initiative, Betty Irene Moore Children's Heart Center,
26 Lucile Packard Children's Hospital, Stanford, CA, USA
- 27 12) College of Information & Computer Sciences at University of Massachusetts Amherst,
28 Amherst, Massachusetts USA
- 29 13) Chan-Zuckerberg Biohub, San Francisco, California, USA
- 30 14) Current affiliation: Tenaya Therapeutics, South San Francisco, California, USA

31
32
33 *These authors contributed equally to this work

34 **ABSTRACT**

35 **Background:** The interventricular septum (IVS) plays a primary role in cardiovascular
36 physiology and a large proportion of genetic risk remains unexplained for structural heart disease
37 involving the IVS such as hypertrophic cardiomyopathy (HCM) and ventricular septal defects
38 (VSD).

39
40 **Objectives:** We sought to develop a reproducible proxy of IVS structure from standard medical
41 imaging, discover novel genetic determinants of IVS structure, and relate these loci to two rare
42 diseases of the IVS.

43
44 **Methods:** We performed machine learning to estimate the cross-sectional area of the
45 interventricular septum (IVS.csad) obtained from the 4-chamber view of cardiac MRI in 32,219
46 individuals from the UK Biobank. Using these extracted measurement of IVS.csad we performed
47 phenome-wide association to relate this proxy measure to relevant clinical phenotypes, followed
48 by genome-wide association studies and Mendelian Randomization.

49
50 **Results:** Automated measures of IVS.csad were highly accurate, and strongly correlated with
51 anthropometric measures, blood pressure, and diagnostic codes related to cardiovascular
52 physiology. A Single nucleotide polymorphism in the intron of *CDKN1A* was associated with
53 IVS.csad (rs2376620, Beta 8.4 mm², 95% confidence intervals (CI) 5.8 to 11.0, p=2.0e-10), and
54 a common inversion incorporating *KANSL1* predicted to disrupt local chromatin structure was
55 associated with an increase in IVS.csad (Beta 8.6 mm², 95% CI 6.3-10.9, p=1.3e-13). Mendelian
56 Randomization suggested that inheritance of a larger IVS.csad was causal for HCM (Beta 2.45

57 log odds ratio (OR) HCM per increase in SD of IVS.csad, standard error (SE) 0.48, pIVW =
58 2.8e-7) while inheritance of a smaller IVS.csad was causal for VSD (Beta -2.06 log odds ratio
59 (OR) VSD per decrease in SD of IVS.csad, SE 0.75, pIVW = 0.006)

60

61 **Conclusion:** Automated derivation of the cross sectional area of the IVS from the 4-chamber
62 view allowed discovery of loci mapping to genes related to cardiac development and Mendelian
63 disease. Inheritance of a genetic liability for either large or small interventricular septum, appears
64 to confer risk for HCM or VSD respectively, which suggests that a considerable proportion of
65 risk for structural and congenital heart disease may be localized to the common genetic
66 determinants of cardiovascular anatomy.

67

68

69 INTRODUCTION

70 Within the cardiac ventricles, the interventricular septum (IVS) separates deoxygenated
71 pulmonary blood flow from oxygenated systemic blood flow and is a foundation of mammalian
72 physiology¹. The IVS is the site of two important diseases; ventricular septal defects (VSD) are
73 among the most common forms of congenital heart disease, where a thin or abnormally
74 developed IVS leaves a communication linking the lumens of the left and right ventricle which
75 when untreated may lead to heart failure and pulmonary hypertension². Alternatively, the
76 increased thickness and altered geometry of the IVS in hypertrophic cardiomyopathy (HCM)
77 may progress to the point of obstruction of the left-ventricular outflow tract.³ While both VSD
78 and HCM are primarily genetic in origin, Mendelian inheritance does not fully account for the
79 risk of HCM⁴ or for individual variability in disease amongst individuals carrying the same
80 pathogenic variant. Monogenic disease causing VSD is very rare, and the genetic basis of VSD is
81 simply not known in the majority of affected individuals⁵.

82
83 The IVS has two primary developmental origins, with the membranous portion arising from
84 mesenchymal lineages in the cardiac cushions⁶ and the muscular portion developing from
85 ingrowth of the primary heart tube⁷. A complex microanatomical orientation of cardiomyocyte
86 fibers through the IVS links the function of the right and left ventricles⁸. After adaptation to
87 postnatal circulation over the first months of life, the mature IVS has a complex relationship with
88 normal physiology, with the thickness of the IVS increased in both exercise capacity⁹ as well as
89 in hypertension¹⁰.

90

91 Here we derive a simple automated measure of the IVS derived from MRI imaging of the heart
92 in the 4-chamber view, IVS cross-sectional area at diastole (IVS.csad). We show that IVS.csad is
93 correlated with standard clinical measures and is useful as suitable proxy for the architecture of
94 the IVS. We report genetic correlates of IVS.csad in 32,219 individuals derived from cardiac
95 MRI of the UK Biobank. Using Mendelian Randomization, we describe new causal relationships
96 linking smaller IVS.csad with risk for VSD and larger IVS.csad with risk for HCM.

97 **METHODS**

98 **Ethics statement.** The NHS National Research Ethics Service (ref: 11/NW/0382) granted ethical
99 approval for the distribution of deidentified imaging, genetic, and medical record data from the
100 UK Biobank (UKB) for any qualified researcher.

101

102 **UK Biobank cohort and IVS cross sectional area measurement from cardiac MRI**

103 We built a U-Net segmentation model of cine images of the 4-chamber (4Ch) view of the heart
104 with pre-trained weights from VGG11, which was further trained on 60 hand-labeled 4Ch
105 images, validated on 20 hand-labeled images, and tested on 20 hand-labeled images¹¹. This
106 approach yielded a validation dice score of 91.2% and a test dice score of 93.8% which were
107 equivalent to differences observed between expert human annotators in an open-source deep
108 learning framework upon which this work is based¹². We then applied the segmentation model to
109 32,219 4Ch images from the UKB and generated masks for all 4 chambers including the left and
110 right ventricles and the left and right atrium. Following generation of the masks from the trained
111 U-Net segmentation model, we used another function to measure the interventricular septal
112 areas. We first located the base of the atriums and apex of the ventricles along the medial axis of
113 the chambers, and secondly we located the annuli of the atrioventricular valves using the
114 intersecting lines of the atriums and the ventricles, then finally we use the contour lines of the
115 masks of atriums and ventricles between the annulus lines and the base of the atriums or the apex
116 of the ventricles to enclose the area of the interventricular septum. The metadata in the dicom
117 measurements were used to convert from pixels² to mm². The frame representing end-diastole
118 was obtained by selecting the image frame with the largest estimate of left ventricular volume
119 provided by the U-Net segmentation algorithm.

120
121 To exclude outliers related to imaging error or methodological inaccuracy, automated
122 measurements were plotted relative to body surface area with standard measures of quality
123 control. We excluded measures +/- 3 standard deviations. Manual annotation of 50 randomly
124 selected images spanning the cardiac cycle was performed by a clinician (JRP) blinded to
125 automated measures or anthropometric characteristics. The percent difference between
126 automated and manual measurements ($[\text{automated measure} - \text{manual measure}] / \text{manual measure}$)
127 was examined for systematic relationships to anthropomorphic predictors (Age in years, body
128 surface area, genetic sex). Body surface area was estimated from height and weight using the
129 Haycock formula¹³.

130
131 **Phenome wide association study**
132 Given that IVS.csad is not a standard clinical measurement, using previously described
133 methods¹⁴ we performed a phenome-wide association studies (PheWAS) to highlight clinical
134 associations with 67 cardiovascular phenotypes aggregated from ICD-10 codes as phecodes¹⁵.
135 For the reporting of PheWAS results, we excluded phenotypes with less than 50 individuals (for
136 continuous traits) or less than 50 cases (for binary-coded traits), and controlled the positive false
137 discovery rate (pFDR). We performed PheWAS for the IVS.csad amongst individuals with MRI
138 data (n=32,219).

139
140 **GWAS, Burden testing, and Annotation**
141 The UK Biobank data release available at the time of analysis included genotypes for 488,377
142 participants, obtained through either the custom UK Biobank Axiom array or the Affymetrix

143 Axiom Array. Genotypes were imputed to the TOPMed panel (Freeze5) at the Michigan
144 imputation server. Only variants with minor allele frequency (MAF) greater than or equal to
145 0.01, and minor allele count (MAC) greater or equal to 5, and variants which have Hardy-
146 Weinberg equilibrium exact test p-value greater than $1e-20$ in the entire MRI dataset and an
147 empirical-theoretical variance ratio (MaCH r^2) threshold above 0.3 were included. The main
148 GWAS was conducted on the largest subset of participants with MRI data from the largest
149 unrelated European-ancestry cohort defined using the variable `in.white.British.ancestry.subset` in
150 the file `ukb_sqc_v2.txt` provided as part of the UKB data release ($n = \sim 27,100$ individuals with
151 estimates derived from imaging data, $age = 55.0 \pm 7.4$). To replicate the findings, we separated the
152 dataset into a discovery set of 22,124 participants and a replication set of 4,899 participants
153 released at a later date. To further explore the findings from the European GWAS, we included
154 three independent sets of other ethnic backgrounds with cardiac MRI who were not included in
155 the discovery set, including African/Afro-Caribbean (AF, $n = 222$, $age = 49.6 \pm 7.0$), East Asian
156 (EAS, $n = 85$, $age = 49.2 \pm 5.5$), and South Asian (SAS, $n = 368$, $age = 52.1 \pm 7.9$). Examination of
157 those samples according the genetic principal components showed that many were mostly of
158 non-European ancestry and were unrelated [Table S1]. Testing of single nucleotide variants and
159 indels was performed using an using linear regression PLINK¹⁶² (v2.00a2LM) additive
160 model^{17,18}, including gender, age, BSA, and genetic principal components 1-4. For genic and
161 regional CNV burden tests, associations were performed using linear regression PLINK2
162 (v2.00a2LM) additive model^{17,18}, including gender, age, BSA, Principal components 1-10 and
163 the length and total number of CNV per individual as covariates as previously described¹⁹.
164 Locuszoom or custom python scripts were used to generate regional association plots²⁰. Trans-

165 ethnic meta-analysis was performed using the METAL software using the standard error analysis
166 scheme.

167

168 Independent SNPs were identified by linkage disequilibrium (LD) r^2 0.6 and were annotated to
169 cardiac eQTL (GTEx v8: heart atrial appendage and heart left ventricle), CADD, RDB, and the
170 GWAS catalog. GWAS of rare variant and gene-based tests were performed using score and
171 SKATO as implemented in the RVTEST package (link). For the RV-GWAS, age, sex, BSA,
172 Principal components 1-10 were used as covariates to calculate the association of 6 million
173 (6,593,945) imputed/genotyped variants of IVS with minor allele frequency between 0.01 and
174 0.0005, minor allele count greater than 15, an imputation quality MachR2 greater than 0.8, and
175 p-value for Hardy Weinberg Equilibrium greater than $1e-20$. For gene-based tests, the selected
176 variants were mapped to the UCSC GRCh38 refGene annotations, where every unique protein-
177 coding gene ($n = 22,933$) listed was tested.

178

179 **In silico analyses for contact prediction and enhancer activity.**

180 To predict changes in 3D chromatin folding brought about by the chr17 inversion, we used
181 Akita²¹, a convolutional neural network model trained to predict Hi-C maps from about one
182 megabase of DNA. We first validated that Akita can reproduce wildtype contact frequencies in
183 the region of [GRCh38/hg38] chr17:45392804-46440468. We then used the inversion sequence
184 as input to generate a mutant Hi-C prediction. Comparing this to the wildtype, we predicted
185 changes in contact frequency between promoters, gene bodies, and candidate enhancer elements
186 in the GeneHancer database²². To determine possible regulatory functions for single variants in
187 GWAS loci, we examined all variants with $P < 10^{-6}$. We intersected these variants with

188 predicted enhancers in 131 cell types identified by the activity-by-contact (ABC) Model, which
189 combines measurements of enhancer activity (based on ATAC-seq, DNase-seq, and H3K27ac
190 ChIP-seq) with estimates of enhancer-promoter 3D contact frequencies (based on Hi-C)^{23,24}. We
191 examined DNase-seq and ATAC-seq data across a range of cardiovascular cell types from the
192 ENCODE Project. Finally, we examined sequence motif predictions for identified variants using
193 a database of transcription factor binding site motifs²⁵.

194

195 **LDSC and genetic correlation analyses**

196 To calculate genetic correlation between polygenic risk score of IVS cross sectional area annular
197 size and other related phenotypes, we obtained summary statistics for cardiac MRI-derived LV
198 measurements (left ventricular end-diastolic volume (LVEDV), left ventricular end-systolic
199 volume (LVESV), stroke volume (SV), the body-surface-area (BSA) indexed versions for
200 cardiovascular traits (LVEDVi, LVESVi, and SVi), and left ventricular ejection fraction
201 (LVEF)²⁶, atrial fibrillation (AF)²⁷, nonischemic cardiomyopathy (NICM)²⁸, heart failure²⁹, heart
202 failure using UK Biobank data²⁸, hypertension³⁰, PR Interval³¹, Myocardial Infarction (MI) and
203 coronary artery disease (CAD)³², heart rate³³, and IVS cross sectional area prolapse (MVP)³⁴.
204 Using these data we performed LD Score regression³⁵ based on the reformatted summary
205 statistics filtered to HapMap3.

206

207 **Mendelian Randomization**

208 We performed two-sample Mendelian randomization (MR) to test for causal relationships
209 between IVS.csad and risk of two structural diseases of the IVS, specifically VSD and HCM.
210 Summary statistics from a VSD GWAS that considered 191 individuals with VSD and 5,159

211 controls were retrieved.²⁶ Summary statistics from a multi-ancestry HCM GWAS that evaluated
212 2,780 HCM cases and 47,486 controls were retrieved³⁶. An instrumental variable (IV), consisting
213 of 22 variants was generated by first selecting variants with p-value < 5e-06 from a standardized
214 trans-ethnic meta-analysis of IVS.csad and then, applying a linkage disequilibrium clumping
215 procedure (r^2 of 0.01 across a window of 1000 kb) via TwoSampleMR
216 (<https://github.com/MRCIEU/TwoSampleMR/>). The F statistic for this 22 SNP instrument was
217 15.1 and accounts for 1.04% SNP heritability.

218 We performed two-sample MR using four different methods, specifically inverse-variance
219 weighted (IVW), weighted median, MR-Egger, and MR-PRESSO (mendelian randomisation
220 pleiotropy residual sum and outlier) using MR-Base^{37,38}. Each method assumes a different set of
221 underlying assumptions; all variants included in the instrument are assumed to be valid in the
222 IVW method, and a fixed effects IVW model assumes each variant confers the same mean effect,
223 with no horizontal pleiotropy. A random effects IVW model, accounts for the possibility that
224 variants included in the instrument yield different mean effects and can provide an unbiased
225 estimate when horizontal pleiotropy is balanced³⁹. The MR-Egger regression estimates the
226 relationship of effects across instruments functioning as a sensitivity analysis for directional
227 pleiotropy⁴⁰, weighted median offers a consistent estimate of effect size when a minimum of
228 50% of the weights come from valid IVs, and the MR-PRESSO model measures and robustly
229 accounts for the presence of horizontal pleiotropy³⁸.

230 RESULTS

231 After exclusion of outliers we obtained automated estimates of the cross-sectional area of the
232 interventricular septum during ventricular diastole (IVS.csad) for 31,587 individuals yielding a
233 mean IVS.csad of 651 mm² (range 207 – 1108 mm², standard deviation 166 mm²) [Fig.1A]. In
234 comparison to blinded manual measurements, automated measures were on average 24 mm² or
235 3.2% smaller than manual measures of IVS.csad with no systematic relationship of measurement
236 error to body surface area (BSA), genetic sex, or age across 50 randomly selected test images
237 [Table S2].

238

239 Given that IVS.csad is not a standard clinical measure in MRI or echocardiography and is
240 infrequently characterized in the literature²⁷ we sought to provide clinical context. Similar to the
241 size of other cardiac structures, IVS.csad scales with BSA in a linear fashion [Fig.1B]. Amongst
242 the 31,587 individuals with an automated estimate, IVS.csad was also correlated with genetic sex
243 and body mass index (BMI) (Pearson's $r > 0.6$), moderately correlated with systolic and diastolic
244 blood pressure (Pearson's $r > 0.3$), and mildly correlated with birthweight (Pearson's $r > 0.1$)
245 [Fig.1C]. Additionally, in a PheWAS of IVS.csad for 67 cardiovascular phenotypes we observed
246 that a larger IVS.csad was positively associated with PheCODES encompassing “abnormal heart
247 sounds” ($p_{\text{fdr}} = 0.003$, Odds Ratio (OR) 2.1 per standard deviation (SD) increase in IVS.csad,
248 ICD10 codes R00 - R01.3), “non-rheumatic aortic valve disease” ($p_{\text{fdr}} = 0.003$, OR 2.4 per SD
249 IVS.csad, ICD10 I08.2, I35, I35.9), and “abnormal functional study of the cardiovascular
250 system” ($p_{\text{fdr}} = 0.02$, OR 2.6 per SD IVS.csad, ICD10 R94.3[Fig.1D]. Importantly, in a random
251 subset of 50 individuals selected for manual measurement, IVS.csad was correlated with

252 maximal interventricular septal thickness (Pearson's $r = 0.62$), a standard clinical measure of IVS
253 structure²⁸.

254
255 Given that IVS.csad was strongly correlated with key clinical measures of cardiovascular
256 structure and function, we performed a GWAS of IVS.csad as a proxy for interventricular septal
257 mass in 26,844 individuals of European ancestry divided into discovery ($n = 21,945$) and
258 replication ($n = 4,899$) subsets [Table S1]. Using the summary statistics, we performed LD-score
259 correlation (LDSC) to relate the genetic determinants of IVS.csad to a set of common
260 cardiovascular traits and diseases to provide further clinical context, which highlighted
261 significant genetic overlap with left ventricular mass (correlation 0.81, $p = 1.6 \times 10^{-24}$) and the
262 PR interval (correlation 0.26, $p = 6.2 \times 10^{-06}$) but did not show strong negative genetic
263 correlations with cardiovascular traits selected for measurement [Fig.S1]. By LDSC, the total
264 observed h^2 for IVS.csad was high at 0.068 (SD 0.0106).

265
266 Within European ancestry individuals there were two loci which were strongly significant [Table
267 1, Fig.2A]. A lead variant on chromosome 6 (rs2376620, $p_{\text{combined}} = 2.00 \times 10^{-10}$, *CDKN1A*
268 intron) which tags a haplotype of linked variants overlapping key open chromatin regions
269 exclusive to ventricular myocardium within the introns of *CDKN1A* [Fig.2C], a canonical
270 regulator of ventricular cardiomyocyte proliferation^{41,42}. The other lead variant on chromosome
271 17 (rs62063281, $p_{\text{combined}} = 1.31 \times 10^{-13}$) appeared to display strong linkage across a large region
272 of 690,200 base pairs in length which represents a duplication flanked inversion (hg38
273 chr17:45571611-46261810) across three genes and is common (minor allele frequency 0.18)
274 within in European populations. Using the Akita model of 3D genome folding²¹, we quantified

275 expected changes in chromatin contact frequency caused by the inversion which predicts a
276 severe disturbance in the local chromatin landscape, including loss of contact between the
277 *KANSL1* promoter and regions including the *MAPT* gene body, *KANSL1-ASI*, and several
278 candidate regulatory enhancers predicted by the GeneHancer database [Fig.2B]. A burden-test of
279 structural variants overlapping genes did not reveal further structural variants¹⁹ [Fig. S2].

280
281 In the European-only GWAS analysis, a number of additional loci did not meet standard
282 genome-wide significance but had strong pre-existing evidence for involvement in
283 cardiovascular biology [Table S4]. Following the European-only analysis, we performed
284 analyses within three additional ethnic strata (South Asian, East Asian, and African/Afro-
285 Caribbean) and a trans-ethnic meta-analysis which confirmed the findings in *CDKN1A* across
286 population strata and identified additional variants on chromosome 3 (rs62253176, $p=2.3 \times 10^{-08}$,
287 *MITF* intron) and chromosome 1 (rs2092867, $p=4.4 \times 10^{-08}$, *NFIA* intron) which were strongly
288 suggestive but were not meet standard significance threshold in the European only analysis
289 [Figs.S3 & S4]. Additional variants of interest on chromosome 2 (rs1368960, $p_{\text{combined}} = 1.3 \times 10^{-07}$,
290 *ASBI* intron) and chromosome 15 (rs11633294, $p_{\text{combined}} = 2.7 \times 10^{-07}$, *IGF1R* intron)
291 displayed clear genomic consequences [Fig.S5] but were not genome-wide significant in
292 European or Trans-ethnic meta-analyses. Additionally we performed both rare variant analyses
293 (MAF 0.01 to 0.001) and gene-burden testing which were unrevealing of genetic associations
294 meeting standard pre-specified significance thresholds [Figs.S6 & S7].

295
296 Common genetic variation is known to underpin susceptibility to a variety of cardiovascular
297 diseases, including those determined by extremes in anatomical size.^{29,30,31} We sought to relate

298 the genetic determinants governing normal variation in the IVS to two rare forms of structural
299 heart disease manifesting as either a small or insufficient interventricular septum (VSD) or as a
300 large interventricular septum (HCM). Using an instrumental variable of 22 variants derived from
301 the IVS.csad GWAS [Table S4], we performed two-sample Mendelian Randomization for VSD
302 (191 VSD cases and 5,159 controls) and HCM (2,780 HCM cases and 47,486 controls) as
303 outcome phenotypes. We observed evidence suggesting a causal effect where inheritance of a
304 smaller IVS.csad corresponded to an increase in risk for VSD (beta coefficient: 2.06 log odds
305 ratio (OR) VSD per decrease in SD of IVS.csad, standard error (SE) 0.75, $p_{IVW} = 0.006$) while
306 inheritance of a larger IVS.csad corresponded to an increase in risk for HCM (beta coefficient:
307 2.45 log OR HCM per increase in SD of IVS.csad, SE: 0.48, $p_{IVW} = 2.8e-7$) [Figure 3].

308
309 When evaluating IVS.csad as the exposure and VSD as the outcome, sensitivity analyses
310 suggested the observed causal effect was uniform across the variants included in the instrumental
311 variable (Q value = 15.8, df=20, p-value=0.73). The MR-Egger test of horizontal pleiotropy was
312 insignificant with broad standard errors around the slope and intercept (beta=-0.62, SE=2.43;
313 intercept=-0.06). The median weighted method, which provides a causal estimate assuming at
314 least 50% of the weight comes from valid instrumental variables, was consistent with the IVW
315 result (beta=-2.40; se=1.05; p=0.02). The MR-PRESSO global test for outliers indicates there
316 was no evidence of horizontal pleiotropy for the outcome of VSD (p=0.78).

317
318 Sensitivity analyses evaluating IVS.csad as the exposure and HCM as the outcome suggested
319 there was evidence of heterogeneity within the instrument (Q value = 84.2, df=20, p-
320 value=7.58E-10, $i^2 = 0.76$) and the MR-PRESSO global test for outliers indicated there was

321 horizontal pleiotropy ($p = 0.001$). The unadjusted MR-PRESSO result (beta coefficient: 2.45,
322 SE: 0.48, $p_{\text{MR-PRESSO}} = 4.36\text{E-}05$) is similar in magnitude to the IVW result, but considering the
323 observed horizontal pleiotropy could be artificially inflated. Performing the outlier-corrected
324 MR-PRESSO test (beta coefficient: 1.62 log OR HCM per increase in SD of IVS.csad, SE: 0.42,
325 $p_{\text{MR-PRESSO}} = 0.001$) shows an effect size comparable to the median (weighted) results (1.47 log
326 OR HCM per increase in SD of IVS.csad, SE 0.47, $p_{\text{median-weighted}} = 0.002$). The estimate of effect
327 size derived from the MR-Egger test is internally consistent (beta=1.82 log OR HCM per
328 increase in SD of IVS.csad, SE=1.67, $p=0.29$, intercept=-0.03).

329
330 Because left-ventricular mass and IVS thickness are known to increase in the setting of chronic
331 hypertension¹⁰, a phenomenon that we observed in our clinical correlates of automated measures
332 of IVS.csad [Fig.1B], we performed an additional sensitivity analysis to rule out confounding for
333 the causal effect observed between IVS.csad and HCM. For the outcome of HCM we repeated
334 the two-sample MR for IVS.csad conditioned upon diastolic blood pressure, which did not
335 significantly change the positive and statistically significant causal estimate for IVS.csad (beta
336 3.08 log OR HCM per increase in SD of IVS.csad, SE 0.48, $p_{\text{IVW}} = 1.05\text{e-}05$) [Table S5].
337

338 **DISCUSSION**

339 Here we report upon the genetic architecture of the cross-section of the interventricular septum
340 derived from a standard view in cardiac imaging. The measured cross-sectional area of the IVS
341 in diastole is correlated with genetic sex and body surface area along with a clinically accepted
342 measure of the interventricular septal thickness⁴³ and genetically correlated with left ventricular
343 mass suggesting that our 2-dimensional automated measure of IVS.csad is a good proxy for the
344 mass of this portion of cardiac anatomy. Genetic associations for IVS.csad are centered upon
345 genetic loci with previously established roles in cardiac development and mendelian forms of
346 cardiac malformations. Additionally, we show a causal relationship between inheritance of a
347 larger interventricular septum with risk of HCM and a smaller interventricular septum with risk
348 of ventricular septal defects.

349
350 Haploinsufficiency or truncating mutations in *KANSL1* are causal for Koolen-de Vries syndrome
351 which includes ventricular septal defects amongst a variety of other phenotypes^{44,45}. We found
352 that a variant associated with IVS.csad, rs62063281, tags a common inversion on chromosome
353 17, referred to as CPX_17_4670 on the gnomAD CNV browser⁴⁶, that encompasses *KANSL1*
354 and *KANSL1-AS1* along with four other protein coding genes, and was associated with an
355 increase in IVS.csad (beta= 8.60 mm², p=1.31e-13). The variant rs62063281 which tags the
356 inversion appears to be a strong tissue-specific splice QTL (Normalized expression -1.5, p =
357 1.7e-88) for *KANSL1*, In addition to disruption of the interaction of the *KANSL1* promoter and
358 predicted regulatory elements, genes that contact *KANSL1-AS1* (including *ARL17B*, *LRRC37A*,
359 *NSFPI*, *LRRC37A2*, *NSF*) in the reference configuration are also expected to lose this interaction
360 as a result of the common inversion, which potentially may additionally result in disrupted

361 expression of these other genes. As a group histone, modifying genes including the WDR5-
362 MLL1 complex of which KANSL1 is a component are well-recognized to be involved in cardiac
363 development and diverse forms of congenital heart disease^{5,47}. Together these data suggest that
364 common alterations to dosage and transcription of *KANSL1* impact the development and growth
365 of the interventricular septum.

366
367 The *CDKN1A* locus has been specifically identified as a target of HIF-1a within the developing
368 ventricular septum⁴⁸ and has been experimentally described as a key mediator of the
369 hypertrophic response within postnatal life^{41,49}. The variant rs2376620 sits within an intron of
370 *CDKN1A*, a locus well known to play a role in cardiomyocyte growth⁴² recently identified in
371 heart failure²⁹. Together these data suggest a specific role for *CDKN1A* in normal development
372 of the interventricular septum along with physiological function throughout postnatal life.

373 Among the other identified in the trans-ethnic analysis, the variant rs62253176 occurs within an
374 intron of the transcription factor *MITF*, an important mediator of beta-adrenergic induced
375 hypertrophy via the renin-angiotensin system expressed in cardiomyocytes⁵⁰. Importantly, *MITF*
376 binds the promoter of canonical cardiac development transcription factor *GATA4* and
377 knockdowns in human embryonic stem cell derived cardiomyocytes specifically reduce the
378 expression of genes in the sarcomere^{51,52}. The role of *NFIA* in cardiac development and function
379 is less well defined though it is associated with a variety of electrocardiographic traits^{53,54} and the
380 variant rs2092867 occurs on a haplotype associated with a recently reported novel trait related to
381 the presence of cardiac trabeculae derived from the same UK Biobank imaging dataset⁵⁵.

382

383 It is well established that rare causal variants confer large effects and produce extreme
384 phenotypes. However, many individuals possessing an extreme phenotype do not yield such
385 causal variants when subject to clinical genetic testing⁵⁶. Instead, there is increasing evidence
386 that such extreme phenotypes may also be attributable to the aggregate burden of common
387 variants associated with the disease or a causal trait⁵⁷. As such, extremes in polygenic risk may
388 explain variability in cardiovascular anatomy and represent an important disease risk factor for a
389 variety of different diseases. For instance, we have previously shown that polygenic risk scores
390 influence ascending aorta size (PRS_{Aorta}); increases in PRS_{Aorta} are associated with risk of
391 developing thoracic aortic aneurysm and dissection⁵⁸, and decreases in PRS_{Aorta} are associated
392 with left ventricular outflow tract congenital heart disease⁵⁹.

393
394 Using two-sample Mendelian randomization, we show that inheritance of a smaller $IVS.csad$
395 results in an increase in ventricular septal defect risk, and conversely, inheritance of a larger
396 $IVS.csad$ confers increased risk in HCM. Identifying risk of ventricular septal defects is
397 increased through the inheritance of a smaller interventricular septum, as measured by $IVS.csad$,
398 supports the need for systematic inquiry into the relationship between normal variation in the
399 size and shape of cardiovascular anatomy and congenital heart disease, for which the majority of
400 genetic risk remains unexplained. For HCM, rare pathogenic variants within core sarcomere
401 genes account for 40% of HCM cases, but with unexplained phenotypic heterogeneity. Recent
402 data suggest common genetic variants also underpin HCM risk³⁶. The analyses presented here
403 suggest that a proportion of HCM risk is attributable to normal variation in the size and mass of
404 the interventricular septum as measured by $IVS.csad$ [Fig.S2].

405

406 There are limitations to each aspect of our study. While we have created an automated measure
407 of the interventricular septum which is simple in conception and shown correlation with a variety
408 of traits, the measure lacks phenotypic context provided by many years of use in standard clinical
409 imaging. The cross-sectional area of the interventricular septum is a reductive two-dimensional
410 measure of a complex three-dimensional structure which changes dynamically over the course of
411 the cardiac cycle⁶⁰. More complex measures of the septum are likely to provide nuanced
412 correlations with clinical phenotypes, patient outcomes, as well as substrate for discovery of
413 novel genetic determinants of structure and function⁵⁵. Our analyses were limited by the small
414 proportion of individuals of African/Afro-Caribbean, South and East Asian descent, which
415 reduced the power to detect genetic signal within these groups for this trait and underscores the
416 urgent need to diversify genetic studies of all human diseases and traits⁵³. While the inclusion of
417 structural variation in our association studies is novel and our finding is supported by the
418 association of *KANSL1* with Mendelian disease and mouse models of cardiac development,
419 there are no cohorts available for external confirmation which include both CNV calls and
420 readily available bespoke measures of cardiac anatomy. Although our Mendelian randomization
421 analyses yielded clear and consistent results relating the heritable architecture of septal
422 architecture to rare cardiovascular diseases, they are limited in scope by the relatively
423 underpowered outcome GWAS for VSD and the clinical utility of these findings require further
424 contextualization in larger cohorts.
425

426 Overall, the data and analyses presented here illustrate the power of combining genetics with
427 phenotyping of cardiac imaging accelerated by machine learning to identify new loci related to
428 cardiovascular development and pathology. The study of cardiovascular anatomy measured as a
429 continuous trait offers the potential to disentangle complex risk factors and reveal previously
430 unrecognized heritability for rare forms of structural heart disease. While the risk is primarily
431 thought to be inherited, the genetic architecture of congenital heart disease has been difficult to
432 establish. In addition other evidence⁵⁹, the relationship between a small IVS to increased risk of
433 VSD may suggest a that a proportion of genetic risk for other forms of congenital heart disease
434 can be localized to heritable extremes in size of cardiovascular anatomy which arise primarily
435 from common genetic variation.

436

437

438 **ACKNOWLEDGEMENTS**

439 The authors wish to thank Dan Bernstein, James Pirruccello, and Euan Ashley for helpful
440 commentary on the manuscript.

441

442 **FUNDING**

443 NIH NHGRI R35HG011324 (to JME), Gordon and Betty Moore and BASE Initiative at the
444 Lucile Packard Children’s Hospital at Stanford University (to JME), American Heart Association
445 / Children’s Heart Foundation Congenital Heart Defect Research Award (to DA), Stanford
446 Maternal and Child Health Research Institute (to DA). JRP has been supported by the Stanford
447 University Department of Pediatrics, National Heart Lung and Blood Institute (R00 HL130523),
448 and Chan-Zuckerberg Biohub.

449 **REFERENCES**

- 450 1. Koshiha-Takeuchi K, Mori AD, Kaynak BL, Cebra-Thomas J, Sukonnik T, Georges RO,
451 Latham S, Beck L, Henkelman RM, Black BL, Olson EN, Wade J, Takeuchi JK, Nemer
452 M, Gilbert SF, Bruneau BG. Reptilian heart development and the molecular basis of
453 cardiac chamber evolution. *Nature* [Internet]. 2009 [cited 2020 Oct 3];461:95–98.
454 Available from: [/pmc/articles/PMC2753965/?report=abstract](https://pubmed.ncbi.nlm.nih.gov/19111111/)
- 455 2. Penny DJ, Vick GW. Ventricular septal defect [Internet]. In: *The Lancet*. Lancet; 2011
456 [cited 2020 Sep 28]. p. 1103–1112. Available from:
457 <https://pubmed.ncbi.nlm.nih.gov/21349577/>
- 458 3. Gersh BJ, Maron BJ, Bonow RO, Dearani JA, Fifer MA, Link MS, Naidu SS, Nishimura
459 RA, Ommen SR, Rakowski H, Seidman CE, Towbin JA, Udelson JE, Yancy CW. 2011
460 ACCF/AHA guideline for the diagnosis and treatment of hypertrophic cardiomyopathy:
461 Executive summary: A report of the American College of cardiology
462 foundation/American heart association task force on practice guidelines. *Circulation*.
463 2011;124:2761–2796.
- 464 4. Canepa M, Fumagalli C, Tini G, Vincent-Tompkins J, Day SM, Ashley EA, Mazzarotto F,
465 Ware JS, Michels M, Jacoby D, Ho CY, Olivotto I. Temporal Trend of Age at Diagnosis
466 in Hypertrophic Cardiomyopathy: An Analysis of the International Sarcomeric Human
467 Cardiomyopathy Registry. *Circ Heart Fail* [Internet]. 2020 [cited 2020 Sep
468 28];13:e007230. Available from:
469 <https://www.ahajournals.org/doi/10.1161/CIRCHEARTFAILURE.120.007230>
- 470 5. Jin SC, Homsy J, Zaidi S, Lu Q, Morton S, DePalma SR, Zeng X, Qi H, Chang W, Sierant
471 MC, Hung W-C, Haider S, Zhang J, Knight J, Bjornson RD, Castaldi C, Tikhonova IR,

- 472 Bilguvar K, Mane SM, Sanders SJ, Mital S, Russell MW, Gaynor JW, Deanfield J,
473 Giardini A, Porter GA, Srivastava D, Lo CW, Shen Y, Watkins WS, Yandell M, Yost HJ,
474 Tristani-Firouzi M, Newburger JW, Roberts AE, Kim R, Zhao H, Kaltman JR, Goldmuntz
475 E, Chung WK, Seidman JG, Gelb BD, Seidman CE, Lifton RP, Brueckner M.
476 Contribution of rare inherited and de novo variants in 2,871 congenital heart disease
477 probands. *Nat Genet* [Internet]. 2017 [cited 2019 Apr 29];49:1593–1601. Available from:
478 <http://www.nature.com/articles/ng.3970>
- 479 6. Camenisch TD, Runyan RB, Markwald RR. Molecular Regulation of Cushion
480 Morphogenesis. In: *Heart Development and Regeneration*. Elsevier Inc.; 2010. p. 363–
481 387.
- 482 7. Anderson RH, Webb S, Brown NA, Lamers W, Moorman A. Development of the heart:
483 (2) Septation of the atriums and ventricles. *Heart* [Internet]. 2003 [cited 2020 Sep
484 21];89:949–958. Available from:
485 <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1767797/>
- 486 8. Saleh S, Liakopoulos OJ, Buckberg GD. The septal motor of biventricular function
487 [Internet]. *Eur. J. Cardio-thoracic Surg*. 2006 [cited 2020 Sep 28];29. Available from:
488 <https://pubmed.ncbi.nlm.nih.gov/16564701/>
- 489 9. Gerling S, Pollinger T, Michel H, Dechant MJ, Melter M, Krutsch W. Z-score values of
490 left ventricular dimensions in adolescent elite male soccer players. *Eur J Pediatr*
491 [Internet]. 2020 [cited 2020 Sep 28]; Available from:
492 <https://pubmed.ncbi.nlm.nih.gov/32705342/>
- 493 10. Eliakim-Raz N, Prokupetz A, Gordon B, Shochat T, Grossman A. Interventricular Septum
494 and Posterior Wall Thickness Are Associated With Higher Systolic Blood Pressure. *J Clin*

- 495 *Hypertens* [Internet]. 2016 [cited 2020 Sep 28];18:703–706. Available from:
496 <https://pubmed.ncbi.nlm.nih.gov/26607051/>
- 497 11. Yu M, Tcheandjieu C, Georges A, Xiao K, Tejada H, Dina C, Le Tourneau T, Fiterau I,
498 Judy R, Tsao N, Amgalan D, Munger CJ, Engreitz JM, Damrauer S, Bouatia-Naji N,
499 Priest JR. Computational estimates of mitral annular diameter in systole and diastole
500 cardiac cycle reveal novel genetic determinants of valve function and disease [Internet].
501 medRxiv. 2020 [cited 2021 Mar 12];2020.12.02.20242206. Available from:
502 <https://doi.org/10.1101/2020.12.02.20242206>
- 503 12. Bai W, Sinclair M, Tarroni G, Oktay O, Rajchl M, Vaillant G, Lee AM, Aung N,
504 Lukaschuk E, Sanghvi MM, Zemrak F, Fung K, Paiva JM, Carapella V, Kim YJ, Suzuki
505 H, Kainz B, Matthews PM, Petersen SE, Piechnik SK, Neubauer S, Glocker B, Rueckert
506 D. Automated cardiovascular magnetic resonance image analysis with fully convolutional
507 networks. *J Cardiovasc Magn Reson* [Internet]. 2018 [cited 2019 Jul 18];20:65. Available
508 from: <http://www.ncbi.nlm.nih.gov/pubmed/30217194>
- 509 13. Haycock GB, Schwartz GJ, Wisotsky DH. Geometric method for measuring body surface
510 area: A height-weight formula validated in infants, children, and adults. *J Pediatr*
511 [Internet]. 1978 [cited 2021 Mar 12];93:62–66. Available from:
512 <https://pubmed.ncbi.nlm.nih.gov/650346/>
- 513 14. Cordova-Palomera A, Tcheandjieu C, Fries J, Varma P, Chen V, Fiterau M, Xiao K,
514 Tejada H, Keavney B, Cordell H, Tanigawa Y, Venkataraman G, Rivas M, Re C, Ashley
515 E, Priest J. Cardiac imaging of aortic valve area from 26,142 UK Biobank participants
516 reveal novel genetic associations and shared genetic comorbidity with multiple disease
517 phenotypes. *medRxiv* [Internet]. 2020 [cited 2020 Sep 15];2020.04.09.20060012.

- 518 Available from: <https://doi.org/10.1101/2020.04.09.20060012>
- 519 15. Wu P, Gifford A, Meng X, Li X, Campbell H, Varley T, Zhao J, Carroll R, Bastarache L,
520 Denny JC, Theodoratou E, Wei WQ. Mapping ICD-10 and ICD-10-CM codes to
521 phecodes: Workflow development and initial evaluation. *J Med Internet Res* [Internet].
522 2019 [cited 2020 Sep 22];21. Available from: <https://pubmed.ncbi.nlm.nih.gov/31553307/>
- 523 16. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide
524 association scans. *Bioinformatics* [Internet]. 2010 [cited 2019 Sep 18];26:2190–1.
525 Available from: <http://www.ncbi.nlm.nih.gov/pubmed/20616382>
- 526 17. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar
527 P, de Bakker PIW, Daly MJ, Sham PC. PLINK: a tool set for whole-genome association
528 and population-based linkage analyses. *Am J Hum Genet* [Internet]. 2007;81:559–75.
529 Available from: <http://linkinghub.elsevier.com/retrieve/pii/S0002929707613524>
- 530 18. PLINK 2.0 [Internet]. [cited 2019 Sep 18]; Available from: <https://www.cog->
531 [genomics.org/plink/2.0/](https://www.cog-genomics.org/plink/2.0/)
- 532 19. Aguirre M, Rivas MA, Priest J. Phenome-wide Burden of Copy-Number Variation in the
533 UK Biobank. *Am J Hum Genet* [Internet]. 2019 [cited 2019 Sep 9];105:373–383.
534 Available from: <http://www.ncbi.nlm.nih.gov/pubmed/31353025>
- 535 20. Pruim RJ, Welch RP, Sanna S, Teslovich TM, Chines PS, Gliedt TP, Boehnke M,
536 Abecasis GR, Willer CJ, Frishman D. LocusZoom: Regional visualization of genome-
537 wide association scan results [Internet]. In: *Bioinformatics*. Oxford University Press; 2011
538 [cited 2021 Jan 20]. p. 2336–2337. Available from:
539 <https://pubmed.ncbi.nlm.nih.gov/20634204/>
- 540 21. Fudenberg G, Kelley DR, Pollard KS. Predicting 3D genome folding from DNA sequence

- 541 with Akita. *Nat Methods* [Internet]. 2020 [cited 2021 Mar 22];17:1111–1117. Available
542 from: <https://pubmed.ncbi.nlm.nih.gov/33046897/>
- 543 22. Fishilevich S, Nudel R, Rappaport N, Hadar R, Plaschkes I, Iny Stein T, Rosen N, Kohn
544 A, Twik M, Safran M, Lancet D, Cohen D. GeneHancer: genome-wide integration of
545 enhancers and target genes in GeneCards. *Database (Oxford)* [Internet]. 2017 [cited 2021
546 Mar 22];2017. Available from: <https://pubmed.ncbi.nlm.nih.gov/28605766/>
- 547 23. Fulco CP, Nasser J, Jones TR, Munson G, Bergman DT, Subramanian V, Grossman SR,
548 Anyoha R, Doughty BR, Patwardhan TA, Nguyen TH, Kane M, Perez EM, Durand NC,
549 Lareau CA, Stamenova EK, Aiden EL, Lander ES, Engreitz JM. Activity-by-contact
550 model of enhancer–promoter regulation from thousands of CRISPR perturbations. *Nat.*
551 *Genet.* 2019;51.
- 552 24. Nasser J, Bergman DT, Fulco CP, Guckelberger P, Doughty BR, Patwardhan TA, Jones
553 TR, Nguyen TH, Ulirsch JC, Natri HM, Weeks EM, Munson G, Kane M, Kang HY, Cui
554 A, Ray JP, Eisenhaure TM, Mualim K, Collins RL, Dey K, Price AL, Epstein CB,
555 Kundaje A, Xavier RJ, Daly MJ, Huang H, Finucane HK, Hacohen N, Lander ES,
556 Engreitz JM. Genome-wide maps of enhancer regulation connect risk variants to disease
557 genes. *bioRxiv* [Internet]. 2020 [cited 2020 Dec 1];2020.09.01.278093. Available from:
558 <https://doi.org/10.1101/2020.09.01.278093>
- 559 25. Kulakovskiy I V., Vorontsov IE, Yevshin IS, Sharipov RN, Fedorova AD, Rumynskiy EI,
560 Medvedeva YA, Magana-Mora A, Bajic VB, Papatsenko DA, Kolpakov FA, Makeev VJ.
561 HOCOMOCO: Towards a complete collection of transcription factor binding models for
562 human and mouse via large-scale ChIP-Seq analysis. *Nucleic Acids Res.* 2018;46.
- 563 26. Pirruccello JP, Bick A, Wang M, Chaffin M, Friedman S, Yao J, Guo X, Venkatesh BA,

- 564 Taylor KD, Post WS, Rich S, Lima JAC, Rotter JI, Philippakis A, Lubitz SA, Ellinor PT,
565 Khera A V., Kathiresan S, Aragam KG. Analysis of cardiac magnetic resonance imaging
566 in 36,000 individuals yields genetic insights into dilated cardiomyopathy. *Nat Commun*
567 [Internet]. 2020 [cited 2020 Sep 14];11. Available from:
568 <https://pubmed.ncbi.nlm.nih.gov/32382064/>
- 569 27. Roselli C, Chaffin MD, Weng LC, Aeschbacher S, Ahlberg G, Albert CM, Almgren P,
570 Alonso A, Anderson CD, Aragam KG, Arking DE, Barnard J, Bartz TM, Benjamin EJ,
571 Bihlmeyer NA, Bis JC, Bloom HL, Boerwinkle E, Bottinger EB, Brody JA, Calkins H,
572 Campbell A, Cappola TP, Carlquist J, Chasman DI, Chen LY, Chen YDI, Choi EK, Choi
573 SH, Christophersen IE, Chung MK, Cole JW, Conen D, Cook J, Crijns HJ, Cutler MJ,
574 Damrauer SM, Daniels BR, Darbar D, Delgado G, Denny JC, Dichgans M, Dörr M,
575 Dudink EA, Dudley SC, Esa N, Esko T, Eskola M, Fatkin D, Felix SB, Ford I, Franco
576 OH, Geelhoed B, Grewal RP, Gudnason V, Guo X, Gupta N, Gustafsson S, Gutmann R,
577 Hamsten A, Harris TB, Hayward C, Heckbert SR, Hernesniemi J, Hocking LJ, Hofman A,
578 Horimoto ARVR, Huang J, Huang PL, Huffman J, Ingelsson E, Ipek EG, Ito K, Jimenez-
579 Conde J, Johnson R, Jukema JW, Kääb S, Kähönen M, Kamatani Y, Kane JP, Kastrati A,
580 Kathiresan S, Katschnig-Winter P, Kavousi M, Kessler T, Kietselaer BL, Kirchhof P,
581 Kleber ME, Knight S, Krieger JE, Kubo M, Launer LJ, Laurikka J, Lehtimäki T,
582 Leineweber K, Lemaitre RN, Li M, Lim HE, et al. Multi-ethnic genome-wide association
583 study for atrial fibrillation. *Nat Genet* [Internet]. 2018 [cited 2020 Sep 14];50:1225–1233.
584 Available from: <https://pubmed.ncbi.nlm.nih.gov/29892015/>
- 585 28. Aragam KG, Chaffin M, Levinson RT, McDermott G, Choi SH, Shoemaker MB, Haas
586 ME, Weng LC, Lindsay ME, Smith JG, Newton-Cheh C, Roden DM, London B, Wells

- 587 QS, Ellinor PT, Kathiresan S, Lubitz SA. Phenotypic Refinement of Heart Failure in a
588 National Biobank Facilitates Genetic Discovery. *Circulation* [Internet]. 2019 [cited 2020
589 Sep 14];139:489–501. Available from: <https://pubmed.ncbi.nlm.nih.gov/30586722/>
- 590 29. Shah S, Henry A, Roselli C, Lin H, Sveinbjörnsson G, Fatemifar G, Hedman ÅK, Wilk
591 JB, Morley MP, Chaffin MD, Helgadóttir A, Verweij N, Dehghan A, Almgren P,
592 Andersson C, Aragam KG, Ärnlöv J, Backman JD, Biggs ML, Bloom HL, Brandimarto J,
593 Brown MR, Buckbinder L, Carey DJ, Chasman DI, Chen X, Chen X, Chung J, Chutkow
594 W, Cook JP, Delgado GE, Denaxas S, Doney AS, Dörr M, Dudley SC, Dunn ME,
595 Engström G, Esko T, Felix SB, Finan C, Ford I, Ghanbari M, Ghasemi S, Giedraitis V,
596 Giulianini F, Gottdiener JS, Gross S, Guðbjartsson DF, Gutmann R, Haggerty CM, van
597 der Harst P, Hyde CL, Ingelsson E, Jukema JW, Kavousi M, Khaw KT, Kleber ME,
598 Køber L, Koekemoer A, Langenberg C, Lind L, Lindgren CM, London B, Lotta LA,
599 Lovering RC, Luan J, Magnusson P, Mahajan A, Margulies KB, März W, Melander O,
600 Mordi IR, Morgan T, Morris AD, Morris AP, Morrison AC, Nagle MW, Nelson CP,
601 Niessner A, Niiranen T, O’Donoghue ML, Owens AT, Palmer CNA, Parry HM, Perola M,
602 Portilla-Fernandez E, Psaty BM, Abecasis G, Backman J, Bai X, Balasubramanian S,
603 Banerjee N, Baras A, Barnard L, Beechert C, Blumenfeld A, Cantor M, Chai Y, et al.
604 Genome-wide association and Mendelian randomisation analysis provide insights into the
605 pathogenesis of heart failure. *Nat Commun* [Internet]. 2020 [cited 2020 Sep 14];11.
606 Available from: <https://pubmed.ncbi.nlm.nih.gov/31919418/>
- 607 30. Wojcik GL, Graff M, Nishimura KK, Tao R, Haessler J, Gignoux CR, Highland HM,
608 Patel YM, Sorokin EP, Avery CL, Belbin GM, Bien SA, Cheng I, Cullina S, Hodonsky
609 CJ, Hu Y, Huckins LM, Jeff J, Justice AE, Kocarnik JM, Lim U, Lin BM, Lu Y, Nelson

610 SC, Park SSL, Poisner H, Preuss MH, Richard MA, Schurmann C, Setiawan VW, Sockell
611 A, Vahi K, Verbanck M, Vishnu A, Walker RW, Young KL, Zubair N, Acuña-Alonso V,
612 Ambite JL, Barnes KC, Boerwinkle E, Bottinger EP, Bustamante CD, Caberto C,
613 Canizales-Quinteros S, Conomos MP, Deelman E, Do R, Doheny K, Fernández-Rhodes
614 L, Fornage M, Hailu B, Heiss G, Henn BM, Hindorff LA, Jackson RD, Laurie CA, Laurie
615 CC, Li Y, Lin DY, Moreno-Estrada A, Nadkarni G, Norman PJ, Pooler LC, Reiner AP,
616 Romm J, Sabatti C, Sandoval K, Sheng X, Stahl EA, Stram DO, Thornton TA, Wassel
617 CL, Wilkens LR, Winkler CA, Yoneyama S, Buyske S, Haiman CA, Kooperberg C, Le
618 Marchand L, Loos RJF, Matise TC, North KE, Peters U, Kenny EE, Carlson CS. Genetic
619 analyses of diverse populations improves discovery for complex traits. *Nature* [Internet].
620 2019 [cited 2020 Sep 14];570:514–518. Available from:
621 <https://pubmed.ncbi.nlm.nih.gov/31217584/>

622 31. Ntalla I, Weng LC, Cartwright JH, Hall AW, Sveinbjornsson G, Tucker NR, Choi SH,
623 Chaffin MD, Roselli C, Barnes MR, Mifsud B, Warren HR, Hayward C, Marten J,
624 Cranley JJ, Concas MP, Gasparini P, Boutin T, Kolcic I, Polasek O, Rudan I, Araujo NM,
625 Lima-Costa MF, Ribeiro ALP, Souza RP, Tarazona-Santos E, Giedraitis V, Ingelsson E,
626 Mahajan A, Morris AP, Del Greco M F, Foco L, Gögele M, Hicks AA, Cook JP, Lind L,
627 Lindgren CM, Sundström J, Nelson CP, Riaz MB, Samani NJ, Sinagra G, Ulivi S,
628 Kähönen M, Mishra PP, Mononen N, Nikus K, Caulfield MJ, Dominiczak A,
629 Padmanabhan S, Montasser ME, O’Connell JR, Ryan K, Shuldiner AR, Aeschbacher S,
630 Conen D, Risch L, Thériault S, Hutri-Kähönen N, Lehtimäki T, Lyytikäinen LP, Raitakari
631 OT, Barnes CLK, Campbell H, Joshi PK, Wilson JF, Isaacs A, Kors JA, van Duijn CM,
632 Huang PL, Gudnason V, Harris TB, Launer LJ, Smith A V., Bottinger EP, Loos RJF,

633 Nadkarni GN, Preuss MH, Correa A, Mei H, Wilson J, Meitinger T, Müller-Nurasyid M,
634 Peters A, Waldenberger M, Mangino M, Spector TD, Rienstra M, van de Vegte YJ, van
635 der Harst P, Verweij N, Kääh S, Schramm K, Sinner MF, Strauch K, Cutler MJ, Fatkin D,
636 London B, et al. Multi-ancestry GWAS of the electrocardiographic PR interval identifies
637 202 loci underlying cardiac conduction. *Nat Commun* [Internet]. 2020 [cited 2020 Sep
638 14];11. Available from: <https://pubmed.ncbi.nlm.nih.gov/32439900/>

639 32. Nikpay M, Goel A, Won H-H, Hall LM, Willenborg C, Kanoni S, Saleheen D, Kyriakou
640 T, Nelson CP, Hopewell JC, Webb TR, Zeng L, Dehghan A, Alver M, Armasu SM, Auro
641 K, Bjornes A, Chasman DI, Chen S, Ford I, Franceschini N, Gieger C, Grace C,
642 Gustafsson S, Huang J, Hwang S-J, Kim YK, Kleber ME, Lau KW, Lu X, Lu Y,
643 Lyytikäinen L-P, Mihailov E, Morrison AC, Pervjakova N, Qu L, Rose LM, Salfati E,
644 Saxena R, Scholz M, Smith A V, Tikkanen E, Uitterlinden A, Yang X, Zhang W, Zhao W,
645 de Andrade M, de Vries PS, van Zuydam NR, Anand SS, Bertram L, Beutner F,
646 Dedoussis G, Frossard P, Gauguier D, Goodall AH, Gottesman O, Haber M, Han B-G,
647 Huang J, Jalilzadeh S, Kessler T, König IR, Lannfelt L, Lieb W, Lind L, Lindgren CM,
648 Lokki M-L, Magnusson PK, Mallick NH, Mehra N, Meitinger T, Memon F-U-R, Morris
649 AP, Nieminen MS, Pedersen NL, Peters A, Rallidis LS, Rasheed A, Samuel M, Shah SH,
650 Sinisalo J, Stirrups KE, Trompet S, Wang L, Zaman KS, Ardissino D, Boerwinkle E,
651 Borecki IB, Bottinger EP, Buring JE, Chambers JC, Collins R, Cupples LA, Danesh J,
652 Demuth I, Elosua R, Epstein SE, et al. A comprehensive 1,000 Genomes-based genome-
653 wide association meta-analysis of coronary artery disease. *Nat Genet* [Internet]. 2015
654 [cited 2019 Apr 29];47:1121–1130. Available from:
655 <http://www.nature.com/articles/ng.3396>

- 656 33. den Hoed M, Eijgelsheim M, Esko T, Brundel BJJM, Peal DS, Evans DM, Nolte IM,
657 Segrè A V, Holm H, Handsaker RE, Westra H-J, Johnson T, Isaacs A, Yang J, Lundby A,
658 Zhao JH, Kim YJ, Go MJ, Almgren P, Bochud M, Boucher G, Cornelis MC, Gudbjartsson
659 D, Hadley D, van der Harst P, Hayward C, den Heijer M, Igl W, Jackson AU, Kutalik Z,
660 Luan J, Kemp JP, Kristiansson K, Ladenvall C, Lorentzon M, Montasser ME, Njajou OT,
661 O'Reilly PF, Padmanabhan S, St. Pourcain B, Rankinen T, Salo P, Tanaka T, Timpson NJ,
662 Vitart V, Waite L, Wheeler W, Zhang W, Draisma HHM, Feitosa MF, Kerr KF, Lind PA,
663 Mihailov E, Onland-Moret NC, Song C, Weedon MN, Xie W, Yengo L, Absher D, Albert
664 CM, Alonso A, Arking DE, de Bakker PIW, Balkau B, Barlassina C, Benaglio P, Bis JC,
665 Bouatia-Naji N, Brage S, Chanock SJ, Chines PS, Chung M, Darbar D, Dina C, Dörr M,
666 Elliott P, Felix SB, Fischer K, Fuchsberger C, de Geus EJC, Goyette P, Gudnason V,
667 Harris TB, Hartikainen A-L, Havulinna AS, Heckbert SR, Hicks AA, Hofman A,
668 Holewijn S, Hoogstra-Berends F, Hottenga J-J, Jensen MK, Johansson Å, Junttila J, Kääb
669 S, Kanon B, Ketkar S, Khaw K-T, et al. Identification of heart rate-associated loci and
670 their effects on cardiac conduction and rhythm disorders. *Nat Genet* [Internet]. 2013 [cited
671 2018 Jun 20];45:621–631. Available from:
672 <http://www.ncbi.nlm.nih.gov/pubmed/23583979>
- 673 34. Toomer KA, Yu M, Fulmer D, Guo L, Moore KS, Moore R, Ka'la DD, Glover J, Peterson
674 N, Ramos-Ortiz S, Drohan A, Catching BJ, Stairley R, Wessels A, Lipschutz JH, Delling
675 FN, Jeunemaitre X, Dina C, Collins RL, Brand H, Talkowski ME, Del Monte F,
676 Mukherjee R, Awgulewitsch A, Body S, Hardiman G, Starr Hazard E, Da Silveira WA,
677 Wang B, Leyne M, Durst R, Markwald RR, Le Scouarnec S, Hagege A, Le Tourneau T,
678 Kohl P, Rog-Zielinska EA, Ellinor PT, Levine RA, Milan DJ, Schott JJ, Bouatia-Naji N,

- 679 Slaugenhaupt SA, Norris RA. Primary cilia defects causing mitral valve prolapse. *Sci*
680 *Transl Med* [Internet]. 2019 [cited 2020 Sep 14];11. Available from:
681 <https://pubmed.ncbi.nlm.nih.gov/31118289/>
- 682 35. Bulik-Sullivan B, Loh PR, Finucane HK, Ripke S, Yang J, Patterson N, Daly MJ, Price
683 AL, Neale BM, Corvin A, Walters JTR, Farh KH, Holmans PA, Lee P, Collier DA,
684 Huang H, Pers TH, Agartz I, Agerbo E, Albus M, Alexander M, Amin F, Bacanu SA,
685 Begemann M, Belliveau RA, Bene J, Bergen SE, Bevilacqua E, Bigdeli TB, Black DW,
686 Bruggeman R, Buccola NG, Buckner RL, Byerley W, Cahn W, Cai G, Cairns MJ,
687 Champion D, Cantor RM, Carr VJ, Carrera N, Catts S V., Chambert KD, Chan RCK, Chen
688 RYL, Chen EYH, Cheng W, Cheung EFC, Chong SA, Cloninger CR, Cohen D, Cohen N,
689 Cormican P, Craddock N, Crespo-Facorro B, Crowley JJ, Curtis D, Davidson M, Davis
690 KL, Degenhardt F, Del Favero J, DeLisi LE, Demontis D, Dikeos D, Dinan T, Djurovic S,
691 Donohoe G, Drapeau E, Duan J, Dudbridge F, Durmishi N, Eichhammer P, Eriksson J,
692 Escott-Price V, Essioux L, Fanous AH, Farrell MS, Frank J, Franke L, Freedman R,
693 Freimer NB, Friedl M, Friedman JI, Fromer M, Genovese G, Georgieva L, Gershon ES,
694 Giegling I, Giusti-Rodríguez P, Godard S, Goldstein JI, Golimbet V, Gopal S, Gratten J,
695 De Haan L, Hammer C, Hamshere ML, Hansen M, et al. LD score regression
696 distinguishes confounding from polygenicity in genome-wide association studies. *Nat*
697 *Genet* [Internet]. 2015 [cited 2020 Sep 14];47:291–295. Available from:
698 <https://pubmed.ncbi.nlm.nih.gov/25642630/>
- 699 36. Harper AR, Goel A, Grace C, Thomson KL, Petersen SE, Xu X, Waring A, Ormondroyd
700 E, Kramer CM, Ho CY, Neubauer S, Kolm P, Kwong R, Dolman SF, Desvigne-Nickens
701 P, Dimarco JP, Geller N, Kim DY, Zhang C, Weintraub W, Abraham T, Anderson L,

- 702 Appelbaum E, Autore C, Berry C, Biagini E, Bradlow W, Bucciarelli-Ducci C, Chiribiri
703 A, Choudhury L, Crean A, Dawson D, Desai MY, Elstein E, Flett A, Friedrich M, Heitner
704 S, Helms A, Jacoby DL, Kim H, Kim B, Larose E, Mahmood M, Mahrholdt H, Maron M,
705 McCann G, Michels M, Mohiddin S, Nagueh S, Newby D, Olivotto I, Owens A, Pierre-
706 Mongeon F, Prasad S, Rimoldi O, Salerno M, Schulz-Menger J, Sherrid M, Swoboda P,
707 van Rossum A, Weinsaft J, White J, Williamson E, Tadros R, Ware JS, Bezzina CR,
708 Farrall M, Watkins H. Common genetic variants and modifiable risk factors underpin
709 hypertrophic cardiomyopathy susceptibility and expressivity. *Nat Genet* [Internet]. 2021
710 [cited 2021 Mar 12];53:135–142. Available from:
711 <https://pubmed.ncbi.nlm.nih.gov/33495597/>
- 712 37. Hemani G, Zheng J, Elsworth B, Wade KH, Haberland V, Baird D, Laurin C, Burgess S,
713 Bowden J, Langdon R, Tan VY, Yarmolinsky J, Shihab HA, Timpson NJ, Evans DM,
714 Relton C, Martin RM, Davey Smith G, Gaunt TR, Haycock PC. The MR-Base platform
715 supports systematic causal inference across the human phenome. *Elife* [Internet]. 2018
716 [cited 2018 Jun 15];7. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/29846171>
- 717 38. Verbanck M, Chen CY, Neale B, Do R. Detection of widespread horizontal pleiotropy in
718 causal relationships inferred from Mendelian randomization between complex traits and
719 diseases. *Nat Genet* [Internet]. 2018 [cited 2021 Mar 12];50:693–698. Available from:
720 <https://www.nature.com/articles/s41588-018-0099-7>
- 721 39. Hemani G, Zheng J, Elsworth B, Wade KH, Haberland V, Baird D, Laurin C, Burgess S,
722 Bowden J, Langdon R, Tan VY, Yarmolinsky J, Shihab HA, Timpson NJ, Evans DM,
723 Relton C, Martin RM, Davey Smith G, Gaunt TR, Haycock PC. The MR-base platform
724 supports systematic causal inference across the human phenome. *Elife*. 2018;7.

- 725 40. Burgess S, Thompson SG. Interpreting findings from Mendelian randomization using the
726 MR-Egger method. *Eur J Epidemiol* [Internet]. 2017 [cited 2021 Mar 12];32:377–389.
727 Available from: <http://link.springer.com/10.1007/s10654-017-0255-x>
- 728 41. Hauck L, Grothe D, Billia F. p21CIP1/WAF1-dependent inhibition of cardiac hypertrophy
729 in response to Angiotensin II involves Akt/Myc and pRb signaling. *Peptides* [Internet].
730 2016 [cited 2021 Jan 20];83:38–48. Available from:
731 <https://pubmed.ncbi.nlm.nih.gov/27486069/>
- 732 42. Garnatz AS, Gao Z, Broman M, Martens S, Earley JU, Svensson EC. FOG-2 mediated
733 recruitment of the NuRD complex regulates cardiomyocyte proliferation during heart
734 development. *Dev Biol* [Internet]. 2014 [cited 2021 Mar 31];395:50–61. Available from:
735 </pmc/articles/PMC4190033/>
- 736 43. Toshima H, Koga Y, Uemura S, Zinnouchi J, Kimura N, Nakakura S. Echocardiographic
737 Study on Hypertrophic Cardiomyopathy. *Jpn Heart J* [Internet]. 1976 [cited 2021 Mar
738 12];17:275–289. Available from: <https://pubmed.ncbi.nlm.nih.gov/133253/>
- 739 44. Zollino M, Marangi G, Ponzi E, Orteschi D, Ricciardi S, Lattante S, Murdolo M, Battaglia
740 D, Contaldo I, Mercuri E, Stefanini MC, Caumes R, Edery P, Rossi M, Piccione M,
741 Corsello G, Monica M Della, Scarano F, Priolo M, Gentile M, Zampino G, Vijzelaar R,
742 Abdulrahman O, Rauch A, Oneda B, Deardorff MA, Saitta SC, Falk MJ, Dubbs H, Zackai
743 E. Intragenic KANSL1 mutations and chromosome 17q21.31 deletions: Broadening the
744 clinical spectrum and genotype-phenotype correlations in a large cohort of patients. *J Med*
745 *Genet* [Internet]. 2015 [cited 2021 Jan 20];52:804–814. Available from:
746 <https://pubmed.ncbi.nlm.nih.gov/26424144/>
- 747 45. Koolen DA, Pfundt R, Linda K, Beunders G, Veenstra-Knol HE, Conta EH, Fortuna AM,

748 Gillessen-Kaesbach G, Dugan S, Halbach S, Abdul-Rahman OA, Winesett HM, Chung
749 WK, Dalton M, Dimova PS, Mattina T, Prescott K, Zhang HZ, Saal HM, Hehir-Kwa JY,
750 Willemsen MH, Ockeloen CW, Jongmans MC, Van Der Aa N, Failla P, Barone C, Avola
751 E, Brooks AS, Kant SG, Gerkes EH, Firth H V., Unap K, Bird LM, Masser-Frye D,
752 Friedman JR, Sokunbi MA, Dixit A, Splitt M, Kukolich MK, McGaughran J, Coe BP,
753 Flórez J, Nadif Kasr N, Brunner HG, Thompson EM, Gecz J, Romano C, Eichler EE, De
754 Vries BBA. The Koolen-de Vries syndrome: A phenotypic comparison of patients with a
755 17q21.31 microdeletion versus a KANSL1 sequence variant. *Eur J Hum Genet* [Internet].
756 2016 [cited 2021 Jan 20];24:652–659. Available from:
757 <https://pubmed.ncbi.nlm.nih.gov/26306646/>

758 46. Collins RL, Brand H, Karczewski KJ, Zhao X, Alföldi J, Francioli LC, Khera A V.,
759 Lowther C, Gauthier LD, Wang H, Watts NA, Solomonson M, O'Donnell-Luria A,
760 Baumann A, Munshi R, Walker M, Whelan CW, Huang Y, Brookings T, Sharpe T, Stone
761 MR, Valkanas E, Fu J, Tiao G, Laricchia KM, Ruano-Rubio V, Stevens C, Gupta N,
762 Cusick C, Margolin L, Alföldi J, Armean IM, Banks E, Bergelson L, Cibulskis K, Collins
763 RL, Connolly KM, Covarrubias M, Cummings B, Daly MJ, Donnelly S, Farjoun Y,
764 Ferriera S, Francioli L, Gabriel S, Gauthier LD, Gentry J, Gupta N, Jeandet T, Kaplan D,
765 Karczewski KJ, Laricchia KM, Llanwarne C, Minikel E V., Munshi R, Neale BM, Novod
766 S, O'Donnell-Luria AH, Petrillo N, Poterba T, Roazen D, Ruano-Rubio V, Saltzman A,
767 Samocha KE, Schleicher M, Seed C, Solomonson M, Soto J, Tiao G, Tibbetts K, Tolonen
768 C, Vittal C, Wade G, Wang A, Wang Q, Ware JS, Watts NA, Weisburd B, Whiffin N,
769 Salinas CAA, Ahmad T, Albert CM, Ardissino D, Atzmon G, Barnard J, Beaugerie L,
770 Benjamin EJ, Boehnke M, Bonnycastle LL, Bottinger EP, Bowden DW, Bown MJ,

- 771 Chambers JC, Chan JC, Chasman D, Cho J, Chung MK, Cohen B, et al. A structural
772 variation reference for medical and population genetics. *Nature* [Internet]. 2020 [cited
773 2021 Jan 20];581:444–451. Available from: <https://pubmed.ncbi.nlm.nih.gov/32461652/>
- 774 47. Kulkarni SS, Khokha MK. WDR5 regulates left-right patterning via chromatin-dependent
775 and-independent functions. *Dev* [Internet]. 2018 [cited 2021 Mar 31];145. Available from:
776 <https://dev.biologists.org/content/145/23/dev159889>
- 777 48. Guimarães-Camboa N, Stowe J, Aneas I, Sakabe N, Cattaneo P, Henderson L, Kilberg
778 MS, Johnson RS, Chen J, McCulloch AD, Nobrega MA, Evans SM, Zambon AC. HIF1 α
779 Represses Cell Stress Pathways to Allow Proliferation of Hypoxic Fetal Cardiomyocytes.
780 *Dev Cell* [Internet]. 2015 [cited 2021 Jan 20];33:507–521. Available from:
781 <https://pubmed.ncbi.nlm.nih.gov/26028220/>
- 782 49. Liu ZP, Olson EN. Suppression of proliferation and cardiomyocyte hypertrophy by
783 CHAMP, a cardiac-specific RNA helicase. *Proc Natl Acad Sci U S A* [Internet]. 2002
784 [cited 2021 Jan 20];99:2043–2048. Available from:
785 <https://pubmed.ncbi.nlm.nih.gov/11854500/>
- 786 50. Tshori S, Gilon D, Beeri R, Nechushtan H, Kaluzhny D, Pikarsky E, Razin E.
787 Transcription factor MITF regulates cardiac growth and hypertrophy. *J Clin Invest*
788 [Internet]. 2006 [cited 2021 Jan 20];116:2673–2681. Available from:
789 <https://pubmed.ncbi.nlm.nih.gov/16998588/>
- 790 51. Fu K, Nakano H, Morselli M, Chen T, Pappoe H, Nakano A, Pellegrini M. A temporal
791 transcriptome and methylome in human embryonic stem cell-derived cardiomyocytes
792 identifies novel regulators of early cardiac development. *Epigenetics*. 2018;13:1013–1026.
- 793 52. Mehta G, Kumarasamy S, Wu J, Walsh A, Liu L, Williams K, Joe B, de la Serna IL.

- 794 MITF interacts with the SWI/SNF subunit, BRG1, to promote GATA4 expression in
795 cardiac hypertrophy. *J Mol Cell Cardiol.* 2015;88:101–110.
- 796 53. Wojcik GL, Graff M, Nishimura KK, Tao R, Haessler J, Gignoux CR, Highland HM,
797 Patel YM, Sorokin EP, Avery CL, Belbin GM, Bien SA, Cheng I, Cullina S, Hodonsky
798 CJ, Hu Y, Huckins LM, Jeff J, Justice AE, Kocarnik JM, Lim U, Lin BM, Lu Y, Nelson
799 SC, Park SSL, Poisner H, Preuss MH, Richard MA, Schurmann C, Setiawan VW, Sockell
800 A, Vahi K, Verbanck M, Vishnu A, Walker RW, Young KL, Zubair N, Acuña-Alonso V,
801 Ambite JL, Barnes KC, Boerwinkle E, Bottinger EP, Bustamante CD, Caberto C,
802 Canizales-Quinteros S, Conomos MP, Deelman E, Do R, Doheny K, Fernández-Rhodes
803 L, Fornage M, Hailu B, Heiss G, Henn BM, Hindorff LA, Jackson RD, Laurie CA, Laurie
804 CC, Li Y, Lin DY, Moreno-Estrada A, Nadkarni G, Norman PJ, Pooler LC, Reiner AP,
805 Romm J, Sabatti C, Sandoval K, Sheng X, Stahl EA, Stram DO, Thornton TA, Wassel
806 CL, Wilkens LR, Winkler CA, Yoneyama S, Buyske S, Haiman CA, Kooperberg C, Le
807 Marchand L, Loos RJF, Matise TC, North KE, Peters U, Kenny EE, Carlson CS. Genetic
808 analyses of diverse populations improves discovery for complex traits. *Nature.*
809 2019;570:514–518.
- 810 54. Evans DS, Avery CL, Nalls MA, Li G, Barnard J, Smith EN, Tanaka T, Butler AM,
811 Buxbaum SG, Alonso A, Arking DE, Berenson GS, Bis JC, Buyske S, Carty CL, Chen W,
812 Chung MK, Cummings SR, Deo R, Eaton CB, Fox ER, Heckbert SR, Heiss G, Hindorff
813 LA, Hsueh WC, Isaacs A, Jamshidi Y, Kerr KF, Liu F, Liu Y, Lohman KK, Magnani JW,
814 Maher JF, Mehra R, Meng YA, Musani SK, Newton-Cheh C, North KE, Psaty BM,
815 Redline S, Rotter JI, Schnabel RB, Schork NJ, Shohet R V., Singleton AB, Smith JD,
816 Soliman EZ, Srinivasan SR, Taylor HA, Van Wagoner DR, Wilson JG, Young T, Zhang

- 817 ZM, Zonderman AB, Evans MK, Ferrucci L, Murray SS, Tranah GJ, Whitsel EA, Reiner
818 AP, Sotoodehnia N. Fine-mapping, novel loci identification, and SNP association
819 transferability in a genome-wide association study of QRS duration in African Americans.
820 *Hum Mol Genet.* 2016;25:4350–4368.
- 821 55. Meyer H V., Dawes TJW, Serrani M, Bai W, Tokarczuk P, Cai J, de Marvao A, Henry A,
822 Lumbers RT, Gierten J, Thumberger T, Wittbrodt J, Ware JS, Rueckert D, Matthews PM,
823 Prasad SK, Costantino ML, Cook SA, Birney E, O’Regan DP. Genetic and functional
824 insights into the fractal structure of the heart. *Nature.* 2020;584:589–594.
- 825 56. Thomson KL, Ormondroyd E, Harper AR, Dent T, McGuire K, Baksi J, Blair E, Brennan
826 P, Buchan R, Bueser T, Campbell C, Carr-White G, Cook S, Daniels M, Deevi SVV,
827 Goodship J, Hayesmoore JBG, Henderson A, Lamb T, Prasad S, Rayner-Matthews P,
828 Robert L, Sneddon L, Stark H, Walsh R, Ware JS, Farrall M, Watkins HC. Analysis of 51
829 proposed hypertrophic cardiomyopathy genes from genome sequencing data in sarcomere
830 negative cases has negligible diagnostic yield. *Genet Med* [Internet]. 2019 [cited 2021 Jan
831 24];21:1576–1584. Available from: <https://pubmed.ncbi.nlm.nih.gov/30531895/>
- 832 57. Khera A V, Chaffin M, Aragam KG, Haas ME, Roselli C, Choi SH, Natarajan P, Lander
833 ES, Lubitz SA, Ellinor PT, Kathiresan S. Genome-wide polygenic scores for common
834 diseases identify individuals with risk equivalent to monogenic mutations. *Nat Genet*
835 [Internet]. 2018 [cited 2019 Apr 29];50:1219–1224. Available from:
836 <http://www.nature.com/articles/s41588-018-0183-z>
- 837 58. Tcheandjieu C, Xiao K, Tejada H, Lynch J, Ruotsalainen S, Bellomo T, Palnati M, Judy
838 R, Kember R, Klarin D, Kember R, Verma S, Center RG, Program VMV, Project F,
839 Palotie A, Daly M, Ritchie M, Rader D, Rivas MA, Assimes T, Tsao P, Damrauer S,

- 840 Priest J. High heritability of ascending aortic diameter and multi-ethnic prediction of
841 thoracic aortic disease. *medRxiv* [Internet]. 2020 [cited 2020 Jul
842 21];14:2020.05.29.20102335. Available from:
843 <https://www.medrxiv.org/content/10.1101/2020.05.29.20102335v1>
- 844 59. Tcheandjieu C, Zanetti D, Yu M, Priest JR. Inherited Extremes of Aortic Diameter Confer
845 Risk for a Specific Class of Congenital Heart Disease [Internet]. *Circ. Genomic Precis.
846 Med.* 2020 [cited 2021 Jan 20];13:724–727. Available from:
847 <https://www.ahajournals.org/doi/10.1161/CIRCGEN.120.003170>
- 848 60. Bidviene J, Muraru D, Maffessanti F, Ereminiene E, Kovács A, Lakatos B, Vaskelyte J-J,
849 Zaliunas R, Surkova E, Parati G, Badano LP. Regional shape, global function and
850 mechanics in right ventricular volume and pressure overload conditions: a three-
851 dimensional echocardiography study. *Int J Cardiovasc Imaging* [Internet]. 2021 [cited
852 2021 Jan 20]; Available from: <http://www.ncbi.nlm.nih.gov/pubmed/33389362>
853

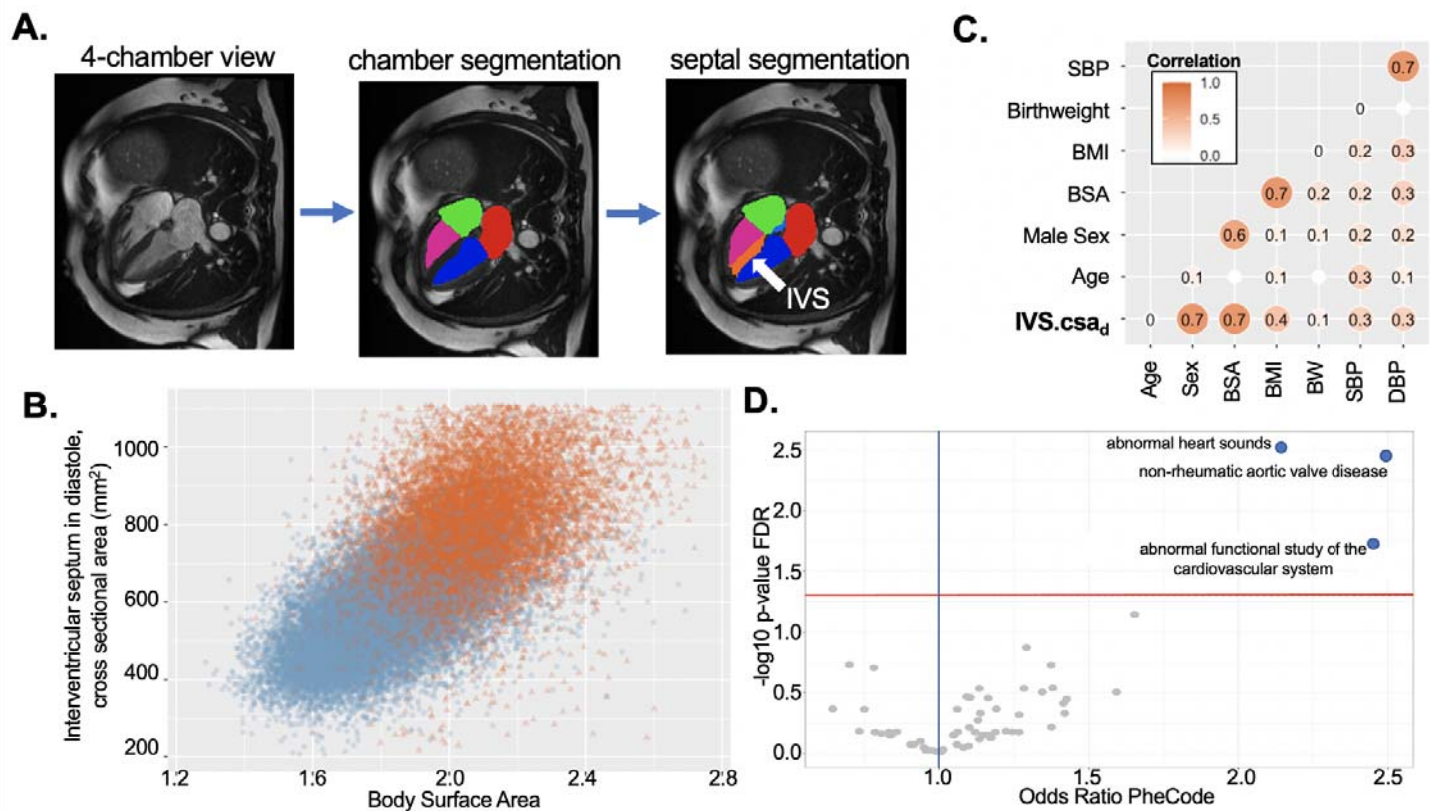


Figure 1. Cross-sectional area in diastole is a new proxy for structure and function of the interventricular septum. **A.** Schema of cardiovascular measurements. Still frame of a cardiac MRI in the four-chamber view (left panel) demonstrating segmentation of lumen of the cardiac chambers (middle panel) with the left ventricle indicated by the blue region, left atrium indicated by the red region, right atrium indicated by the green region, and right ventricle in magenta. Chamber segmentation is followed by segmentation of the interventricular septum (orange, indicated by white arrow). **B.** The distribution of interventricular septal cross-sectional area in diastole (IVS.csa_d). Values are plotted in mm² relative to body surface area (BSA), orange and blue points represent male and female participants respectively. In a simple linear model including Sex, Age, systolic and diastolic blood pressure and body mass index, BSA explains 44.8% of the observed variation in IVS.csa_d (284 mm² IVS.csa_d per m² BSA, $p < 2e-16$). **C.** Correlation plot of IVS.csa_d versus a standard set of clinical measures. IVS.csa_d is most strongly correlated with genetic sex and BSA, moderately correlated with BMI and blood pressure, and slightly correlated with birthweight. Notably there was no strong correlation with age. All non-zero Pearson's correlations reported are strongly statistically significant ($p < 5e-08$). **D.** Volcano plot representing the PheWAS of IVS.csa_d for 67 cardiovascular phenotypes grouped as PheCODES. The X-axis represents the association with the PheCODE per standard deviation change in the IVS.csa_d, and the absolute value of Y-axis represents the negative logarithm of the p-value adjusted with a standard false discovery rate. Across an unbiased assessment of clinical phenotypes, increased IVS.csa_d was positively associated with the clinical findings of abnormal heart sounds, non-rheumatic aortic valve disease, and abnormal functional studies of the cardiovascular system.

CHR	POS	SNP	REF	RA	RA_FREQ	Combined Discovery & Replication		Discovery set		Replication Set	
						Beta [95%CI]	P-value	Beta [95%CI]	P-value	Beta [95%CI]	P-value
17	45961419	rs62063281	A	G	0.22	8.60 [6.33, 10.88]	1.31×10^{-13}	7.72 [5.20, 10.2]	1.89×10^{-09}	12.39 [7.07, 17.71]	5.11×10^{-06}
6	36681816	rs2376620	A	G	0.16	8.38 [5.80, 10.96]	2.00×10^{-10}	7.72 [4.87, 10.5]	1.14×10^{-07}	11.29 [5.24, 17.34]	2.58×10^{-04}

Table 1. Genome wide associations from measurements of IVS.csad in European ancestry individuals.
RA: risk allele, FREQ: frequency

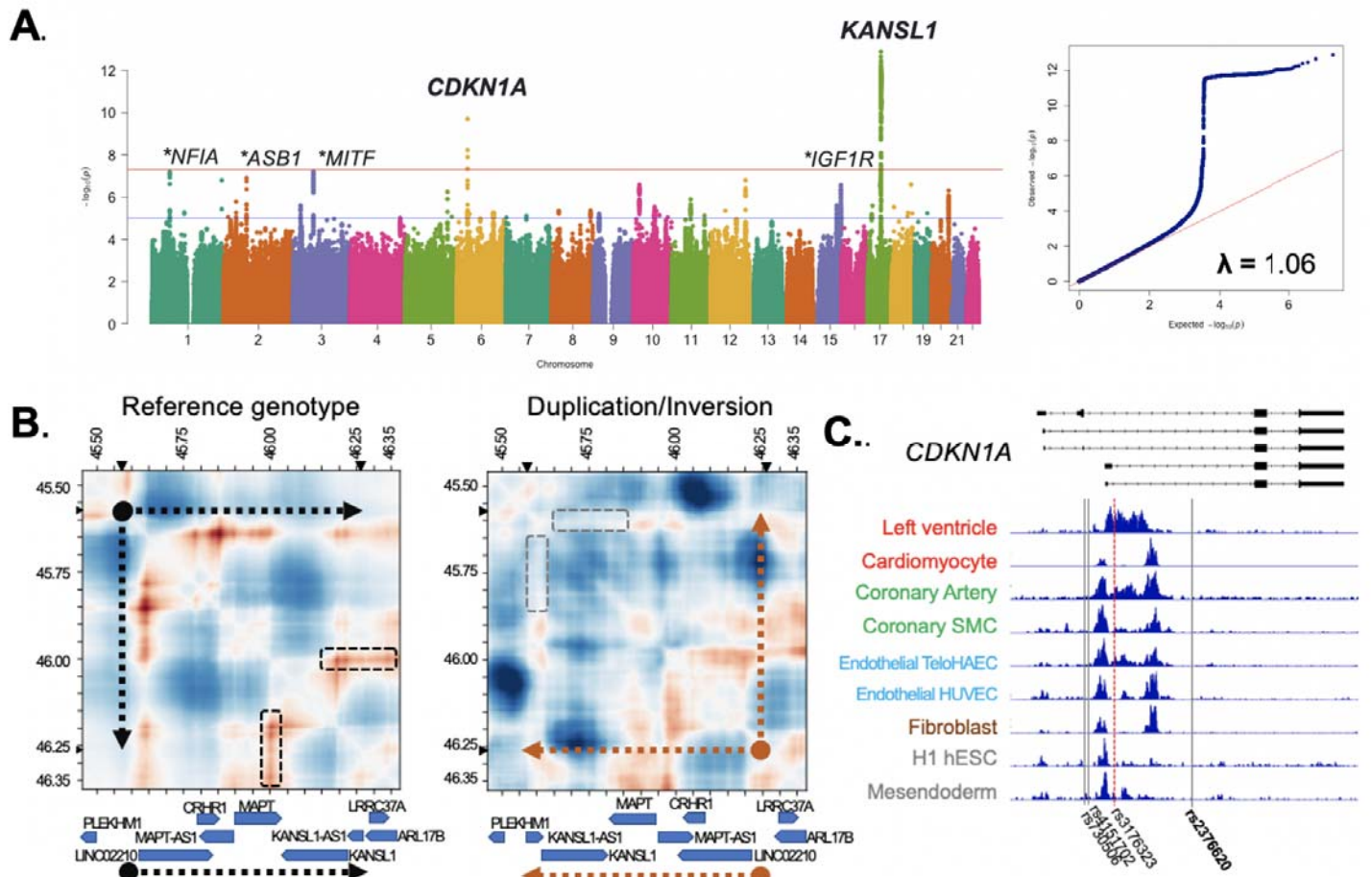


Figure 2. GWAS of IVS.csad identifies two loci meeting standard significance thresholds. A. Manhattan plot with two loci at *KANSL1* and *CDKN1A*, along with loci below $5e-8$ marked with a star. Quantile-quantile plot indicates an absence of systematic inflation. **B.** Two-dimensional plots of the Akita model of genome folding with GRCh38 coordinates on the X and Y axis with relevant genes and strand orientation indicated by blue boxes. Predicted frequency of interaction between regions of DNA is indicated by color, where red/blue shows a higher/lower frequency of interaction relative to what is expected given linear genomic distance. The reference genotype predicts a strong chromatin interaction (black boxes) between the *KANSL1* promoter and regions including the *MAPT* gene body and *KANSL1-AS1*. The common duplication/inversion 17:45571611-46261810 (orientation indicated by the orange arrow) is predicted to completely disrupt the interaction (corresponding locations within the inversion indicated by the grey boxes). The lead variant *rs62063281* is an eQTL for multiple genes across the locus and a strong splicing QTL for *MAPT* and *KANSL1* (-1.5 normalized expression, $p=1.7e-88$) within the left ventricle [Table S3]. **C.** Plot of tracks showing output of the activity-by-contact (ABC) which identifies *CDKN1A* as the gene target for the lead variant *rs2376620*, which is the marker for a haplotype inclusive of three other variants localized in open chromatin enhancer regions present most prominently in the left ventricle. Displayed for contrast are DNase-seq for cardiomyocyte, human umbilical vein endothelial cell (HUVEC), fibroblast of fetal lung, H1 human embryonic stem cells (hESCs), and H1-derived mesendodermal precursors; or ATAC-seq for heart left ventricle (LV), coronary artery, coronary smooth muscle cell (SMC), and telomerase immortalized human aortic endothelial cells (TeloHAEC).

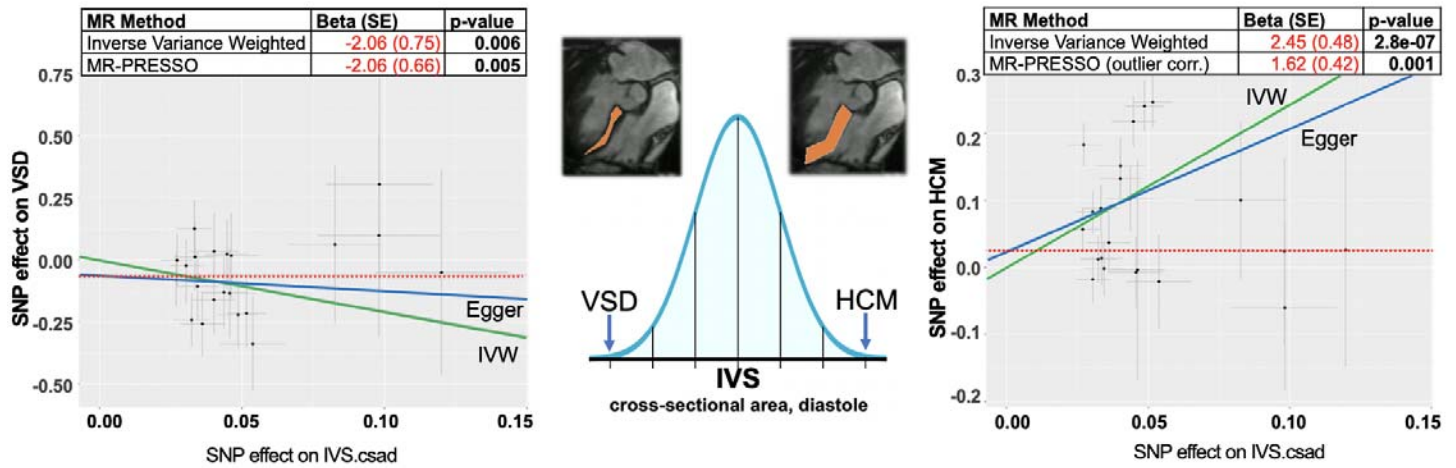


Figure 3. Mendelian Randomization suggests that inheritance of increased IVS.csad is causal for HCM while inheritance of decreased IVS.csad is causal for VSD. Twenty two loci were selected for use as instrumental variables for interventricular septum, cross-sectional area as the exposure for two-sample Mendelian Randomization analysis of ventricular septal defects (VSD) and hypertrophic cardiomyopathy (HCM). Regression lines for the MR-Egger and Inverse Variants Weighted (IVW) methods are displayed in blue and green respectively, with the intercept from the MR-Egger method in dotted red for both graphs. The causal relationships of IVS.csad and risks for VSD and HCM are not driven by a single variant when analyzed individually or in a “leave-one out” sensitivity analysis (Figure S6).