Transcriptome- and proteome-wide Mendelian randomization to prioritize therapeutic targets for coronary heart disease

2 3

1

4 Liam Gaziano PhD#^{1,2}, Elias Allara MD, PhD#^{2,3,4}, Claudia Giambartolomei PhD^{5,6}, David

5 Stacey PhD^{7,8}, Jing Hua Zhao PhD^{2,3}, Hesam Dashti PhD^{9,10}, Tao Jiang MSc^{2,11}, Scott C. Ritchie

- 6 PhD^{2,3,12,13,14,15,16}, Brian R Ferolito MS⁹, Danielle Rasooly PhD^{9,17}, Gina M. Peloso
- 7 PhD^{9,18}, Emanuele Di Angelantonio MD, PhD, FMedSci^{2,3,4,5,12,13}, Eleanor Wheeler PhD¹⁹, Maik
- 8 Pietzner PhD^{19,20,21}, Themistocles L Assimes MD, PhD^{22,23}, Peter WF Wilson MD^{24,25}, Kelly
- 9 Cho PhD, MPH^{26,27,28}, Krishna G Aragam MD, MS^{29,30}, Stephen Burgess PhD^{2,31}, John Danesh
- 10 FMedSci^{2,3,4,12,13,32}, Claudia Langenberg MD, PhD^{19,20,21}, Juan Pablo Casas MD, PhD³³, J
- Michael Gaziano MD, MPH^{26,27,28}, Alexandre C Pereira MD, PhD^{+9,27,28,34}, Adam S Butterworth
 PhD^{+2,3,4,12,13}
- 13
- 14 #Contributed equally. †Contributed equally.
- 15
- ¹⁶ ¹Massachusetts Veterans Epidemiology Research and Information Center (MAVERIC), Veterans
- 17 Affairs Healthcare System, Boston, MA, USA
- 18 ²British Heart Foundation Cardiovascular Epidemiology Unit, Department of Public Health and
- 19 Primary Care, University of Cambridge, Cambridge, UK
- ³Victor Phillip Dahdaleh Heart and Lung Research Institute, University of Cambridge,
- 21 Cambridge, UK
- ⁴NIHR Blood and Transplant Research Unit in Donor Health and Behaviour, Department of
- 23 Public Health and Primary Care, University of Cambridge, Cambridge, UK
- ⁵Health Data Science Centre, Human Technopole, Milan, Italy
- ⁶Million Veteran Program (MVP) Consultant, Veterans Affairs Healthcare System, Boston, MA,
 USA
- 27 ⁷Australian Centre for Precision Health, Unit of Clinical and Health Sciences, University of
- 28 South Australia, Adelaide, South Australia, Australia
- ⁸South Australian Health and Medical Research Institute, Adelaide, South Australia, Australia
- ⁹Million Veteran Program (MVP) Coordinating Center, Veterans Affairs Healthcare System,
- 31 Boston, MA, USA
- ¹⁰Broad Institute of MIT and Harvard, Cambridge, MA, USA
- ³³ ¹¹Victor Philip Dahdaleh Heart and Lung Research Institute, University of Cambridge,
- 34 Cambridge, UK
- ¹²British Heart Foundation Centre of Research Excellence, University of Cambridge, Cambridge,
 UK
- ¹³Health Data Research UK Cambridge, University of Cambridge, Cambridge, UK
- ¹⁴Health Data Research UK Cambridge, Wellcome Genome Campus, Hinxton, UK
- ¹⁵Cambridge Baker Systems Genomics Initiative, Department of Public Health and Primary
- 40 Care, University of Cambridge, Cambridge, UK
- 41 ¹⁶Cambridge Baker Systems Genomics Initiative, Baker Heart & Diabetes Institute, Melbourne,
- 42 Victoria, Australia
- 43 ¹⁷Division of Aging, Brigham and Women's Hospital, Harvard Medical School, Boston, MA,
- 44 USA
- ¹⁸Department of Biostatistics, Boston University School of Public Health, Boston, MA, USA
- 46 ¹⁹MRC Epidemiology Unit, University of Cambridge, Cambridge, UK

- 47 ²⁰Precision Healthcare University Research Institute, Oueen Mary University of London,
- 48 London, UK
- ²¹Computational Medicine, Berlin Institute of Health at Charité Universitätsmedizin Berlin, 49
- 50 Berlin, Germany
- ²²VA Palo Alto Health Care System, Palo Alto, CA, USA 51
- ²³Department of Medicine, Stanford University School of Medicine, Stanford, CA, USA 52
- 53 ²⁴Emory University School of Medicine (Cardiology), Emory University, Atlanta, GA, USA
- 54 ²⁵Internal Medicine, VA Atlanta Healthcare System, Decatur, GA, USA
- 55 ²⁶Million Veteran Program (MVP) Coordinating Center, VA Boston Healthcare System, Boston, 56 MA, USA
- 57 ²⁷Division of Aging, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, 58 USA
- 59 ²⁸Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, 60 MA, USA
- ²⁹Cardiology Division, Massachusetts General Hospital, Boston, MA, USA 61
- 62 ³⁰Cardiovascular Disease Initiative, Broad Institute of MIT and Harvard, Cambridge, MA, USA
- ³¹MRC Biostatistics Unit, University of Cambridge, Cambridge, UK 63
- 64 ³²Human Genetics, Wellcome Sanger Institute, Hinxton, UK
- ³³Biomarker Development/Translational Medicine, Novartis Institutes for BioMedical Research, 65
- 66 Cambridge, MA, USA
- ³⁴Laboratory of Genetics and Molecular Cardiology, Heart Institute, University of Sao Paulo, 67
- Sao Paulo, Brazil 68

69		
70	Corresponding author:	Professor Adam Butterworth
71		Department of Public Health and Primary Care
72		University of Cambridge
73		Victor Phillip Dahdaleh Heart & Lung Research Institute,
74		Papworth Road, Cambridge, CB2 0BB
75		Tel. 0044 1223 748673
76		E-mail: <u>asb38@medschl.cam.ac.uk</u>
77		
78		

- 79
- 80 Manuscript word count: 3703
- 81
- 82
- 83

84

85 Abstract

86 Despite widespread use of drugs targeting traditional cardiovascular risk factors such as lipids and 87 blood pressure, a high burden of coronary heart disease (CHD) remains, hence novel therapeutics 88 are needed for people who harbor residual risk. Using transcriptomic and proteomic data to 89 instrument 15,527 genes or proteins, we conducted systematic *cis*-Mendelian randomization (MR) 90 and conditional colocalization analyses with a genetic meta-analysis involving nearly 300,000 91 CHD cases. We identified 567 targets with putative causal relevance to CHD, of which 69 were 92 not identified in previous genetic discovery or MR studies and were the sole causal signal in that 93 genomic region. To aid translation of our findings, we annotated results with up-to-date 94 information on drugs acting on these targets. Our results revealed opportunities for drug 95 repurposing and development prioritization. For example, we provide evidence that cilostazol, a 96 drug that targets *PDE3A* and is currently used for claudication, could be repurposed for prevention 97 of CHD.

98

99

100 Introduction

101

There is a need to develop novel therapeutics to prevent CHD, which remains a leading cause of death in adults across the globe¹. Although several safe and effective drugs exist for primary prevention of CHD, many adults experience residual hypertension or dyslipidemia following standard therapies^{2,3}, suggesting a need to identify new approaches to target these traditional cardiovascular risk factors. Furthermore, an increasing proportion of myocardial infarctions occur in patients without traditional cardiovascular risk factors⁴, highlighting the need to identify and modulate novel pathways implicated in coronary atherosclerosis and thrombosis.

109

Around 50% of drug development failures occur due to lack of efficacy⁵, suggesting that the chosen target is not causally related to the outcome. Mendelian randomization (MR)⁶ analyses exploit the random allocation of genetic alleles through the population and avoid the pitfalls of relying on pre-clinical models that may not translate to human biology. Indeed, drugs that target proteins with supportive evidence from human genetic studies have a higher chance of gaining regulatory approval⁷.

116

Recently, large-scale proteomic profiling has led to the creation of extensive catalogs of genetic variants that influence circulating protein levels⁸⁻¹¹. Similarly, the GTEx consortium¹² has measured transcriptome-wide RNA levels in nearly 50 tissues to identify associated genetic variants. In combination, these resources can provide tissue-specific instruments to evaluate the causal relevance of the human proteome and transcriptome on CHD. To achieve this goal, we leverage transcriptomic and proteomic summary statistics to perform Mendelian randomization and colocalization analyses between more than 15,000 genes or proteins and the risk of CHD,

- 124 utilizing a *de novo* genome-wide association study (GWAS) meta-analysis involving nearly
- 125 300,000 cases from the Million Veteran Program¹³, CARDIoGRAMplusC4D¹⁴ and FinnGen¹⁵.

126

127 **Results**

We took a similar approach as our previous transcriptomic and proteomic MR analyses¹⁶, with 128 the overall study design for the current analysis shown in Figure 1. To minimize the potential for 129 130 horizontal pleiotropy (i.e. the CHD-associated variants acting through a different gene/protein to 131 the one being instrumented), we used locally-acting (cis), conditionally-independent, protein 132 quantitative trait loci (*cis*-pQTLs) to instrument levels of 1,497 plasma proteins assayed using the SomaScan V4 platform in the Fenland study¹⁷ (Supplementary Table 1). We also selected 133 134 conditionally-independent, cis gene expression QTLs (cis-eQTLs) from Version 8 of the Genotype-Tissue Expression (GTEx) resource¹² as instruments for 15,412 protein-coding genes 135 136 (Supplementary Table 2). With these instruments, we performed MR using a GWAS meta-137 analysis of CHD that we conducted using results from European ancestry participants of the Million Veteran Program (MVP)¹⁸, the latest CARDIoGRAMplusC4D Consortium GWAS¹⁴, and 138 the FinnGen biobank¹⁵ involving 296,537 CHD cases among 1,593,766 participants 139 140 (Supplementary Table 3). 15,527 genes or proteins had at least one instrumental variant in at least one tissue, hence we set a Bonferroni-adjusted significance threshold at $P=3.22\times10^{-6}$ 141 142 (0.05/15,527). A full list of MR results can be found in Supplementary Table 4. To reduce the 143 possibility that MR associations arise due to confounding by linkage disequilibrium (LD), we additionally performed a pairwise conditional colocalization¹⁹ analysis between pQTL or eQTL 144 145 summary statistics and CHD summary statistics for all significant MR results. We annotated 146 results (those that passed both Bonferroni-adjusted P-value for MR and colocalization 147 [PPH4>0.80] thresholds) with information about drugs that act on these targets using the OpenTargets²⁰ and DrugBank²¹ resources (Supplementary Table 5-7), and assessed novelty and 148 149 potential for horizontal pleiotropy (Supplementary Table 8). For specific targets of interest, we

performed phenome-wide association scans of instrumental variants and colocalization to identify
putative mechanisms, safety signals and alternative indications (Supplementary Table 9). Lastly,
we conducted enrichment analyses to assess whether the potential causal genes for CHD occur
more frequently in certain biological pathways (Supplementary Table 10).

154

Overall, we found 567 unique genes/proteins colocalizing with CHD. Using eQTL instruments,
736 genes passed the MR significance threshold (*P*<3.22×10⁻⁶) with 551 (74.9%) colocalizing
(PPH4>0.8). Using pQTL instruments, 39 proteins passed MR significance with 24 colocalizing.
Twenty of these were also tested using eQTL instruments, of which nine (*ABO*, *GSTT2B*, *IL6R*, *MST1*, *MXRA7*, *PCSK9*, *PDE5A*, *TMEM106A*, *TMEM106B*) also passed the colocalization
threshold, and seven (*ABO*, *GSTT2B*, *MST1*, *MXRA7*, *PCSK9*, *PDE5A*, *TMEM106A*) had an eQTL
signal in at least one tissue in the same direction as the plasma pQTL.

162

163 **Replication of past Mendelian randomization studies**

164 We assessed whether we could replicate results from two transcriptome-/proteome-wide MR 165 studies that have included CHD as an outcome, both of which used outcome summary statistics 166 from the 2015 CARDIoGRAMplusC4D GWAS that included 60,801 CHD cases²². First, Zheng et al.¹⁹ used five^{9,23-26} publicly available pQTL resources to select instrumental variants, 167 168 identifying LPA and IL6R with MR $P < 3.22 \times 10^{-6}$ when restricting to results that used cisinstruments and colocalized. Here, we also identified *IL6R* ($P=2.5\times10^{-22}$, highest PPH4 in any 169 170 tissue is 1.00) but could not assess LPA, as it is not measured on the SomaScan V4 panel. Second, Richardson et al.²⁷ used cis-eQTL variants from GTEx Version 7 and eQTLGen²⁸ as instruments 171 and found 21 protein-coding genes with MR $P < 3.22 \times 10^{-6}$ and P_{HEIDI} (a test for confounding by 172

173 LD) >0.05, of which 20 were replicated here (ABO, AIDA, ATP2B1, CARF, CELSR2, FAM117B, 174 FES, GGCX, GUCYIAI, ICAIL, JCAD, LIPA, MIA3, MORF4L1, MRAS, NBEAL1, PHACTR1, 175 *PSRC1, SORT1, SWAP70*), suggesting that our framework identifies robust results. Only *VAMP8* 176 from Richardson *et al.* did not colocalize in our analysis (highest PPH4 in any tissue was 0.33). In 177 the VAMP8 region, we instead identified GGCX ($P=1.7 \times 10^{-52}$, PPH4=1.00 in a ortic artery tissue), 178 which is targeted by the Anisindione, an anticoagulant occasionally used in those who cannot 179 tolerate warfarin²⁹. It is reassuring that our approach identified GGXC over VAMP8 in that region, 180 considering the biological role of GGCX in coagulation by activating vitamin K-dependent 181 proteins and the relevance of coagulation to CHD.

182

183 Assessment of horizontal pleiotropy

184 We considered an instrumental variant that influences expression of multiple genes / proteins as 185 having the potential to introduce a form of horizontal pleiotropy that arises from shared regulation of nearby genes. Therefore, we grouped results with instrumental variants that are in close 186 187 proximity (250kb) or are correlated (r^2 >0.2 in 1000 Genomes EUR). The 567 genes/proteins, 188 hereafter referred to as "targets", that passed MR and colocalization thresholds lie within 283 189 genomic regions. We allocated results into tiers based on the number of targets within a genomic 190 region: Tier 1 includes 165 targets that are the sole result within a region to pass MR and 191 colocalization thresholds, while the remaining 404 results (located within 116 regions) were 192 annotated as Tier 2 (Supplementary Table 8). Tier 1 genes included several well-known to relate 193 to CHD risk, such as APOB, IL6R, IL1RN, LIPA, and others previously identified by GWAS as 194 the most likely causal gene in the region, such as ITGA1, MYO9B, and SERPINA1¹⁴. Because we 195 tested all protein-coding genes in GTEx, Tier 1 results that are the only significant signal in the

196 region are less likely to exhibit this form of horizontal pleiotropy. Conversely, since Tier 2 results 197 contain multiple targets, it can be difficult to disentangle the gene responsible for driving the 198 signal(s) in each region using MR and colocalization alone. For example, in one region on 199 chromosome 1 (region #13 in Supplementary Table 8), there are four targets (SARS1, CELSR2, 200 *PSRC1* and *SORT1*) that pass MR and colocalization thresholds. However, previous functional work has shown that *SORT1* is likely driving the signal in the 1p13.3 locus³⁰, meaning results for 201 202 SARS1, CELSR2, and PSRC1 likely arose due to horizontal pleiotropy (Supplementary Figure 203 1).

204

205 Novel putative therapeutic targets

We classified results as novel if they (1) were based on instrumental variants that were not correlated ($r^2 < 0.2$ in 1000 Genomes EUR) or in close proximity (± 250 kb) with genome-wide significant ($P < 5 \times 10^{-8}$) variants for CHD reported in Tcheandjieu *et al.*¹⁸, Aragam *et al.*¹⁴, or Supplementary Table 2 of Aragam *et al.*¹⁴ which includes 216 CHD-associated variants previously reported in the literature and (2) were not reported in Zheng *et al.*¹⁹ or Richardson *et al.*²⁷. We classified 184 targets as novel, 69 of which were classified also as Tier 1 (**Figure 2**, **Supplementary Table 8**).

213

214 Targets with existing drugs

Of the 567 targets that passed MR and colocalization thresholds, 30 had drugs developed that act on these targets according to DrugBank²¹ (**Supplementary Table 5**) and 32 according to OpenTargets²⁰ (**Supplementary Table 6**), of which 19 from both. Forty-three targets had an existing drug according to either DrugBank or OpenTargets and the remaining 524 targets were

classified as not currently having a drug (Supplementary Table 7). Of the 43 targets with existing
drugs, several are involved in well-known vascular pathways such as lipid metabolism (e.g., *PCSK9, LPL, APOC3, NPC1L1, CETP, ABCA1*), hemostasis (e.g., *F11, FN1, PDE3A, PIK3CB, TFPI, KLKB1, PDE8A, GGCX*) and inflammation (e.g. *IL6R, IL23A, MIF*). Some of these - such
as *PCSK9*³¹, *NP1CL1*³², and *CETP*³³ - have already been shown to be effective therapeutic targets
for CHD in clinical trials, while others like *IL6R* are currently being tested³⁴.

225

226 Repurposing

227 Our results also reveal notable repurposing opportunities. An example is the novel, Tier 1 result 228 for PDE3A, which is targeted by several inhibitory drugs for a wide range of indications, including 229 thrombotic events and respiratory diseases (Supplementary Table 5 and 6). Higher genetically-230 predicted expression of *PDE3A* was positively associated with CHD ($P=2.25\times10^{-7}$) using a 231 colocalizing (PPH4=0.99) eQTL instrumental variant (rs7488772) in subcutaneous adipose tissue, 232 suggesting that therapeutic inhibition might be beneficial. A phenome-wide association scan with 233 colocalization (PheWAS-coloc) for rs7488772 showed that higher genetically-predicted 234 expression of PDE3A was associated with higher levels of several cardiometabolic risk factors, 235 including systolic blood pressure, triglycerides, and glycated hemoglobin (Figure 3 and 236 Supplementary Table 9).

237

238 Identifying safety concerns for non-cardiovascular drugs

239 We also found an inverse association between higher genetically-predicted expression of *CXCR4*

240 (instrumented by rs10171574) and risk of CHD ($P=9.52 \times 10^{-7}$) (Supplementary Table 5).

241 PheWAS-coloc showed associations of the CHD-increasing (and CXCR4-decreasing) allele with

higher neutrophil percentage of white blood cells and lower lymphocyte percentage of white blood cells (**Supplementary Figure 2** and **Supplementary Table 9**). These associations are consistent with the role of *CXCR4* inhibitors such as plerixafor, which is commonly used in stem cell transplantation to mobilize granulocytes and their precursors into the blood stream³⁵. Our findings suggest that *CXCR4* inhibitors (e.g. plerixafor and uloplucumab) may increase risk of CHD in an on-target fashion, raising potential safety concerns for long-term use.

248

249 Disentangling effects for drugs with multiple targets

250 One therapeutic target, *PLA2G5*, which encodes secretory phospholipase A2 group V (sPLA2-V), 251 is inhibited by Varespladib methyl. We found that higher genetically-predicted PLA2G5 252 expression, a Tier 1 result, was associated with higher risk of CHD ($P=3.18\times10^{-6}$) (Figure 4) using 253 an instrumental variant (rs12408798) in aortic artery tissue (PPH4=0.99). However, Varespladib 254 methyl did not reduce coronary events in ~5000 patients with acute coronary syndrome in the 255 VISTA-16 trial³⁶, and even showed an increased risk of myocardial infarction. Varespladib methyl 256 also targets isoenzymes sPLA2-IIA (encoded by PLA2G2A) and sPLA2-X (encoded by 257 PLA2G10)³⁷. Here, eQTL instruments for PLA2G2A expression appeared in multiple tissues, none 258 of which were associated with CHD (P>0.52), while there was no eQTL instrument for PLA2G10 259 in any tissue nor a pQTL instrument.

260

261 **Targets currently without a drug**

The 524 targets for which drugs do not currently exist can help prioritize drug development programs. For example, we found positive associations for genetically-predicted *ABCA8* expression with risk of CHD ($P=6.50\times10^{-7}$), suggesting ATP binding cassette subfamily A

265 member 8 as a potential therapeutic target. PheWAS-coloc for *ABCA8* (instrumented by 266 rs34931250 in several tissues) showed positive associations with LDL-cholesterol concentration 267 and Apolipoprotein B, and inverse associations with HDL-cholesterol and Apolipoprotein-A1 268 (**Figure 5** and **Supplementary Table 9**), suggesting pro-atherogenic lipids and poor cholesterol 269 efflux as the potential mediators.

270

271 Pathway and gene ontology term enrichment analysis

272 To determine whether our set of 567 putative causal genes for CAD was over-represented in 273 specific biological mechanisms or pathways, we performed enrichment analyses using Enrichr³⁸⁻ 274 ⁴⁰. Overall, we found significant ($P_{adj} < 0.05$) evidence of enrichment for 25 Kyoto Encyclopedia 275 of Genes and Genomes (KEGG) pathways or gene ontology (GO) terms, with the highest 276 significance observed for lipid-related categories (Supplementary Table 10). This included the KEGG pathways "cholesterol metabolism" ($P=8.85 \times 10^{-10}$, odds ratio [OR]=12.32) and "fat 277 digestion and absorption" ($P=2.37\times10^{-5}$, OR=7.95), as well as the GO terms "lipoprotein particle 278 279 remodeling" ($P=3.35\times10^{-5}$, OR=12.24) and "sterol transfer activity" ($P=1.49\times10^{-4}$, OR=12.36). 280 Furthermore, Enricht also highlighted ($P_{adj} < 0.05$) GO terms relating to endothelial cell migration 281 $(P=2.71\times10^{-5}, OR=5.19)$, focal adhesion $(P=3.00\times10^{-4}, OR=2.33)$, and cell-substrate junctions $(P=4.03 \times 10^{-4}, OR=2.28)$. Thus, these analyses revealed a highly significant enrichment of genes 282 283 involved in CAD-related biological pathways and mechanisms.

284

We then performed a separate enrichment analysis focusing specifically on our set of 165 Tier 1 putative causal genes. These analyses highlighted 51 significant (P_{adj} <0.05) KEGG pathways or GO terms (**Supplementary Table 11**), double the number observed in the previous enrichment

288 analysis of all 567 candidates and which again included the KEGG pathways "cholesterol 289 metabolism" ($P=3.39\times10^{-6}$, OR=16.97) and "fat digestion and absorption" ($P=4.29\times10^{-4}$, 290 OR=12.61). The Tier 1 analysis also highlighted a major theme related to cell migration, motility, 291 and chemotaxis, while reiterating an enrichment of gene products localized to focal adhesions $(P=4.18 \times 10^{-6}, OR=4.84)$ and cell-substrate junctions $(P=5.29 \times 10^{-6}, OR=4.73)$. Together, this 292 293 analysis, confined to the 165 Tier 1 genes, highlights pathways that influence CAD risk by 294 regulating spatial arrangements and interactions between cells at the arterial wall, which is 295 consistent with current models of atherosclerosis⁴¹.

296

297 **Discussion**

298

299 Using instruments derived from gene expression in multiple tissues and plasma protein levels, we 300 identified 567 targets within 283 genomic regions with a putative causal relevance to CHD. We 301 annotated the findings with information on existence of current targeting drugs to aid translation 302 of our findings into actionable interventions and followed-up specific results with phenome-wide 303 scans and colocalization to identify possible mechanisms or on-target safety concerns. While 304 previous studies have used transcriptomic and proteomic genetic instruments on CHD^{19,27}, none 305 have been applied to a GWAS of this size. As a result, this analysis increased the number of 306 genes/proteins with putative causal relevance to the development of CHD, with 69 novel, Tier 1, 307 eQTL-CHD or pQTL-CHD pairs.

308

309 These results can help establish novel therapeutics for prevention of CHD in several ways. First, 310 we identified possible opportunities for drugs currently on the market, that might be repurposed. 311 PDE3A is targeted by multiple small molecules, two of which (milrinone and amrinone) were 312 tested in overt heart failure but displayed null or harmful effects⁴²⁻⁴⁴, and one, cilostazol, which is 313 currently used for intermittent claudication. Cilostazol has antiplatelet and vasodilatory effects⁴⁵ 314 by inhibiting PDE3A's catalysis of inactive adenosine monophosphate (AMP) from cyclic 315 (cAMP). We found higher genetically-predicted *PDE3A* expression to be positively associated 316 with CHD and a number of cardiovascular risk factors. Some of the PheWAS-coloc results for the 317 PDE3A instrumental variant have been additionally corroborated by human trials of cilostazol. For 318 example, trials have found that cilostazol reduces triglycerides and increases HDL-cholesterol^{46,47}. 319 Cilostazol, in addition to its antiplatelet effects, also acts as a vasodilator, so the PheWAS-coloc 320 result for systolic blood pressure is perhaps not surprising, especially considering cilostazol has

been shown in trials to reduce ankle brachial index⁴⁸. The congruence between existing trial data and the PheWAS-coloc analysis provides evidence that the instrumental variant for *PDE3A* expression is accurately mimicking the on-target effects of cilostazol. Our findings additionally corroborate previous findings for PDE3B (a protein with similar function as PDE3A and is also targeted by Cilostazol⁴⁹), which found loss of function mutations in *PDE3B* are associated with reductions in CHD, decreased triglycerides, and increased HDL-C⁵⁰.

327

Cilostazol has also been shown to reduce secondary ischemic stroke by 32% in a meta-analysis of trials⁵¹. It is often tested against antiplatelet drugs, like aspirin, clopidogrel, or a combination of the two, and has shown superiority in preventing bleeding events as well, including a 57% reduction in hemorrhagic stroke⁵¹. While aspirin and clopidogrel reduce CHD in primary prevention, there is debate about whether the increased bleeding risk outweighs the benefits⁵². Therefore, large trials of cilostazol for primary prevention of CHD may be warranted to determine if it provides a safer and more effective antiplatelet strategy than aspirin or clopidogrel.

335

336 Second, our findings can help clear up questions surrounding drugs that target multiple proteins. 337 Varespladib inhibits isoenzymes sPLA2-V (encoded by PLA2G5), sPLA2-IIA (encoded by 338 PLA2G2A) and sPLA2-X (encoded by PLA2G10). All three of these genes have been assessed previously in MR studies for CHD and all have been null⁵³⁻⁵⁵. The previous study instrumented 339 340 gene expression of *PLA2G5* with rs525380 in aorta adventitia tissue from 133 participants of the 341 Advanced Study of Aortic Pathology (ASAP). ASAP participants were genotyped using the 342 Illumina Human 610 W-Quad Bead array without imputation. The variant from Holmes et al., rs525380, is in weak LD (r^2 =0.28 in 1000 Genomes European ancestry) with the one used here for 343

PLA2G5, and this may explain the different results (**Figure 4**). We replicated a separate study that found no association between genetically-predicted *PLA2G2A* and CHD⁵⁴ that used an identical instrumental variant as the present analysis (rs11573156). A previous MR analysis on sPLA2-X suffered from weak instruments for *PLA2G10* expression⁵⁵ (similarly, there were no significant eQTLs for *PLA2G10* in GTEx), and without strong genetic instruments it is difficult to understand its causal relevance to CHD through MR.

350

351 The phase 3 VISTA-16 trial found that Varespladib had no effect on a composite outcome of 352 cardiovascular mortality, nonfatal MI, nonfatal stroke, or unstable angina in those with acute 353 coronary syndrome³⁶ but did find an increase in myocardial infarction. This may be explained by 354 Varespladib's inhibition of sPLA2-X, which showed antiatherogenic effects in mouse models⁵⁶. 355 Indeed, Varespladib is five times more active for sPLA2-X than sPLA2-V, with half maximal inhibitory concentrations of 15 nM for sPLA2-X versus 77 nM for sPLA2-V⁵⁷. In aggregate, these 356 357 results suggest that specific inhibition of sPLA2-V without inhibition of sPLA2-IIA or sPLA2-X 358 could be an effective intervention for CHD prevention.

359

Third, MR results for non-druggable targets can help to indicate which targets should be carried forward for development. We found that higher levels of *ABCA8* gene expression decrease HDLcholesterol levels and increase LDL-cholesterol levels and risk of CHD, suggesting that modulating ABCA8 could prevent CHD, potentially through either reductions in Apo-Bcontaining lipoproteins or improved cholesterol efflux. Our findings for HDL-cholesterol are discordant with studies of deleterious *ABCA8* mutations and *ABCA8* knock-out mice, which have shown *decreased* HDL-cholesterol levels ⁵⁸⁻⁶⁰. Combined with the biological role of ABCA8 and

other ATP binding cassette proteins⁵⁸, and the congruence between the effect of the instrumental 367 368 variant on ApoB-containing lipoproteins and CHD, these findings suggest that it is actually 369 reduced ABCA8 protein levels that increase ApoB-containing lipoproteins and risk of CHD. This 370 highlights one of the challenges of transcriptomic-derived instruments, in that there are several 371 reasons why expression of a gene may be elevated. For example, gene expression may be 372 upregulated as a compensatory mechanism for a variant that causes decreased function of a protein 373 through changes in key amino acids, splicing, or other mechanisms. It is possible such a 374 phenomenon is occurring for ABCA8, particularly because the instrumental variant lies within a 375 splice region⁶¹. More in-depth evaluation of ABCA8 and the many other potentially causal targets 376 instrumented by gene expression are warranted to determine the desired direction of therapeutic 377 modulation.

378

379 Lastly, our findings can help identify on-target effects of drugs that may increase the risk of CHD, 380 suggesting potential safety concerns We found an association between greater expression of 381 CXCR4 and lower risk of CHD, which is consistent with a previous study in hyperlipidemic mice⁶². 382 Additionally, CXCL12, the natural ligand for CXCR4, shows a positive, Tier 1, association for 383 CHD in our analysis. CXCR4, which encodes C-X-C chemokine receptor type 4, is the target of 384 multiple marketed medications, including Plerixafor and Ulocuplumab, and our findings suggest 385 that these drugs could increase the risk of CHD in an on-target fashion. CXCR4 inhibitors are 386 currently used in stem cell transplant donors to elicit more stem cells into the peripheral blood and 387 this transient use is unlikely to have a lasting harmful impact on CHD risk. However, they have 388 been considered for more long-term use for other conditions⁶³⁻⁶⁵, and our findings recommend 389 evaluation of cardiovascular safety in such settings. Other drug development programs can

390 similarly use our findings to better predict whether a drug for any disease could have a harmful391 on-target effect on CHD.

392

393 The major limitation of our analytical approach is that within 116 loci there were multiple targets 394 that passed MR and colocalization thresholds. Tier 2 results that lie within the same genomic 395 region, like the SORT1, CELSR2 and PSRC1 locus, provide evidence that even cis-instruments 396 may exhibit horizontal pleiotropy, perhaps through shared regulation of multiple local genes or 397 linkage disequilibrium, making it difficult to determine the true causal target(s) in a region. 398 Although the approach taken here equally prioritizes these genes, functional work has suggested 399 that *SORT1* is likely driving the signal in the 1p13.3 locus³⁰. Conversely, though, due to our 400 comprehensive testing of all protein-coding genes in GTEx, Tier 1 results that are the only result 401 in the region are more likely to be the causal gene driving the association.

402

There are other limitations as well. This study was undertaken in almost exclusively European ancestry, so other studies in diverse ancestries are warranted. The pQTL MR only covers a small proportion of the human proteome and instrumenting plasma protein levels may not always provide the most relevant tissue in the context of CHD. The eQTL MR analysis used instruments for gene expression which may not translate to effects on protein levels in a relevant cell type. Lastly it can be difficult through MR, especially with gene expression data, to determine how much inhibition or activation of a certain protein is required for a meaningful reduction in CHD.

410

411 Our study has several strengths. We conducted a *de novo* GWAS meta-analysis involving nearly
412 300,000 CHD cases, providing a well-powered outcome dataset. We performed conditional

413 colocalization analysis to robustly test for pleiotropy, avoiding the one-signal-per-locus limitation 414 of the traditional colocalization approaches. Annotations of results with up-to-date information on 415 the druggable genome help facilitate translation into actionable interventions. Further 416 investigations with phenome-wide scans with colocalization on instrumental variants can identify 417 possible on-target safety concerns as well as elucidate potential mechanisms for CHD 418 pathogenesis.

419

In summary, our transcriptomics and proteomics analysis identified 69 novel, Tier 1 targets with putative causal relevance for CHD. We show opportunities for drug repurposing and development prioritization. Additionally, our results can be used for understanding CHD safety of current and future drugs indicated for non-cardiovascular diseases.

424

425

426 **References**

- Global burden of 369 diseases and injuries in 204 countries and territories, 1990-2019: a
 systematic analysis for the Global Burden of Disease Study 2019. *Lancet* 396, 1204-1222
 (2020).
- 430 2. Reith, C. & Armitage, J. Management of residual risk after statin therapy. *Atherosclerosis*431 245, 161-70 (2016).
- 432 3. Noubiap, J.J. *et al.* Global prevalence of resistant hypertension: a meta-analysis of data
 433 from 3.2 million patients. *Heart* 105, 98-105 (2019).
- 434 4. Vernon, S.T. *et al.* Increasing proportion of ST elevation myocardial infarction patients
 435 with coronary atherosclerosis poorly explained by standard modifiable risk factors. *Eur J*436 *Prev Cardiol* 24, 1824-1830 (2017).
- 437 5. Harrison, R.K. Phase II and phase III failures: 2013-2015. *Nat Rev Drug Discov* 15, 817438 818 (2016).
- 439 6. Burgess, S., Butterworth, A. & Thompson, S.G. Mendelian randomization analysis with 440 multiple genetic variants using summarized data. *Genet Epidemiol* **37**, 658-65 (2013).
- 441 7. Minikel, E.V., Painter, J.L., Dong, C.C. & Nelson, M.R. Refining the impact of genetic
 442 evidence on clinical success. *Nature* (2024).
- 8. Pietzner, M. *et al.* Mapping the proteo-genomic convergence of human diseases. *Science*374, eabj1541 (2021).
- 445 9. Sun, B.B. *et al.* Genomic atlas of the human plasma proteome. *Nature* **558**, 73-79 (2018).
- 44610.Ferkingstad, E. *et al.* Large-scale integration of the plasma proteome with genetics and
disease. *Nat Genet* 53, 1712-1721 (2021).
- 448 11. Folkersen, L. *et al.* Genomic and drug target evaluation of 90 cardiovascular proteins in
 30,931 individuals. *Nat Metab* 2, 1135-1148 (2020).
- 450 12. The Genotype-Tissue Expression (GTEx) project. *Nat Genet* **45**, 580-5 (2013).
- 451 13. Gaziano, J.M. *et al.* Million Veteran Program: A mega-biobank to study genetic influences
 452 on health and disease. *J Clin Epidemiol* **70**, 214-23 (2016).
- 453 14. Aragam, K.G. *et al.* Discovery and systematic characterization of risk variants and genes
 454 for coronary artery disease in over a million participants. *Nat Genet* 54, 1803-1815 (2022).
- Mars, N. *et al.* Polygenic and clinical risk scores and their impact on age at onset and prediction of cardiometabolic diseases and common cancers. *Nat Med* 26, 549-557 (2020).
- 457 16. Gaziano, L. *et al.* Actionable druggable genome-wide Mendelian randomization identifies
 458 repurposing opportunities for COVID-19. *Nature Medicine* (2021).
- 459 17. Pietzner, M. *et al.* Synergistic insights into human health from aptamer- and antibody460 based proteomic profiling. *Nat Commun* 12, 6822 (2021).
- 461 18. Tcheandjieu, C. *et al.* Large-scale genome-wide association study of coronary artery
 462 disease in genetically diverse populations. *Nat Med* 28, 1679-1692 (2022).
- 463 19. Zheng, J. *et al.* Phenome-wide Mendelian randomization mapping the influence of the plasma proteome on complex diseases. *Nat Genet* 52, 1122-1131 (2020).
- 465 20. Ochoa, D. *et al.* Open Targets Platform: supporting systematic drug–target identification
 466 and prioritisation. *Nucleic Acids Research* 49, D1302-D1310 (2020).
- 467 21. Wishart, D.S. *et al.* DrugBank 5.0: a major update to the DrugBank database for 2018.
 468 *Nucleic Acids Res* 46, D1074-d1082 (2018).
- 469 22. Nikpay, M. *et al.* A comprehensive 1,000 Genomes-based genome-wide association meta470 analysis of coronary artery disease. *Nat Genet* 47, 1121-1130 (2015).

- 471 23. Suhre, K. *et al.* Connecting genetic risk to disease end points through the human blood
 472 plasma proteome. *Nat Commun* 8, 14357 (2017).
- 473 24. Folkersen, L. *et al.* Mapping of 79 loci for 83 plasma protein biomarkers in cardiovascular
 474 disease. *PLoS Genet* 13, e1006706 (2017).
- 475 25. Yao, C. *et al.* Genome-wide mapping of plasma protein QTLs identifies putatively causal genes and pathways for cardiovascular disease. *Nat Commun* 9, 3268 (2018).
- 477 26. Emilsson, V. *et al.* Co-regulatory networks of human serum proteins link genetics to disease. *Science* 361, 769-773 (2018).
- 479 27. Richardson, T.G., Hemani, G., Gaunt, T.R., Relton, C.L. & Davey Smith, G. A
 480 transcriptome-wide Mendelian randomization study to uncover tissue-dependent
 481 regulatory mechanisms across the human phenome. *Nat Commun* 11, 185 (2020).
- 482 28. van der Wijst, M. et al. The single-cell eQTLGen consortium. Elife 9(2020).
- 483 29. Spyropoulos, A.C., Hayth, K.A. & Jenkins, P. Anticoagulation with anisindione in a patient
 484 with a warfarin-induced skin eruption. *Pharmacotherapy* 23, 533-6 (2003).
- 485 30. Musunuru, K. *et al.* From noncoding variant to phenotype via SORT1 at the 1p13
 486 cholesterol locus. *Nature* 466, 714-9 (2010).
- 487 31. Sabatine, M.S. *et al.* Evolocumab and Clinical Outcomes in Patients with Cardiovascular
 488 Disease. *N Engl J Med* 376, 1713-1722 (2017).
- 489 32. Cannon, C.P. *et al.* Ezetimibe Added to Statin Therapy after Acute Coronary Syndromes.
 490 *N Engl J Med* 372, 2387-97 (2015).
- 491 33. Bowman, L. *et al.* Effects of Anacetrapib in Patients with Atherosclerotic Vascular
 492 Disease. *N Engl J Med* 377, 1217-1227 (2017).
- 493 34. Ridker, P.M. From RESCUE to ZEUS: will interleukin-6 inhibition with ziltivekimab 494 prove effective for cardiovascular event reduction? *Cardiovasc Res* **117**, e138-e140 (2021).
- 495 35. De Clercq, E. Mozobil®(Plerixafor, AMD3100), 10 years after its approval by the US Food
 496 and Drug Administration. *Antiviral Chemistry and Chemotherapy* 27, 2040206619829382
 497 (2019).
- 49836.Nicholls, S.J. *et al.* Varespladib and cardiovascular events in patients with an acute499coronary syndrome: the VISTA-16 randomized clinical trial. *Jama* **311**, 252-62 (2014).
- S00 37. Carvalho-Silva, D. *et al.* Open Targets Platform: new developments and updates two years
 S01 on. *Nucleic Acids Res* 47, D1056-d1065 (2019).
- 50238.Chen, E.Y. *et al.* Enrichr: interactive and collaborative HTML5 gene list enrichment503analysis tool. *BMC Bioinformatics* 14, 128 (2013).
- 50439.Kuleshov, M.V. *et al.* Enrichr: a comprehensive gene set enrichment analysis web server5052016 update. *Nucleic Acids Res* 44, W90-7 (2016).
- 506 40. Xie, Z. *et al.* Gene Set Knowledge Discovery with Enrichr. *Curr Protoc* **1**, e90 (2021).
- 507 41. Engelen, S.E., Robinson, A.J.B., Zurke, Y.X. & Monaco, C. Therapeutic strategies
 508 targeting inflammation and immunity in atherosclerosis: how to proceed? *Nat Rev Cardiol*509 19, 522-542 (2022).
- 510 42. Cuffe, M.S. *et al.* Short-term intravenous milrinone for acute exacerbation of chronic heart
 511 failure: a randomized controlled trial. *Jama* 287, 1541-7 (2002).
- 43. Packer, M. *et al.* Effect of oral milrinone on mortality in severe chronic heart failure. The
 PROMISE Study Research Group. *N Engl J Med* 325, 1468-75 (1991).
- Massie, B. *et al.* Long-term oral administration of amrinone for congestive heart failure:
 lack of efficacy in a multicenter controlled trial. *Circulation* **71**, 963-71 (1985).

- 516 45. Zhang, W. & Colman, R.W. Thrombin regulates intracellular cyclic AMP concentration in 517 human platelets through phosphorylation/activation of phosphodiesterase 3A. *Blood* 110, 518 1475-82 (2007).
- 519 46. Thompson, P.D., Zimet, R., Forbes, W.P. & Zhang, P. Meta-analysis of results from eight
 520 randomized, placebo-controlled trials on the effect of cilostazol on patients with
 521 intermittent claudication. *Am J Cardiol* **90**, 1314-9 (2002).
- 47. Rizzo, M., Corrado, E., Patti, A.M., Rini, G.B. & Mikhailidis, D.P. Cilostazol and
 atherogenic dyslipidemia: a clinically relevant effect? *Expert Opin Pharmacother* 12, 64755 (2011).
- 525 48. Bedenis, R. *et al.* Cilostazol for intermittent claudication. *Cochrane Database Syst Rev*526 2014, Cd003748 (2014).
- 527 49. Sudo, T. *et al.* Potent effects of novel anti-platelet aggregatory cilostamide analogues on recombinant cyclic nucleotide phosphodiesterase isozyme activity. *Biochem Pharmacol* 529 59, 347-56 (2000).
- 530 50. Klarin, D. *et al.* Genetics of blood lipids among ~300,000 multi-ethnic participants of the
 531 Million Veteran Program. *Nat Genet* 50, 1514-1523 (2018).
- 532 51. McHutchison, C. *et al.* Cilostazol for Secondary Prevention of Stroke and Cognitive
 533 Decline: Systematic Review and Meta-Analysis. *Stroke* 51, 2374-2385 (2020).
- 534 52. Patrono, C. & Baigent, C. Role of aspirin in primary prevention of cardiovascular disease.
 535 *Nat Rev Cardiol* 16, 675-686 (2019).
- 536 53. Holmes, M.V. *et al.* Novel genetic approach to investigate the role of plasma secretory
 phospholipase A2 (sPLA2)-V isoenzyme in coronary heart disease: modified Mendelian
 randomization analysis using PLA2G5 expression levels. *Circ Cardiovasc Genet* 7, 144539 50 (2014).
- 540 54. Holmes, M.V. *et al.* Secretory phospholipase A(2)-IIA and cardiovascular disease: a mendelian randomization study. *J Am Coll Cardiol* **62**, 1966-1976 (2013).
- 542 55. Guardiola, M. *et al.* PLA2G10 Gene Variants, sPLA2 Activity, and Coronary Heart 543 Disease Risk. *Circ Cardiovasc Genet* **8**, 356-62 (2015).
- 54456.Ait-Oufella, H. et al. Group X secreted phospholipase A2 limits the development of545atherosclerosis in LDL receptor-null mice. Arterioscler Thromb Vasc Biol 33, 466-73546(2013).
- 547 57. Rosenson, R.S. *et al.* Effects of 1-H-indole-3-glyoxamide (A-002) on concentration of
 548 secretory phospholipase A2 (PLASMA study): a phase II double-blind, randomised,
 549 placebo-controlled trial. *Lancet* 373, 649-58 (2009).
- 550 58. Trigueros-Motos, L. *et al.* ABCA8 Regulates Cholesterol Efflux and High-Density
 551 Lipoprotein Cholesterol Levels. *Arterioscler Thromb Vasc Biol* 37, 2147-2155 (2017).
- 552 59. Wang, C.Y. *et al.* A Novel Nonsense Mutation of ABCA8 in a Han-Chinese Family With
 553 ASCVD Leads to the Reduction of HDL-c Levels. *Front Genet* 11, 755 (2020).
- 55460.Motazacker, M.M. *et al.* Evidence of a polygenic origin of extreme high-density555lipoprotein cholesterol levels. Arterioscler Thromb Vasc Biol 33, 1521-8 (2013).
- 556 61. Zeng, T. & Li, Y.I. Predicting RNA splicing from DNA sequence using Pangolin. *Genome* 557 *Biol* 23, 103 (2022).
- 558 62. Döring, Y. *et al.* Vascular CXCR4 Limits Atherosclerosis by Maintaining Arterial
 559 Integrity. *Circulation* 136, 388-403 (2017).

- 63. Ghobrial, I.M. *et al.* A Phase Ib/II Trial of the First-in-Class Anti-CXCR4 Antibody
 561 Ulocuplumab in Combination with Lenalidomide or Bortezomib Plus Dexamethasone in
 562 Relapsed Multiple Myeloma. *Clin Cancer Res* 26, 344-353 (2020).
- 563 64. McDermott, D.H. *et al.* Plerixafor for the Treatment of WHIM Syndrome. *N Engl J Med*564 380, 163-170 (2019).
- 565 65. McDermott, D.H. *et al.* A phase III randomized crossover trial of plerixafor versus G-CSF
 566 for treatment of WHIM syndrome. *J Clin Invest* 133(2023).
- 567 66. Hemani, G. *et al.* The MR-Base platform supports systematic causal inference across the human phenome. *Elife* **7**(2018).
- 569 67. Pruim, R.J. *et al.* LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics* 26, 2336-7 (2010).
- 68. O'Connor, L., Brage, S., Griffin, S.J., Wareham, N.J. & Forouhi, N.G. The cross-sectional association between snacking behaviour and measures of adiposity: the Fenland Study, UK. *British journal of nutrition* 114, 1286-1293 (2015).
- 574 69. Willer, C.J., Li, Y. & Abecasis, G.R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190-1 (2010).
- 576 70. Yang, J. *et al.* Conditional and joint multiple-SNP analysis of GWAS summary statistics
 577 identifies additional variants influencing complex traits. *Nat Genet* 44, 369-75, s1-3 (2012).
- 578 71. Auton, A. *et al.* A global reference for human genetic variation. *Nature* **526**, 68-74 (2015).
- 579 72. Giambartolomei, C. *et al.* Bayesian test for colocalisation between pairs of genetic
 580 association studies using summary statistics. *PLoS Genet* 10, e1004383 (2014).
- 73. Robinson, J.W. *et al.* An efficient and robust tool for colocalisation: Pair-wise Conditional and Colocalisation (PWCoCo). *bioRxiv*, 2022.08.08.503158 (2022).
- 583
- 584
- 585

586 Author contribution

- 587 L.G., J.P.C., A.C.P., and A.S.B. conceptualized the study. E.W. M.P. and C.L. generated pQTL
- 588 information. S.C.R., T.J., K.A., T.J.A., P.W.F.W., K.C. assisted in preparation of CHD summary
- 589 statistics. L.G. performed the Mendelian randomization analysis. C.G. performed conditional
- 590 colocalization analysis on CHD. E.A. performed phenome-wide scans of instrumental variants
- 591 with conditional colocalization. D.S. performed pathway enrichment analysis. L.G., J.H.Z., and
- 592 E.A. generated figures. L.G., E.A., A.C.P., and A.S.B. wrote the manuscript. A.S.B. oversaw all
- analyses. All authors contributed to study design and provided editorial comments.
- 594

595 **Competing interests**

J.P.C. moved to work with Novartis Institute for Biomedical Research during the completion of this project. E.W. is now an employee of AstraZeneca. A.S.B. reports institutional grants from AstraZeneca, Bayer, Biogen, BioMarin, Bioverativ, Novartis, Regeneron and Sanofi. J. Danesh serves on scientific advisory boards for AstraZeneca, Novartis and the UK Biobank and has received multiple grants from academic, charitable and industry sources outside of the submitted work.

602 Figure Legends

603

Figure 1. Overall study summary. We used *cis*-pQTLs to instrument 1,497 plasma proteins and *cis*-eQTLs to instrument 15,412 protein-coding genes to perform MR on a GWAS meta-analysis of CHD totaling 296,537 cases. For significant ($P < 3.22 \times 10^{-6}$) MR results, we additionally performed colocalization using a pairwise conditional and colocalization (PWCoCo)¹⁹ approach between pQTL or eQTL summary statistics and CHD. Lastly, we annotated results (those that

- passed both Bonferroni-adjusted P-value for MR and colocalization [PP.H4>0.80] thresholds)
 according to whether an existing drug according status using the OpenTargets³⁷ or DrugBank²¹.
- 611

612 **Figure 2. Circos plot from Mendelian randomization and colocalization analysis.** Targets

- 613 that are labeled are Tier 1 and novel results that passed MR ($P < 3.22 \times 10^{-6}$) and colocalization
- 614 thresholds (PPH4>0.8). Tier 1 results are defined as those that are not in close proximity
- 615 (+250kb) or correlated (r^2 >0.2 in 1000 Genomes EUR) with any other result that passed MR and
- 616 colocalization thresholds. Results were defined as novel if they (1) were based on instrumental
- 617 variants that were not correlated ($r^2 < 0.2$ in 1000 Genomes EUR) or in close proximity (± 250 kb)
- 618 with genome-wide significant ($P < 5 \times 10^{-8}$) variants for CHD reported in Aragam *et al.*¹⁴ or
- 619 Tcheandjieu *et al.*¹⁸ or Supplementary Table 2 of Aragam *et al.*¹⁴ which includes CHD-
- 620 associated variants previously reported in the literature and (2) were not previously reported in 621 Zheng *et al.*¹⁹ or Richardson *et al.*²⁷ Targets with existing drug according to either OpenTargets³⁷
- 622 or DrugBank²¹ are labeled in red, targets without a current drug are labeled in black.
- 623

624 Figure 3. Phenome-wide scan and colocalization results for PDE3A. Results are shown for 625 scans in MR Base⁶⁶ of rs7488772, the instrumental variant for PDE3A in adipose subcutaneous 626 tissue. (a) results for all genetic associations with $P < 1 \times 10^{-5}$ (aligned by CHD-increasing allele: 627 positive betas are shown in blue and negative betas in red; if the same trait is estimated in more 628 than one dataset, this plot shows the one with the smallest p-value) and conditional 629 colocalization PPH4 20.8 (if a trait colocalizes with more than one variant after conditional analysis, this plot shows the highest PPH4), and (b) LocusZoom⁶⁷ plots for selected 630 631 cardiometabolic traits. Results from colocalization analyses using marginal summary statistics 632 are labeled as 'unconditioned'. Results from conditional colocalization show the conditioning

633

variant in the plot title.

634

Figure 4. Results for PLA2G5 and PLA2G2A. (a) Mendelian randomization results for 635 636 PLA2G5 and PLA2G2A in all tissues where an instrument exists. Only results for PLA2G5 637 in aortic artery tissue (highlighted in blue) passed the P-value significance threshold and was tested for colocalization. All results are from eQTL except for PLA2F2A / blood 638 639 plasma which are derived from pQTL. Regional association statistics for (b) gene 640 expression of *PLA2G5* in a ortic artery tissue from GTEx and (c) coronary heart disease. (a) and (b) show the pairwise correlation (1000G European ancestry) for each variant with 641 642 rs12408798, the instrumental variant used in the present study. (a) and (b) colocalize at 643 PPH4=0.98. (d) Local association plot of the *PLA2G5* region for coronary heart disease, showing the correlation for rs525380, the instrumental variant for PLA2G5 expression used 644 in the previous study Holmes *et al.*⁵³ The association for rs525380 and CHD was P=0.20645 in Holmes *et al.*⁵³ and P = 0.009 in the present study. 646

647

648 Figure 5. Phenome-wide scan and colocalization results for ABCA8. Results are shown for

scans in MR Base⁶⁶ of rs34931250, the instrumental variant for *ABCA8*. (a) results for all genetic 649

associations with $P < 1 \times 10^{-5}$ (aligned by CHD-increasing allele; positive betas are shown in blue 650

651 and negative betas in red; if the same trait is estimated in more than one datasets, this plot shows the one with the smallest p-value) and conditional colocalization PPH4 ≥ 0.8 (if a trait colocalizes

- 652 653 with more than one variant after conditional analysis, this plot shows the highest PPH4), and (b)
- LocusZoom plots for selected cardiovascular traits. Results from colocalization analyses using
- 654
- marginal summary statistics are labeled as 'unconditioned'. 655
- 656
- 657

658 Methods

659 Selecting genetic instruments

660 First, we selected conditionally-independent cis-pQTLs for plasma proteins reported in Pietzner et al. (www.omicscience.org)⁸ Briefly, protein levels were measured by SomaLogic Inc (Boulder, 661 662 Colorado, US) with the SomaScan® V4 platform in 10,708 European-ancestry individuals from the Fenland cohort^{17,68}. This technique uses DNA-aptamers called SOMAmers, oligonucleotide 663 664 strands that fold and bind to proteins with high specificity, and association analyses were 665 performed for 5,210 SOMAmers. We removed pQTLs with minor allele frequency (MAF)<0.01, 666 trans-pQTLs and SOMAmers that measured more than one protein. 667 668 For eQTL instruments, we downloaded a file provided by the Genotype-Tissue Expression (GTEx) 812 669 Consortium version 670 (https://storage.googleapis.com/gtex analysis v8/single tissue qtl data/GTEx Analysis v8 eQ 671 TL_independent.tar) that included all conditionally independent cis-eQTLs for expression of all 672 genes in 49 tissues. We excluded eQTLs for non-protein-coding genes (i.e. those that do not have a UniProt ID), $P > 5 \times 10^{-8}$ and MAF<0.01. We also removed eOTLs, pOTLs or genes that lie 673 within the human leukocyte antigen (HLA) region (chr6:29691116-33054976), due to the 674 675 complicated LD structure in that region.

676

677 Generation of outcome summary statistics

To generate outcome summary statistics for CHD, we meta-analyzed results from three nonoverlapping sources: The Million Veteran Program (MVP)¹⁸, CARDIoGRAMplusC4D Consortium¹⁴, and the FinnGen biobank¹⁵. MVP is an ongoing biobank recruiting from 63

Veterans Affairs (VA) facilities across the United States¹³. In MVP, we utilized a GWAS that 681 682 consisted of European ancestry individuals. The CARDIoGRAMplusC4D Consortium is a GWAS meta-analysis on CHD from many contributing studies, primarily of European ancestry¹⁴. FinnGen 683 684 is a consortium of biobanks within Finland that includes participants linked to electronic health 685 data. For FinnGen, we downloaded summary statistics for the I9 ISCHHEART outcome 686 (gs://finngen-public-data-r3/summary_stats/finngen_r3_I9_ISCHHEART.gz) from release 3, 687 which defined CHD as ICD-10 I20-I25. We meta-analyzed summary statistics from the three data sources using METAL software⁶⁹ with fixed-effects and inverse-variance weighting. 688

689

690 Mendelian Randomization

We used the TwoSampleMR R package (<u>https://github.com/MRCIEU/TwoSampleMR</u>) to estimate MR effects between instruments and CHD. Inverse variance-weighted MR with fixedeffects was used for instruments with more than one variant, and the Wald ratio method for instruments with one variant. For instruments with more than one variant, we additionally assessed heterogeneity across the variant-level MR estimates and reported the Cochran Q *P* value.

696

697 Conditional Colocalization

Significant MR associations can arise from a strong association between an instrumental variant and an outcome that is not causal, but in LD with a peak variant (i.e. an MR estimate can be due to confounding by LD and not due to a shared causal variant). To address this, we performed colocalization between the pQTL or eQTL summary statistics and CHD for all results that passed the MR *P*-value threshold. Additionally, when performing colocalization, false negatives can arise from interference of a strong independent signal, because conventional colocalization methods

704 traditionally assume one causal variant within the tested region. To address this, we performed pairwise conditional colocalization¹⁹, which requires conditional summary statistics for each 705 region. Conditional analyses for pOTL data from Pietzner et al.⁸ were conducted using GCTA-706 707 COJO⁷⁰ and a reference panel of genotypes from the Fenland cohort. GCTA-COJO is a stepwise 708 model selection procedure to select independently associated SNPs using the LD structure to 709 account for the correlation between SNPs. For eQTLs, we generated summary statistics using raw 710 genotype data from European ancestry participants following the procedure outlined by the GTEx 711 Consortium. Briefly, after filtering the genotypes (MAF<0.01, HWE<0.000001, and no ambiguous 712 SNPs), GWAS between variants and gene expression was performed adjusting for the top 5 713 principal components, PEER factors, sequencing platform and sex. Conditional analysis for gene 714 expression were performed with raw genotype data by iteratively adjusting for the most strongly 715 associated variant in a region. Conditional analysis for CHD summary statistics were performed using GCTA-COJO with a 1000G⁷¹ European ancestry reference panel. 716

717

Pairwise colocalization was performed using the COLOC R package⁷² on regions ± 200 kb around the instrumental variant with at least 50 SNPs matching across the two datasets. The prior probability of each SNP that is causal to only one of the traits (PPH3) was set to 1×10^{-5} and causal to both traits (PPH4) was set to 1×10^{-6} . A posterior probability of a shared causal variant (PPH4) of $\geq 80\%$ for any of the conditional iterations was considered strong evidence of colocalization.

724

725 **Drugs against identified targets**

Targets that passed MR and colocalization thresholds were annotated with information on drugs
 against targets from by OpenTargets²⁰ and DrugBank²¹. We used DrugBank release 5.1.10,
 (<u>https://go.drugbank.com/releases</u>) and restricted to known pharmacologically active drug-target
 pairs.

730

731 Assessment of horizontal pleiotropy

732 Valid instrumental variants for MR must meet the exclusion restriction requirement, which 733 assumes that the variant only affects the outcome through the instrumented exposure. A variant is 734 an invalid instrument if it alters expression of other genes or proteins that affect the outcome 735 through distinct biological/causal pathways (horizontal pleiotropy). In the case that results for 736 multiple genes/proteins in a region colocalize with the same CHD signal, it can be difficult to 737 disentangle which is truly influencing CHD risk. Therefore, we considered a variant that influences 738 expression of nearby genes through shared regulation as possibly introducing horizontal pleiotropy 739 and grouped results into regions based on instrumental variants that are in close proximity 740 (+250 kb) or correlated $(r^2 > 0.2 \text{ in } 1000 \text{ Genomes EUR})$. Because we tested all protein-coding genes 741 in GTEx, results that are the only significant signal in a region are less likely to exhibit horizontal 742 pleiotropy. Therefore, we allocated results into tiers based on the number of results within a 743 genomic region, where Tier 1 results are the sole result within a region to pass MR and 744 colocalization thresholds, and Tier 2 results contain multiple signals within a region.

745

746 Phenome-wide scans of instrumental variants

For some results that passed both MR and colocalization thresholds, we performed Phenome-wide
association scan with colocalization (PheWAS) of instrumental variants. Associations between

749 instrumental variants with disease and traits can mimic on-target effects of therapeutic targeting 750 observed in trials. This can provide confidence that instrumental variants are correctly mimicking 751 on-target effects of a drug. Additionally, a PheWAS analysis can identify possible harmful effects 752 of protein modulation or provide mechanistic insights. Therefore, we queried instrumental variants 753 for associations with diseases and traits in the MRbase PheWAS platform⁶⁶, a database that allows 754 for PheWAS across a wide range of publicly available datasets. For PheWAS associations with P $<1\times10^{-5}$ we also performed pairwise conditional colocalization between eQTL and traits regional 755 756 summary stats, to help ensure those associations are not due to confounding by LD. For this, we 757 utilized the PwCoCo v. 1.0 (https://github.com/jwr-git/pwcoco/tree/pwcoco-v1.0) tool⁷³, using LD 758 information from the UK Biobank after restricting the analysis sample to 367,641 unrelated 759 European-ancestry participants. We set the prior probabilities that each SNP was causal to either of the traits (PPH3) and causal to both traits (PPH4) to 1×10^{-5} and 1×10^{-6} , respectively. We 760 761 considered a posterior probability of a shared causal variant (PPH4) of \geq 80% as strong evidence 762 LocusZoom⁶⁷ of colocalization. inspection with For 763 (http://genome.sph.umich.edu/wiki/LocusZoom_Standalone), we also exported conditioned 764 genetic associations using the "--out-cond" option in PwCoCo.

765

766 Pathway and gene ontology term enrichment analysis

Pathway and functional term enrichment analyses were performed with Enrichr (https://maayanlab.cloud/Enrichr/)³⁸⁻⁴⁰ by inputting the HGNC symbols for all putative causal targets as a single gene set. Briefly, Enrichr uses the Fisher's exact test to assess whether the putative causal gene set was enriched with genes assigned to a given GO term or KEGG pathway relative to a permuted background gene set. An odds ratio significantly greater than one is

- indicative of an enriched term or pathway. Correction for multiple testing (P_{adj}) was conducted
- vising the Benjamini-Hochberg procedure.

774

775

776

777 Acknowledgements

778 Scott C. Ritchie was funded by a British Heart Foundation (BHF) Programme Grant 779 (RG/18/13/33946) and by the National Institute for Health and Care Research (NIHR) Cambridge 780 BRC (BRC-1215-20014; NIHR203312) [*] * The views expressed are those of the authors and 781 not necessarily those of the NIHR or the Department of Health and Social Care. Emanuele Di 782 Angelantonio holds an NIHR Senior Investigator Award. Stephen Burgess is supported by the 783 Wellcome Trust (225790/Z/22/Z) and the United Kingdom Research and Innovation Medical 784 Research Council (MC_UU_00002/7). John Danesh holds a British Heart Foundation 785 Professorship and an NIHR Senior Investigator Award. Adam Butterworth and Claudia 786 Langenberg were supported by BHF-DZHK grant funding "Prot4CVD: Translational proteomics 787 for cardiovascular diseases: from population prediction to clinical and therapeutic applications" 788 (FKZ 81X2100281). The BHF Cardiovascular Epidemiology Unit was supported by core funding 789 from the NIHR Blood and Transplant Research Unit in Donor Health and Genomics (NIHR 790 BTRU-2014-10024), the NIHR BTRU in Donor Health and Behaviour (NIHR203337), the UK 791 Medical Research Council (MR/L003120/1), the British Heart Foundation (SP/09/002; 792 RG/13/13/30194; RG/18/13/33946) and NIHR Cambridge BRC (BRC-1215-20014). This work 793 was also supported by Health Data Research UK, which is funded by the UK Medical Research 794 Council, the Engineering and Physical Sciences Research Council, the Economic and Social 795 Research Council, the Department of Health and Social Care (England), the Chief Scientist Office 796 of the Scottish Government Health and Social Care Directorates, the Health and Social Care 797 Research and Development Division (Welsh Government), the Public Health Agency (Northern 798 Ireland), the British Heart Foundation and Wellcome. Some analyses were performed using 799 resources provided by the Cambridge Service for Data Driven Discovery (CSD3) operated by the

- 800 University of Cambridge Research Computing Service (https://www.csd3.cam.ac.uk), provided
- 801 by Dell EMC and Intel using tier 2 funding from the Engineering and Physical Sciences Research
- 802 Council (capital grant EP/P020259/1), and DiRAC funding from the Science and Technology
- 803 Facilities Council (https://www.dirac.ac.uk).

804

805

806



PDE3A expression in adipose subcutaneous: unconditioned

Systolic blood pressure: 12:20586395

b

Glycated haemoglobin: unconditioned

20.6

20.7

20.8

20.5

20.4

Figure 3

а

Figure 4

а

Trait		Beta	P-value	H3	H4
Kidney injury molecule 1		0.16	2.02e-11	0.02	0.98
Statin medication		0.11	6.32e-09	0.00	1.00
Phospholipids to total lipids ratio in large VLDL		0.06	4.90e-11	0.00	1.00
Triglycerides to total lipids ratio in medium HDL		0.06	1.10e-11	0.00	1.00
Total lipids in very small VLDL		0.05	3.90e-08	0.00	0.91
Cholesteryl esters in very large VLDL		0.05	2.20e-08	0.00	1.00
Cholesteryl esters in large VLDL	_ 	0.05	2.00e-08	0.00	1.00
Cholesterol in very large VLDL	(0.04	3.40e-08	0.00	1.00
Triglycerides to total lipids ratio in large HDL		0.04	4.00e-08	0.00	1.00
Phospholipids in VLDL	_ 	0.04	2.50e-07	0.00	1.00
Free cholesterol in very large VLDL	_ 	0.04	1.00e-07	0.00	1.00
Phospholipids to total lipids ratio in IDL	_ 	0.04	4.90e-07	0.00	1.00
Cholesterol in very small VLDL	- -	0.04	1.50e-06	0.00	1.00
Phospholipids in very large VLDL	(0.04	3.70e-07	0.00	1.00
Free cholesterol in large VLDL		0.04	8.20e-07	0.00	1.00
Concentration of small VLDL particles		0.04	1.50e-06	0.01	0.99
Phospholipids in large VLDL		0.04	1.10e-06	0.00	1.00
Phospholipids to total lipids ratio in very small VLDL		0.04	1.60e-06	0.00	1.00
Total lipids in VLDL	_ 	0.04	2.30e-06	0.00	0.99
Cholesterol in chylomicrons and extremely large VLDL	- - -	0.04	2.30e-06	0.00	0.99
Free cholesterol in chylomicrons and extremely large VLDL	- - -	0.04	3.30e-06	0.00	0.99
Cholesteryl esters in chylomicrons and extremely large VLDL		0.04	3.90e-06	0.00	0.99
Free cholesterol to total lipids ratio in very large HDL		0.04	5.50e-06	0.00	1.00
Triglycerides in very small VLDL		0.04	5.00e-06	0.00	0.97
Low density lipoprotein cholesterol	-	0.04	5.10e-19	0.01	0.99
Concentration of very large VLDL particles	- -	0.04	5.30e-06	0.00	0.98
Phospholipids in chylomicrons and extremely large VLDL		0.04	7.30e-06	0.00	0.98
Concentration of chylomicrons and extremely large VLDL particles	_ I	0.04	9.60e-06	0.00	0.97
Total lipids in very large VLDL		0.04	7.90e-06	0.00	0.98
Apolipoprotein B	-	0.02	2.30e-08	0.04	0.94
Triglycerides		0.02	7.30e-08	0.00	1.00
Forced expiratory volume in 1-second	-	0.02	9.30e-06	0.00	0.98
Free cholesterol in large HDL	-	-0.04	3.70e-06	0.00	1.00
Cholesteryl esters to total lipids ratio in large HDL	-	-0.04	4.50e-06	0.00	0.99
Bioavailable testosterone	-	-0.04	3.10e-15	0.00	1.00
Concentration of large HDL particles	-	-0.04	1.00e-07	0.00	1.00
Phospholipids in large HDL	-	-0.04	3.50e-08	0.00	1.00
Cholesteryl esters to total lipids ratio in very large HDL	-	-0.04	4.70e-07	0.00	1.00
Free cholesterol in HDL	-	-0.04	3.90e-08	0.00	1.00
High density lipoprotein cholesterol	-	-0.05	1.90e-29	0.00	1.00
Apolipoprotein A–I	-	-0.05	5.20e-30	0.00	1.00
Free cholesterol in medium HDL	-	-0.05	1.30e-09	0.00	1.00
Apolipoprotein A	-	-0.05	3.34e-25	0.00	1.00
Concentration of medium HDL particles	-	-0.05	7.80e-10	0.00	1.00
Concentration of HDL particles	-	-0.05	1.30e-09	0.00	1.00
Total testosterone	-	-0.05	3.20e-21	0.00	1.00
Cholesteryl esters in medium HDL	-	-0.06	5.70e-13	0.00	1.00
Cholesteryl esters to total lipids ratio in medium HDL	-	-0.06	7.50e-13	0.00	1.00
-0.1	l 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.				

Beta and 95% confidence intervals per CHD-increasing allele

b

ABCA8 expression in colon sigmoid: unconditioned

10

Cholesteryl esters in very large VLDL: unconditioned

Triglycerides: unconditioned

Concentration of HDL particles: unconditioned

← ABCA8 ← ABCA9 ← ABCA6

ABCA9–AS1→

MR4524B→

67.1

+- MIR4524A

LINC01482-

66.7

66.8

66.9 Position on chr17 (Mb)

Cholesteryl esters in medium HDL: unconditioned

Figure 5