

1 **Contribution of leukocyte telomere length to major cardiovascular diseases onset:**
2 **phenotypic and genetic insights from a large-scale genome-wide cross-trait**
3 **analysis**

4
5 Jun Qiao^{1#}, Qian Wang^{2#}, Yuhui Zhao^{2#}, Minjing Chang¹, Liuyang Cai¹, Feng Liu³, Kaixin Yao²,
6 Leilei Zheng², Ning Tan⁴, Pengcheng He⁴, Anil G. Jegga^{5,6,7}, Siim Pauklin^{7*}, Lei Jiang^{4*}, Yining
7 Yang^{8,9*}, Yuliang Feng^{1*}

8 ¹ Department of Pharmacology, School of Medicine, Southern University of Science and
9 Technology, Shenzhen, China.

10 ² Shanxi Key Laboratory of Big Data for Clinical Decision, Shanxi Medical University, Taiyuan,
11 China.

12 ³ Botnar Research Centre, Nuffield Department of Orthopaedics, Rheumatology and
13 Musculoskeletal Sciences, University of Oxford, Headington, Oxford, UK.

14 ⁴ Department of Cardiology, Guangdong Cardiovascular Institute, Guangdong Provincial
15 People's Hospital (Guangdong Academy of Medical Sciences), Southern Medical University,
16 Guangzhou, China.

17 ⁵ Department of Pediatrics, University of Cincinnati College of Medicine, Cincinnati, USA.

18 ⁶ Department of Computer Science, University of Cincinnati College of Engineering,
19 Cincinnati, USA.

20 ⁷ Department of Internal Medicine, College of Medicine, University of Cincinnati, Cincinnati,
21 USA.

22 ⁸ Department of Cardiology, People's Hospital of Xinjiang Uygur Autonomous Region, Urumqi,
23 China.

24 ⁹ Xinjiang Key Laboratory of Cardiovascular Homeostasis and Regeneration Research, Urumqi,
25 China.

26 [#]These authors contributed equally to this work.

27

28 *Correspondence to:

29 Yuliang Feng MD, PhD

30 Associate Professor

31 Department of Pharmacology, School of Medicine; Southern University of Science and
32 Technology, Shenzhen, Guangdong, 518055, China.

33 Tel:+86 (755)-88012564

34 Email: fengyl@sustech.edu.cn

35

36 Yining Yang MD, PhD

37 Professor

38 Department of Cardiology, People's Hospital of Xinjiang Uygur Autonomous Region, Urumqi, ,

39 Xinjiang Uygur Autonomous Region, 830000, China.

40 Xinjiang Key Laboratory of Cardiovascular Homeostasis and Regeneration Research, Urumqi,

41 Xinjiang Uygur Autonomous Region, 830001, China.

42 Tel:+86 (0991)-960200

43 Email: yangyn5126@xjrmyy.com

44

45 Lei Jiang MD, PhD

46 Associate Professor

47 Department of Cardiology; Guangdong Cardiovascular Institute, Guangdong Provincial

48 People's Hospital (Guangdong Academy of Medical Sciences); Southern Medical University,

49 Guangzhou, Guangdong, 510080, China.

50 Tel:+86 (20)-83837812

51 Email: jianglei0731@gmail.com

52

53 Siim Pauklin PhD

54 Group leader and CRUK Career Development Fellow

55 Botnar Research Centre, Nuffield Department of Orthopaedics, Rheumatology and

56 Musculoskeletal Sciences, University of Oxford, Headington, Oxford OX3 7LD, UK.

57 Tel: +44 (0)-1865226492

58 Email: siim.pauklin@ndorms.ox.ac.uk

59

60 **Abstract**

61 Telomere shortening, a marker of cellular aging and genomic instability, has been
62 epidemiologically linked to an increased risk of various cardiovascular diseases (CVDs).
63 However, shared genetic determinants involved in these associations remain unclear. We
64 composed an atlas of the shared genetic associations between leukocyte telomere length (LTL)
65 and six major CVDs by investigating shared genetic elements, encompassing SNPs, genes,
66 biological pathways, and protein targets with pleiotropic implications. Extensive genetic
67 overlaps beyond genetic correlations were observed, but no causal relationships were
68 established. We identified 248 independent pleiotropic genomic risk loci, implicating 50 unique
69 genes in two or more trait pairs, especially the *SH2B3* gene, which was further validated by a
70 proteome-wide Mendelian Randomization study. Functional analysis demonstrated a link to
71 both DNA biosynthetic processes and telomere maintenance mechanisms. These findings
72 suggest a genetic link between LTL and CVDs, highlighting a shared genetic basis crucial for
73 developing future interventions and therapeutic targets.

74 **Introduction**

75 Telomeres, DNA-protein complexes located at the ends of linear eukaryotic chromosomes, play
76 a critical role in genome protection and act as indicators of biological aging.¹ With each cell
77 division, telomere progressively shortens due to the increased rate of somatic cell turnover and
78 aging, eventually reaching a critical threshold known as the Hayflick limit. Beyond this point,
79 DNA damage and cellular senescence begin, marking the onset of genomic instability. This
80 process of telomere attrition is particularly relevant in cardiovascular diseases (CVDs), where
81 the accumulation of senescent cells leads to tissue inflammation and matrix degradation. These
82 changes contribute to the thinning of the fibrous cap, thereby heightening the risk of CVDs.
83 Previous meta-analyses have identified an inverse association between leukocyte telomere
84 length (LTL) and the risk of most CVDs, particularly coronary artery disease (CAD) and heart
85 failure (HF), independent of conventional vascular risk factors, although the relationship with
86 other CVDs remains more ambiguous²⁻⁴. This body of evidence highlighted the potential
87 impact of LTL on the development and progression of various CVDs. Moreover, inflammation
88 and oxidative stress central to the pathogenesis of CVDs further accelerate telomere shortening
89 and cellular senescence. This complex interplay highlights the significant connection between
90 LTL and CVDs, underscoring the need for further research into their joint mechanisms.

91
92 Previous epidemiological studies have demonstrated that telomere shortening was associated
93 with an elevated risk of various CVDs. One plausible explanation is the overlap of genetic
94 determinants for LTL and CVDs. LTL varies significantly among individuals from birth and
95 throughout their lifespan, exhibiting high heritability with estimates ranging from 44% to
96 86%^{5,6}. The largest genome-wide association study (GWAS) to date on LTL, using UK Biobank
97 data, identified 138 associated loci, revealing links between genetically determined LTL⁷ and
98 multiple CVD phenotypes, such as CAD. The recent accessibility of GWAS data for a diverse
99 spectrum of CVD phenotypes offers a valuable opportunity to investigate the genetic basis
100 between LTL and CVDs⁸⁻¹³. The shared genetic foundations may be understood as genetic
101 variants influencing multiple complex phenotypic traits through both vertical and horizontal
102 pleiotropy. Briefly, vertical pleiotropy emerges when a genetic variation affects one phenotype,
103 which in turn influences the occurrence of another phenotype—a concept primarily addressed

104 by Mendelian Randomization (MR) studies. Several previous MR studies have suggested that
105 the associations of shorter LTL with CAD and Stroke were genetically causal, although
106 evidence for an association with AF and HF was less certain, presenting confused results¹⁴⁻¹⁷.
107 The causality of the relationships between LTL and other CVD phenotypes has not been
108 adequately demonstrated. Conversely, horizontal pleiotropy occurs when a single genetic
109 variation simultaneously impacts multiple phenotypes, which may highlight potential shared
110 biological pathways among complex traits. Recent advances in genomics statistical tools have
111 unveiled the vital role of horizontal pleiotropy beyond vertical pleiotropy in shared genetic
112 foundations across complex traits, as highlighted in a study by Gong *et al*¹⁸. However, prior
113 research on horizontal pleiotropy has primarily concentrated on exploring the relationship
114 between LTL and severe mental disorders¹⁹. Therefore, in light of the widespread associations
115 reported in epidemiological studies and the existing evidence of a shared genetic basis
116 LTL and CVDs, the necessity for a comprehensive, large-scale analysis is clear.

117

118 Recognizing the knowledge gap, our study aims to provide a thorough analysis of shared
119 genetic architectures between LTL and CVDs by encompassing the unprecedentedly large
120 available GWAS datasets in individuals of European ancestry. First, by charting the landscape
121 of genetic overlap beyond genetic correlation, we provide more granular insights into the
122 unique and shared genetic architectures between LTL and CVDs. Then, we employed
123 statistical techniques to capture diverse forms of genetic pleiotropy, followed by thorough analyses
124 to link the genomic findings to biological pathways, yielding profound implications for
125 conceptualizing shared genetic risk for both LTL and CVDs.

126

127 **Results**

128 **Genetic overlap beyond genetic correlation between LTL and six major CVDs**

129 Following the harmonization and filtering of SNPs shared across GWAS summary statistics,
130 we employed cross-trait linkage disequilibrium (LD) score regression (LDSC) to calculate
131 SNP-based heritability (h^2_{SNP}) and to assess genome-wide genetic correlation (r_g) between
132 LTL and six major CVDs. Univariate LDSC revealed that the estimated h^2_{SNP} for LTL was
133 2.76% (SE = 0.30%). On average, the estimated h^2_{SNP} was nearly threefold higher for AF

134 ($h^2_{SNP} = 2.50\%$, SE = 0.33%), CAD ($h^2_{SNP} = 3.25\%$, SE = 0.19%), and VTE ($h^2_{SNP} = 1.82\%$, SE
135 = 0.23%), in comparison to HF ($h^2_{SNP} = 0.80\%$, SE = 0.06%), PAD ($h^2_{SNP} = 0.94\%$, SE =
136 0.13%), and Stroke ($h^2_{SNP} = 0.60\%$, SE = 0.05%) (Supplementary Fig. 1a and Supplementary
137 Table 2a). The results of bivariate LDSC indicated a range of weak to moderate genome-wide
138 r_g between LTL and CVDs, excluding LTL-AF. Notably, the most pronounced negative r_g
139 were observed for LTL-PAD ($r_g = -0.250$, SE = 0.040, $P = 3.83 \times 10^{-10}$) and CAD ($r_g = -0.171$,
140 SE = 0.025, $P = 4.65 \times 10^{-12}$), while smaller but statistically significant r_g were noted for
141 Stroke ($r_g = -0.104$, SE = 0.037, $P = 4.60 \times 10^{-3}$) and VTE ($r_g = -0.072$, SE = 0.025, $P =$
142 4.20×10^{-3}). LTL was only moderately genetically correlated to HF ($r_g = -0.145$, SE = 0.037, $P =$
143 8.20×10^{-5}). In contrast, no significant r_g was observed between LTL and AF (Supplementary
144 Fig. 1b and Supplementary Table 2b). Although genome-wide r_g offered valuable insight into
145 the genetic overlap between phenotypes, it could not distinguish genetic overlap resulting
146 from a mixture of concordant and discordant effects from the absence of genetic overlap,
147 potentially yielding an estimated r_g near zero in both scenarios. Therefore, using multiple
148 methods with different model assumptions to identify and understand this "missing
149 dimension" of genetic overlap was essential for comprehensively characterizing the shared
150 genetic foundations across phenotypes.

151

152 Despite minimal r_g estimated by LDSC, the causal mixture modeling approach (MiXeR) was
153 then applied to elucidate extensive genetic overlap beyond genetic correlation by determining
154 the number of overlapping variants between LTL and six major CVDs, irrespective of the
155 direction of their effects. Univariate MiXeR revealed that LTL exhibited a lower degree of
156 polygenicity (N = 0.380K, SD = 0.026K). Among the six major CVDs, HF (N = 2.305K
157 'causal' variants explaining 90% of HF's h^2_{SNP} , SD = 0.213K) was the most polygenic,
158 followed by CAD (N = 1.528K, SD = 0.311K) and Stroke (N = 1.055K, SD = 0.117K). VTE,
159 PAD, and AF demonstrated lower polygenicity, associated with 0.308K to 0.504K variants at
160 90% h^2_{SNP} (Supplementary Table 3a). These findings highlighted a pattern of polygenicity
161 distinct from h^2_{SNP} estimates.

162

163 The result of bivariate MiXeR revealed substantial but distinct patterns of polygenic overlap

164 between LTL and CVDs. Given the low polygenicity of LTL and these CVD phenotypes
165 (including AF, PAD, and VTE), LTL was found to share less proportion of causal variants
166 with these CVDs, ranging from 18.30% in AF to 22.63% in PAD. However, relatively large
167 genetic overlaps were also observed between LTL and these CVDs (Fig. 2a, Supplementary
168 Fig. 2, Supplementary Table 3b). For example, polygenic overlap between LTL and PAD was
169 particularly striking (Dice coefficient $\hat{\rho} = 0.229$, $SD = 0.015$), with 0.086K ($SD = 0.007K$)
170 shared variants, representing 22.63% LTL-influencing variants and 23.24% PAD-influencing
171 variants, consistent with the strongest negative genome-wide genetic correlation ($r_g = -0.223$,
172 $SE = 0.014$) and genetic correlation of shared variants ($r_{gs} = -0.970$, $SE = 0.022$). A total of
173 0.069K ($SD = 0.019$) variants were estimated to be shared between LTL and AF,
174 representing 18.30% LTL-influencing variants and 13.79% AF-influencing variants, despite
175 weak negative genetic correlation. This pattern of extensive genetic overlap but weak r_g
176 indicated a predominance of mixed effect directions, supported by the MiXeR-estimated
177 proportion of shared 'causal' variants with concordant effects (0.416, $SD = 0.029$).
178 Considering the low polygenicity of LTL and high polygenic diseases such as CAD, HF, and
179 Stroke, significant disparities were observed in the number of shared and unique "causal"
180 variants. In particular, MiXeR estimated that of the 0.380K LTL-influencing variants, 48.69%,
181 39.82%, and 28.07% also influence HF, CAD, and Stroke, respectively. For example, LTL and
182 HF shared the largest number of variants ($N = 0.185K$, $SD = 0.024K$), with many more
183 unique variants of HF ($N = 2.12K$, $SD = 0.208K$) than unique variants of LTL (0.195K, $SD =$
184 0.030K), representing 48.69% LTL-influencing variants and 8.02% HF-influencing variants.
185 While they were moderately correlated at the genome-wide level ($r_g = -0.180$, $SE = 0.017$),
186 shared variants were strongly correlated ($r_{gs} = -0.913$, $SE = 0.087$). A similar, although less
187 pronounced, relationship was evident in LTL-CAD and LTL-Stroke. Furthermore, the
188 overlapping variants demonstrated a low level of effect direction concordance, highlighting
189 the prevalence of mixed effect directions between LTL and CVDs. These observations
190 suggested that the extent of polygenic overlap between LTL and CVDs were likely
191 underestimated by genome-wide genetic correlations (Fig. 2b).
192
193 Local genetic correlations provide a more effective means of capturing genetic associations

194 with mixed effect directions. Specifically, a pair of traits may display no genome-wide r_g due
195 to an equal number of positive and negative (opposite effect directions) local genetic
196 correlations with comparable magnitudes. To prevent the potential masking of local genetic
197 correlations when evaluating r_g at the genome-wide level, we applied Local Analysis of
198 [co]Variant Annotation (LAVA) to perform local genetic correlations ($loc-r_g$ s) between LTL
199 and six major CVDs at loci, where both phenotypes had heritability estimates significantly
200 different from zero (Supplementary Table 4). Overall, 45 local genomic regions were found
201 significant for bivariate analysis after correcting for multiple testing using FDR ($FDR < 0.05$,
202 Fig. 2a and Supplementary Table 5), with 62% and 38% of the partitions showing negative
203 and positive $loc-r_g$ s, respectively. Corroborating the MiXeR findings, LAVA estimated
204 correlated loci of LTL-PAD (2 positively correlated and 3 negatively correlated loci),
205 LTL-VTE (3 positively correlated and 4 negatively correlated loci), and LTL-AF (9 positively
206 correlated and 4 negatively correlated loci), adding further support for a shared genetic basis.
207 Local correlations for LTL-CAD, comprising 14 negatively and 3 positively correlated loci,
208 and LTL-Stroke, with 3 negatively correlated loci and none positively correlated, were
209 inconsistent with the mixed effect directions estimated by MiXeR. The absence of significant
210 loci between LTL and HF may be attributed to LAVA's propensity for identifying loci with
211 extreme correlations, in contrast to MiXeR, thereby highlighting loci more likely to be
212 statistically significant.

213

214 Interestingly, our investigation also identified a solitary region (LD block 1,841 on
215 chromosome 12, ranges from 111,592,382 to 113,947,983) displaying significant correlations
216 for a majority of the trait pairs, with uniform negative correlation values between -0.530 and
217 -0.692. Subsequent analyses utilizing Hypothesis Prioritisation in Multi-trait Colocalization
218 (HyPrColoc) revealed robust colocalization evidence for this locus between LTL and all
219 CVDs, excluding AF, HF and PAD, with a posterior probability (PP) higher than 0.7, which
220 encompassed the shared causal SNP (rs10774625, an intronic variant of the ataxin 2 (*ATXN2*)
221 gene on 12q24.12). The *ATXN2* rs10774625 polymorphism has been associated with various
222 CVDs, notably CAD, alongside cardiometabolic markers such as blood pressure and blood
223 lipids^{20,21}.

224

225 **The Causal inference between LTL and six major CVDs**

226 Despite these findings substantiating a shared genetic foundation between LTL and six major
227 CVDs, there was uncertainty in relation to whether the complex interplay predominantly
228 reflected horizontal pleiotropy or potentially involved a causal relationship (referred to as
229 ‘vertical pleiotropy’). Mendelian randomization (MR) harnesses vertical pleiotropy to deduce
230 potential causal relationships, excluding any SNPs indicative of horizontal pleiotropy.
231 Therefore, latent causal variable (LCV) analysis was utilized to elucidate the possible causal
232 relationships underlying the genetic correlations observed. Notably, none of the trait pairs
233 demonstrated tendencies indicative of partial genetic causation (Supplementary Fig. 3a,
234 Supplementary Table 6a). To ascertain the reproducibility of the partial causal associations
235 between trait pairs, we utilized the Latent Heritable Confounder Mendelian Randomization
236 (LHC-MR) approach, taking into account factors such as sample overlap, bidirectional causal
237 associations, and unobserved heritable confounders. Consistently, no evidence was found to
238 support a putative causal effect of LTL on CVDs (Supplementary Fig. 3b, Supplementary
239 Table 6b). Only the possible weak positive causal impact of CAD on LTL was observed, yet
240 lacked confirmation from conventional bidirectional MR analyses (Supplementary Table 6c),
241 emphasizing the need for cautious interpretation. Overall, MR analysis revealed that the
242 shared genetic foundation between LTL and CVDs cannot be ascribed to vertical pleiotropy.

243

244 **Pleiotropic genomic loci identified for LTL and CVDs**

245 The observed comorbidity between LTL and six major CVDs suggests that instead of a
246 predominance of trait-specific risk variants, there may be a set of pleiotropic variants
247 influencing the risk of both LTL and CVDs (i.e., horizontal pleiotropy). To investigate this,
248 we employed the Pleiotropic Analysis under a Composite Null Hypothesis (PLACO) to
249 pinpoint potential pleiotropic variants shared between LTL and CVDs, resulting in the
250 identification of 12,604 SNPs comprising 10,008 unique variants. Functional Mapping and
251 Annotation (FUMA) further delineated 248 independent genomic risk loci as pleiotropic,
252 spanning 122 unique chromosomal regions (Fig. 3, Supplementary Fig 4, Supplementary
253 Table 7). Among these, 194 loci were associated with LTL and 80 loci with CVDs. In

254 aggregate, 32 loci overlapped between LTL and CVDs, accounting for 16.49% and 40.00% of
255 the total number of loci linked to these respective categories. A total of 188 pleiotropic loci
256 exhibited genetic signals for multiple trait pairs, with 74 of them (39.36%) spanning 16
257 unique chromosomal regions, demonstrating this phenomenon in over half of the investigated
258 trait pairs. For example, the pleiotropic locus 16q22.1 (mapped gene: *TMED6*) was jointly
259 associated with LTL and all CVDs. A mixture of concordant and discordant allelic effects
260 existed in these pleiotropic loci. Remarkably, the selected effect alleles at the top SNPs within
261 or near 117 loci (47.2%) exhibited inconsistent effects on two traits within a pair of traits.
262 Essentially, these variants may concurrently increase LTL and diminish the risk of developing
263 CVDs, aligning with their robust genome-wide genetic correlation. Conversely, the remaining
264 SNPs demonstrated concordant associations with both LTL and CVDs, implying that these
265 SNPs might influence the risk of both traits in the same direction.

266

267 ANNOVAR category annotation of candidate SNPs shared between LTL and CVDs revealed
268 that 67 (27.0%) were in intergenic regions, 114 (45.9%) were in intronic regions, and only 14
269 (5.6%) were in exonic regions. For example, the index SNP rs1566452 at 16q22.1 locus
270 ($P_{PLACO} = 4.46 \times 10^{-8}$ for LTL-HF) was associated with artery coronary and artery tibial
271 eQTLs ($P_{Artery_Coronary} = 4.06 \times 10^{-5}$, $P_{Artery_Tibial} = 3.32 \times 10^{-10}$, Supplementary Table 9) for WW
272 domain-containing E3 ubiquitin protein ligase 2 (*WWP2*) gene encoding one of the E3
273 ubiquitin ligases, which critically participate in the development and progression of
274 cardiovascular diseases²². Besides, numerous ubiquitin E3 ligases have also been documented
275 to promote the degradation of human telomerase reverse transcriptase (hTERT), thereby
276 reducing telomerase activity and potentially leading to decreased telomere length²³.
277 Furthermore, we identified 21 top SNPs with combined annotation-dependent deletion
278 (CADD) scores exceeding 12.37, and 7 mRNA exonic variants had higher CADD scores,
279 which was indicative of potentially deleterious effects. Notably, rs11556924 within the zinc
280 finger C3HC-type containing 1 (*ZC3HC1*) gene represents an exonic non-synonymous variant
281 with a CADD score of 28 (variants with scores surpassing 20 are predicted to be among the
282 1.0% most deleterious substitutions in the human genome). Furthermore, nine SNPs were
283 assigned RegulomeDB scores of 1f, 1d, or 1a, indicating a likely influence on binding sites.

284 We followed up this finding using the GTEx database to investigate the gene regulatory
285 effects. For example, rs11779558 at 8p21.3 locus was significantly associated with eQTL
286 functionality in artery tibial ($P_{Artery_Tibial} = 5.79 \times 10^{-5}$) for exportin 7 (*XPO7*) gene.

287

288 Colocalization analysis further revealed 22 out of 248 potential pleiotropic loci with PP.H4
289 greater than 0.7, wherein 14 top SNPs at the corresponding loci were identified as
290 candidate-shared causal variants (Fig. 3, Supplementary Fig. 5, Supplementary Table 7).
291 Notably, the 12q24.12 locus, identified as pleiotropic for all correlated trait pairs except for
292 LTL-AF, exhibited strong evidence of colocalization between these trait pairs (PP.H4 ranging
293 from 0.729 to 0.998). HyPrColoc analysis further revealed strong colocalization evidence for
294 this locus between LTL and all CVDs except AF and PAD, with a PP exceeding 0.7, which
295 encompassed rs10774625 (an intronic variant of the *ATXN2* gene on 12q24.12) as the
296 potential shared causal variant. Moreover, 40 pleiotropic loci were identified with PP.H3
297 exceeding 0.7, indicating the possibility of different causal variants within these loci.

298

299 **Pleiotropic genes associated with LTL and multiple CVDs**

300 Despite the success of the above analyses in identifying disease risk loci, the biological
301 significance of most identified variants remains unknown. To achieve a more comprehensive
302 understanding of how genetic variation influences disease risk, we adopted approaches
303 integrating SNPs across a spectrum of association significance to construct a cohort of
304 predicted genes that could subsequently be mapped to functional pathways for analysis. We
305 employed two distinct strategies for mapping SNPs to genes: Firstly, a genome-wide
306 gene-based association study (GWGAS) in MAGMA and positional mapping in FUMA were
307 utilized, mapping SNPs to genes based on their physical position in the genome. Secondly, an
308 eQTL-informed GWGAS in e-MAGMA and eQTL mapping in FUMA were employed to
309 map SNPs to genes through their eQTL associations.

310

311 MAGMA analysis, utilizing 557 potential pleiotropic genes located within or overlapping
312 with 248 pleiotropic loci, identified 478 significant pleiotropic genes (323 unique), in which
313 244 genes were detected in two or more trait pairs (Fig. 4, Supplementary Table 11). For

314 example, *SH3PXD2A*, *SH2B3*, *BRAP*, *ATXN2*, *PTPN11*, *NAA25*, *ALDH2*, and *ACAD10* were
315 identified as significant pleiotropic genes in five pairs of traits. Remarkably, seven of eight
316 genes (excluding *SH3PXD2A*) were located on the 12q24.12 locus, identified in all trait pairs
317 except for LTL-AF. Of the pleiotropic genes identified, 98 (20.50%) were novel for LTL and
318 258 (53.97%) for CVDs. Only one pleiotropic gene, CD19 molecule (*CD19*), had not
319 previously been reported to be associated with both traits. Furthermore, 472 genes (98.74%)
320 identified by MAGMA were confirmed by FUMA positional mapping (Supplementary Table
321 9).

322

323 To pinpoint tissues potentially integral to the biological processes of LTL and six major CVDs,
324 we utilized LDSC-SEG for tissue-specific enrichment analysis using single trait GWAS
325 summary statistics. Regarding multi-tissue gene expression, we found that expressions of
326 LTL-associated loci were significantly enriched in spleen tissues, surpassing an FDR > 0.05
327 threshold (Fig. 5, Supplementary Table 14). Additionally, AF showed significant enrichment in
328 heart-related tissues, including heart left ventricle and heart atrial appendage, while CAD
329 demonstrated enrichment in artery-related tissues, such as artery tibial, artery aorta, and artery
330 coronary. Conversely, no significant tissue-specific enrichment was observed for VTE, HF,
331 PAD, and Stroke. These findings were corroborated by multi-tissue chromatin interaction
332 results.

333

334 Given that MAGMA assigns SNPs to the nearest genes based on arbitrary genomic windows,
335 and considering that the effects of a locus don't always operate through the nearest gene, a
336 critical need remains to functionally link SNPs to genes (e.g., through genetic regulation) to
337 enhance our understanding of potential underlying mechanisms. Consequently, e-MAGMA
338 was conducted to uncover functional gene associations potentially overlooked by the
339 proximity-based SNP assignment in MAGMA, thereby illuminating alternative causal
340 pathways from SNPs to traits. E-MAGMA analysis revealed 1,844 significant tissue-specific
341 pleiotropic genes (419 unique) after Bonferroni correction, each strongly enriched in at least
342 one tissue (Supplementary Table 15). Of these, 918 tissue-specific pleiotropic genes (75
343 unique) were significantly identified across multiple trait-related tissues in at least two trait

344 pairs. For example, *MAPKAPK5*, *TMEM116*, *HECTD4*, *ALDH2*, and *ACAD10*, all located on
345 the 12q24.12 locus, were recognized as significant tissue-specific pleiotropic genes in all trait
346 pairs except for LTL-AF. Notably, transmembrane protein 116 (*TMEM116*) gene was found to
347 be highly tissue-specific, showing enrichment in eight trait-related tissues, including artery
348 tibial, artery coronary, adipose visceral (omentum), adipose subcutaneous, heart atrial
349 appendage, heart left ventricle, liver, and whole blood. *TMEM116* encoded a transmembrane
350 protein involved in blood coagulation and had been identified as a potential risk gene for
351 coronary atherosclerosis in previous studies²⁴. However, its relationship with LTL remained
352 less understood. In comparison with the transcriptome-wide association study (TWAS) results
353 for single-trait GWAS, we identified 415 tissue-specific pleiotropic genes as novel for LTL
354 and 851 for CVDs (Supplementary Table 16). Finally, we successfully replicated 1,117 genes
355 (60.57%) using FUMA eQTL mapping (Supplementary Table 9).

356

357 Finally, 289 pleiotropic genes (207 unique) were jointly identified by MAGMA and
358 e-MAGMA analysis, in which 50 unique genes were detected in 2 or more trait pairs, further
359 suggesting the tissue specificity of these pleiotropic genes (Supplementary Table 11). For
360 example, *ACAD10* (12q24.12), *ALDH2* (12q24.12), *HECTD4* (12q24.12), *MAPKAPK5*
361 (12q24.12), *NAA25* (12q24.12), *SH2B3* (12q24.12), *TMEM116* (12q24.12), *SERPINF1*
362 (17p13.3), *TMED6* (16q22.1), and *XPO7* (8p21.3) were identified as significant pleiotropic
363 genes in more than half of the trait pairs. Remarkably, two of ten genes (including *ALDH2*
364 and *ACAD10*) were located at the 12q24.12 locus, identified in five pairs of traits except for
365 LTL-AF. Acyl-CoA dehydrogenase family member 10 (*ACAD10*), a gene encoding an
366 enzyme crucial for fatty acid beta-oxidation in mitochondria, critically regulates cellular lipid
367 synthesis with significant expression in the human brain²⁵. Previous data show that
368 homozygous loss of function of *ACAD10* results in perturbed lipid synthesis that potentially
369 influences the development of CVDs, whereas common gene variants have been associated
370 with CAD, Stroke, and hypertension (a common CVD risk factor). Additionally, aldehyde
371 dehydrogenase 2 family member (*ALDH2*) gene encoding a vital mitochondrial enzyme
372 critical for cardiac function, has been associated with exacerbated myocardial remodeling and
373 contractile dysfunction in aging. This association is possibly through mitochondrial damage

374 mediated by the AMPK/Sirt1 pathway.

375

376 **Shared biological pathways between LTL and six major CVDs**

377 To investigate the concept that a group of genes might work collectively to fulfill specific
378 biological functions through shared pathways or functional category enrichments, we
379 employed various analytical strategies, including gene-set analysis of genes identified via
380 MAGMA and functional enrichment analysis targeting tissue-specific genes. After rigorously
381 adjusting for 7,744 gene sets (biological processes sets from the Molecular Signatures
382 Database (MSigDB, v.2023.1; C5: GO BP)), we noted minimal overlap in gene sets between
383 LTL and six major CVDs. Only three gene sets associated with chromatin organization, the
384 negative regulation of nucleobase-containing compound metabolic processes, and the
385 negative regulation of miRNA maturation were enriched across several trait pairs, and no
386 gene set demonstrated significant enrichment across more than one trait (Supplementary
387 Table 17a). Remarkably, massive genes shared between LTL and AF were most significantly
388 linked to the 'negative regulation of the nucleobase-containing compound metabolic process.'
389 Subsequently, we identified several biological processes overrepresented among the 50
390 pleiotropic genes detected by MAGMA and e-MAGMA analysis in 2 or more trait pairs
391 shared between LTL and CVDs using the ToppGene Functional Annotation tool (ToppFun)
392 (Supplementary Fig. 6, Supplementary Table 17b). All identified significant biological
393 processes, with the exception of the DNA biosynthetic process, were directly associated with
394 telomere maintenance processes such as 'telomere maintenance via telomerase,' 'telomere
395 maintenance via telomere lengthening,' 'telomere capping,' and 'telomere organization.'
396 Interestingly, the nucleobase-containing compound metabolic process identified in the
397 MAGMA gene-set analysis encompassed both DNA biosynthetic processes and telomere
398 maintenance mechanisms, implying a key role in shared biological pathways between LTL
399 and CVDs.

400

401 **Shared causal proteins between LTL and six major CVDs**

402 We undertook a comprehensive analysis of the Mendelian randomization (MR) associations
403 between 1,922 unique proteins and the risk of LTL and six major CVDs using the summary

404 data-based Mendelian randomization (SMR), each protein with index cis-acting variants
405 (cis-pQTL) obtained from the UK Biobank Pharma Proteomics Project (UKB-PPP).
406 Following the exclusion of associations failing the HEIDI test, and after conducting
407 sensitivity analysis with multi-SNPs-SMR and applying multiple testing corrections via
408 Bonferroni adjustment, the genetically predicted levels of 85 proteins were found to be
409 significantly associated with the risk of LTL and CVDs (Fig. 6, Supplementary Table 18).
410 Specifically, 12, 9, 26, 24, 2, 6, and 6 proteins were significantly associated with LTL, AF,
411 CAD, VTE, HF, PAD, and Stroke, respectively. Notably, SH2B3 emerged as significantly
412 associated with both LTL-CAD and LTL-VTE, also showing strong colocalization evidence in
413 HyPrColoc analysis, with rs10774625 pinpointed as a shared causal variant. The index SNP
414 rs10774625, located at the 12q24.12 locus (an intronic variant of the *ATXN2* gene), was
415 associated with eQTLs in whole blood ($P_{Whole_Blood} = 3.48 \times 10^{-4}$) and was also linked to pQTLs
416 in whole blood ($P_{Whole_Blood} = 1.08 \times 10^{-3}$) for the SH2B adaptor protein 3 (SH2B3). SH2B3
417 acted as an adaptor protein, playing a crucial role in negatively regulating cytokine signaling
418 and cell proliferation. Prior research indicated that SH2B3 was associated with longevity,
419 with missense alleles within SH2B3 influencing life expectancy through predisposition to
420 cardiovascular events²⁶. Enhancing SH2B3 function has demonstrated therapeutic potential in
421 hypertension and other cardiovascular conditions.

422

423 **Discussion**

424 In this extensive genome-wide pleiotropy association study, we systematically elucidated the
425 shared genetic architecture beyond genome-wide genetic correlations between LTL and six
426 major CVDs, uncovering no causal relationships between them. Subsequent in-depth analyses
427 identified pleiotropic genetic variants and loci, pleiotropic genes, biological pathways, and
428 protein targets, all reinforcing the involvement of the DNA biosynthesis and telomere
429 maintenance in the shared genetic etiology of these traits. Overall, these findings offer novel
430 insights into the relationship and shared genetic mechanisms underlying LTL and six major
431 CVDs.

432

433 Consistent with prior robust epidemiological evidence, our findings reveal weak to moderate,

434 yet significant, negative genome-wide genetic correlations between LTL and six major CVDs,
435 except for LTL-AF. Beyond mere genetic correlations, our analysis uncovers a more extensive
436 degree of genetic overlap between LTL and CVDs using MiXeR and LAVA, involving a
437 broad mixture of both concordant and discordant effect sizes. Employing MiXeR, we
438 demonstrated that the polygenicity of LTL and CVDs presents fundamental distinctions
439 beyond SNP-based heritability. Briefly, LTL was substantially less polygenic than three CVD
440 phenotypes (CAD, HF, and Stroke), yet exhibited similar polygenicity to other CVD
441 phenotypes, including AF, VTE, and PAD. Despite notable differences in polygenicity, we
442 observed extensive genetic overlaps between LTL and all CVDs, supported by LAVA local
443 correlations. This pattern emerged in cases of weak or non-significant genome-wide genetic
444 correlations, such as between LTL and AF, and in strong genome-wide correlations, such as
445 between LTL and PAD. For example, despite the absence of genome-wide genetic
446 correlations, a pronounced fraction of the genetic risk underlying LTL overlaps with AF was
447 indicated using MiXeR. The findings correspond with the discovery of a similar number of
448 positively and negatively correlated genomic regions between LTL and AF by LAVA,
449 alongside further detecting much more shared pleiotropic loci below the genome-wide
450 significance threshold. Although the local genetic correlations encompassed genomic regions
451 averaging approximately 1 megabase (Mb) in width, this resolution required refinement to
452 minimize the impact of heterogeneous effects on estimation accuracy. For example, we noted
453 inconsistent effects across lead variants in the LTL and AF pleiotropic analysis. Specifically,
454 out of 47 top lead variants, 28 demonstrated the same effect direction for LTL and AF,
455 whereas the remaining 19 exhibited opposite effect directions. This indicates that pleiotropic
456 analysis at the level of single variants was necessary to offer further, more detailed insights
457 into the shared genetic underpinnings across complex traits. These findings indicate that
458 genetic overlap between LTL and CVDs was greatly underestimated due to the patterns of
459 mixed effect directions concealed by estimates of genome-wide genetic correlations.

460

461 The intricate relationship between LTL and six major CVDs may be further elucidated by
462 examining the vertical and horizontal pleiotropy mechanisms that underpin their shared
463 genetic basis. The causal relationships between LTL and six major CVDs were then

464 predominantly elucidated by the effects attributed to vertical pleiotropy. Using more recent
465 GWAS summary data, we discovered minimal evidence supporting the causal effects of LTL
466 on CVDs and vice versa using LCV and LHC-MR. This finding stood in stark contrast to
467 existing research that suggested such effects were present. Previous two-sample MR studies
468 have reported associations between genetic liability to LTL shortening and the increased risk
469 of CAD and Stroke, but results for the risk of HF have been inconsistent. Another study even
470 demonstrated that genetically predicted AF contributes to the shortening of LTL rather than
471 the reverse. Indeed, the majority of GWAS meta-analyses of CVDs have incorporated
472 samples from the UK Biobank, resulting in considerable sample overlap with GWAS studies
473 on LTL, which might contravene the basic principles of two-sample MR. In addition, the
474 interpretation of the MR estimate in this case was quite complicated by the fact that both LTL
475 and CVDs were time-varying outcomes with a late age of onset. Consequently, our study's
476 findings suggest that the association between LTL and CVDs may have been greatly
477 overestimated in prior research, potentially as a result of sample overlap, reverse causation, or
478 unmeasured heritable confounding factors, indicating that pleiotropic and common biological
479 pathways may be a better explanation for their association.

480

481 The horizontal pleiotropic analyses, encompassing different levels such as pleiotropic genetic
482 variants and loci, pleiotropic genes, biological pathways, and protein targets, revealed
483 significant genetic overlaps between LTL and CVDs from another perspective from another
484 perspective. At the SNP level, pleiotropic variants linking LTL and CVDs were broadly
485 distributed, with a notable emphasis on shared pleiotropic loci among certain trait pairs,
486 including 16q22.1 (*TMED6*), 8p21.3 (*XPO7*), 17p13.3 (*SERPINF1*), and 12q24.12 (*ATXN2*).
487 For example, the transmembrane emp24 domain-containing protein 6 precursor (*TMED6*),
488 highly and selectively expressed in pancreatic islets, was found to be associated with LTL and
489 all CVDs, which belonged to the EMP24_GP25L superfamily and played a crucial role in
490 protein trafficking and secretion. Knockdown of the *TMED6* gene in Min6 β -cells and INS1
491 cells led to a reduction in glucose-stimulated insulin secretion, suggesting that dysregulation
492 of *TMED6* may play a critical role in the onset of type 2 diabetes^{27,28}. To date, there have been
493 no reports of an association between *TMED6* and either LTL or CVDs, hinting at a potential

494 biological mechanism that may be mediated through CVD risk factors, particularly type 2
495 diabetes. Exportin-7 (*XPO7*) is a bidirectional transporter regulating the nuclear-cytoplasmic
496 shuttling of a wide array of substrates, yet its function remains relatively obscure.²⁹ Recent
497 research has identified *XPO7* as a novel regulator of cellular senescence. The depletion of
498 *XPO7* leads to decreased levels of *TCF3* (transcription factor 3, also known as *E2A*) and
499 impaired induction of the cyclin-dependent kinase inhibitor p21^{CIP1}, which was critical during
500 oncogene-induced senescence³⁰. The role and malfunction of the Serpin Family F Member 1
501 (*SERPINF1*), also known as the pigment epithelium-derived factor (*PEDF*) gene, in the aging
502 process has currently been a hot topic. Briefly, the reduction in *PEDF* expression levels, both
503 directly and indirectly, can prompt cellular senescence via the modulation of various signaling
504 pathways^{31,32}. Studies have also demonstrated that *PEDF* possesses insulin-sensitizing effects
505 in the liver and adipose tissues and exhibits anti-inflammatory, anti-thrombogenic, and
506 vasculoprotective properties in vivo, offering protection against metabolic syndrome and
507 cardiovascular diseases^{33,34}. Overall, this study extends previous findings on shared genetic
508 architecture by offering a more comprehensive characterization of specific pleiotropic loci.

509

510 At the gene level, we employed two gene-mapping strategies to identify credible mapped
511 genes for all jointly associated pleiotropic loci. Seven genes, namely *ALDH2*, *ACAD10*,
512 *TMEM116*, *SH2B3* (all located at 12q24.12), *TMED6* (16q22.1), *SERPINF1* (17p13.3), and
513 *XPO7* (8p21.3), have emerged as the most pleiotropic; they are thought to have important
514 regulatory functions, influencing over half of the trait pairs examined. For example, among
515 these genes, *ALDH2* and *ACAD10*, located at the 12q24.12 locus, were associated with LTL
516 and all CVDs except AF. Aldehyde dehydrogenase 2 (*ALDH2*) is the gene with the highest
517 number of genetic polymorphisms in humans; it is involved in encoding a mitochondrial
518 enzyme critical for detoxifying reactive aldehydes. For example, the *ALDH2* rs671
519 inactivating polymorphism, found in up to 8% of the global population and up to 50% of the
520 East Asian population, is associated with an elevated risk of several CVDs, such as CAD.
521 While numerous studies have connected aldehyde accumulation, due to alcohol consumption,
522 ischemia, or heightened oxidative stress, to elevated CVD risk, this accumulation alone does
523 not fully account for their complex interactions^{35,36}. Moreover, previous studies indicated that

524 the *ALDH2* enzyme might also exacerbate myocardial remodeling and contractile dysfunction
525 during aging, potentially via AMPK/Sirt1-mediated mitochondrial damage. Acyl-CoA
526 dehydrogenase family member 10 (*ACAD10*) gene encoded an enzyme involved in fatty acid
527 beta-oxidation in mitochondria³⁷. *ACAD10* is predominantly expressed in the human brain
528 and is believed to play a role in physiological functions within the central nervous system,
529 such as the regulation of cellular lipid synthesis. *ACAD10* has previously been identified as
530 one of the putative causal genes for CAD, stroke, and hypertension, a common risk factor for
531 CVDs³⁸. Collectively, these findings imply that *ACAD10* has a significant role in regulating
532 lipid synthesis through distinct molecular and cellular pathways, potentially influencing the
533 development of CVDs. Further research is required to elucidate the precise mechanisms of
534 interaction between *ACAD10* and both LTL or CVDs.

535

536 The 12q24.12 locus was a top hit region, which identified as pleiotropic for all correlated trait
537 pairs except for LTL-AF and exhibited strong evidence of colocalization between these trait
538 pairs. Besides, HyPrColoc further revealed robust colocalization evidence for this locus
539 between LTL and all CVDs, excluding AF and PAD, encompassing the shared causal SNP
540 (i.e., rs10774625). LAVA results supported these findings and showed consistent negative
541 local genetic correlations between LTL and all CVDs, except for AF, HF and PAD. The index
542 SNP rs10774625 polymorphism has been reported to be associated with a variety of CVDs,
543 notably CAD, alongside cardiometabolic markers such as blood pressure and blood lipids.
544 The index SNP rs10774625, located at the 12q24.12 locus (an intronic variant of the *ATXN2*
545 gene), was associated with eQTLs and pQTLs in whole blood for the SH2B adaptor protein 3
546 (*SH2B3*). The proteome-wide Mendelian Randomization study revealed that the genetically
547 predicted levels of SH2B3 protein were significantly associated with both LTL-CAD and
548 LTL-VTE. *SH2B3* is a member of the adapter protein family and plays a critical role in
549 negatively regulating cytokine signaling and cell proliferation. It was originally described as a
550 regulator of hematopoietic and lymphocyte differentiation and was implicated in the
551 transduction and regulation of growth factors and inflammation-related cytokine
552 receptor-mediated signaling. *SH2B3* missense variants may affect lifespan through
553 cardiovascular disease³⁹. Specifically, *SH2B3* can lead to increased production of IFN γ , which

554 acts as a pro-inflammatory mediator to induce the polarization of macrophages into different
555 states. This process is crucial for mediating inflammatory regulation and fibrosis
556 post-myocardial infarction⁴⁰. On the other hand, IFN γ is released and activated by CD4 T
557 helper cells—specifically Th1 cells—which are instrumental in coordinating immune cell
558 infiltration and inflammation. The above immune cells and inflammatory signals are pivotal
559 in the progression of non-ischemic heart failure in patients⁴¹. Notably, increased
560 cardiomyocyte size and fibrosis are critical characteristics of cardiac hypertrophy and
561 remodeling, which ultimately lead to heart failure⁴². Further studies have demonstrated that
562 cardiac-specific *SH2B3* overexpression exacerbates pressure overload, leading to cardiac
563 hypertrophy, fibrosis, and dysfunction by activating focal adhesion kinase, which
564 subsequently triggers the downstream phosphoinositide 3-kinase-AKT-target of Programmed
565 Death-1 (PD-1). Significant overexpression of the *SH2B3* gene promotes the activation of the
566 Akt signaling pathway, which can promote cardiac hypertrophy and fibrosis and lead to the
567 deterioration of cardiac function⁴³. Meanwhile, age-related telomere dysfunction is a core
568 driver of inflammation⁴⁴. Therefore, *SH2B3* may become one of the most promising
569 therapeutic targets for CVDs. In contrast, *TMEM116*, a member of the TMEM family of
570 proteins that spans the plasma membrane of cells to facilitate intercellular communication,
571 demonstrates a different aspect of disease association. Multiple members of the TMEM
572 family may be up or down-regulated in tumor tissues, and some of them are used as cancer
573 prognostic biomarkers⁴⁵. However, in studies of cardiovascular diseases, this protein has only
574 been found to be related to coronary atherosclerosis⁴⁶, and the specific mechanism is unclear
575 and requires further research.

576

577 At the pathway level, functional analyses of the pleiotropic loci between LTL and CVDs have
578 implicated genes involved in the metabolic processes of nucleobase-containing compounds,
579 including DNA biosynthesis and telomere maintenance. There is a close relationship between
580 telomere maintenance, telomerase expression, and extension of cell lifespan. Telomeres
581 undergo shortening during repeated cell divisions, and when their length diminishes to a
582 critical point, the resultant genomic instability can lead to further genetic abnormalities that
583 promote cell death or apoptosis, which is a hallmark of cellular senescence. Estrogen, stress

584 accumulation from oxidative damage, hypertension, etc., are believed to significantly impact
585 telomere homeostasis and contribute to the development of CVDs⁴⁷. This finding bolsters the
586 theory that progressive telomere shortening contributes to the pathogenesis of age-related
587 human diseases such as CVDs. Specifically, Minano et al. documented the presence of
588 vascular endothelial cells exhibiting age-related phenotypes within human atherosclerotic
589 lesions⁴⁸. Excessive vascular smooth muscle cells are stimulated to proliferate and migrate,
590 leading to the growth of atherosclerosis. Besides, telomerase activation and telomere
591 maintenance are critical in increasing the proliferation and growth of vascular smooth muscle
592 cells. Activation through the telomerase reverse transcriptase component (TERT) extends the
593 lifespan of cultured vascular smooth muscle cells. Conversely, telomerase inhibition can
594 extend the lifespan and reduce the proliferation of cultured vascular smooth muscle cells,
595 thereby decreasing the risk of atherosclerosis⁴⁹. Therefore, these findings highlight the role of
596 telomeres in cardiovascular health, suggesting that methods to modulate the balance of
597 telomerase activity may be critical for developing effective interventions for related CVDs.

598

599 There were some limitations to the current study. Firstly, the main analysis focused solely on
600 individuals of European ancestry due to the scarcity of sufficiently powered GWAS involving
601 other ancestries. Nevertheless, we utilized GWAS summary data from East Asian ancestries
602 for replication, partially confirming the consistency of the genetic foundation identified in the
603 European sample. Future studies with a trans-ancestral approach are necessary to evaluate the
604 universality of these findings. Secondly, the analysis focused on common genetic variants that
605 account for only a small fraction of overall disease risk. The remaining variance was likely
606 attributable to many undetected SNPs, rare variants, or gene interactions. Therefore, they need
607 to be further studied to achieve a more complete understanding. Thirdly, although we
608 uncovered the potential shared genetic architecture, the mechanisms of shared biological
609 pathways still require further experimental validation. Finally, our analysis of GWAS
610 summary data encompassed six major CVDs, representing a significant portion of the genetic
611 risk architecture for these conditions, though not comprehensively. As GWAS datasets expand,
612 it will be crucial to undertake cross-trait analyses incorporating more varied datasets and
613 additional diseases to enhance our understanding.

614

615 **Conclusion**

616 In conclusion, we found extensive polygenic overlap between LTL and CVDs, with distinct patterns
617 of genetic correlations and effect directions, uncovering no causal relationships. It was proved that
618 LTL and six major CVDs share pleiotropic genetic variants, loci, genes, biological pathways, and
619 protein targets, underscoring the role of DNA biosynthesis and telomere maintenance in their
620 common genetic etiology. These findings elucidate the interconnected mechanisms between LTL
621 and CVDs, potentially guiding targeted therapies and clinical practice.

622

623 **Methods**

624 **Data Sources and Quality Control**

625 Figure 1 outlines the workflow for our study. Due to the confounding effects of ancestral
626 differences in linkage disequilibrium (LD) structure and the scarcity of sufficiently large
627 multi-ancestry samples, we limited our main analysis to individuals of European ancestry. We
628 sourced genome-wide association study (GWAS) summary statistics from the most
629 comprehensive and recent publicly available datasets of European ancestry. Specifically,
630 GWAS summary statistics for leukocyte telomere length (LTL) were derived from a published
631 GWAS comprising 464,716 individuals of European ancestry from the UK Biobank⁷. LTL
632 was quantified as the ratio of telomere repeats copy number (T) to a single copy gene (S) in a
633 mixed leukocyte population, measured via a multiplex quantitative polymerase chain reaction
634 (qPCR) assay, and subsequently log-transformed to achieve an approximation to a normal
635 distribution. Our selection criteria for GWAS included studies with sample sizes exceeding
636 50,000 to ensure adequate statistical power. Accordingly, we included GWAS summary
637 statistics for six major cardiovascular diseases (CVDs): atrial fibrillation (AF)⁸, coronary
638 artery disease (CAD)⁹, venous thromboembolism (VTE)¹⁰, heart failure (HF)¹¹, peripheral
639 artery disease (PAD)¹², and stroke¹³. AF GWAS summary statistics were sourced from a
640 genome-wide meta-analysis of six studies (The Nord-Trøndelag Health Study [HUNT],
641 deCODE, the Michigan Genomics Initiative [MGI], DiscovEHR, UK Biobank, and the Atrial
642 Fibrillation Genetics [AFGen] Consortium), encompassing 60,620 AF cases and 970,216
643 controls of European ancestry. For CAD, we utilized GWAS summary statistics from a

644 genome-wide meta-analysis by the CARDIoGRAMplusC4D Consortium and the UK
645 Biobank, which included 181,522 cases and 984,168 controls. VTE GWAS summary statistics
646 were extracted from a meta-analysis of 81,190 cases and 1,419,671 controls of European
647 ancestry across 7 cohorts (the Copenhagen Hospital Biobank Cardiovascular Disease Cohort
648 [CHB-CVDC], Danish Blood Donor Study [DBDS], deCODE, Intermountain Healthcare, UK
649 Biobank, FinnGen, and Million Veterans Program [MVP] Consortium). GWAS summary
650 statistics for HF came from the Heart Failure Molecular Epidemiology for Therapeutic
651 Targets (HERMES) Consortium, including 47,309 cases and 930,014 controls. GWAS
652 summary statistics for PAD were derived from a genome-wide meta-analysis of 11
653 independent GWASs, totaling 12,086 cases and 499,548 controls. Lastly, GWAS summary
654 statistics for Stroke were obtained from the GIGASTROKE consortium, which comprised
655 73,652 cases and 1,234,808 controls of European ancestry. Detailed information about these
656 GWAS summary statistics and their original publication sources is available in Supplementary
657 Table 1.

658

659 Prior to further analysis, stringent quality control measures were applied to the GWAS
660 summary statistics, encompassing several key steps: (i) alignment with the hg19 genome
661 build, referencing the 1000 Genomes Project Phase 3 Europeans; (ii) restriction of the
662 analysis to autosomal chromosomes; (iii) removal of single nucleotide polymorphisms (SNPs)
663 lacking a rsID or presenting duplicated rsIDs; and (iv) exclusion of rare or low-frequency
664 variants, defined by a minor allele frequency (MAF) less than 1%. To ensure robust and
665 interpretable comparisons between LTL and CVDs, we standardized the summary statistics to
666 include only SNPs present across all analyzed phenotypes, resulting in a cohesive dataset of
667 6,923,146 SNPs. Additionally, in subsequent analyses, we implemented further data
668 processing techniques tailored to the specific requirements of various statistical tools.

669

670 **Genetic overlap**

671 To explore the shared genetic foundations between LTL and six major CVDs, we evaluated
672 genetic overlap across genome-wide, polygenic, and local levels.

673

674 **Genome-wide genetic correlation analysis between LTL and CVDs**

675 At the genome-wide level, we analyzed SNP-level heritability for each trait and the genetic
676 correlations (r_g) between LTL and six major CVDs using cross-trait linkage disequilibrium
677 (LD) score regression (LDSC)^{50,51}. LDSC facilitates the estimation of the average genetic
678 effect sharing across the entire genome between two traits, leveraging GWAS summary
679 statistics. This includes the contribution of SNPs below the threshold of genome-wide
680 significance and accounts for potential confounding factors such as polygenicity, sample
681 overlap, and population stratification. First, SNP-based heritability (h^2_{SNP} , representing the
682 fraction of phenotypic variation explained by common genetic variations included in the
683 study) for LTL and six major CVDs was estimated using univariate LDSC. This analysis
684 utilized pre-computed LD scores from the European reference panel in the 1000 Genomes
685 Project Phase 3, excluding SNPs that did not overlap with the reference panel. Notably, the
686 major histocompatibility complex (MHC) region (chr 6: 25-35 Mb), known for its intricate
687 LD structure, was omitted from the main analysis. Second, bivariate LDSC analysis estimated
688 the genetic correlations between LTL and the six major CVDs. This method utilizes a
689 weighted linear model, where it regresses the product of Z-statistics from two traits against
690 the LD score across all genetic variants genome-wide. Genetic correlations with P -values
691 below the Bonferroni-adjusted threshold ($P = 0.05 / \text{number of trait pairs} = 0.05 / 6 =$
692 8.33×10^{-3}) were deemed statistically significant.

693

694 To elucidate the biological underpinnings of the shared genetic predisposition to LTL and six
695 major CVDs, we employed stratified LDSC applied to specifically expressed genes
696 (LDSC-SEG) to identify relevant tissue and cell types. This analysis incorporated tissue and
697 cell type-specific expression data from the Genotype-Tissue Expression (GTEx) project and
698 the Franke lab, covering 53 tissues and 152 cell types. Additionally, we utilized
699 chromatin-based annotations associated with six epigenetic marks (DNase hypersensitivity,
700 H3K27ac, H3K4me1, H3K4me3, H3K9ac, and H3K36me3) for validation purposes. These
701 annotations included 93 labels from the Encyclopedia of DNA Elements (ENCODE) project
702 and 396 labels from the Roadmap Epigenomics database. For the identified relevant tissues or
703 cell types, we adjusted the P -values for the significance of the coefficients using the False

704 Discovery Rate (FDR) method. An FDR threshold of < 0.05 was established as the criterion
705 for statistical significance.

706

707 **Polygenic overlap analysis between LTL and CVDs**

708 To augment the genome-wide genetic correlation analysis, we engaged the causal mixture
709 modeling approach (MiXeR) to quantify the polygenic overlap between LTL and six major
710 CVDs, independent of genetic correlation directions. MiXeR estimates the quantity of shared
711 and phenotype-specific "causal" variants that exert non-zero additive genetic effects,
712 accounting for 90% of SNP-heritability in each trait⁵². This 90% SNP-heritability threshold
713 minimizes the influence of variants with negligible effects. Importantly, MiXeR's capacity to
714 evaluate polygenic overlap without regard to the directional effects of variants offers a
715 nuanced view of local genetic associations that might be obscured in traditional genome-wide
716 genetic correlation estimations due to opposing variant effects. Initially, univariate MiXeR
717 analyses estimated the "causal" variant count for LTL and six major CVDs, assessing both
718 polygenicity and the average magnitude of additive genetic effects among these variants. The
719 LD structure was determined using the genotype reference panel from the 1000 Genomes
720 Project Phase 3. The MHC region (chr 6: 25-35 Mb), known for its intricate LD structure, was
721 excluded from the main analysis. Subsequently, bivariate MiXeR analysis quantified the
722 polygenic overlap between LTL and the six CVDs. This analysis delineated the additive
723 genetic effects across four categories: (i) SNPs with zero effect on both traits, (ii)
724 trait-specific SNPs with non-zero effects on the first trait, (iii) trait-specific SNPs with
725 non-zero effects on the second trait, and (iv) SNPs affecting both traits. The Dice coefficient
726 (DC), representing the proportion of shared SNPs between two traits relative to the total
727 number of SNPs associated with either trait, was used to estimate polygenic overlap.
728 Additionally, MiXeR calculated the overall genetic correlations (r_g), the correlation of effect
729 sizes within the shared genetic component ($r_{g,s}$), and the fraction of variants with concordant
730 effects in the shared component. The model's predictive accuracy was evaluated through
731 comparison of the modeled versus actual data, utilizing conditional quantile-quantile (Q-Q)
732 plots, log-likelihood plots, and the Akaike information criterion (AIC). Positive AIC
733 differences are interpreted as evidence that the best-fitting MiXeR estimates are

734 distinguishable from the reference model. A negative AIC value indicates that the MiXeR
735 model fails to distinguish effectively between maximal and minimal overlap scenarios.

736

737 **Local genetic correlation between LTL and CVDs**

738 To investigate whether there are any genomic loci with pronounced genetic correlations
739 despite negligible genome-wide r_g , we utilized the Local Analysis of [co]Variant Annotation
740 (LAVA) method to estimate the localized genetic correlation between LTL and six major
741 CVDs. LAVA, grounded in a fixed-effects statistical model, enables the estimation of local
742 SNP heritability ($loc-h^2_{SNP}$) and genetic correlations ($loc-r_{g,s}$) across 2,495 semi-independent
743 genetic regions of approximately equal size (~1 Mb). This method is adept at identifying loci
744 with mixed effect directions, offering a nuanced measure of genome-wide genetic overlap,
745 albeit influenced by the statistical power of the underlying GWAS data. For this analysis, we
746 adopted the genomic regions defined by Werme et al. as autosomal LD blocks, characterized
747 by minimal inter-LD block linkage with an average size of 1 million bases, each containing at
748 least 2500 variants. The LD reference panel from the 1000 Genomes Project Phase 3 for
749 European samples was employed, and consistent with LDSC and MiXeR analyses, the MHC
750 region (chr 6: 25-35 Mb) was excluded. Identifying meaningful $loc-r_{g,s}$ requires a significant
751 local genetic signal; thus, a stringent P -value threshold ($P < 1 \times 10^{-4}$) was applied to filter out
752 non-significant loci. Subsequent bivariate testing on selected loci and traits with notable
753 univariate genetic signals was conducted. For these loci, P -values for $loc-r_{g,s}$ were adjusted
754 using the Benjamini-Hochberg FDR method, with an FDR < 0.05 establishing statistical
755 significance. LAVA incorporated an estimate of sample overlap (genetic covariance intercept)
756 from bivariate LDSC analyses to account for potential sample overlap effects.

757

758 In cases where shared risk loci were identified across multiple phenotypes, the Hypothesis
759 Prioritisation for multi-trait Colocalization (HyPrColoc) method was employed to assess
760 whether association signals across more than a pair of traits were colocalized. HyPrColoc, an
761 efficient deterministic Bayesian clustering algorithm, leverages GWAS summary statistics to
762 identify clusters of colocalized traits and potential causal variants within a genomic locus,
763 providing a posterior probability (PP) of colocalization for each cluster. Loci with a PP > 0.7

764 were deemed colocalized, enhancing our understanding of shared genetic architectures across
765 traits.

766

767 **Pleiotropy insights to dissect genetic overlap**

768 **Causal inference between LTL and CVDs**

769 To elucidate the potential causal relationships underlying the genetic correlations observed
770 between LTL and six major CVDs, we employed latent causal variable (LCV) analysis. This
771 method posits that the genetic correlation between two traits operates through a latent factor,
772 allowing for the distinction between genetic causality (vertical pleiotropy) and both correlated
773 and uncorrelated horizontal pleiotropy⁵³. It achieves this by estimating the genetic causality
774 proportion (GCP) across all genetic variants, where GCP quantifies the share of each trait's
775 heritability explained by a mutual latent factor. It is essential to emphasize that GCP does not
776 indicate the magnitude of causal effects but only implies a causal relationship between traits.
777 This method offers insights to evaluate whether the impact of one trait on the second exceeds
778 the evidence in the reverse direction. The sign of genetic correlation can be employed to infer
779 the consequence of the partial genetic causality of one trait on another. A GCP greater than
780 zero suggests a partial genetic causal relationship from trait 1 to trait 2 and vice versa, with
781 values closer to $|GCP| = 1$ indicating stronger evidence of vertical pleiotropy. Conversely, a
782 GCP of zero implies horizontal pleiotropy. We considered an absolute GCP estimate ($|GCP|$)
783 greater than 0.60 as indicative of substantial genetic causality, applying a
784 Bonferroni-corrected significance threshold of $P < 8.33 \times 10^{-3}$ to account for multiple
785 comparisons across trait pairs. The LCV analysis, while robust, assumes a singular,
786 unidirectional latent variable driving the genetic correlation, a premise potentially confounded
787 by bidirectional causal effects or multiple latent factors.

788

789 Addressing the limitations inherent to LCV, we also conducted an analysis using Latent
790 Heritable Confounder Mendelian Randomization (LHC-MR)⁵⁴. This advanced Mendelian
791 randomization approach leverages genome-wide variants to explore bi-directional causal
792 associations between complex traits, thus enhancing the capacity to discern bi-directional
793 genetic effects, direct heritabilities, and confounder effects, even amidst sample overlap. By

794 modeling an unobserved heritable confounder affecting both exposure and outcome traits,
795 LHC-MR assures the critical assumption of exchangeability⁵⁵. Its comprehensive modeling of
796 potential SNP effects-direct, indirect, or null-on the traits offers more accurate causal effect
797 estimates compared to traditional MR methods, such as MR Egger, weighted median, inverse
798 variance weighted (IVW), simple mode, and weighted mode⁵⁶⁻⁵⁸. For statistical significance,
799 we adjusted for multiple testing with a threshold of $P < 4.17 \times 10^{-3}$, considering both the
800 number of trait pairs and the number of tests conducted.

801

802 **Pairwise Pleiotropic analysis between LTL and CVDs**

803 To delve into the role of horizontal pleiotropy between LTL and six major CVDs, we utilized
804 the Pleiotropic Analysis under the Composite Null Hypothesis (PLACO) method to conduct a
805 comprehensive genome-wide identification of pleiotropic SNPs that concurrently influence
806 the risk of both traits simultaneously. PLACO operates on the composite null hypothesis that
807 asserts a genetic variant is either associated with just one or neither trait, thereby
808 distinguishing between pleiotropic effects and singular trait associations⁵⁹. The method
809 evaluates this hypothesis by examining the product of the Z statistics derived from the GWAS
810 summary statistics of both traits, formulating a null distribution of the test statistic as a
811 mixture distribution. This allows for the acknowledgment of SNPs that may be linked to only
812 one or none of the phenotypes under study. The threshold for identifying pleiotropic SNPs
813 with significant evidence of genome-wide pleiotropy was set at $P_{PLACO} < 5 \times 10^{-8}$. Additionally,
814 to adjust for possible sample overlap, we de-correlated the Z-scores using a correlation matrix
815 directly estimated from the GWAS summary statistics, ensuring a more accurate interpretation
816 of pleiotropic effects.

817

818 **Characterization of pleiotropic loci and functional annotation**

819 The Functional Mapping and Annotation of Genome-Wide Association Studies (FUMA)
820 platform was employed to identify independent genomic loci and conduct functional
821 annotation for pleiotropic SNPs revealed by PLACO analysis⁵⁹. FUMA, leveraging data from
822 18 biological databases and analytical tools, annotates GWAS findings to highlight probable
823 causal genes through positional and eQTL mapping⁶⁰. The 1000 Genomes Project Phase 3

824 European-based LD reference panels were utilized for LD structure correction. Initially,
825 FUMA distinguishes independent significant SNPs (meeting genome-wide significance at $P <$
826 5×10^{-8} and $r^2 < 0.6$), further defining a subset as lead SNPs based on mutual independence (r^2
827 < 0.1). LD blocks within 500 kb of lead SNPs are merged to delineate distinct genomic loci,
828 with the SNP exhibiting the lowest P -value in each locus designated as the top lead SNP. The
829 analysis then assesses directional effects between LTL and six major CVDs by comparing
830 Z-scores of these top lead SNPs. SNPs achieving genome-wide significance ($P < 5 \times 10^{-8}$) in
831 individual GWAS for each trait were annotated using FUMA for comparative analysis. The
832 identified pleiotropic loci were considered novel if they did not coincide with the loci
833 previously reported in the original GWAS for LTL or any of the six major CVDs. In other
834 words, to be deemed 'novel,' a locus identified through FUMA should not have exhibited
835 statistical significance in the single trait GWAS.

836

837 To elucidate the biological underpinnings of the observed statistical associations, lead SNPs
838 were annotated using Annotate Variation (ANNOVAR)⁶¹ for their proximity to genes and
839 potential impact on gene function. The Combined Annotation-Dependent Depletion (CADD)
840 score⁶², which aggregates insights from 67 annotation resources, was used to assess the
841 deleteriousness of variants. Variants with CADD scores greater than 12.37 were deemed
842 likely to exert deleterious effects. Furthermore, the RegulomeDB score provided a categorical
843 assessment of an SNP's regulatory potential based on expression quantitative trait loci (eQTL)
844 and chromatin marks, ranging from 1 (strong evidence of regulatory functionality) to 7
845 (minimal evidence)⁶³. The highest regulatory potential is indicated by a score of 1a, while a
846 score of 7 suggests the least regulatory significance. Chromatin states, determined by
847 ChromHMM using data from 127 epigenomes and five chromatin marks, revealed the
848 genomic regions' accessibility, categorized into 15 states. For the identification of putative
849 causal genes, SNPs were mapped using two approaches: positional mapping within a 10-kb
850 window around the SNP and eQTL mapping.

851

852 **Colocalization analysis**

853 For pleiotropic loci identified and annotated by FUMA, we conducted a colocalization

854 analysis using COLOC to pinpoint potential shared causal variants across pairwise traits
855 within each locus. COLOC evaluates five mutually exclusive hypotheses for each pair of
856 traits at a locus⁶⁴: H0 posits no association with either trait; H1 and H2 suggest an association
857 with only one of the traits; H3 indicates that both traits are associated due to different causal
858 variants; and H4 implies a shared association for both traits stemming from the same causal
859 variant. The analysis was performed using default COLOC prior probabilities: p_1 and p_2 ,
860 each set at 1×10^{-4} for an SNP's association with the first and second trait, respectively, and
861 p_{12} at 1×10^{-5} for an SNP associated with both traits. A Posterior Probability for Hypothesis 4
862 (PP.H4) greater than 0.7 was considered strong evidence for colocalization, suggesting the
863 presence of shared causal variants at the locus. The SNP exhibiting the highest PP.H4 was
864 identified as a candidate causal variant.

865

866 **Gene level analyses**

867 Building on the insights from PLACO, we delved into the shared biological processes and
868 pathways involving the identified pleiotropic loci. Through gene-level analysis using
869 Multi-marker Analysis of GenoMic Annotation (MAGMA), we assessed genes within or
870 intersecting the pleiotropic loci, integrating data from both PLACO and single-trait GWAS.
871 Unlike permutation-based approaches, MAGMA employs a multiple regression model that
872 incorporates principal component analysis to evaluate gene associations. This model
873 calculates a p-value for each gene, aggregating the impact of all SNPs linked to that gene
874 while considering gene size, SNP count per gene, and linkage disequilibrium (LD) among the
875 markers. SNPs were attributed to genes based on their location within the gene body or within
876 a 10 kb range upstream or downstream. The LD calculations leveraged the 1000 Genomes
877 Project Phase 3 European population as the reference panel, with SNP locations determined
878 using the human genome Build 37 (GRCh37/hg19) and focusing on 17,636 autosomal
879 protein-coding genes. A gene was deemed significant if its p-value was below 0.05 after
880 applying a Bonferroni correction for the total number of protein-coding genes and the six trait
881 pairs analyzed ($P = 0.05 / 17,636 / 6 = 4.73 \times 10^{-7}$). Due to complex LD patterns, the MHC
882 region (chr6: 25-35 Mb) was excluded from MAGMA's gene-based analysis.

883

884 To overcome the limitations of MAGMA, which assigns SNPs to their nearest genes based on
885 arbitrary genomic windows potentially missing functional gene associations due to long-range
886 regulatory effects, we employed eQTL-informed MAGMA (e-MAGMA) for a more nuanced
887 investigation of tissue-specific gene involvement based on PLACO results. e-MAGMA
888 retains the statistical framework of MAGMA, using a multiple linear principal component
889 regression model, but enhances gene-based association analysis by incorporating
890 tissue-specific cis-eQTL information for SNP assignment to genes, which yields more
891 biologically relevant and interpretable findings. For our analysis, we utilized eQTL data from
892 47 tissues provided by the GTEx v8 reference panel, as available on the e-MAGMA website.
893 Guided by the principle that analyses focused on disease-relevant tissues yield more pertinent
894 insights, we selected ten relevant tissues for our study. These included three artery tissues,
895 two adipose tissues, two heart tissues, and three additional tissues (whole blood, liver, and
896 EBV-transformed lymphocytes), chosen based on their significant enrichment in the
897 LDSC-SEG analysis. The LD reference data for our analysis came from the 1000 Genomes
898 Phase 3 European panel. We calculated tissue-specific p-values for each gene across the
899 selected tissues, with significance determined post-Bonferroni correction for the number of
900 tissue-specific protein-coding genes and trait pairs examined. For instance, the significance
901 threshold for adipose subcutaneous tissue was set at $P = 0.05 / 9,613 / 6 = 8.67 \times 10^{-7}$. Similar
902 to MAGMA, e-MAGMA analysis results within the MHC region (chr6: 25-35 Mb) were
903 excluded to avoid confounding due to complex LD patterns. Additionally, we conducted a
904 transcriptome-wide association study (TWAS) based on single-trait GWAS results using the
905 functional summary-based imputation software, FUSION, applying tissue-specific Bonferroni
906 corrections to determine significance. The FUSION approach integrates GWAS summary
907 statistics with pre-computed gene expression weights, referencing the same tissues analyzed
908 in the e-MAGMA study from the GTEx v8 dataset. This integration takes into account the LD
909 structures to identify significant relationships between gene expression levels and specific
910 traits.

911

912 **Pathway level analyses**

913 To investigate the genetic pathways underlying the comorbidity of LTL and six major CVDs,

914 we utilized MAGMA for gene-set analysis. This analysis employs a competitive approach,
915 where test statistics for all genes within a gene set, such as a biological pathway, are
916 aggregated to derive a joint association statistic. Gene sets were sourced from the Gene
917 Ontology biological processes (GO_BP) via the Molecular Signatures Database (MsigDB),
918 with gene definitions and association signals derived from MAGMA gene-based analysis. We
919 adjusted for multiple testing using a Bonferroni correction, setting the threshold at $P = 0.05 /$
920 $7,744 / 6 = 1.08 \times 10^{-6}$. We then conducted functional enrichment analysis on the genes
921 overlapping across more than one trait pair, which were significantly identified by both
922 MAGMA and e-MAGMA analyses. For this purpose, the ToppGene Functional Annotation
923 tool (ToppFun) was employed to identify significantly represented biological processes and
924 enriched signaling pathways, considering the entire genome as the background. ToppFun
925 performs Functional Enrichment Analysis (FEA) on the specified gene list, leveraging a broad
926 spectrum of data sources, including transcriptomics, proteomics, regulomics, ontologies,
927 phenotypes, pharmacogenomics, and bibliographic data. The list of candidate genes was
928 submitted to the ToppFun tool within the ToppGene Suite, with an FDR < 0.05 established as
929 the threshold for statistical significance.

930

931 **Proteome-wide Mendelian Randomization study analysis**

932 To explore potential common causal factors at the proteomic level, we utilized Summary
933 data-based Mendelian Randomization (SMR) to examine associations between protein
934 abundance and disease phenotype. This analysis leveraged index cis-acting variants
935 (cis-pQTLs) identified in the UK Biobank Pharma Proteomics Project (UKB-PPP), which
936 includes plasma samples from 34,557 European individuals with the measurement of 2,940
937 plasma proteins using the Olink Explore platform. Cis-pQTLs were defined as SNPs located
938 within a 1Mb radius from the transcription start site (TSS) of the gene encoding the protein.
939 Only index cis-pQTLs associated with plasma protein levels at a genome-wide significance
940 threshold ($P < 5 \times 10^{-8}$) were considered for inclusion in the SMR analysis. Summary
941 data-based Mendelian Randomization (SMR) is designed to prioritize genes for which
942 expression levels are potentially causally linked to an outcome trait, utilizing summary
943 statistics within a Mendelian Randomization framework. To differentiate between pleiotropy

944 and linkage (where protein abundance and a phenotype manifestation could be influenced by
945 two separate causal variants in strong linkage disequilibrium with one another), the
946 Heterogeneity in Dependent Instrument (HEIDI) test was employed. A HEIDI test *p*-value
947 below 0.01 signifies the presence of two distinct genetic variants in high linkage
948 disequilibrium, explaining the observed associations. Additionally, to address potential biases
949 from analyzing single SNPs, a multi-SNP approach (multi-SNPs-SMR) was utilized as a
950 sensitivity analysis, enhancing the reliability of the statistical evidence. A *p*-value less than
951 0.05 in the multi-SNPs-SMR analysis was deemed significant. Furthermore, HyPrColoc
952 analysis was employed to ascertain whether the associations identified between proteins and
953 various diseases stemmed from the same causal variant or were due to linkage disequilibrium.
954 A posterior probability of a shared causal variant (PP.H4) greater than 0.7 signifies strong
955 evidence of colocalization between proteins and multiple diseases.

956

957 **Reporting summary**

958 Further information on research design is available in the Nature Portfolio Reporting
959 Summary linked to this article.

960

961 **Data availability**

962 The study used only openly available GWAS summary statistics on leukocyte telomere length
963 and six major cardiovascular diseases that have originally been conducted using human data.
964 GWAS summary statistics on LTL are available at
965 <https://figshare.com/s/caa99dc0f76d62990195>. GWAS summary statistics on AF, HF, and
966 Stroke are available at the GWAS Catalog (GCST90104539, GCST009541, and
967 GCST90104539). GWAS summary statistics on CAD and PAD are publicly available for
968 download at the Cardiovascular Disease Knowledge Portal (CVDKP) website:
969 <https://cvd.hugeamp.org/datasets.html>. GWAS summary statistics on VTE are obtained from
970 the deCODE genetics website: <https://www.decode.com/summarydata/>. Blood-based
971 cis-pQTL from UKB-PPP are obtained from [https://](https://www.synapse.org/#!Synapse:syn51365303)
972 www.synapse.org/#!Synapse:syn51365303, respectively.

973

974 **Code availability:**

975 All software used to conduct the analyses in this paper are freely available online. Software
976 (version, where applicable) and sources are listed below: LDSC (v1.0.1;
977 <https://github.com/bulik/ldsc>), MiXeR (v1.3; <https://github.com/precimed/mixer>), LAVA
978 (v0.1.0; <https://github.com/josefin-werme/LAVA>), LCV ([https://github.com/lukejconnor/](https://github.com/lukejconnor/LCV)
979 LCV); LHC-MR (v0.0.0.9000; <https://github.com/LizaDarrow/lhcMR>), PLACO (v0.1.1;
980 <https://github.com/RayDebashree/PLACO>), FUMA (v1.5.4; <http://fuma.ctglab.nl/>),
981 HyPrColoc(v1.0; <https://github.com/jrs95/hyprcoloc>), MAGMA (v.1.08; [https://ctg.cncr.nl/](https://ctg.cncr.nl/software/magma)
982 software/magma), e-MAGMA (<https://github.com/eskederks/eMAGMA-tutorial>), TWAS
983 (<http://gusevlab.org/projects/fusion/>), SMR (v1.31; [https://yanglab.westlake.edu.cn/](https://yanglab.westlake.edu.cn/software/smr/)
984 software/smr/), COLOC (v5.2.1; <https://github.com/chr1swallace/coloc>), and R (v.4.1.3;
985 <https://www.r-project.org/>).

986

987 **Acknowledgements**

988 This study was supported by the Natural Science Foundation of China Excellent Young
989 Scientists Fund (Overseas) (Grant no. K241141101), Guangdong Basic and Applied Basic
990 Research Foundation for Distinguished Young Scholars (Grant no. 24050000763), Shenzhen
991 Pengcheng Peacock Plan, Shenzhen Basic Research General Projects of Shenzhen Science and
992 Technology Innovation Commission (Grant no. JCYJ20230807093514029) (To Y.F.), National
993 Natural Science Foundation (Grant no. 82170339 and 82270241), NSFC Incubation Project of
994 Guangdong Provincial People's Hospital (Grant no. KY0120220021), Natural Science
995 Foundation of Guangdong Province (Grant no. 2023B1515020082) (To L.J.), National Natural
996 Science Foundation of China (Grant no. 82260073); Tianshan Talent Cultivation Program
997 Project of Xinjiang Uygur Autonomous Region (Grant no. 2022TSYCLJ0028) (To Y.Y.), and
998 Center for Computational Science and Engineering at Southern University of Science and
999 Technology. The funder had no role in the design, implementation, analysis, interpretation of
1000 the data, approval of the manuscript, and decision to submit the manuscript for publication.

1001

1002 **Author contributions**

1003 J.Q., Y. F., Y.Y., L.J., and S.P. conceptualized and supervised this project and wrote the

1004 manuscript. J.Q., Q.W., and Y.Z. performed the main analyses and wrote the manuscript. J.Q.,
1005 M.C., L.C., and F.L. performed the statistical analysis and assisted with interpreting the
1006 results. K.Y., L.Z., N.T., P.H., and A.J. provided expertise in cardiovascular biology and
1007 GWAS summary statistics. All authors discussed the results and commented on the paper.

1008

1009 **Competing interests**

1010 All authors declare no competing interests.

1011 **References**

- 1012 1 Blasco, M. A. Telomeres and human disease: ageing, cancer and beyond. *Nat Rev*
1013 *Genet* **6**, 611-622, doi:10.1038/nrg1656 (2005).
- 1014 2 van der Harst, P. *et al.* Telomere length of circulating leukocytes is decreased in
1015 patients with chronic heart failure. *J Am Coll Cardiol* **49**, 1459-1464,
1016 doi:10.1016/j.jacc.2007.01.027 (2007).
- 1017 3 Zheng, Y. *et al.* Association between leucocyte telomere length and the risk of atrial
1018 fibrillation: An updated systematic review and meta-analysis. *Ageing Res Rev* **81**,
1019 101707, doi:10.1016/j.arr.2022.101707 (2022).
- 1020 4 Haycock, P. C. *et al.* Leucocyte telomere length and risk of cardiovascular disease:
1021 systematic review and meta-analysis. *Bmj* **349**, g4227, doi:10.1136/bmj.g4227
1022 (2014).
- 1023 5 Njajou, O. T. *et al.* Telomere length is paternally inherited and is associated with
1024 parental lifespan. *Proc Natl Acad Sci U S A* **104**, 12135-12139,
1025 doi:10.1073/pnas.0702703104 (2007).
- 1026 6 Broer, L. *et al.* Meta-analysis of telomere length in 19,713 subjects reveals high
1027 heritability, stronger maternal inheritance and a paternal age effect. *Eur J Hum Genet*
1028 **21**, 1163-1168, doi:10.1038/ejhg.2012.303 (2013).
- 1029 7 Codd, V. *et al.* Polygenic basis and biomedical consequences of telomere length
1030 variation. *Nature genetics* **53**, 1425-1433, doi:10.1038/s41588-021-00944-6 (2021).
- 1031 8 Nielsen, J. B. *et al.* Biobank-driven genomic discovery yields new insight into atrial
1032 fibrillation biology. *Nature genetics* **50**, 1234-1239, doi:10.1038/s41588-018-0171-3
1033 (2018).
- 1034 9 Aragam, K. G. *et al.* Discovery and systematic characterization of risk variants and
1035 genes for coronary artery disease in over a million participants. *Nature genetics* **54**,
1036 1803-1815, doi:10.1038/s41588-022-01233-6 (2022).
- 1037 10 Ghouse, J. *et al.* Genome-wide meta-analysis identifies 93 risk loci and enables risk
1038 prediction equivalent to monogenic forms of venous thromboembolism. *Nature*
1039 *genetics* **55**, 399-409, doi:10.1038/s41588-022-01286-7 (2023).
- 1040 11 Shah, S. *et al.* Genome-wide association and Mendelian randomisation analysis

- 1041 provide insights into the pathogenesis of heart failure. *Nature communications* **11**,
1042 163, doi:10.1038/s41467-019-13690-5 (2020).
- 1043 12 van Zuydam, N. R. *et al.* Genome-Wide Association Study of Peripheral Artery
1044 Disease. *Circ Genom Precis Med* **14**, e002862, doi:10.1161/circgen.119.002862
1045 (2021).
- 1046 13 Mishra, A. *et al.* Stroke genetics informs drug discovery and risk prediction across
1047 ancestries. *Nature* **611**, 115-123, doi:10.1038/s41586-022-05165-3 (2022).
- 1048 14 Haycock, P. C. *et al.* Association Between Telomere Length and Risk of Cancer and
1049 Non-Neoplastic Diseases: A Mendelian Randomization Study. *JAMA Oncol* **3**,
1050 636-651, doi:10.1001/jamaoncol.2016.5945 (2017).
- 1051 15 Deng, Y. *et al.* Telomere length and the risk of cardiovascular diseases: A Mendelian
1052 randomization study. *Front Cardiovasc Med* **9**, 1012615,
1053 doi:10.3389/fcvm.2022.1012615 (2022).
- 1054 16 Sha, Z. *et al.* Causal relationship between atrial fibrillation and leukocyte telomere
1055 length: A two sample, bidirectional Mendelian randomization study. *Front*
1056 *Cardiovasc Med* **10**, 1093255, doi:10.3389/fcvm.2023.1093255 (2023).
- 1057 17 (!!! INVALID CITATION !!!).
- 1058 18 Gong, W. *et al.* Role of the Gut-Brain Axis in the Shared Genetic Etiology Between
1059 Gastrointestinal Tract Diseases and Psychiatric Disorders: A Genome-Wide
1060 Pleiotropic Analysis. *JAMA Psychiatry* **80**, 360-370,
1061 doi:10.1001/jamapsychiatry.2022.4974 (2023).
- 1062 19 Pisanu, C. *et al.* Dissecting the genetic overlap between severe mental disorders and
1063 markers of cellular aging: Identification of pleiotropic genes and druggable targets.
1064 *Neuropsychopharmacology*, doi:10.1038/s41386-024-01822-5 (2024).
- 1065 20 Levy, D. *et al.* Genome-wide association study of blood pressure and hypertension.
1066 *Nature genetics* **41**, 677-687, doi:10.1038/ng.384 (2009).
- 1067 21 Soranzo, N. *et al.* A genome-wide meta-analysis identifies 22 loci associated with
1068 eight hematological parameters in the HaemGen consortium. *Nature genetics* **41**,
1069 1182-1190, doi:10.1038/ng.467 (2009).
- 1070 22 Chen, X., Ma, J., Wang, Z. W. & Wang, Z. The E3 ubiquitin ligases regulate

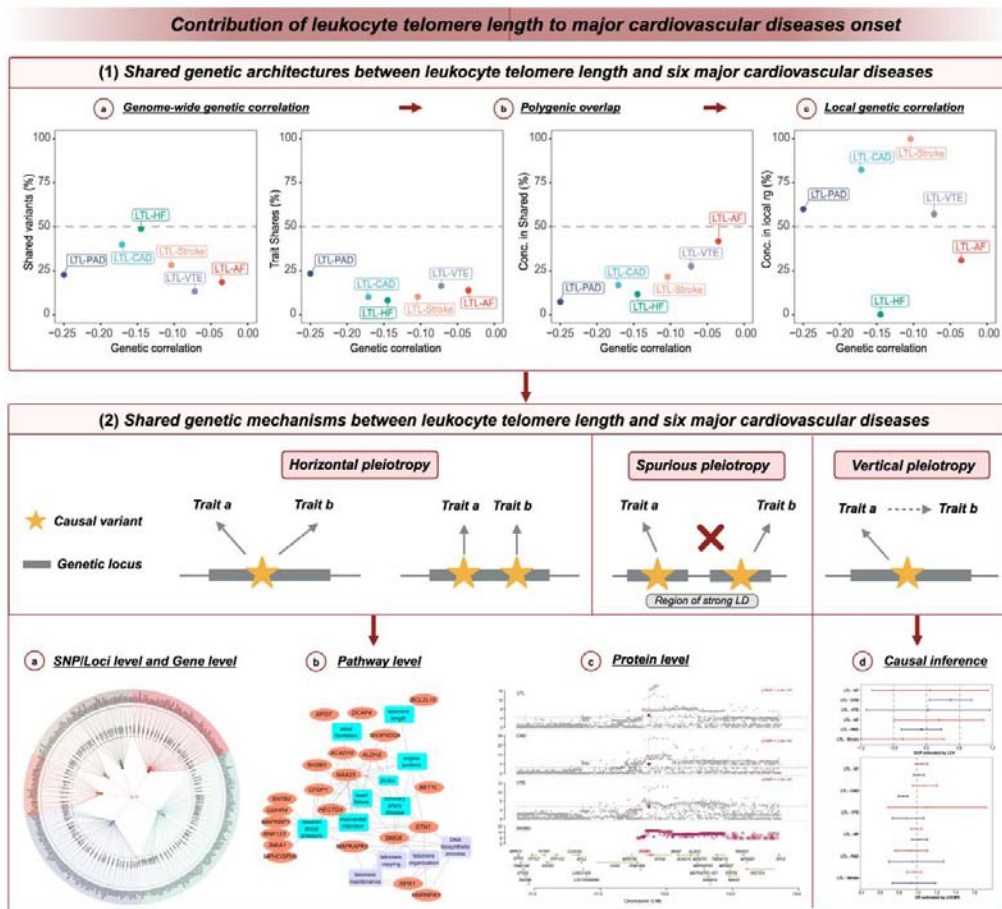
- 1071 inflammation in cardiovascular diseases. *Semin Cell Dev Biol* **154**, 167-174,
1072 doi:10.1016/j.semcdb.2023.02.008 (2024).
- 1073 23 Choi, J. & Baek, K. H. Cellular functions of stem cell factors mediated by the
1074 ubiquitin-proteasome system. *Cell Mol Life Sci* **75**, 1947-1957,
1075 doi:10.1007/s00018-018-2770-7 (2018).
- 1076 24 Brænne, I. *et al.* Prediction of Causal Candidate Genes in Coronary Artery Disease
1077 Loci. *Arterioscler Thromb Vasc Biol* **35**, 2207-2217, doi:10.1161/atvbaha.115.306108
1078 (2015).
- 1079 25 He, M. *et al.* Identification and characterization of new long chain acyl-CoA
1080 dehydrogenases. *Mol Genet Metab* **102**, 418-429, doi:10.1016/j.ymgme.2010.12.005
1081 (2011).
- 1082 26 Joshi, P. K. *et al.* Genome-wide meta-analysis associates HLA-DQA1/DRB1 and LPA
1083 and lifestyle factors with human longevity. *Nature communications* **8**, 910,
1084 doi:10.1038/s41467-017-00934-5 (2017).
- 1085 27 Wang, X. *et al.* Transmembrane emp24 protein transport domain 6 is selectively
1086 expressed in pancreatic islets and implicated in insulin secretion and diabetes.
1087 *Pancreas* **41**, 10-14, doi:10.1097/MPA.0b013e318223c7e4 (2012).
- 1088 28 Fadista, J. *et al.* Global genomic and transcriptomic analysis of human pancreatic
1089 islets reveals novel genes influencing glucose metabolism. *Proc Natl Acad Sci U S A*
1090 **111**, 13924-13929, doi:10.1073/pnas.1402665111 (2014).
- 1091 29 Aksu, M. *et al.* Xpo7 is a broad-spectrum exportin and a nuclear import receptor. *J*
1092 *Cell Biol* **217**, 2329-2340, doi:10.1083/jcb.201712013 (2018).
- 1093 30 Innes, A. J. *et al.* XPO7 is a tumor suppressor regulating p21(CIP1)-dependent
1094 senescence. *Genes Dev* **35**, 379-391, doi:10.1101/gad.343269.120 (2021).
- 1095 31 Borg, M. L. *et al.* Pigment epithelium-derived factor regulates lipid metabolism via
1096 adipose triglyceride lipase. *Diabetes* **60**, 1458-1466 (2011).
- 1097 32 Yamagishi, S.-i. & Matsui, T. Pigment epithelium-derived factor (PEDF) and
1098 cardiometabolic disorders. *Current pharmaceutical design* **20**, 2377-2386 (2014).
- 1099 33 Yamagishi, S. & Matsui, T. Pigment epithelium-derived factor (PEDF) and
1100 cardiometabolic disorders. *Curr Pharm Des* **20**, 2377-2386,

- 1101 doi:10.2174/13816128113199990473 (2014).
- 1102 34 Ma, S., Wang, S., Li, M., Zhang, Y. & Zhu, P. The effects of pigment
1103 epithelium-derived factor on atherosclerosis: putative mechanisms of the process.
1104 *Lipids Health Dis* **17**, 240, doi:10.1186/s12944-018-0889-z (2018).
- 1105 35 Zhang, J. *et al.* The role of aldehyde dehydrogenase 2 in cardiovascular disease. *Nat*
1106 *Rev Cardiol* **20**, 495-509, doi:10.1038/s41569-023-00839-5 (2023).
- 1107 36 Chen, C. H., Sun, L. & Mochly-Rosen, D. Mitochondrial aldehyde dehydrogenase
1108 and cardiac diseases. *Cardiovasc Res* **88**, 51-57, doi:10.1093/cvr/cvq192 (2010).
- 1109 37 Lee, J. Y. *et al.* A genome-wide association study of a coronary artery disease risk
1110 variant. *J Hum Genet* **58**, 120-126, doi:10.1038/jhg.2012.124 (2013).
- 1111 38 Islam, T., Rahman, M. R., Khan, A. & Moni, M. A. Integration of Mendelian
1112 randomisation and systems biology models to identify novel blood-based biomarkers
1113 for stroke. *J Biomed Inform* **141**, 104345, doi:10.1016/j.jbi.2023.104345 (2023).
- 1114 39 Kuo, C. L. *et al.* The Longevity-Associated SH2B3 (LNK) Genetic Variant: Selected
1115 Aging Phenotypes in 379,758 Subjects. *J Gerontol A Biol Sci Med Sci* **75**, 1656-1662,
1116 doi:10.1093/gerona/glz191 (2020).
- 1117 40 Goswami, S. K., Ranjan, P., Dutta, R. K. & Verma, S. K. Management of
1118 inflammation in cardiovascular diseases. *Pharmacol Res* **173**, 105912,
1119 doi:10.1016/j.phrs.2021.105912 (2021).
- 1120 41 Nevers, T. *et al.* Th1 effector T cells selectively orchestrate cardiac fibrosis in
1121 nonischemic heart failure. *J Exp Med* **214**, 3311-3329, doi:10.1084/jem.20161791
1122 (2017).
- 1123 42 Flister, M. J. *et al.* SH2B3 Is a Genetic Determinant of Cardiac Inflammation and
1124 Fibrosis. *Circ Cardiovasc Genet* **8**, 294-304, doi:10.1161/circgenetics.114.000527
1125 (2015).
- 1126 43 Zhu, X. *et al.* Exacerbating Pressure Overload-Induced Cardiac Hypertrophy: Novel
1127 Role of Adaptor Molecule Src Homology 2-B3. *Hypertension* **66**, 571-581,
1128 doi:10.1161/hypertensionaha.115.05183 (2015).
- 1129 44 Nassour, J., Schmidt, T. T. & Karlseder, J. Telomeres and Cancer: Resolving the
1130 Paradox. *Annu Rev Cancer Biol* **5**, 59-77,

- 1131 doi:10.1146/annurev-cancerbio-050420-023410 (2021).
- 1132 45 Zhang, S. *et al.* TMEM116 is required for lung cancer cell motility and metastasis
1133 through PDK1 signaling pathway. *Cell Death Dis* **12**, 1086,
1134 doi:10.1038/s41419-021-04369-1 (2021).
- 1135 46 Sinner, M. F. *et al.* Integrating genetic, transcriptional, and functional analyses to
1136 identify 5 novel genes for atrial fibrillation. *Circulation* **130**, 1225-1235,
1137 doi:10.1161/circulationaha.114.009892 (2014).
- 1138 47 Fuster, J. J. & Andrés, V. Telomere biology and cardiovascular disease. *Circ Res* **99**,
1139 1167-1180, doi:10.1161/01.Res.0000251281.00845.18 (2006).
- 1140 48 Minamino, T. *et al.* Endothelial cell senescence in human atherosclerosis: role of
1141 telomere in endothelial dysfunction. *Circulation* **105**, 1541-1544,
1142 doi:10.1161/01.cir.0000013836.85741.17 (2002).
- 1143 49 Farzaneh-Far, R. *et al.* Telomere length trajectory and its determinants in persons with
1144 coronary artery disease: longitudinal findings from the heart and soul study. *PLoS*
1145 *One* **5**, e8612, doi:10.1371/journal.pone.0008612 (2010).
- 1146 50 Bulik-Sullivan, B. K. *et al.* LD Score regression distinguishes confounding from
1147 polygenicity in genome-wide association studies. *Nature genetics* **47**, 291-295,
1148 doi:10.1038/ng.3211 (2015).
- 1149 51 Bulik-Sullivan, B. *et al.* An atlas of genetic correlations across human diseases and
1150 traits. *Nature genetics* **47**, 1236-1241, doi:10.1038/ng.3406 (2015).
- 1151 52 Frei, O. *et al.* Bivariate causal mixture model quantifies polygenic overlap between
1152 complex traits beyond genetic correlation. *Nature communications* **10**, 2417,
1153 doi:10.1038/s41467-019-10310-0 (2019).
- 1154 53 O'Connor, L. J. & Price, A. L. Distinguishing genetic correlation from causation
1155 across 52 diseases and complex traits. *Nature genetics* **50**, 1728-1734,
1156 doi:10.1038/s41588-018-0255-0 (2018).
- 1157 54 Darrous, L., Mounier, N. & Kutalik, Z. Simultaneous estimation of bi-directional
1158 causal effects and heritable confounding from GWAS summary statistics. *Nature*
1159 *communications* **12**, 7274, doi:10.1038/s41467-021-26970-w (2021).
- 1160 55 Sanderson, E. *et al.* Mendelian randomization. *Nat Rev Methods Primers* **2**,

- 1161 doi:10.1038/s43586-021-00092-5 (2022).
- 1162 56 Burgess, S., Scott, R. A., Timpson, N. J., Davey Smith, G. & Thompson, S. G. Using
1163 published data in Mendelian randomization: a blueprint for efficient identification of
1164 causal risk factors. *Eur J Epidemiol* **30**, 543-552, doi:10.1007/s10654-015-0011-z
1165 (2015).
- 1166 57 Bowden, J., Davey Smith, G. & Burgess, S. Mendelian randomization with invalid
1167 instruments: effect estimation and bias detection through Egger regression. *Int J*
1168 *Epidemiol* **44**, 512-525, doi:10.1093/ije/dyv080 (2015).
- 1169 58 Hartwig, F. P., Davey Smith, G. & Bowden, J. Robust inference in summary data
1170 Mendelian randomization via the zero modal pleiotropy assumption. *Int J Epidemiol*
1171 **46**, 1985-1998, doi:10.1093/ije/dyx102 (2017).
- 1172 59 Ray, D. & Chatterjee, N. A powerful method for pleiotropic analysis under composite
1173 null hypothesis identifies novel shared loci between Type 2 Diabetes and Prostate
1174 Cancer. *PLoS genetics* **16**, e1009218, doi:10.1371/journal.pgen.1009218 (2020).
- 1175 60 Watanabe, K., Taskesen, E., van Bochoven, A. & Posthuma, D. Functional mapping
1176 and annotation of genetic associations with FUMA. *Nature communications* **8**, 1826,
1177 doi:10.1038/s41467-017-01261-5 (2017).
- 1178 61 Wang, K., Li, M. & Hakonarson, H. ANNOVAR: functional annotation of genetic
1179 variants from high-throughput sequencing data. *Nucleic Acids Res* **38**, e164,
1180 doi:10.1093/nar/gkq603 (2010).
- 1181 62 Kircher, M. *et al.* A general framework for estimating the relative pathogenicity of
1182 human genetic variants. *Nature genetics* **46**, 310-315, doi:10.1038/ng.2892 (2014).
- 1183 63 Boyle, A. P. *et al.* Annotation of functional variation in personal genomes using
1184 RegulomeDB. *Genome Res* **22**, 1790-1797, doi:10.1101/gr.137323.112 (2012).
- 1185 64 Giambartolomei, C. *et al.* Bayesian test for colocalisation between pairs of genetic
1186 association studies using summary statistics. *PLoS genetics* **10**, e1004383,
1187 doi:10.1371/journal.pgen.1004383 (2014).
- 1188

Figure1



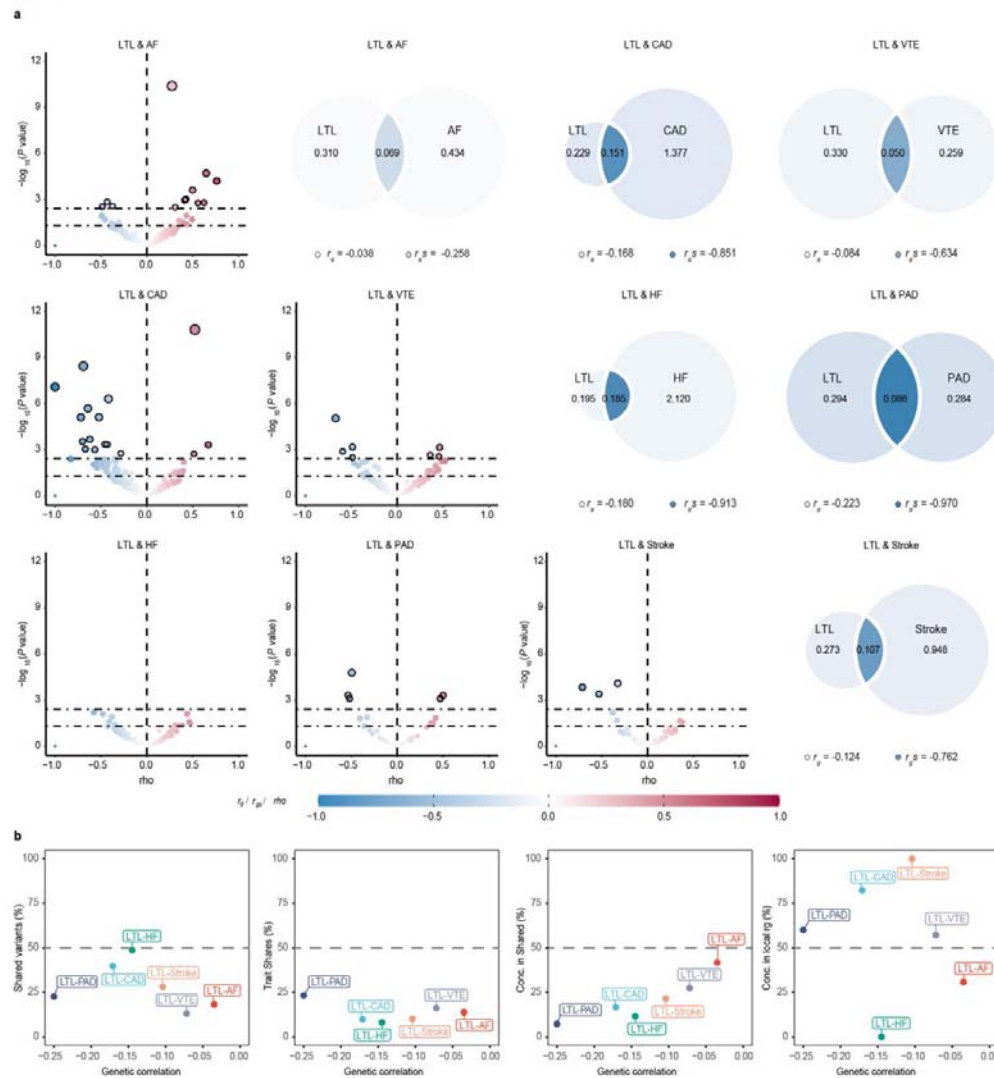
1189

1190 **Figure 1: Schematic representation of analyses performed for leukocyte telomere length**
 1191 **and six major cardiovascular diseases in the current study.**

1192 This figure demonstrates the comprehensive pleiotropic analysis conducted for LTL and six
 1193 major CVDs from multiple perspectives within this study. We first investigated the shared
 1194 genetic architectures between LTL and six major CVDs by assessing pairwise genetic overlap
 1195 beyond correlation. Extensive analyses were then conducted to investigate two types of
 1196 pleiotropy: horizontal pleiotropy, whereby causal variants for two traits colocalize in the same
 1197 locus, and vertical pleiotropy, whereby a variant exerts an effect on one trait through another.
 1198 Notably, spurious pleiotropy was excluded from the analysis, whereby causal variants for two
 1199 traits fall into distinct loci but are in LD with a variant associated with both traits. Therefore,
 1200 We applied Mendelian randomization to evaluate the pairwise causal associations between
 1201 LTL and the CVDs, primarily elucidating the contributions from vertical pleiotropy. We used
 1202 novel statistical tools to capture horizontal pleiotropy by characterizing the shared loci and

1203 their implications on genes, tissues, biological functions, and protein targets. This
1204 comprehensive pleiotropic analysis allowed us to construct an atlas of the shared genetic
1205 associations, enhancing our understanding of the complex interactions between LTL and
1206 cardiovascular health. The diagram was generated using BioRender (www.biorender.com) and
1207 has been included with permission for publication. LTL, leukocyte telomere length; AF, Atrial
1208 fibrillation; CAD, Coronary artery disease; VTE, Venous thromboembolism; HF, Heart failure;
1209 PAD, Peripheral artery disease.

Figure 2



1210

1211 **Figure 2: Genetic overlap between leukocyte telomere length and six major**
 1212 **cardiovascular diseases beyond genome-wide genetic correlation.**

1213 (a) MiXeR-modeled genome-wide genetic overlap and genetic correlations (top right) and

1214 LAVA local correlations (bottom left) between LTL and six major CVDs. Top right: MiXeR

1215 Venn diagrams showing the number (in thousands) of estimated 'causal' variants that are

1216 unique to LTL (left circle), six major CVDs (non-overlapping part of the right circle) or

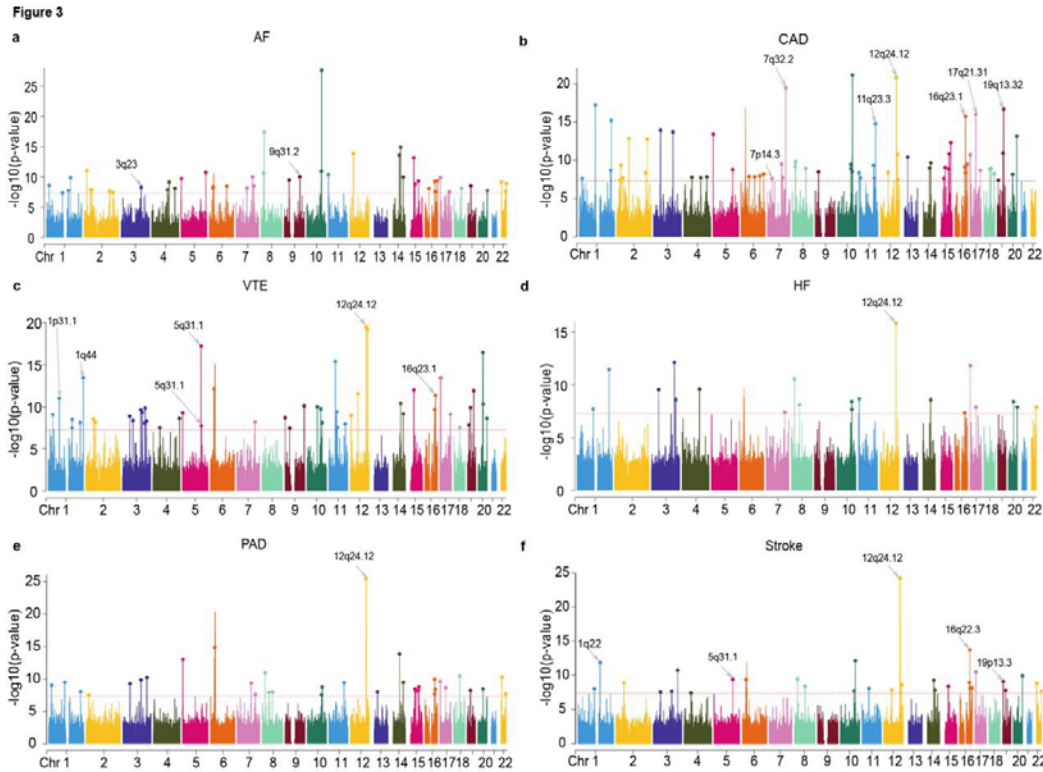
1217 shared between LTL and six major CVDs (overlapping part of circles). Genome-wide genetic

1218 correlation (r_g) and genetic correlation of shared variants (r_{gs}) are represented by the color of the

1219 trait-specific (r_g) and shared regions (r_{gs}), respectively. The circle size represents the

1220 extent of polygenicity of each trait, with larger circles corresponding to greater polygenicity

1221 and vice versa. Bottom left: Volcano plots of LAVA local genetic correlation coefficients (ρ ,
1222 y-axis) against $-\log_{10}$ (p-values) for each pairwise analysis per locus. Larger dots with black
1223 circles represent significantly correlated loci after FDR correction ($FDR < 0.05$). MiXeR
1224 estimated r_g and r_{gS} , and LAVA estimated ρ are represented on the same blue to red color
1225 scale. Note that the volcano plots were plotted at p-values truncated by 1×10^{-F12} for better
1226 visualization, thus excluding a region (LD block 1,581 on chromosome 10, ranges from
1227 104,206,838 to 106,142,283) influencing LTL - CAD from volcano plots of LAVA. (b)
1228 Genetic correlation estimated by LDSC (x-axis) against the percentage of LTL variants that
1229 are shared with CVDs as estimated by MiXeR (first plot), the percentage of CVD variants
1230 that are shared with LTL (second plot), and the percentage of CVD variants that are shared
1231 with LTL that have concordant effect directions (third plot). The fourth plot shows the
1232 percentage of local genetic correlations from LAVA with concordant effect directions on the
1233 y-axis. LTL, leukocyte telomere length; AF, Atrial fibrillation; CAD, Coronary artery disease;
1234 VTE, Venous thromboembolism; HF, Heart failure; PAD, Peripheral artery disease.

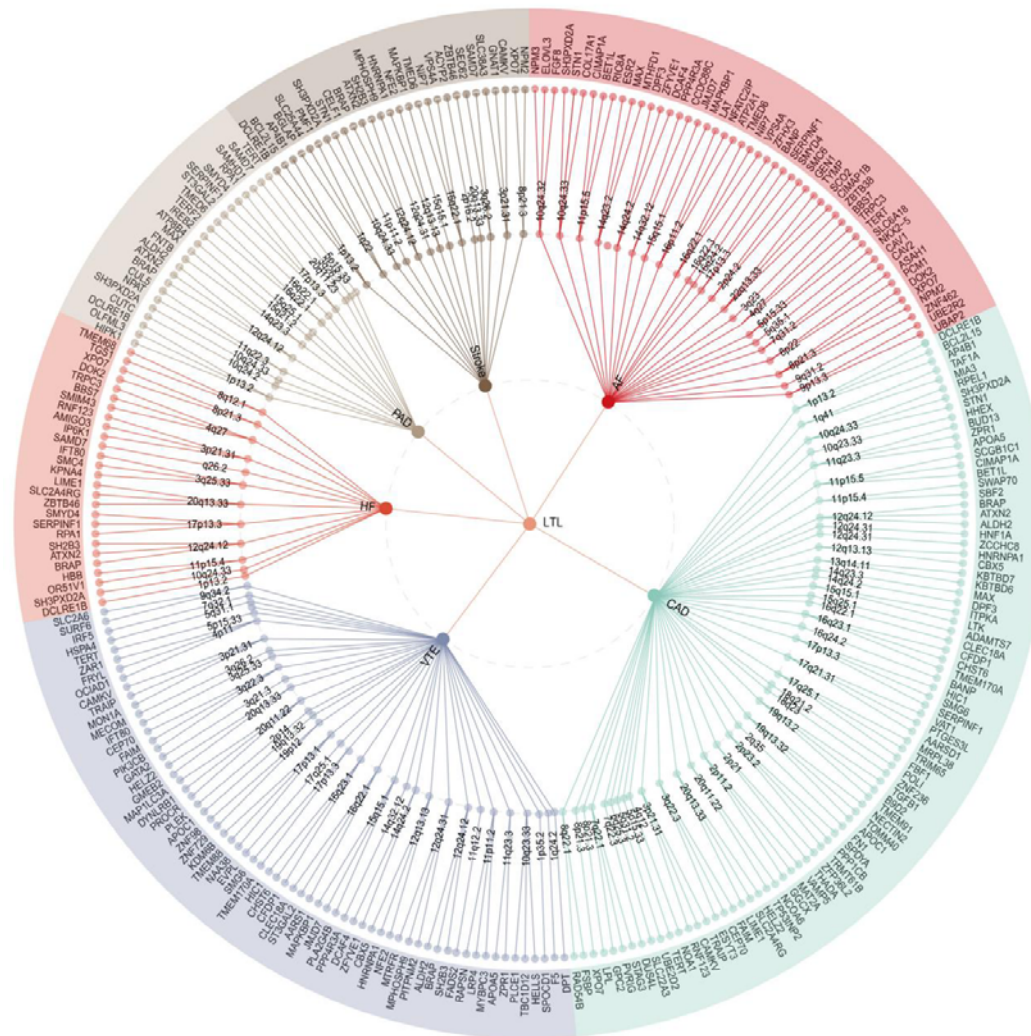


1235

1236 **Figure 3: Manhattan plots for the PLACO results of leukocyte telomere length and six**
1237 **major cardiovascular diseases.**

1238 The x-axis reflects the chromosomal position, and the y-axis reflects negative log10
1239 transformed *P*-values for each SNP. The horizontal dashed red line indicates the genome-wide
1240 significant *P*-value of $-\log_{10}(5 \times 10^{-8})$. The independent genome-wide significant associations
1241 with the smallest *P*-value (Top lead SNP) are encircled in a colorful circle. Only SNPs shared
1242 across all summary statistics were included. Labels are the chromosome regions where
1243 genomic risk loci with strong evidence for colocalization (PP.H4 > 0.7) are located. LTL,
1244 leukocyte telomere length; AF, Atrial fibrillation; CAD, Coronary artery disease; VTE,
1245 Venous thromboembolism; HF, Heart failure; PAD, Peripheral artery disease.

Figure 4

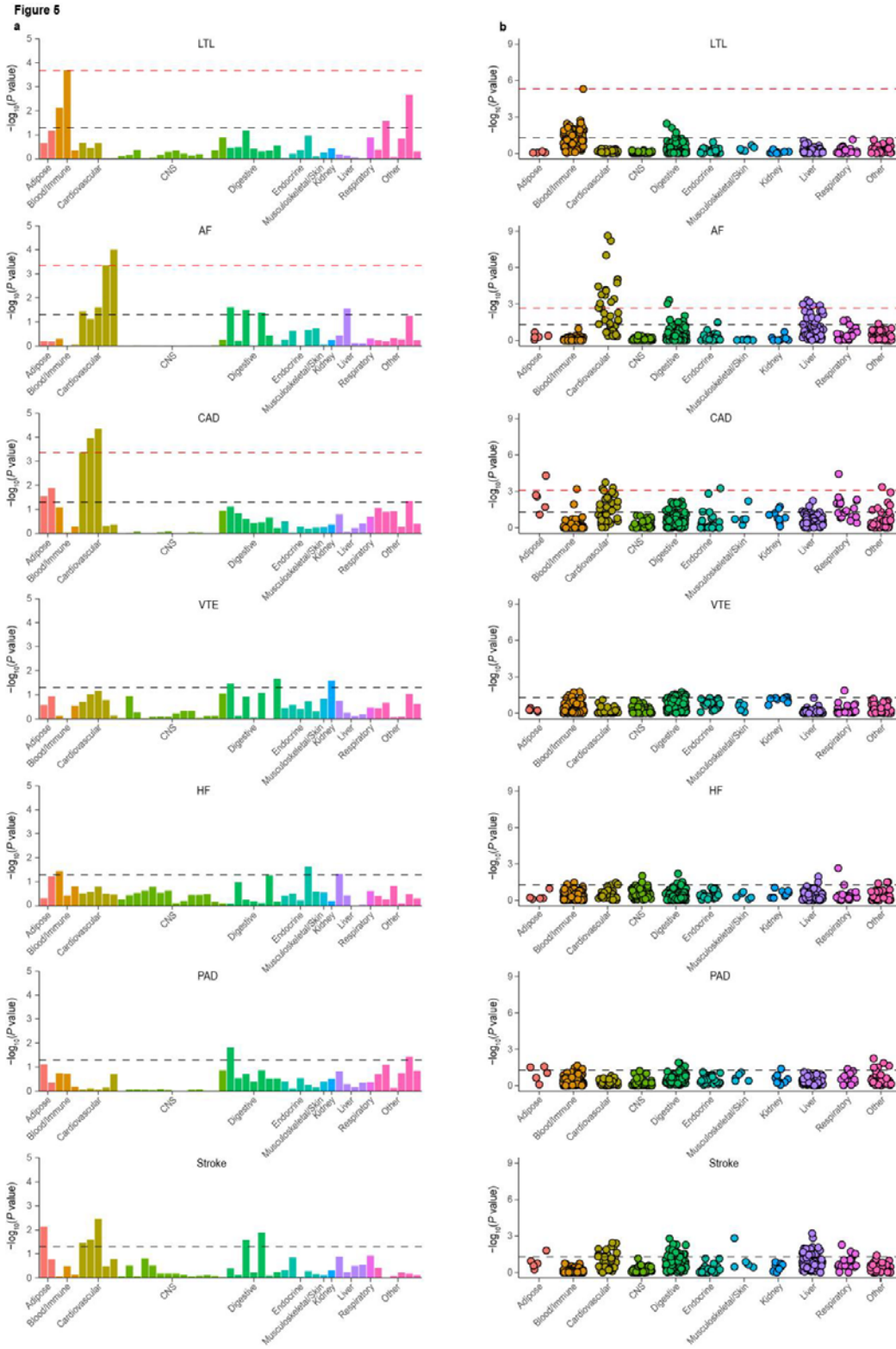


1246

1247 **Figure 4: The Overall landscape of the pleiotropic associations across leukocyte telomere**
1248 **length and six msjor cardiovascular diseases.**

1249 A circular dendrogram showing the shared genes between LTL (center circle) and each of six
1250 CVDs (first circle), resulting in six pairs. A total of 248 shared loci were identified across six
1251 trait pairs, mapped to 478 significant pleiotropic genes (323 unique) identified by multimarker
1252 analysis of GenoMic annotation (MAGMA). For the trait pairs with more than three
1253 pleiotropic genes, we only showed the top 3 pleiotropic genes according to the prioritization
1254 of candidate pleiotropic genes (fourth circle). LTL, leukocyte telomere length; AF, Atrial
1255 fibrillation; CAD, Coronary artery disease; VTE, Venous thromboembolism; HF, Heart failure;
1256 PAD, Peripheral artery disease.

1257



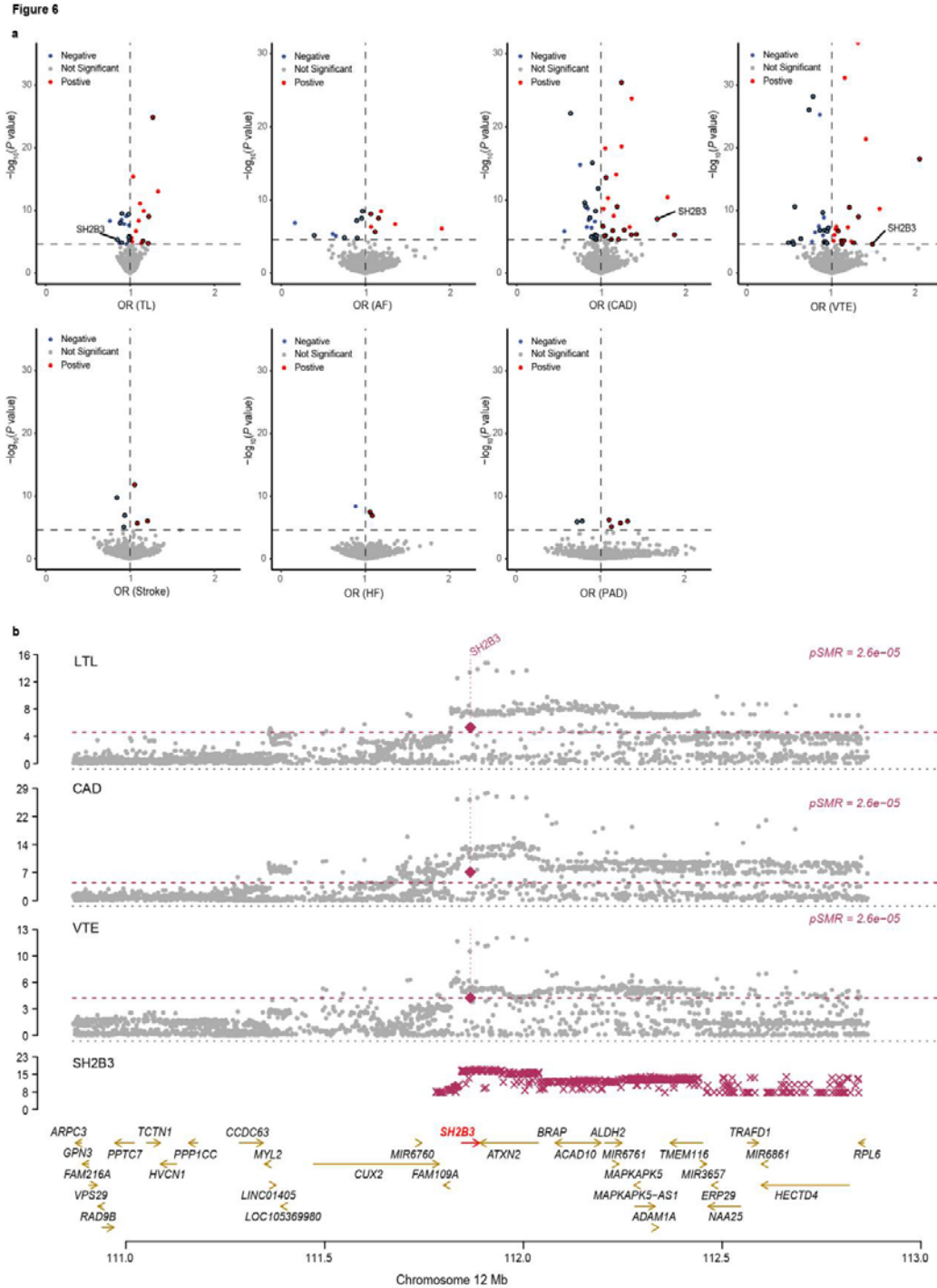
1258

1259

1260

Figure 5: The results of multiple-tissue analysis using gene expression data and chromatin data for leukocyte telomere length and six cardiovascular diseases.

1261 (a) Tissue type-specific enrichment of single nucleotide polymorphism (SNP) heritability for
1262 LTL and CVDs in 49 tissues from GTEx v8 estimated using stratified LDSC applied to
1263 specifically expressed genes (LDSC-SEG). Each bar represents a tissue from the GTEx
1264 dataset. The x-axis reflects tissue types, and the y-axis reflects negative log₁₀ transformed
1265 P-values. (b) Each point represents a peak for DNase I hypersensitivity site (DHS) or histone
1266 marks (including H3K27ac, H3K36me3, H3K4me1, H3K4me3, and H3K9ac) in a tissue or
1267 cell type. Tissues or cell types were classified into 11 distinct categories, i.e., 'Adipose,'
1268 'Blood/Immune,' 'Cardiovascular,' 'CNS,' 'Digestive,' 'Endocrine,' 'Musculoskeletal/Skin,'
1269 'kidney,' 'Liver,' 'Respiratory,' and 'Other.' The black dotted line represents the significance
1270 threshold of $P < 0.05$, and the red line indicates the significant P-value after conducting FDR
1271 correction ($FDR < 0.05$). The color of the dots in a and b indicates different tissue types. LTL,
1272 leukocyte telomere length; AF, Atrial fibrillation; CAD, Coronary artery disease; VTE,
1273 Venous thromboembolism; HF, Heart failure; PAD, Peripheral artery disease.



1274

1275 **Figure 6: Result summary of Proteome-wide MR analysis on the associations between**
 1276 **leukocyte telomere length and six cardiovascular diseases.**

1277 **(a)** Volcano plots based on SMR showing circulating proteins (red, blue, or gray dots) with
 1278 the associations between circulating protein levels and each of the traits (x-axis) and the
 1279 corresponding P-value (y-axis). Red dots indicate positive causal relations. Blue dots indicate

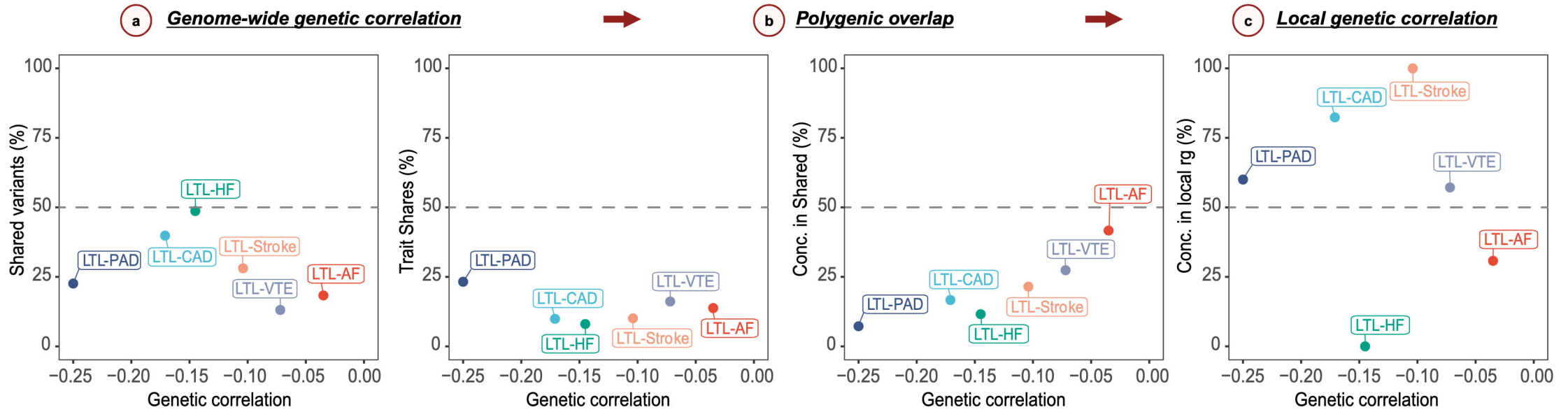
1280 negative causal relations. Gray dots indicate insignificant causal relations. Protein-trait
1281 associations passing Bonferroni correction ($P < 2.60 \times 10^{-5}$) are outlined in black. **(b)** Miami
1282 plots of the 12q24.12 region. Genes that were mapped under the locus at 12q24.12 were
1283 highlighted in the analysis. The protein SH2B3 is labeled as significantly associated with LTL,
1284 CAD, and VTE. The red horizontal dashed line corresponds to Bonferroni correction ($P <$
1285 2.60×10^{-5}). LTL, leukocyte telomere length; AF, Atrial fibrillation; CAD, Coronary artery
1286 disease; VTE, Venous thromboembolism; HF, Heart failure; PAD, Peripheral artery disease.

Figure 1

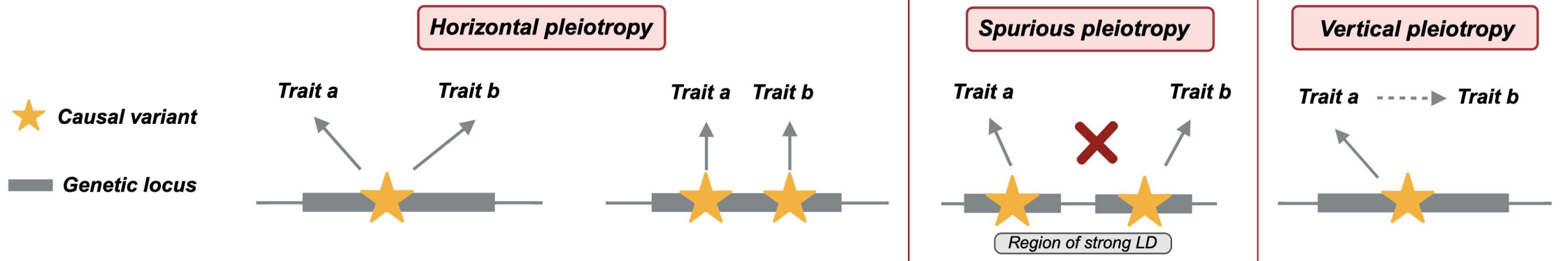
medRxiv preprint doi: <https://doi.org/10.1101/2024.05.20.24307614>; this version posted May 23, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Contribution of leukocyte telomere length to major cardiovascular diseases onset

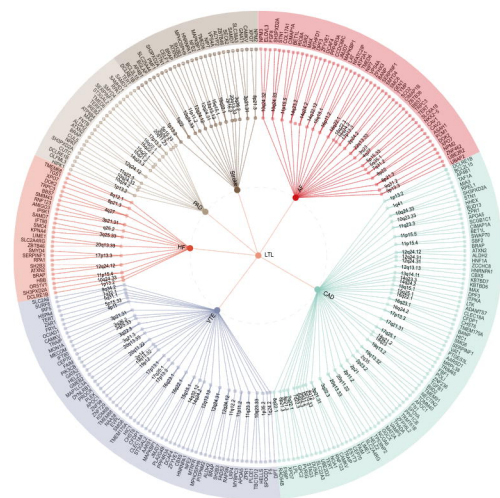
(1) Shared genetic architectures between leukocyte telomere length and six major cardiovascular diseases



(2) Shared genetic mechanisms between leukocyte telomere length and six major cardiovascular diseases



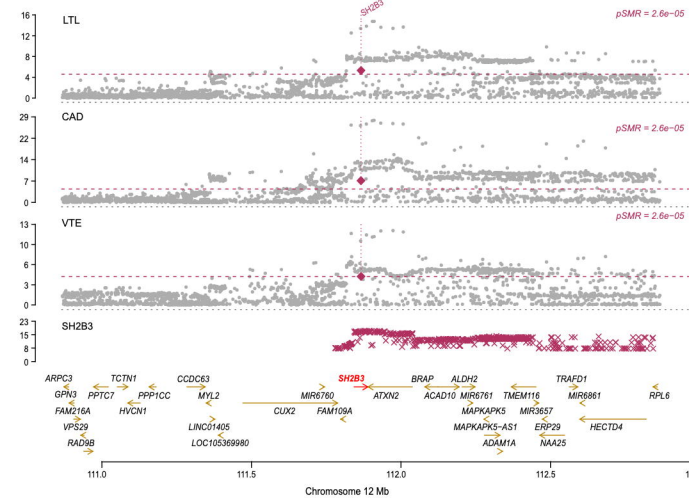
a SNP/Loci level and Gene level



b Pathway level



c Protein level



d Causal inference

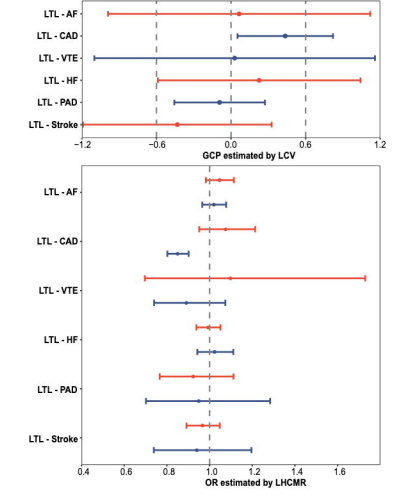


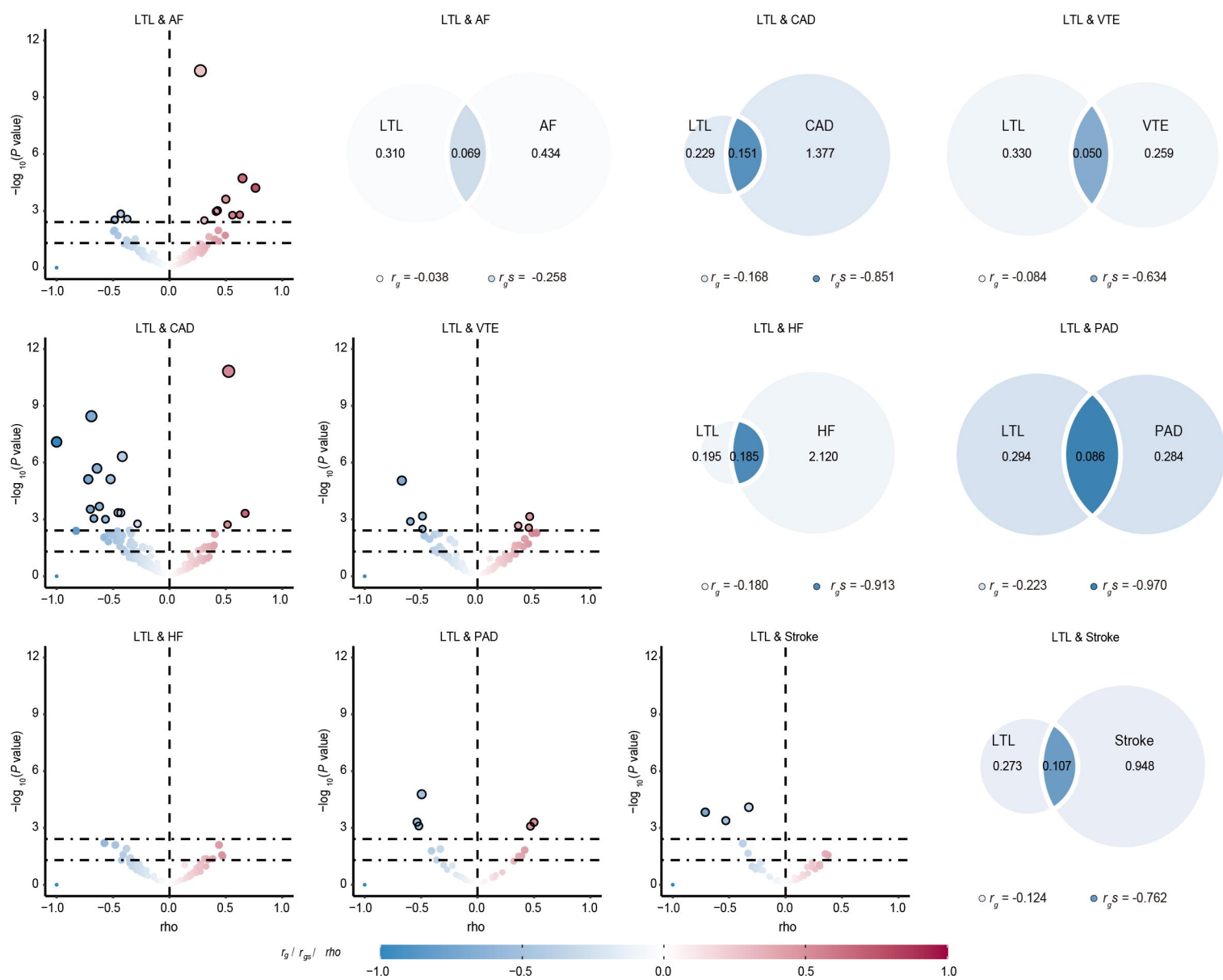
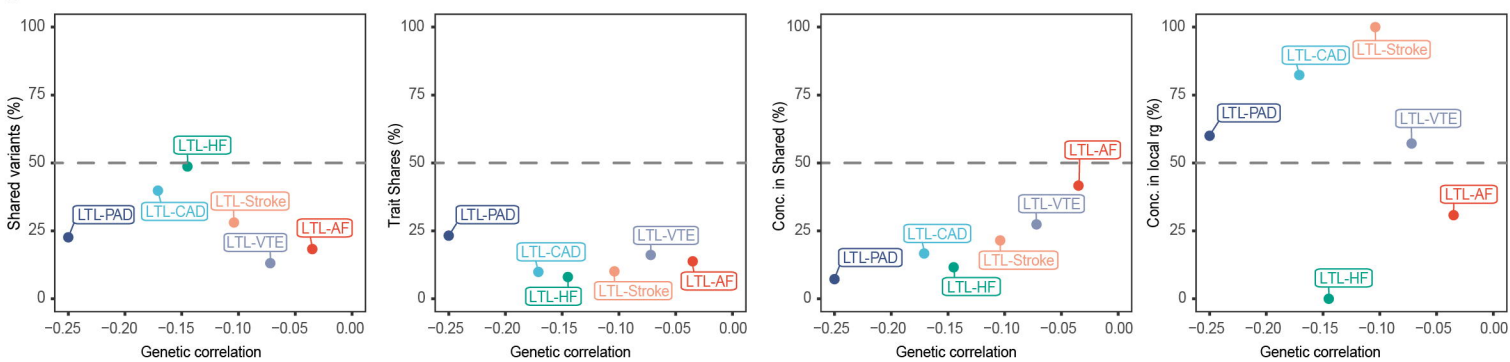
Figure 2**a****b**

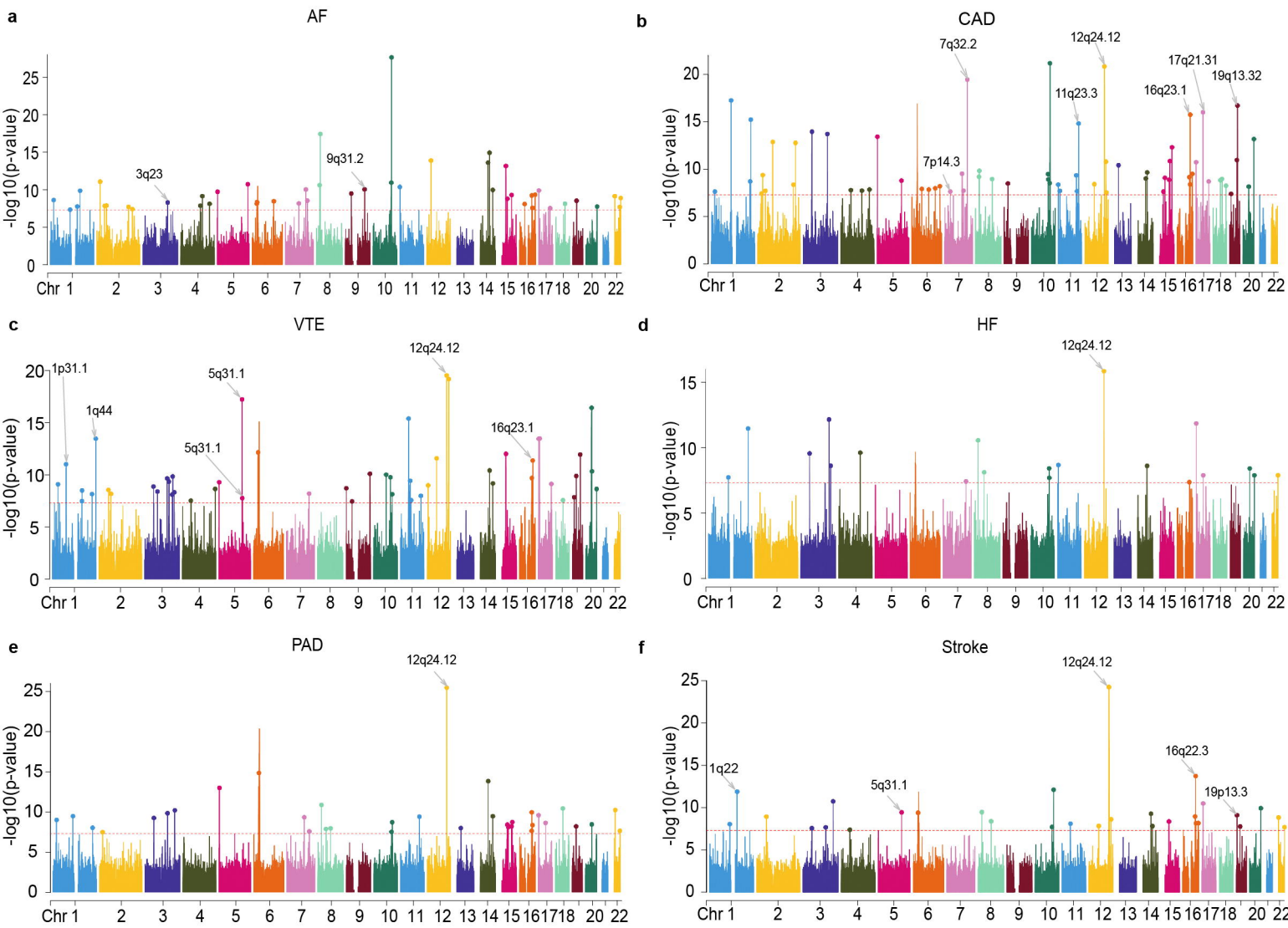
Figure 3

Figure 6