# Genetic Architecture And Clinical Outcomes Of The Fredrickson-Levy-Lees Dyslipoproteinemias

- 23
- 4 Thomas Gilliland, MD (1-3), Jaqueline S. Dron, PhD (1-2), Margaret Sunitha Selvaraj, PhD (1-
- 5 3), Mark Trinder, BA (4-5), Kaavya Paruchuri, MD (1-3), Sarah M. Urbut, MD PhD (1-3), Sara
- 6 Haidermota, BA (1-2), Rachel Bernardo, BS (1-2), Md Mesbah Uddin, PhD (1-2), Michael C.
- 7 Honigberg, MD MPP (1-3), Gina Peloso, PhD (6), Pradeep Natarajan, MD, MMSc (1-3)
- 8 9
- (1) Cardiovascular Research Center, Massachusetts General Hospital, Boston, MA
- (2) Program in Medical and Population Genetics and the Cardiovascular Disease Initiative,
   Broad Institute of Harvard and MIT, Cambridge, MA
- 12 (3) Department of Medicine, Harvard Medical School, Boston, MA
- 13 (4) Centre for Heart Lung Innovation, University of British Columbia, Vancouver.
- (5) Program in Medical and Population Genetics and the Cardiovascular Disease Initiative,
   Broad Institute of Harvard, Cambridge, Massachusetts.
  - (6) Department of Biostatistics, Boston University School of Public Health, Boston, MA
- 16 17
- 18 Please address correspondence to:
- 19 Pradeep Natarajan, MD, MMSc
- 20 55 Fruit Street
- 21 Boston, Massachusetts 02145
- 22 pnatarajan@mgh.harvard.edu
- 23 617-726-1843

#### 24 Abstract

25 Background and Aims: The genetic basis and clinical relevance of the classical Fredrickson-26 Levy-Lees (FLL) dyslipoproteinemia classifications has not been studied in general population-27 based cohorts. We aimed to evaluate the phenotypic and genetic characteristics of FLL disorders. 28 Methods: Among UK Biobank participants free of prevalent coronary artery disease (CAD), we 29 used blood lipids and apolipoprotein B concentrations to infer FLL classes (Types I, IIa, IIb, III, 30 IV, and V). For each FLL class, Cox proportional hazards regression estimated risk of incident 31 CAD. Phenome-wide association testing was performed. GWAS were performed, followed by in 32 silico causal gene prioritization and heritability analyses. Prevalence of disruptive Mendelian 33 lipid variants was assessed from whole exome sequencing. 34 Results: Of 450,636 individuals, 259,289 (57.5%) met criteria for a FLL dyslipoproteinemia: 63 35 (0.01%) type I; 40,005 (8.9%) type IIa; 94,785 (21.0%) type IIb; 13,998 (3.1%) type III; 110,389 36 (24.5%) type IV; and 49 (0.01%) type V. Over median 11.1 years follow-up, compared to 37 normolipidemics the type IIb pattern conferred the highest hazard of incident CAD overall (HR 38 1.92, 95% CI 1.84-2.01, P<0.001) and in meta-analysis across matched non-HDL-C strata (HR 39 1.45, 95% CI 1.30-1.60). GWAS revealed 250 loci associated with FLL, of which 13 were 40 shared across all classes; compared to GWAS of isolated lipid traits, 72 additional loci were 41 detected. Mendelian lipid variants were rare (2%), but polygenic heritability was high, ranging 42 from 23% (type III) to 54% (type IIb). 43 Conclusions: FLL classes have distinct genetic architectures yielding new insights for 44 cardiometabolic disease beyond single lipid analyses.

- **Keywords** ApoB: apolipoprotein B FLL: Fredrickson-Levy-Lees GWAS: genome-wide association study

#### 51 Introduction:

52 Over 50 years ago, Fredrickson, Levy, and Lees (FLL) introduced a classification scheme 53 for lipid disorders based on patterns of circulating lipoproteins(1). In their seminal series of 54 papers, the authors describe six dyslipoproteinemia types involving excess of one or more 55 lipoprotein fractions. These were termed types I, IIa, IIb, III, IV, and V, reflecting chylomicron 56 (CM) excess, low-density lipoprotein (LDL) excess, combined LDL and very low-density 57 lipoprotein (VLDL) excess, remnant cholesterol-rich lipoprotein excess, VLDL excess, and 58 combined CM and VLDL excess, respectively (**Table 1**). The authors observed distinguishing 59 clinical characteristics and patterns of heritability for each disorder and hypothesized 60 pathophysiologic mechanisms.

Owing to the cost and complexity of laboratory techniques used for diagnosis, the FLL 61 62 phenotypes have been studied primarily in specialized lipid clinics and research cohorts. These 63 settings may tend to capture extreme phenotypes, and conclusions drawn may not generalize to 64 contemporary general populations beset with growing obesity and diabetes epidemics. Thus, the 65 proposed clinical rationale for FLL classification beyond considering conventional lipid 66 measures alone remains theoretical. Sniderman and colleagues developed and validated a 67 diagnostic algorithm for the FLL disorders based on serum total cholesterol (TC), triglycerides 68 (TG), and apolipoprotein B (apoB)(2), allowing FLL diagnosis with routine lipid measurements. Application to contemporary Western populations reveals a high prevalence of FLL 69 70 dyslipoproteinemia with relationships to subclinical atherosclerosis(3,4). Associations with 71 clinical outcomes in large contemporary cohorts are presently unknown. 72 Monogenic factors have largely been studied as the basis of FLL classes but comprise a 73 very small fraction of individuals with dyslipoproteinemias. Inheritance patterns for FLL

74	disorders are believed to be variably monogenic, oligogenic, and polygenic, but characterizations					
75	have largely been restricted to the former. Furthermore, phenotypic expression may depend on					
76	the interactions between genetic and environmental factors(5)'(6). Expression of the type III					
77	pattern in patients with APOE2 homozygosity is modified by mutations in other lipid-related					
78	genes(7,8) and likely requires comorbid disease or exposure – hyperinsulinemia, alcohol excess,					
79	or the postmenopausal state(9). Further, clinical type III can occur independently of APOE2(10)					
80	While the type II phenotype with isolated hypercholesterolemia (type IIa) is classically					
81	monogenic (i.e., familial hypercholesterolemia [FH]), FH represents a very small fraction of type					
82	II pattern(11,12). FLL type IIb is believed to be polygenic, but the genetic factors are not well					
83	understood(6)'(13).					
84	Here we apply the apoB-based classification scheme to a general population-based cohort					
85	to evaluate the clinical characteristics and cardiovascular outcomes of the FLL phenotypes. We					
86	complement our findings with genome-wide association studies of the FLL disorders and					
87	perform unbiased phenotype association testing to glean novel insights into lipid metabolic					
88	disturbances.					
89						
90	Methods					

91 Study participants.

92 The UK Biobank (UKBB) is a cohort of ~500,000 volunteer participants living in the 93 United Kingdom at ages 40-69 years recruited across 22 sites between 2006 and 2010, with 94 ongoing follow up(14). Extensive genetic and phenotypic data is available as previously 95 described. Laboratory measurements, anthropometrics, and clinical histories were ascertained at 96 study enrollment. This research was conducted under UK Biobank application number 7089. All 97 study participants provided written and informed consent. The Mass General Brigham

98 Institutional Review Board approved secondary use of these data (2013P001840).

99

#### 100 Biochemical measurement and definition of Fredrickson-Levy-Lees classes.

101 Among UKBB participants, non-fasting blood lipids and apoB were measured at 102 baseline. ApoB was quantified by immuno-turbimetric assay, TC and TG by enzymatic assay, 103 high-density lipoprotein cholesterol (HDL-C) by enzyme immune-inhibition assay, and direct 104 low-density lipoprotein cholesterol (LDL-C) by enzymatic selective protection assay, all using 105 the Beckman Coulter AU5800 analytical platform, as previously reported (15). The presence of 106 lipid-lowering medications was ascertained at baseline. Similar to prior studies, in the presence 107 of a statin prescription, total cholesterol (TC) was divided by 0.8 and apoB and LDL-C were 108 divided by 0.7(16,17). 109

Among 450,636 UKBB participants free of CAD at baseline, FLL phenotype was determined by an algorithm developed by Sniderman and colleagues, which utilizes apoB, TG and the ratios of TG/apoB and TC/apoB to group individuals into one of six dyslipoproteinemias (**Table 1**)(2). Individuals not meeting criteria for FLL were categorized as normolipidemic.

113 Sensitivity analyses were performed among individuals who were not prescribed lipid-lowering

114 medications (including statins, ezetimibe, fibrates, or niacin) (N = 382,159).

115

116 Genotyping.

117 UKBB samples were genotyped using the UK Biobank Lung Exome Variant Evaluation
118 (UK BiLEVE) or Applied Biosystems UK Biobank Axiom Array. Genetic data was imputed

using either the Haplotype Reference Consortium panel or the UK10K + 1000 Genomespanel(14,18).

121

122 Clinical phenotypes.

UKBB study subjects were free of CAD at the time of enrollment. The CAD definition was based on the appearance of a qualifying International Classification of Diseases (ICD) code corresponding to acute myocardial infarction and coronary artery revascularization, physician, or patient report, as used previously (**Supplementary Table 1**)(19–21). Similarly, the covariates of hypertension and type 2 diabetes were based on a combination qualifying ICD code or physician or patient report.

We coded 1584 phenome-wide phenotypes in UKBB from both prevalent and incident hospital episodes using ICD-9 and -10 diagnosis codes grouped into phecodes(22–24). Phecodes that were sex-specific were coded as not applicable for the opposite sex. Clinical phenotypes were defined using the Phecode Map 1.2 ICD-9 and ICD-10 phenotype groupings(25). ICD codes were queried up to March 2020 in UKBB.

134

#### 135 Monogenic dyslipidemia variant annotation

Monogenic dyslipidemia variant identification was performed on the subset of ~200,000
whole exome sequences (WES) available for UKBB participants who were also included in our
study(26).

139 We extracted variants disrupting genes associated with monogenic dyslipidemia

140 phenotypes, including high LDL-C, low HDL-C, and high TG. Each gene was defined on

141 whether the resultant phenotype followed an autosomal dominant (AD) or autosomal recessive

143	LDLRAP1 (AR), ABCG5 (AR), and ABCG8 (AR). Low HDL-C genes included: APOA1 (AD),
144	ABCA1 (AR), and LCAT (AR). High TG genes included: LPL (AD), APOA5 (AD), APOC2
145	(AR), GPIBHP1 (AR), LMF1 (AR), and APOE (AR). Individuals were considered as having a
146	consistent Mendelian lipid variant if they were either homozygous or heterozygous for the

(AR) inheritance pattern. High LDL-C genes included LDLR (AD), APOB (AD), PCSK9 (AD),

147 alternative allele in AD genes, or homozygous for the alternative allele in AR genes.

Variants with minor allele frequencies of <1% and <10% were considered in AD and AR</li>
genes, respectively. This list was further restricted based on multiple variant annotations;

150 variants were included if they were defined as high-confidence loss-of function by the Ensemble

151 Variant Effect Predictor (VEP) LOFTEE plugin, had a ClinVar interpretation of "pathogenic" or

152 "pathogenic/likely pathogenic" (last checked November 2021), or were predicted as "damaging"

153 by metaSVM(27–29). All variants were annotated following canonical gene transcripts.

154 Further, we determined genotypes for *APOE* (E2/E3/E4) based on allelic combinations

155 from rs429358 and rs7412 (where rs429358 and rs7412 are T and T, respectively, for APOE E2;

156 T and C for *APOE* E3; and C and C for *APOE* E4).

157 We used PLINK (version 2.0)(30) to perform variant extractions.

158

142

159 Statistical analysis.

160 For descriptive analyses, categorical variables are presented as frequencies and

161 proportions and compared using chi square or Fisher's exact test, and continuous variables are

162 presented as mean with standard deviation or median with interquartile range and compared with

163 Student's T-test or Wilcoxon rank sum test.

164	In UKBB, incident CAD was compared between FLL phenotypes using Cox proportional
165	hazards regression ('survival' R software package, version 3.2-13) relative to normolipidemic
166	individuals. The model accounted for study enrollment to March 31 <sup>st</sup> , 2020. Participants were
167	censored if loss-to-follow-up or death occurred prior to March 31 <sup>st</sup> , 2020. Covariates included
168	age, sex, smoking status, genotyping array, the first five principal components of genetic
169	ancestry, hypertension, type 2 diabetes, BMI, and statin prescription. In secondary analysis,
170	individuals were sub-grouped according to non-HDL-C concentration (130-160 mg/dl, 160-190
171	mg/dl, 190-220 mg/dl, and 220-250 mg/dl); individuals with FLL types IIa, IIb, III, and IV were
172	compared separately against a reference group of stratum-matched normolipidemic individuals
173	for incident CAD using Cox proportional hazards regression. Heterogeneity testing was used to
174	assess significant differences within non-HDL-C groups, and meta-analysis was performed
175	across non-HDL-C groups to calculate an overall FLL phenotype effect ('meta' R software
176	package, version 4.19-0)(31). Statistical significance was assigned at $P < 0.05$ .
177	We tested the phenome-wide associations between FLL disorders and combined
178	prevalent and incident hospital episode statistic phenotypes using the PheWAS version 0.99.5-4
179	package for R (version 4.0.2). In a single model with normolipidemics as reference, the
180	association between FLL class and each phecode was assessed using logistic regression with
181	adjustments for age, sex, genotyping array, and the first 5 principal components of genetic
182	ancestry. Accounting for multiple hypothesis-testing, statistical significance was assigned at
183	$P < 3.2 \times 10^{-5}$ (0.05/1584 phecodes).
184	We performed genome-wide association studies (GWAS) using REGENIE software

(version 1.0.2) across four FLL phenotypes against normolipidemic controls(32). Third-degree
related individuals (N=34828) were excluded by the KING method(33). REGENIE step 1 was

187 performed on UKBB array genotypes. We excluded variants with minor allele frequency (MAF) 188 <1%, minor allele count <100, genotype missingness >10%, or Hardy-Weinberg equilibrium pvalue  $>10^{-15}$ . Additionally, 6 samples (of 488377 used in step 1) were excluded due to genotype 189 190 missingness >10%. In REGENIE step 2, Logistic regression was performed on genotyped and 191 imputed variants (N=19475096) meeting MAF and imputation quality filters (MAF >=0.01% 192 and INFO  $\geq$ =0.3), with Firth correction applied to SNPs where association p-value from standard 193 logistic regression was <0.01. Covariates included were age, sex, genotyping array, and the first 10 principal components of ancestry. SNPs and indels with association p-values  $<5x10^{-8}$  were 194 195 considered statistically significant. In a secondary analysis examining the effect of fasting status, 196 we repeated the above GWAS additionally adjusting for fasting time, restricted to the top 197 independent SNPs for each phenotype. Using LD score regression version 1.01 (restricted to 198 SNPs present in HapMap3), we assessed the SNP heritability and pair-wise genetic correlation of 199 FLL types IIa, IIb, III, and IV(34,35). We utilized a newly developed similarity-based method 200 for gene prioritization, Polygenic Priority Score (PoPS), to investigate protein-coding gene 201 contributions to each FLL phenotype(36). PoPS was performed on summary level GWAS data 202 for each phenotype using 1000 Genome Project as a reference(37) to yield a ranked set of 203 candidate causal genes.

Analyses were performed using R version 4.0.2 (R Core Team, 2020) unless otherwise
specified.

206

207 **Results** 

208 Baseline characteristics of UKBB Participants.

209	Overall, 450,636 UKBB participants with non-missing apoB and lipids, free of prevalent
210	CAD, and with active participant consent were included in the analysis. Mean age was 56.3 (8.1)
211	years, 44.4% were male, and 94.3% were white (Table 2). FLL dyslipoproteinemia was
212	common, with 259,289 (57.5%) participants meeting criteria for one of the FLL types. The most
213	prevalent FLL type was type IV (110,389, 24.5% of the overall population), followed by type IIb
214	(94,785, 21.0%), type IIa (40,005, 8.9%), type I (63, 0.01%), and type V (49, 0.01%)
215	(Supplementary Figure 1). Mean BMI was lowest among participants without
216	dyslipoproteinemia (25.9 kg/m <sup>2</sup> ), intermediate in type IIa individuals (26.9 kg/m <sup>2</sup> ), and highest in
217	FLL types I, IIb, IV, and V (28.9, 28.9, 28.8, and 29.3 kg/m <sup>2</sup> , respectively). Type 2 diabetes was
218	more common in among individuals with FLL disorders I, IIb, IV, and V $-$ 9.5%, 2.5%, 3.3%,
219	and 10.2%, respectively – compared to normolipidemics or FLL disorders IIa and III – $1.4\%$ ,
220	1.2%, and 1.2%, respectively. Statin treatment was most common in participants with FLL types
221	IIa and IIb (25.1% and 24.3%, respectively), compared to 8.3% of normolipidemics. The
222	distribution of lipids adjusted for lipid-lowering treatment (see Methods) reflected the
223	assignments according to the apoB-based algorithm (Figure 1).

# 225 Risk of incident CAD by FLL dyslipoproteinemia class.

Over median 11.1 years of follow-up, adjusted for age, sex, smoking status, BMI, hypertension, type 2 diabetes mellitus, statin prescription, genotyping array, and the first five principal components of ancestry, participants with FLL type IIb had the highest risk of incident CAD (HR 1.92, 95% CI 1.84-2.01, P<0.001) compared to normolipidemics, followed by type IIa (1.51, 95% CI 1.42-1.61, P<0.001) and type IV (HR 1.28, 95% CI 1.22-1.34, P<0.001) (**Figure 231 2, Supplementary Table 2**). Type III individuals were not at statistically increased risk relative

232	to normolipidemics (HR 1.05, 95% CI 0.93-1.18, P=0.44). Adjusting the model further for either
233	apoB or non-HDL-C attenuates the effect of FLL class on incident CAD, with preservation of
234	increased hazard among individuals with the type IIb phenotype compared to other common FLL
235	classes (Supplementary Tables 3 and 4 for models additionally adjusting for apoB and non-
236	HDL-C, respectively). Type III reaches nominal significance only in the model adjusted for
237	apoB (HR 1.08, 95% CI 1.02-1.29, P=0.02). When patients prescribed lipid-lowering
238	medications at baseline were excluded from the primary survival analysis, a similar pattern was
239	observed of increased hazard for type IIb versus the other FLL disorders against
240	normolipidemics (Supplementary Table 5).
241	Across strata of non-HDL-C (130-160, 160-190, 190-220, and 220-250 mg/dL), the type
242	IIb pattern was associated with increased risk of incident CAD compared to types IIa, III, and IV
243	when compared to stratum-matched normolipidemic individuals. In random-effects meta-
244	analysis across the non-HDL-C strata, type IIb conferred a 1.45-fold independent hazard (95%
245	CI 1.30-1.60) for incident CAD (Figure 3). Risk was elevated for types IV and IIa (HR 1.18,
246	95% CI 1.10-1.30 and HR 1.17, 95% CI 1.04-1.30, respectively) but not for type III individuals
247	(Supplementary Figure 2).

# 249 *Phenome-wide association testing in UK Biobank.*

Unbiased phenome-wide association testing was performed to screen for associations between FLL classes and combined prevalent or incident phenotypes among UKBB participants, adjusted for age, sex, genotyping array, and the first five principal components of ancestry. A total of 1584 phenotypes were considered in the UKBB. Effect estimates (odds ratio) between FLL type and phecode for the union of the top 50 associations for each FLL type are shown in Supplementary Figure 3. Representative top associations across the four phenotypes are shown in Figure 4. While type IIa was strongly associated with atherosclerotic cardiovascular disease (ASCVD) phenotypes, effects were greater for type IIb, which was also accompanied by strong associations with diabetes mellitus, liver disease, and obesity. Type IV was also similarly associated with diabetes mellitus, liver disease, and obesity, but effects for ASCVD were less. Type III was most strongly associated with alcohol-related disorders and malnutrition along with liver disease but without significant associations for diabetes or CAD.

262

#### 263 Genomic architecture of FLL dyslipoproteinemias.

264 Among 98,587 individuals with any FLL dyslipoproteinemia with WES, only 2,020 265 (2.05%) carried a disruptive variant in a monogenic dyslipidemia gene. Type IIa had the greatest 266 prevalence (479 of 15,309 [3.1%]) of a disruptive monogenic dyslipidemia gene, most 267 commonly in LDLR (Supplementary Table 6, Supplementary Figure 4). We also assessed 268 APOE genotype status among study participants. Notably, the APOE-2/2 genotype, which is 269 classically associated with remnant hypercholesterolemia, was markedly enriched among FLL 270 type III individuals (11.4% prevalence) relative to the other FLL phenotypes and was rare among 271 normolipidemics (0.5% prevalence) (Supplementary Table 7, Supplementary Figure 5). 272 Common variants were assessed for relevance separately for FLL types IIa, IIb, III, and 273 IV (types I and V were excluded due to low case counts) against normolipidemic individuals 274 through GWAS. Among 418603 unrelated individuals included, there were 57 with type I, 275 37,211 with type IIa, 88,007 with type IIb, 13,030 with type III, 102,677 with type IV, 45 with 276 type V, and 177,576 normolipidemics. These independent GWAS identify 77, 210, 31, and 99 277 variants contributing to FLL types IIa, IIb, III, and IV, respectively (Supplementary Tables 8278 11). Taken together, these represent a total of 250 unique genome-wide significant ( $P < 5 \times 10^{-8}$ ) 279 loci contributing to any FLL class (Figure 5, Supplementary Table 12). Of the total associated 280 loci, 13 were shared across all phenotypes (Supplementary Table 13), all of which are 281 previously described lipid-associated genes. Unique loci for types IIa, IIb, III, and IV, number 282 12, 97, 5, and 21, respectively (Supplementary Figure 7, Supplementary Table 14). 283 When compared to a recent GWAS of similar size in the Million Veterans Program for 284 individual lipid concentrations (inclusive of HDL-C, LDL-C, TC, and TG), several loci (72 of 285 250 [28.8%]) are novel(38). When further compared against published studies in the GWAS 286 Catalog(39), 14 of the total loci are not reported to be associated with any of HDL-C, LDL-C, 287 TG, total-C, or apoB concentrations (Supplementary Table 15). Examples include rs17576576, 288 which associates with the FLL IIa phenotype and is upstream of peroxisome proliferator-289 activated receptor gamma, coactivator 1 alpha (PPARGC1A), and rs3856522, which associates 290 with the FLL IIb phenotype and is an exonic variant of ATPase homolog 2 pseudogene 291 (AHSA2P). Rs141955254 is the sole novel purported association with the Type III phenotype and 292 is annotated as an intergenic variant in proximity to BMP/retinoic acid inducible neural specific 293 1 (*BRINP1*). 294 Independent variants reported in the recent CARDIoGRAM-C4D Consortium CAD 295 GWAS meta-analysis were extracted from FLL GWAS summary statistics for each 296 phenotype(40). There was substantial overlap between CAD and FLL phenotype variants. Of the 297 genome-wide significant CAD variants extracted from FLL GWAS summary statistics, 42.8% of 298 variants were nominally significant at P < 0.05 for type IIb (16.9% genome-wide), 32.2% for type

299 IIa (12.7%), 27.1% for type IV (11.4%), and 24.2% for type III (6.4%) (**Supplementary Figure** 

300 8). Direction of effect was generally shared between CAD variants and FLL variants for

phenotypes IIa and IIb but not for phenotypes III and IV (Supplementary Figure 9). For
example, *APOB* locus variants increasing the risk of CAD increased the risk of FLL types IIa
and IIb but decreased the risk of phenotypes III and IV. Further, we observe that certain loci
strongly affect risk of one condition while only modestly affecting the comparator phenotype. *LPA* variants have a strong influence on CAD risk but more modest influence on the FLL
phenotypes.

307 Polygenic priority scoring ranked protein-coding genes from the common variants
308 GWAS results for FLL types IIa, IIb, III, and IV using external multi-omics data.

309 Supplementary Figure 10 shows the union of the top 20 genetic contributors to each phenotype

310 according to PoPS score. *APOE* was the top-ranked gene across all phenotypes, and common

311 variants linked to Mendelian lipid-related genes were prioritized for all phenotypes, including

312 LDLR, APOB, SCARB1, and the APOA5/A4/C3/A1 locus. LPL was prioritized more highly for

313 the hypertriglyceridemic types (i.e., IIb, III, and IV) than for type IIa. Some prioritized novel

314 genes are platelet derived growth factor subunit B (*PDGFB*) for type IV, BRCA1 DNA repair

associated (BRCA1) and beta-1,4-galactosyltransferase 1 (B4GALT1) for type III, and insulin

316 receptor substrate 1 (*IRS1*) and insulin like growth factor 1 (*IGF1*) for type IV.

LD score regression estimated heritabilities and genetic correlations among the FLL dyslipoproteinemias (**Supplementary Figure 11**). Observed SNP heritabilities (standard errors [SE]) for FLL types IIa, IIb, III, and IV were 0.37 (0.1), 0.54 (0.09), 0.23 (0.29), and 0.28 (0.04), respectively. Pairwise genetic correlations were greatest between FLL types IV and III (0.96, SE 0.03), types IIb and III (0.92, SE 0.05), and types IIb and IV (0.82, SE 0.02). The least genetic correlations were observed between types IIa and III (0.02, SE 0.05) and IIa and IV (0.09, SE 0.01).

#### 324 Sensitivity analysis: Accounting for fasting time

A minority of individuals (N=18488) fasted for greater than or equal to 8 hours prior to baseline blood draw. Utilizing the same algorithm for FLL classification, a similar prevalence of dyslipoproteinemia was observed: 52.3% were dyslipoproteinemic among fasting individuals compared to 57.5% among all individuals. No significant different in the prevalences of FLL types I, IIb, III, and V was noted, whereas types IIa and IV were less common among fasting individuals (**Supplementary Table 17**).

Repeating GWAS with fasting time as a covariate, restricted to top independent SNPs for each phenotype, yielded no significant different in effect estimates or association P-value for any of the FLL disorders (**Supplementary Figure 12**).

334

#### 335 Discussion

336 In a large contemporary prospective cohort, we identified individuals with one of six FLL 337 dyslipoproteinemias using an apoB-based classification algorithm. The prevalence of FLL 338 disorders was high, affecting more than half of individuals. In both primary analysis and 339 sensitivity analyses controlling for apoB or non-HDL-C, the type IIb phenotype conferred the 340 highest risk of incident CAD compared to the remaining FLL dyslipoproteinemias; importantly, 341 risk was also elevated for types IIa and IV in all analyses. FLL patterns exhibited diverse 342 phenotypic relationships, including CAD for types IIa, IIb, and IV; diabetes and liver disease for 343 types IIb and IV; and alcohol-related disorders and liver disease for type III. Common variant 344 analysis using GWAS for the four prevalent FLL disorders revealed multiple genetic 345 contributions to each, with novel loci detected not previously associated with individual lipids. 346 Monogenic dyslipidemia variants were most common among type IIa individuals, usually

affecting *LDLR*, but were rare overall (i.e., 2% prevalence), and heritability analyses highlighted
the polygenicity of the FLL phenotypes.

349 Our study adds to the growing evidence that FLL dyslipoproteinemias are 350 underrecognized in contemporary Western populations(3,4). We show that identification of 351 combinations of lipid derangement – as initially conceived by Fredrickson and colleagues – 352 allows for refinement of CAD risk, as seen with the type IIb pattern when traditional lipid 353 parameters are controlled for. Characterizing combinations of lipids furthermore yields distinct 354 patterns of cardiometabolic risk not captured by single lipid parameters. While both types IIa and 355 IIb are associated with CAD risk, only the latter carries significant risk independent of apoB 356 concentration and also carries risk for diabetes and liver disease. While current guidelines 357 prioritize FH, a pattern consistent with type IIa, for early preventive therapies, our observations 358 support extension to type IIb alongside strategies to prevent diabetes onset.

359 The FLL type III phenotype was not associated with CAD in our study, in contrast to 360 prior descriptions. As previously reported, the apoB-based scheme may be more sensitive and 361 less specific for severe versions of type III dysbetalipoproteinemia in comparison to 362 ultracentrifugation(4,41). Further, given that type III is reported to be a labile finding even 363 among APOE-2/2 carriers and likely highly influenced by environmental exposures, follow-up 364 time in our model may not accurately reflect total remnant cholesterol exposure. Finally, the 365 subclinical atherosclerosis described in contemporary cohorts with FLL type III individuals may 366 not manifest as clinical ASCVD in the relatively young and healthy UKBB cohort.

Our genetic studies and downstream analysis revealed several interesting findings. First,
canonical lipid loci are important contributors to each of the four most common FLL phenotypes,
but non-lipid loci appear to be associated as well. For example, we identify the *IRS1* locus as an

370 important contributor to both type IIb and IV phenotypes. A variant ~500kb upstream of IRS1 371 (rs2943641) is previously reported to be associated with a 35% increased risk of type 2 diabetes, 372 the pathophysiologic basis of which appears to be impaired insulin sensitivity (and thus 373 increased serum insulin) rather than altered beta cell homeostasis(42). This variant is strongly 374 associated with both type IIb and IV phenotypes and is in close proximity and LD with lead 375 CAD risk variant rs2943634(43). Thus, our strategy of combining multiple lipid derangements 376 allows us to link insulin resistance, lipid abnormalities, and CAD risk for a variant likely driving 377 CAD risk via a mechanism distinct from hypercholesterolemia. Second, while there may be 378 important instances of monogenic inheritance for each of the FLL types, polygenicity is a far 379 more common pattern of inheritance. This is consistent with prior genetic studies of individuals 380 with severe hypercholesterolemia (akin to FLL type IIa) and with combined hyperlipidemia (akin 381 to FLL type IIb)(13). In the case of the type IIb phenotype, our demonstration of high 382 concordance between CAD risk loci and type IIb risk loci is consistent with the finding of high 383 atherogenicity. Further, we find that, though concentrated among type III individuals, the APOE-384 2/2 genotype is not necessary for its expression; rather, polygenic susceptibility appears 385 sufficient. Third, SNP heritability is high, and genetic correlation analysis reveals clusters of 386 similarity – the hypertriglyceridemic conditions (types IIb, III, and IV) share greater genetic 387 overlap than either types IIa and IIb or IIa with the other phenotypes.

388 Phenotype association testing reveals both previously reported and novel disease 389 associations. As expected, type IIa was associated with traditional ASCVD phenotypes, and 390 hypertriglyceridemic types IIb and IV were associated with obesity and diabetes mellitus, with 391 type IV being associated with more biliary disorders and less strongly for ASCVD, suggesting 392 that these may reflect a spectrum of a similar dyslipoproteinemia associated with metabolic

393 syndrome. Interestingly, FLL type III was associated with alcohol-related disorders and both
394 malnutrition and hyperalimentation. This finding suggests that the type III pattern may be related
395 to an exposure not well captured by these descriptors (e.g., liver disease) or simply cannot be
396 explained by a single genetic by exposure axis.

397 Combining our genetic and phenotypic findings reveals patterns of association for 398 specific FLL classes and highlights opportunities for precision medicine approaches for 399 dyslipidemia treatment. For example, we find that variants at *RBM47* are uniquely associated 400 with the type IIb phenotype. The *RBM47* locus is recently reported as a putative novel CAD risk 401 locus among individuals of Middle Eastern ancestry(44). Further, in exome sequencing of 402 diverse ancestry populations there was an association between the burden of deleterious variants 403 in RBM47 and TG/HDL-C ratio, and specific variants have been shown to significantly alter 404 apoC-III concentrations(45). This may indicate that the type IIb phenotype would benefit from 405 apo-CIII-lowering therapies in addition to LDL-C-lowering therapies. We also found the 406 FAM13A locus to be uniquely associated with the type IIb phenotype (lead variant rs6824451, 407 OR 1.04,  $P=7.84 \times 10^{-10}$ ). This variant also increased risk for the type IV phenotype, although it 408 did not reach genome-wide significance (OR 1.03, P=7.48x10<sup>-7</sup>). FAM13A encodes family with 409 sequence similarity 13 member A, which has been implicated in body fat distribution in humans 410 and visceral to subcutaneous adipose tissue ratio in mice via effects on adipocyte 411 differentiation(46); a functional genomic study in adipocytes prioritized FAM13A as a causal 412 gene in insulin resistance(47). Overall, these findings point to a common mechanism connecting 413 abnormal adjocyte homeostasis with lipid phenotype among hypertriglyceridemic FLL types IIb 414 and IV.

415 This study has several limitations. First, the UKBB is a predominantly European ancestry 416 cohort, and studies in ancestrally diverse populations are necessary to test whether these findings 417 generalize to non-European groups. Second, we relied on non-fasting lipids (the UKBB Biobank 418 standard) to assign Fredrickson class. Reassuringly, we found that when we limited the 419 population to those with  $\geq 8$  hour fast prior to the baseline lab draw, a similar proportion of 420 participants met criteria for a FLL disorder, suggesting this is simply a high prevalence 421 population. Further evidence of high baseline dyslipoproteinemia is that type IIa prevalence is 422 double that in other contemporary cohorts (8.9% versus 4-5%), despite apoB not being 423 significantly influenced by the post-prandial state(22,48). Importantly, though isolated 424 hypertriglyceridemia (FLL type IV) is classically defined in the fasting state, post-prandial 425 hypertriglyceridemia has equal implications for cardiovascular risk (49). Our approach of 426 applying the FLL framework to non-fasting samples permits clinical and genetic discovery on a 427 scale not feasible when restricting to fasting-only samples, and we show that genetic results are 428 unaffected by fasting time. 429 By accounting for patterns of change in multiple lipoprotein subfractions, the FLL 430 classification scheme describes an underlying pathophysiologic disturbance with specificity that

individual lipoprotein fractions cannot. Individuals with the type IIb pattern are at elevated risk
of ASCVD relative to other dyslipoproteinemic individuals at equivalent non-HDL-C and thus
merit intensive risk reduction and attention to the associated disease states that may contribute to
the phenotype. Differential genetic risk factors underlying the FLL types highlights opportunities
for precision medicine strategies aimed at lipid patterns beyond LDL-C alone.

436

#### 437 Funding

438 Dr. Gilliland is partially supported by the National Heart, Lung, and Blood Institute T32

439 grant 5T32HL125232. Dr. Dron is supported by the Canadian Institute of Health Research as a

440 Banting Postdoctoral Research Fellow. Drs. Peloso and Natarajan are supported by grants from

- the National Heart, Lung, and Blood Institute (R01HL142711, R01HL127564). In addition, Dr.
- 442 Natarajan is supported by grants from the National Heart Lung and Blood Institute
- 443 (R01HL148050, R01HL151283, R01HL148565, R01HL135242, R01HL151152), National
- 444 Institute of Diabetes and Digestive and Kidney Diseases (R01DK125782), Fondation Leducq
- 445 (TNE-18CVD04), and Massachusetts General Hospital (Paul and Phyllis Fireman Endowed
- 446 Chair in Vascular Medicine).

# 447 Acknowledgements

- 448 The authors would like to acknowledge and thank the staff, investigators, and participants
- 449 of the UK Biobank.

# 450 **Disclosures**

- 451 Dr. Natarajan reports grant support from Amgen, Apple, AstraZeneca, Boston Scientific,
- 452 and Novartis, spousal employment and equity at Vertex, consulting income from Apple,
- 453 AstraZeneca, Novartis, Genentech / Roche, Blackstone Life Sciences, Foresite Labs, and
- 454 TenSixteen Bio, and is a scientific advisor board member and shareholder of TenSixteen Bio and
- 455 geneXwell, all unrelated to this work. The other authors report no disclosures.

# 457 Data Availability

- 458 The data underlying this article were accessed from the UK Biobank under Application
- 459 number 7089. UK Biobank individual-level data are available for request by application
- 460 (www.ukbiobank.ac.uk).

### 461 **References**

- Fredrickson DS, Levy RI, Lees RS. Fat Transport in Lipoproteins An Integrated
   Approach to Mechanisms and Disorders. New England Journal of Medicine. 1967 Jan
   5;276(1):34–44.
- 465 2. de Graaf J, Couture P, Sniderman A. A diagnostic algorithm for the atherogenic
  466 apolipoprotein B dyslipoproteinemias. Nat Clin Pract Endocrinol Metab. 2008
  467 Nov;4(11):608–18.
- 468
  468
  469
  469
  469
  469
  470
  470
  470
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
  471
- 4. Sathiyakumar V, Pallazola VA, Park J, Vakil RM, Toth PP, Lazo-Elizondo M, et al. Modern prevalence of the Fredrickson-Levy-Lees dyslipidemias: findings from the Very Large Database of Lipids and National Health and Nutrition Examination Survey. Arch Med Sci. 2020;16(6):1279–87.
- Keulen ET, Voors-Pette C, de Bruin TW. Familial dyslipidemic hypertension syndrome:
  familial combined hyperlipidemia, and the role of abdominal fat mass. Am J Hypertens.
  2001 Apr;14(4 Pt 1):357–63.
- 479 6. Brouwers MCGJ, van Greevenbroek MMJ, Stehouwer CDA, de Graaf J, Stalenhoef AFH.
  480 The genetics of familial combined hyperlipidaemia. Nat Rev Endocrinol. 2012 Feb
  481 14;8(6):352–62.
- 482
  483
  7. Evans D, Seedorf U, Beil FU. Polymorphisms in the apolipoprotein A5 (APOA5) gene and type III hyperlipidemia. Clin Genet. 2005 Oct;68(4):369–72.
- 8. Schaefer JR, Sattler AM, Hackler B, Kurt B, Hackler R, Maisch B, et al. Hyperlipidemia in patients with apolipoprotein E 2/2 phenotype: apolipoprotein A5 S19W mutation as a cofactor. Clin Chem. 2004 Nov;50(11):2214.
- 487 9. de Beer F, Stalenhoef AFH, Hoogerbrugge N, Kastelein JJP, Gevers Leuven JA, van Duijn
  488 CM, et al. Expression of type III hyperlipoproteinemia in apolipoprotein E2 (Arg158 -->
  489 Cys) homozygotes is associated with hyperinsulinemia. Arterioscler Thromb Vasc Biol.
  490 2002 Feb 1;22(2):294–9.
- 491 10. Breckenridge WC, Little JA, Alaupovic P, Wang CS, Kuksis A, Kakis G, et al. Lipoprotein
  492 abnormalities associated with a familial deficiency of hepatic lipase. Atherosclerosis. 1982
  493 Nov;45(2):161–79.
- 494 11. Khera AV, Won HH, Peloso GM, Lawson KS, Bartz TM, Deng X, et al. Diagnostic Yield
  495 and Clinical Utility of Sequencing Familial Hypercholesterolemia Genes in Patients With
  496 Severe Hypercholesterolemia. J Am Coll Cardiol. 2016 Jun 7;67(22):2578–89.

- 497 12. Benn M, Watts GF, Tybjærg-Hansen A, Nordestgaard BG. Mutations causative of familial
  498 hypercholesterolaemia: screening of 98 098 individuals from the Copenhagen General
  499 Population Study estimated a prevalence of 1 in 217. Eur Heart J. 2016 May
  500 1;37(17):1384–94.
- 501 13. Gill PK, Dron JS, Berberich AJ, Wang J, McIntyre AD, Cao H, et al. Combined
   502 hyperlipidemia is genetically similar to isolated hypertriglyceridemia. J Clin Lipidol. 2021
   503 Feb;15(1):79–87.
- 504 14. Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, et al. The UK Biobank
   505 resource with deep phenotyping and genomic data. Nature. 2018 Oct;562(7726):203–9.
- 506 15. serum\_biochemistry.pdf [Internet]. [cited 2021 Dec 2]. Available from:
   507 https://biobank.ndph.ox.ac.uk/showcase/ukb/docs/serum\_biochemistry.pdf
- 508 16. Natarajan P, Peloso GM, Zekavat SM, Montasser M, Ganna A, Chaffin M, et al. Deep509 coverage whole genome sequences and blood lipids among 16,324 individuals. Nat
  510 Commun. 2018 Aug 23;9(1):3391.
- 511 17. Cholesterol Treatment Trialists' (CTT) Collaboration, Baigent C, Blackwell L, Emberson J,
  512 Holland LE, Reith C, et al. Efficacy and safety of more intensive lowering of LDL
  513 cholesterol: a meta-analysis of data from 170,000 participants in 26 randomised trials.
  514 Lancet. 2010 Nov 13;376(9753):1670–81.
- 515 18. Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, et al. Genome-wide genetic
  516 data on ~500,000 UK Biobank participants [Internet]. Genetics; 2017 Jul [cited 2021 Sep 7].
  517 Available from: http://biorxiv.org/lookup/doi/10.1101/166298
- 518 19. Honigberg MC, Zekavat SM, Aragam K, Finneran P, Klarin D, Bhatt DL, et al. Association
  519 of Premature Natural and Surgical Menopause With Incident Cardiovascular Disease.
  520 JAMA. 2019 Dec 24;322(24):2411–21.
- 521 20. Klarin D, Zhu QM, Emdin CA, Chaffin M, Horner S, McMillan BJ, et al. Genetic analysis in
   522 UK Biobank links insulin resistance and transendothelial migration pathways to coronary
   523 artery disease. Nat Genet. 2017 Sep;49(9):1392–7.
- 524 21. Honigberg MC, Zekavat SM, Niroula A, Griffin GK, Bick AG, Pirruccello JP, et al.
  525 Premature Menopause, Clonal Hematopoiesis, and Coronary Artery Disease in
  526 Postmenopausal Women. Circulation. 2021 Feb 2;143(5):410–23.
- 527 22. Denny JC, Ritchie MD, Basford MA, Pulley JM, Bastarache L, Brown-Gentry K, et al.
  528 PheWAS: demonstrating the feasibility of a phenome-wide scan to discover gene-disease
  529 associations. Bioinformatics. 2010 May 1;26(9):1205–10.
- Wei WQ, Bastarache LA, Carroll RJ, Marlo JE, Osterman TJ, Gamazon ER, et al. Evaluating
   phecodes, clinical classification software, and ICD-9-CM codes for phenome-wide
   association studies in the electronic health record. PLoS One. 2017;12(7):e0175508.

- 533 24. Wu P, Gifford A, Meng X, Li X, Campbell H, Varley T, et al. Mapping ICD-10 and ICD-10534 CM Codes to Phecodes: Workflow Development and Initial Evaluation. JMIR Med Inform.
  535 2019 Nov 29;7(4):e14325.
- 536 25. Wu P, Gifford A, Meng X, Li X, Campbell H, Varley T, et al. Mapping ICD-10 and ICD-10537 CM Codes to Phecodes: Workflow Development and Initial Evaluation. JMIR Med Inform.
  538 2019 Nov 29;7(4):e14325.
- 539 26. Szustakowski JD, Balasubramanian S, Kvikstad E, Khalid S, Bronson PG, Sasson A, et al.
  540 Advancing human genetics research and drug discovery through exome sequencing of the
  541 UK Biobank. Nat Genet. 2021 Jul;53(7):942–8.
- 542 27. Liu X, Wu C, Li C, Boerwinkle E. dbNSFP v3.0: A One-Stop Database of Functional
  543 Predictions and Annotations for Human Nonsynonymous and Splice-Site SNVs. Hum
  544 Mutat. 2016 Mar;37(3):235–41.
- 545 28. Landrum MJ, Lee JM, Riley GR, Jang W, Rubinstein WS, Church DM, et al. ClinVar: public
  546 archive of relationships among sequence variation and human phenotype. Nucleic Acids
  547 Res. 2014 Jan;42(Database issue):D980-985.
- 548 29. Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alföldi J, Wang Q, et al. The
  549 mutational constraint spectrum quantified from variation in 141,456 humans. Nature. 2020
  550 May;581(7809):434–43.
- 30. PLINK 2.0 [Internet]. [cited 2021 Dec 31]. Available from: https://www.cog-genomics.org/plink/2.0/
- 31. Schwarzer G, Carpenter JR, Rücker G. Meta-Analysis with R [Internet]. Cham: Springer
  International Publishing; 2015 [cited 2021 Sep 7]. (Use R!). Available from: http://link.springer.com/10.1007/978-3-319-21416-0
- 32. Mbatchou J, Barnard L, Backman J, Marcketta A, Kosmicki JA, Ziyatdinov A, et al.
  Computationally efficient whole-genome regression for quantitative and binary traits. Nat Genet. 2021 Jul;53(7):1097–103.
- 33. Manichaikul A, Mychaleckyj JC, Rich SS, Daly K, Sale M, Chen WM. Robust relationship
  inference in genome-wide association studies. Bioinformatics. 2010 Nov 15;26(22):2867–
  73.
- 34. International HapMap 3 Consortium, Altshuler DM, Gibbs RA, Peltonen L, Altshuler DM,
  Gibbs RA, et al. Integrating common and rare genetic variation in diverse human
  populations. Nature. 2010 Sep 2;467(7311):52–8.
- 35. Bulik-Sullivan BK, Loh PR, Finucane HK, Ripke S, Yang J, Schizophrenia Working Group
  of the Psychiatric Genomics Consortium, et al. LD Score regression distinguishes
  confounding from polygenicity in genome-wide association studies. Nat Genet. 2015
  Mar;47(3):291–5.

- 36. Weeks EM, Ulirsch JC, Cheng NY, Trippe BL, Fine RS, Miao J, et al. Leveraging polygenic
  enrichments of gene features to predict genes underlying complex traits and diseases
  [Internet]. Genetic and Genomic Medicine; 2020 Sep [cited 2021 Sep 7]. Available from:
  http://medrxiv.org/lookup/doi/10.1101/2020.09.08.20190561
- 37. 1000 Genomes Project Consortium, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang
  HM, et al. A global reference for human genetic variation. Nature. 2015 Oct
  1;526(7571):68–74.
- 38. Klarin D, Damrauer SM, Cho K, Sun YV, Teslovich TM, Honerlaw J, et al. Genetics of
  blood lipids among ~300,000 multi-ethnic participants of the Million Veteran Program. Nat
  Genet. 2018 Nov;50(11):1514–23.
- 579 39. GWAS Catalog [Internet]. [cited 2021 Dec 13]. Available from: https://www.ebi.ac.uk/gwas/
- 40. Aragam KG, Jiang T, Goel A, Kanoni S, Wolford BN, Weeks EM, et al. Discovery and
  systematic characterization of risk variants and genes for coronary artery disease in over a
  million participants [Internet]. 2021 Jun [cited 2021 Dec 31] p. 2021.05.24.21257377.
  Available from: https://www.medrxiv.org/content/10.1101/2021.05.24.21257377v1
- 584 41. Sniderman AD, de Graaf J, Thanassoulis G, Tremblay AJ, Martin SS, Couture P. The
   585 spectrum of type III hyperlipoproteinemia. J Clin Lipidol. 2018 Dec;12(6):1383–9.
- 42. Rung J, Cauchi S, Albrechtsen A, Shen L, Rocheleau G, Cavalcanti-Proença C, et al. Genetic
  variant near IRS1 is associated with type 2 diabetes, insulin resistance and
  hyperinsulinemia. Nat Genet. 2009 Oct;41(10):1110–5.
- 589 43. Genomewide Association Analysis of Coronary Artery Disease | NEJM [Internet]. [cited
   590 2022 Jan 7]. Available from: https://www.nejm.org/doi/full/10.1056/NEJMoa072366
- 44. Kunji K, El -Menyar Ayman, Ullah E, Kullo I, Saad M, Al SJ. Whole genome sequencing of
  a middle eastern cohort replicates known loci for coronary heart disease and identifies
  potential novel ones. Journal of the American College of Cardiology. 2021 May
  11;77(18\_Supplement\_1):13–13.
- 45. Sandholm N, Hotakainen R, Haukka JK, Sigfrids FJ, Dahlström EH, Antikainen A, et al.
  Novel protein-altering variants associated with serum apolipoprotein and lipid levels
  [Internet]. 2021 Sep [cited 2022 Jan 10] p. 2021.09.19.21263610. Available from: https://www.medrxiv.org/content/10.1101/2021.09.19.21263610v1
- 46. Fathzadeh M, Li J, Rao A, Cook N, Chennamsetty I, Seldin M, et al. FAM13A affects body
  fat distribution and adipocyte function. Nat Commun. 2020 Mar 19;11(1):1–13.
- 47. Chen Z, Yu H, Shi X, Warren CR, Lotta LA, Friesen M, et al. Functional Screening of
  Candidate Causal Genes for Insulin Resistance in Human Preadipocytes and Adipocytes.
  Circulation Research. 2020 Jan 31;126(3):330–46.

- 48. Mora S, Chang CL, Moorthy MV, Sever PS. Association of Nonfasting vs Fasting Lipid
- Levels With Risk of Major Coronary Events in the Anglo-Scandinavian Cardiac Outcomes
  Trial-Lipid Lowering Arm. JAMA Intern Med. 2019 Jul 1;179(7):898–905.
- 49. Langsted A, Nordestgaard BG. Nonfasting versus fasting lipid profile for cardiovascular risk
   prediction. Pathology. 2019 Feb 1;51(2):131–41.



611 **Figure 1. Distribution of lipids among normolipidemics and FLL classes**. Lipids are

612 presented as mg/dL on the log(10) scale, with vertical dash marking distribution median.

613 Abbreviations: 'ApoB' = apolipoprotein B; 'HDL-C' = high-density lipoprotein cholesterol;

- 614 'LDL-C' = directly measured low-density lipoprotein cholesterol. ApoB, LDL-C, and total
- 615 cholesterol are corrected for baseline statin prescription, as outlined in *Methods*.



618 Figure 2. Survival free from incident CAD according to FLL class. Individuals with FLL

619 dyslipoproteinemias were followed for a median of 11.1 years for incident CAD. Abbreviations:

620 'CAD' = coronary artery disease; 'FLL' = Fredrickson-Levy-Lees; 'Normolipidemic' = not

621 meeting criteria for any FLL disorder.





# 624 Figure 3. FLL type IIb confers elevated CAD risk across non-HDL-C strata. Meta-analysis 625 of FLL type IIb effect compared to normolipidemic individuals across non-HDL-C strata. 626 Individuals meeting criteria for FLL type IIb were compared against stratum-matched normolipidemic individuals across four strata of non-HDL-C. Random-effects meta-analysis 627 628 reported overall type IIb effect. Abbreviations: 'Non-HDL-C' = non-high-density lipoprotein 629 cholesterol; 'HR' = hazard ratio; 'CI' = confidence interval; 'N'=total individuals in a given stratum; $\tau^2$ = effect size variance estimated by restricted maximum likelihood method. 630





# 633 Figure 4. Phenome-wide association of lipid patterns yields distinct phenotypic patterns.

634 PheWAS UKBB of representative phenotypes reaching phenome-wide significance (Bonferroni

635 adjusted  $P < 3.2 \times 10^{-5}$ ) relative to normolipidemics. OR is presented only when nominal P value of

636 P<0.05 is reached and are marked with "\*" if reaching phenome-wide significance. Blank cells

637 reflect exposure-phenotype associations P>0.05.



Figure 5. Shared and distinct genetic loci associated with FLL classes. Independent GWAS were performed using REGENIE software for each of the common FLL phenotypes (IIa, IIb, III, and IV) against normolipidemic individuals. Independent genome-wide significant SNPs define +/- 500kbp loci, which are highlighted in red or blue. Red loci are unique to the phenotype, and blue loci are shared between two or more phenotypes. X-axis reflects genomic location labeled according to chromosome 1-22. Y-axis displays -log<sub>10</sub>(P-value) on the log<sub>10</sub> scale.

Fredrickson-Levy-Lees class definitions								
	ApoB (mg/dL)	TG (mg/dL)	TG/apoB	TC (mg/dL) /apoB	Lipoprotein excess			
Type I	<75	≧ 133	≧ 8.8	-	Chylomicrons			
Type IIa	≧ 120	< 133	-	-	LDL			
Type IIb	≧ 120	≧ 133	-	-	LDL and VLDL			
Type III	< 120	≧ 133	< 8.8	≧ 2.4	Remnant particles			
Type IV	< 120	≧ 133	< 8.8	< 2.4	VLDL			
Type V	$\geq$ 75 and < 120	≧ 133	≧ 8.8	-	Chylomicrons and VLDL			

Table 1. Definition and characteristics of FLL dyslipoproteinemias. FLL disorders were defined as previously based on apoB and TG and the ratios of TG/apoB and TC/apoB. ApoB and TG are expressed in mg/dL for both isolated measurements and in defining the thresholds for the ratios TG/apoB and TC/apoB. Lipoprotein excess refers to the lipoprotein particle(s) found in excess for each phenotype. Abbreviations: 'apoB' = apolipoprotein B; 'TG' = triglycerides; 'TC' = total cholesterol; 'LDL' = low-density lipoprotein; 'VLDL' = very low-density lipoprotein.

	All	<u>Normolipidemic</u>	<u>Type I</u>	<u>Type IIa</u>	<u>Type IIb</u>	<u>Type III</u>	<u>Type IV</u>	<u>Type V</u>
Number (%)	450636 (100)	191347 (42.5)	63 (0.01)	40005 (8.9)	94785 (21.0)	13998 (3.1)	110389 (24.5)	49 (0.01)
Demographics					, 			, 
Age at enrollment, mean (SD), y	56.3 (8.1)	55.0 (8.3)	55.8 (8.6)	58.4 (7.3)	57.9 (7.5)	56.8(7.7)	56.6 (8.1)	52.6 (8.6)
Male sex, n (%)	200353 (44.4)	67992 (35.5)	57 (90.5)	14991 (37.5)	48317 (51.0)	4792 (34.2)	64163 (58.1)	41 (83.7)
White, n (%)	424851 (94.3)	179080 (93.6)	59 (93.7)	37725 (94.3)	90518 (95.5)	13408 (95.8)	104015 (94.2)	46 (93.9)
Black, n (%)	7114 (1.6)	4635 (2.4)	0	847 (2.1)	635 (0.7)	128(0.9)	869 (0.8)	0
Baselin e lipids								
TC, median (IQR), mg/dl	225.8 (199.5- 253.9)	208.4 (186.7- 229.6)	176.4 (155.8- 200.2)	265.5 (247.1- 285.7)	267.4 (248.2- 289.9)	221.6 (197.1- 247.0)	214.5 (195.0- 232.5)	230.4 (206.8- 256.5)
LDL-C, median(IQR), mg/dl	143.0 (122.7- 165.2)	127.6 (111.6- 142.9)	90.5 (76.4- 98.6)	176.2 (166.0- 189.5)	178.5 (166.2- 194.7)	126.1 (109.8- 143.1)	137.2 (123.5- 149.2)	114.2 (107.0- 129.6)
HDL-C, median (IQR), mg/dl	54.5 (45.7- 65.1)	60.8 (51.5-71.2)	32.7 (27.7- 39.0)	60.6 (51.9- 70.7)	49.5 (43.0- 57.2)	64.4 (54.6- 74.1)	46.9 (40.3- 54.5)	32.6 (29.5- 37.4)
TG, median (IQR), mg/dl	130.6 (92.3- 188.9)	90.3 (72.1- 109.4)	605.7 (512.2- 696.1)	105.1 (88.6- 119.5)	205.0 (165.5- 267.6)	179.3 (151.2- 235.5)	181.9 (153.6- 232.7)	839.3 (786.0- 903.3)
АроВ, median (IQR), mg/dl	107.3 (91.9- 123.7)	95.0 (83.5- 106.0)	57.4 (48.4- 68.6)	129.9 (124.2- 139.7)	135.4 (126.7- 148.3)	86.2 (75.0- 97.1)	104.7 (94.8- 112.5)	84.2 (78.4- 94.3)
Lipid-lowering therapy								
Statin, n (%)	62990 (14.0)	15903 (8.3)	5 (7.9)	9984 (25.0)	22888 (24.1)	522 (3.7)	13680 (12.4)	8 (16.3)
Ezetimibe, n (%)	1861 (0.4)	437 (0.2)	1(1.6)	175 (0.4)	623 (0.6)	33 (0.2)	592 (0.5)	0
Fibrate, n (%)	1024 (0.2)	245 (0.1)	2 (3.2)	57 (0.1)	333 (0.4)	30 (0.2)	357 (0.3)	0
Niacin, n (%)	39	14	0 (0%)	1	11	2	11	0 (0%)
Any therapy, n (%)	78477 (17.4)	22718 (11.8)	9 (14.3)	11393 (28.5)	25869 (27.3)	1054 (7.5)	17426 (15.8)	8 (16.3)
Medical history								
Smoking status, n (%)	199707 (44.3)	77564 (40.5)	36 (57.1)	17689 (44.2)	46826 (49.4)	6353 (45.4)	51207 (46.4)	32 (65.3)
Never	248706 (55.2)	112920 (59.0)	27 (42.9)	22145 (55.4)	47466 (50.1)	7573 (54.1)	58559 (53.0)	16 (32.7)
Previous	152861 (33.9)	60920 (31.8)	27 (42.9)	13759 (34.4)	34631 (36.5)	4920 (35.1)	38581 (35.0)	23 (46.9)
Current	46846 (10.4)	16644 (8.7)	9 (14.3)	3930 (9.8)	12195 (12.9)	1433 (10.2)	12626 (11.4)	9 (18.4)
Hypertension, n (%)	126092 (28.0)	41155 (21.5)	30 (47.6)	11699 (29.2)	33336 (35.2)	3573 (25.5)	36279 (32.9)	20 (40.8)
Diabetes mellitus, type 2, n (%)	9316 (2.1)	2624 (1.4)	6 (9.5)	467 (1.2)	2366 (2.5)	174 (1.2)	3674 (3.3)	5 (10.2)
BMI, mean (SD), kg/m <sup>2</sup>	27.4 (4.8)	25.9 (4.4)	28.9 (3.8)	26.9 (4.3)	28.9 (4.5)	26.9 (4.5)	28.8 (4.8)	29.3 (4.1)

655

656 **Table 2. Baseline characteristics of UKBB study participants.** We report demographic

657 characteristics, baseline lipids, lipid-lowering medication use, and relevant medical history for all

658 study participants ('All'), individuals not meeting criteria for an FLL disorder

- 659 ('Normolipidemic'), and individuals meeting criteria for any of the six FLL dyslipoproteinemias.
- 660 Abbreviations: 'SD' = standard deviation; 'IQR' = interquartile range; 'TC' = total cholesterol;
- 661 'LDL-C' = low density lipoprotein cholesterol; 'HDL-C' = high density lipoprotein cholesterol;
- 662 'TG' = triglyceride; 'ApoB' = apolipoprotein B; 'BMI' = body-mass index.