

1 Pairwise genetic interactions modulate lipid plasma levels and 2 cellular uptake

3
4 Magdalena Zimon^{1,2,*}, Yunfeng Huang^{3,*}, Anthi Trasta^{1,2,*}, Jimmy Z. Liu³, Chia-Yen Chen^{3,4},
5 Aliaksandr Halavatyi⁵, Peter Blattmann^{1,2,6}, Bernd Klaus⁷, Christopher D. Whelan³, David
6 Sexton³, Sally John³, Wolfgang Huber⁷, Ellen A. Tsai³, Rainer Pepperkok^{1,2,5,\$,#}, Heiko Runz^{1,3,\$,¶,#}
7
8

9 ¹ Molecular Medicine Partnership Unit (MMPU), University of Heidelberg/EMBL, Heidelberg,
10 Germany

11 ² Cell Biology and Biophysics Unit, European Molecular Biological Laboratory, Heidelberg,
12 Germany

13 ³ Translational Biology, Biogen Inc., Cambridge, MA, USA

14 ⁴ Psychiatric and Neurodevelopmental Genetics Unit, Massachusetts General Hospital, Boston,
15 MA, USA

16 ⁵ Advanced Light Microscopy Facility (ALMF), European Molecular Biological Laboratory,
17 Heidelberg, Germany

18 ⁶ present address: Idorsia Pharmaceuticals Ltd, Basel, Switzerland

19 ⁷ Genome Biology Unit, European Molecular Biological Laboratory, Heidelberg, Germany

20 * These authors contributed equally

21 \$ Senior authors

22 ¶ Lead contact

23 # Correspondence to heiko.runz@gmail.com or pepperko@embl.de

27 SUMMARY

28 Genetic interactions (GIs), the joint impact of different genes or variants on a phenotype, are
29 foundational to the genetic architecture of complex traits. However, identifying GIs through
30 human genetics is challenging since it necessitates very large population sizes, while findings
31 from model systems not always translate to humans. Here, we combined exome-sequencing and
32 genotyping in the UK Biobank with combinatorial RNA-interference (coRNAi) screening to
33 systematically test for pairwise GIs between 30 lipid GWAS genes. Gene-based protein-
34 truncating variant (PTV) burden analyses from 240,970 exomes revealed additive GIs for *APOB*
35 with *PCSK9* and *LPL*, respectively. Both, genetics and coRNAi identified additive GIs for 12
36 additional gene pairs. Overlapping non-additive GIs were detected only for *TOMM40* at the
37 *APOE* locus with *SORT1* and *NCAN*. Our study identifies distinct gene pairs that modulate both,
38 plasma and cellular lipid levels via additive and non-additive effects and nominates drug target
39 pairs for improved lipid-lowering combination therapies.

40

41 **INTRODUCTION**

42 Genome-wide association studies (GWAS) have firmly established that changes in blood lipids
43 and the risk of coronary artery disease (CAD) are heritable. Hundreds of genetic loci have been
44 identified that reach genome-wide significant associations with plasma levels of low-density
45 lipoprotein cholesterol (LDL), high-density lipoprotein cholesterol (HDL), triglycerides (TG), total
46 cholesterol (TC) and CAD^{1–4}. In rare instances, susceptibility to altered blood lipids can be
47 attributed to mutations in individual genes such as *LDLR*, *PCSK9* or *APOB* that lead to familial
48 forms of disease. For the vast majority of dyslipidemic individuals, however, no single-gene
49 mutation can be identified. Instead, recent evidence suggests that in these cases inherited
50 susceptibility is caused by a cumulative effect of numerous common alleles within and across
51 GWAS loci. Individually, such common alleles have only a minor effect, but when summarized in
52 polygenic scores they can modify a phenotype to a similar extent as single high-impact
53 mutations⁵, or further magnify the penetrance of individual mutations causing Mendelian
54 disease⁶. The biological mechanisms behind the cumulative effect of risk alleles in different
55 genes remain unclear.

56 While the refined understanding of the polygenic nature of complex disease is starting to show
57 promise for improved risk prediction and treatment decisions^{7,8}, it has made it increasingly
58 difficult to decide which individual genes could be the most suitable targets for developing new
59 drugs. Drug development is traditionally focused on discrete targets with well-understood
60 biology. For certain diseases, an additive therapeutic benefit has been demonstrated through
61 combination therapies that simultaneously modulate two or more targets at once. For instance,
62 combinations of statins, inhibitors of HMG-CoA-reductase (*HMGCR*), with distinct other
63 cholesterol-lowering medications including *NPC1L1*, *PCSK9* and *APOB* inhibitors have been
64 demonstrated to lower LDL levels and CAD-risk further than statin-treatment alone^{9,10}. Despite
65 such successes, systematic strategies to predict that joint modulation of drug target pairs in
66 combination therapies will show benefit beyond standard of care have yet to be explored.
67 Genetic support for a drug target increases the probability that a medicine directed against the
68 respective target will succeed by several fold^{11,12}. We thus hypothesized that genetics might also

69 assist in nominating drug target pairs that, when addressed jointly, will have a higher probability
70 to reach a desired therapeutic benefit. A particular attractive approach to prioritize optimal target
71 pairs would be to leverage synergistic gene-gene interactions, where genetic variants in two
72 disease risk genes induce a phenotype that is more pronounced than what would be expected
73 from each of the variants' individual effects. Non-additive genetic interactions (naGIs), or
74 epistasis, have been extensively studied in model organisms and cell models with the aim to
75 identify functional relationships among genes and gene products^{13,14}. In humans, however, the
76 contribution of naGIs to the architecture of complex traits has been controversial. While there is
77 increasing evidence for modifier genes that modulate Mendelian phenotypes in non-additive
78 manners¹⁵, most of the variance of complex traits appears to be explained by genes acting
79 additively within or between loci (additive GIs, or aGIs)¹⁶.

80 Here we systematically test for pairwise GIs regulating blood lipid levels by studying interactions
81 between 30 genes prioritized based on known lipid-regulatory functions from GWAS loci using
82 three complementary tools: protein-truncating variants (PTVs) identified through exome
83 sequencing in the UK Biobank; reported GWAS lead SNPs genotyped or imputed in the UK
84 Biobank; and combinatorial RNA-interference (coRNAi) screening measuring LDL-uptake into
85 cultured cells. Our combined genetics and functional genomics approach establishes pairwise
86 additive and non-additive GIs as foundational elements in controlling blood lipid levels and
87 highlights distinct gene pairs as promising targets for lipid lowering combination therapies.

88

89 **RESULTS**

90 **Study outline**

91 To explore pairwise interactions between genes in GWAS loci and how these impact plasma lipid
92 levels and LDL-uptake into cultured cells, we followed three parallel approaches: First, we
93 extracted protein-truncating variants (PTVs) from whole exome sequencing data of 200,654
94 participants of the UK Biobank. Second, we utilized GWAS lead SNPs commonly used to
95 construct polygenic risk scores from the full set of 378,033 unrelated participants of European
96 ancestry in the UK Biobank. And third, we conducted systematic RNAi-based combinatorial

97 knockdown experiments in cells (**Figure 1a**). We focussed our analyses on 30 high-confidence
98 candidate genes from 18 genomic regions associated with blood lipid levels or the risk for CAD
99 (**Table S1**). Twenty-eight of these genes had scored as functional regulators of LDL-uptake,
100 cellular levels of free cholesterol, or LDL-receptor (LDLR) mRNA or protein levels in an earlier
101 study where we had functionally analysed 133 genes at 56 lipid and CAD GWAS loci through
102 RNAi-based knockdown experiments¹⁷. Causality for several of these genes to drive GWAS
103 associations was further supported through systematic colocalization of plasma LDL GWAS lead
104 SNPs with GTEx liver eQTLs¹ (2 genes), cis-pQTL signals¹⁸ (3 genes) and independently
105 reported biological evidence for lipid-relevant functions (15 genes) (**Table S2**). To identify
106 pairwise GIs, we applied four linear regression models (modified from [Axelsson et al., 2011¹⁹](#)) to
107 model the data. For each gene pair, both the additive genetic interaction effect (aGI) (model 3),
108 which measures the sum of effects from each gene or variant individually, as well as the non-
109 additive genetic interaction effect (naGI) (model 4), which measures the difference between the
110 expected additive and the observed combined effect, were calculated, with a naGI being either
111 synergistic or buffering (**Figure 1a** and Methods). Pairwise analyses were conducted for four
112 plasma lipid parameters (LDL, HDL, TG, TC) and CAD as available from UK Biobank²⁰ (see
113 Methods).

114

115 **PTV burden tests in UK Biobank reveal additive genetic interactions for PCSK9-APOB and**
116 **LPL-APOB**

117 We first studied pairwise modifier effects between the 30 candidate genes using high-impact
118 protein-truncating variants (PTVs). PTVs are expected to cause loss-of-function and compared
119 to other types of mutations are rare at the population level due to purifying selection^{21,22}. We
120 sequenced the exomes of 200,654 UK Biobank participants, annotated PTVs using Variant
121 Effect Predictor v96²³ and the LOFTEE plugin²¹, and identified 462,762 high-confidence PTVs in
122 the canonical transcripts of 18,869 genes. Within the 30 lipid GWAS genes, we detected a total
123 of 755 unique rare PTVs (**Table S3**). For instance, we discovered 29 different PTVs in *LDLR*, 47
124 in *PCSK9* and 102 in *APOB*. Most PTVs in these three genes were associated with strongly

125 abnormal plasma LDL levels in heterozygote carriers consistent with Familial
126 Hypercholesterolemia, although only 32 of the PTVs were annotated as pathogenic or likely
127 pathogenic in ClinVar²⁴.
128 Gene-based PTV-burden association analyses were conducted in a cohort of 161,508 unrelated
129 UK Biobank participants of European ancestry. Single-gene PTV-burden testing identified three
130 genes that were significantly associated (Bonferroni-corrected p<0.05) with both LDL and TC
131 (*APOB*, *PCSK9*, *LDLR*), two with HDL (*LPL*, *APOB*) and two with TG (*LPL*, *APOB*), respectively
132 (**Table S4**). Loss-of-function of these genes had already been identified earlier as associated
133 with the respective lipid traits at the population level². Next, we next expanded from these single
134 gene PTV-burden analyses to study pairwise PTV-based GIs, which could be tested for 42 of the
135 435 theoretically possible gene combinations (**Table S5** and Methods). For the two gene pairs
136 that met our stringent criteria to be classified as genetic interactions from this analysis, *PCSK9*-
137 *APOB* and *LPL-APOB*, we conducted replication analyses in an additional 79,462 UK Biobank
138 exomes, bringing the total sample size available for PTV-based GI testing to 240,970 individuals
139 (**Table S6**). *PCSK9-APOB* showed an aGI for both, LDL and TC, reflecting that joint loss-of-
140 function of both genes reduces these two lipid measures more than if only one of the two genes
141 is truncated. For instance, PTVs in *PCSK9* and *APOB* individually reduced mean plasma LDL by
142 34.21 mg/dl and 69.42 mg/dl relative to individuals without PTVs in these genes, consistent with
143 previous reports²⁵⁻²⁷. However, the three UK Biobank participants who carried both, *PCSK9* and
144 *APOB* PTVs, showed on average a further reduction in plasma LDL by 40.01 mg/dl compared to
145 individuals with PTVs in only one of the two genes, and by 90.45 mg/dl compared to individuals
146 with no PTV in either of the two genes (**Figure 1b**), suggesting considerable additional protection
147 from CAD. Additive GIs were further identified between *LPL* and *APOB* for HDL and TG.
148 Individuals who carried PTVs in both, *LPL* and *APOB*, showed consistently higher HDL and TG
149 levels than individuals with no PTVs, or PTVs in only one gene (**Figure 1c**). No naGIs were
150 identified through PTV-based burden tests in up to 240,970 exomes. Our results are consistent
151 with the prediction that for rare variant-based burden analyses very large sample sizes are
152 necessary to robustly detect GIs in the human population^{16,25}.

153

154 **Pairwise genetic interactions between GWAS loci modulate plasma lipid levels**

155 We next tested for GIs using 28 lipid/CAD GWAS lead SNPs representing the 30 loci in 378,033
156 unrelated individuals of European ancestry in the UK Biobank²⁰. Of a total of 1,890 pairwise
157 SNP-SNP interactions tested, 195, 98, 124, 238 and 10 aGIs were identified for LDL, HDL, TG,
158 TC and CAD, respectively (**Figure 2a-e; Table S7**). Interestingly, SNP-based analyses also
159 suggested pairwise effects between GWAS loci that deviated from an additive model and were
160 classified as naGIs. Specifically, we detected ten naGIs for LDL, one for HDL, six for TG, and
161 nine for TC (**Table 1**). No naGI was detected for CAD. The strongest driver of interactions came
162 from the 19q13.32 locus encompassing the *CBLC/BCAM/PVRL2/TOMM40/APOE* gene cluster
163 that was contributing to 19 of the 26 naGIs identified across all traits. Fourteen naGIs were
164 between lead SNPs from within the same GWAS region (“*cis*-naGI”, e.g., *NCAN-TM6SF2*,
165 *BCAM-APOE*, *ZNF259-SIK3*) with nine of them being suggestive *cis*-effects of rs4420638 near
166 *APOE*. However, naGIs were also identified between loci on different chromosomes (“*trans*-
167 naGIs”), such as between *ZNF259* and *APOE*, or *SORT1/CELSR2* and *TOMM40* for LDL and
168 TC, or between *LPL* and *ZNF259*, or *LPL* and *SIK3* for TG. Overall, our data support the
169 hypothesis that aGIs between GWAS loci are pervasive and individually small, yet if summed up
170 across many loci in polygenic scores modulate complex traits⁵. Conversely, naGIs are
171 considerably less prevalent, with the *APOE* locus being a potential contributor to naGIs for lipid
172 traits.

173

174 **Genetic interactions between gene-based PTV-burden and GWAS loci or polygenic scores**

175 Next, we queried for GIs between different types of genetic variation. Pairwise interaction testing
176 between gene-based PTV-burden and GWAS lead SNPs identified one naGI for LDL ($LDLR_{PTV}$ -
177 $PVRL2_{SNP}$), one for HDL ($APOB_{PTV}$ - LPL_{SNP}), three for TC ($LDLR_{PTV}$ - $PVRL2_{SNP}$, $LDLR_{PTV}$ - $SIK3_{SNP}$,
178 $LDLR_{PTV}$ - $PAFAH1B2_{SNP}$), and six for TG (LPL_{PTV} - $SIK3_{SNP}$, LPL_{PTV} - $ZNF259_{SNP}$, LPL_{PTV} -
179 $PAFAH1B2_{SNP}$, $BAZ1B_{PTV}$ - $NCAN_{SNP}$, $BAZ1B_{PTV}$ - $TM6SF2_{SNP}$, $BAZ1B_{PTV}$ - $PAFAH1B2_{SNP}$) (**Table**
180 **S8**). Moreover, 56, 26, 54 and 31 aGIs were identified for LDL, HDL, TC and TG, respectively.

181 These results are consistent with the genetic architecture regulating plasma lipids being
182 continuous between high-impact rare and low-impact common alleles⁴.
183 A recent study⁶ proposed that the penetrance of Mendelian disease, including FH, can be
184 substantially modulated by interactions between the respective mutant gene with common
185 variants (minor allele frequency >0.01) of individually small effect size subsumed in polygenic
186 risk scores (PRS). We created PRS for the four lipid species using PRS-CS²⁶ (and Methods) and
187 tested for GIs between PRS and PTV-burden for each of the 30 genes. Of all combinations
188 tested, only PTV-burden in *LPL*, mostly driven by the frequent p.S447Ter variant, showed
189 evidence for a naGI with the PRS for TG ($p<1.13\times10^{-15}$; beta=-0.04) (**Figure 2f; Table S9**). This
190 supports the hypothesis that a high polygenic risk for elevated TG can be mitigated by a
191 concomitant stop-gain mutation in LPL. Additionally, 10 aGIs were identified between *APOB*_{PTV}
192 with PRS for all four lipid species, *LDLR*_{PTV} and *PCSK9*_{PTV} with PRS for LDL and TC, and *LPL*_{PTV}
193 with PRS for LDL and HDL.

194

195 **RNAi identifies pairwise functional gene interactions modulating cellular LDL-uptake**

196 To gain insights into the functional consequences of GIs, we complemented our genetic
197 analyses with systematic experiments in cells using combinatorial RNAi (coRNAi) (**Figure 3a** and
198 Methods). We applied solid-phase reverse transfection to simultaneously knock down candidate
199 gene pairs in cultured HeLa cells, which we have previously shown to reliably reflect various
200 aspects of LDL biology and lipid homeostasis^{17,27,28}. Each of the 30 lipid genes was profiled with
201 a single siRNA that had previously been validated to significantly enhance or reduce cellular
202 uptake of fluorescent-labelled LDL (Dil-LDL) or free cellular cholesterol levels, and/or to
203 efficiently downregulate mRNA or protein levels of its respective target gene (**Table S2**)¹⁷. The
204 impact of both, single and combinatorial gene knockdown on LDL-uptake per cell was measured
205 and quantified from high-content microscopy images using automated image analysis routines
206 as described (**Figure S1**)^{27,28}. All pairwise knockdown combinations between the 30 lipid genes
207 (435 gene pairs) were assayed in a total of 16,128 experiments (**Figure 3b**). Each combination
208 was tested in at least seven biological replicates. Using BIC-model based robust linear

209 regression fitting analogous to the genetic interaction analyses, we identified 18 aGIs and 33
210 naGIs to differentially impact cellular LDL-uptake (**Table S10**). A similar proportion of GIs was
211 identified using robust linear model fitting and deriving p-values from the linear regression model
212 term describing non-additive effects as an alternative statistical approach (see Methods). This
213 identified 35 naGIs, with 31 naGIs overlapping between both analytical approaches (**Table S11**).
214 The corresponding gene pairs were brought forward to independent liquid-phase based coRNAi
215 replication experiments that validated 20 of these naGIs (**Table 2**, **Table S12**, **Figure S2**). Of the
216 20 validated naGIs identified through coRNAi, seven were classified (according to Horlbeck et
217 al., 2018¹⁴) as *synergistic*, i.e., simultaneous knockdown of both genes magnified the effect size
218 beyond expectations for an aGI; and thirteen naGIs were categorized as *buffering*, i.e., relative to
219 an aGI the joint knockdown mitigated LDL-uptake into cells (**Figure 3c**). For instance,
220 simultaneous knockdown of *HMGCR* and *APOB* enhanced cellular LDL-uptake beyond a mere
221 additive effect expected from knockdown of either of the two genes, proposing a *synergistic* naGI
222 (**Figure 3d**), that is most likely explained by a higher capacity of cells to bind and internalize LDL
223 via increased availability of LDL-receptor at the cell surface (**Figure S3**). Conversely, knockdown
224 of *LDLR* strongly inhibited, whereas partial knockdown of *LDLRAP1* increased cellular LDL-
225 uptake under our experimental conditions. When silencing *LDLR* and *LDLRAP1* jointly, the
226 reduction of LDL-uptake was less attenuated than expected under an additive model, suggesting
227 a *buffering* naGI (**Figure 3e**). Interestingly, reduction of LDL-uptake upon knockdown of *LDLR*
228 was magnified when *LDLR* was jointly silenced with *HAVCR1*, a suggested LDL scavenger
229 receptor that might contribute to maintain the potential of LDLR-depleted cells to internalize
230 LDL²⁹ (**Figure 3f**). Noteworthy, among the remaining validated coRNAi naGIs, simultaneous
231 silencing of *PCSK9* and *TMEM57*, as well as of *SIK3* and *PAFAH1B1* increased cellular LDL-
232 uptake to a similar extent as the simultaneous knockdown of *HMGCR* and *APOB*, although
233 silencing of these genes individually had a significant, yet only modest impact on cellular LDL-
234 uptake. In summary, coRNAi identified aGIs and naGIs between established lipid-regulatory
235 genes, but also proposed combinations of less well characterized genes as potentially important
236 factors in maintaining cellular lipid levels.

237

238 **Integrated analysis highlights GIs supported by both human genetics and cellular
239 function**

240 In order to assess whether GIs identified through either PTV-based gene-burden tests, GWAS
241 lead SNPs, or cell-based coRNAi overlapped, we integrated results from the three approaches
242 (**Figure 4; Table S13**). *LDLR-SIK3* showed an aGI both in coRNAi and PTV-SNP analyses for
243 LDL (**Figure 4a**). Both, coRNAi screening (ΔBIC 16.87, $p\text{Val}(\text{FDR})=1.18\text{E}-07$) and PTV-SNP
244 analyses for LDL and TC proposed a naGI between *LDLR* and *PVLR2* (**Figure 4b**), although this
245 gene pair failed to score as naGI in the independent coRNAi validation experiments. Twelve of
246 the 18 gene pairs nominated by coRNAi as aGIs also scored as aGIs in SNP-based interaction
247 testing for LDL and TC, including *LDLR-SIK3*. Five aGIs involved *HMGCR* and four *LDLRAP1*
248 (**Figure 4c**). Two gene pairs, *SORT1-TOMM40* and *NCAN-TOMM40*, scored as naGIs both in
249 the SNP-based as well as the coRNAi-based interaction testing (**Figure 4d**), with *TOMM40*
250 exerting a buffering naGI in either gene pair (**Figure 4e**) that could not be explained by an off-
251 target effect of *TOMM40* siRNAs on *APOE* as an adjacent gene in the 19q13.32 GWAS locus
252 (**Figure S4**). In conclusion, integrating genetic with functional data validated 12 proposed aGIs
253 and further substantiates a role of the *APOE* locus, and possibly *TOMM40*, as contributing to
254 non-additive genetic interactions.

255

256 **DISCUSSION**

257 Here, we apply whole-exome sequencing, genotyping and coRNAi to systematically test for
258 pairwise GIs between 30 lipid-regulatory genes at lipid and CAD GWAS loci. GIs are considered
259 to be central constituents of biological pathways and complex traits, contributors to human
260 disease, and promising starting points for therapy development^{13,15}. Mapping GIs, and
261 particularly non-additive epistasis, however, has been challenging. GI studies require very large
262 population sizes in order to obtain sufficient statistical power, so that the large number of
263 potential interactions to be evaluated quickly leads to a prohibitive number of statistical tests³⁰.
264 Together with most GI studies to date being limited to just a single datatype, the relative

265 contribution of GIs to variation in human complex traits has been controversial, and the
266 relevance of epistasis potentially overestimated¹⁶.

267 In our study, we have tried to overcome several of these challenges through a systematic
268 approach to GI testing that integrates genetic with functional data and relies on the UK Biobank,
269 a population cohort linking genetic with phenotype data at an unprecedented scale²⁰. To protect
270 against statistical penalties from multiple hypothesis testing we focused on pairwise interaction
271 analyses between 30 candidate genes nominated through GWAS that functional or genetic
272 follow-up studies have proposed as likely causal to confer associations with lipid traits or CAD¹⁷.
273 We assessed these genes for GIs across the allelic spectrum, from rare PTVs ascertained from
274 the exomes of more than 240,000 individuals, to common GWAS lead SNPs. Genetic GI-testing
275 was complemented by functionally knocking down gene pairs with siRNAs and determining the
276 consequence on LDL internalization into cells.

277 Several of the GIs identified in our study can be expected to be high potential starting points for
278 the development of advanced lipid-lowering combination therapies. Lowering LDL with statins is
279 the first-line pharmacological strategy to treat or prevent CAD and ischaemic heart disease as its
280 clinical manifestation. However, many patients do not reach their recommended goals of LDL-
281 lowering through statins alone, or they are intolerant against statins. For these, combination
282 therapies have become available that aim to lower atherogenic lipid levels further. A motivation
283 for this is that every 1 mmol/l (39 mg/dl) reduction in blood LDL is associated with a 19%
284 reduction in coronary mortality and a 21% reduction in major vascular events, supporting that, at
285 least for secondary prevention, the lower blood LDL levels, the better³¹. Among the options that
286 lower atherogenic blood lipids the most successfully are therapeutics against drug targets that
287 when mutated cause familial hypercholesterolemia (FH), such as *NPC1L1*, the target of
288 ezetimibe, or *PCSK9*⁹. Genetic analyses in extreme phenotypes have identified a small number
289 of individuals with concomitant mutations in two distinct FH genes, such as *LDLR* and *APOB*^{32,33},
290 *LDLR* and *LDLRAP1*^{34,35} or *APOB* and *PCSK9*³⁶. However, due to the rarity of highly penetrant
291 FH mutations such findings have thus far remained limited to individual families. Conversely, on
292 a population level, a previous GI analysis based on common alleles from ~24,000 individuals

293 ascertained for lipid traits reported 14 replicated GIs between lipid GWAS loci, most notably, like
294 in our study, with SNPs at the *APOE* locus being a key contributor³⁷. Additional support for the
295 relevance of GIs for modulating lipid traits comes from a recent study that includes a subset of
296 the UK Biobank exomes analysed here and proposes an interplay of genetic variation across the
297 allelic spectrum⁶. Notably, that study reports that carriers of monogenic CAD risk variants show
298 an up to 12.6-fold higher risk to manifest disease if they are in the highest quintile of the
299 polygenic risk distribution.

300 Our analyses here propose distinct gene pairs that modulate plasma and cellular lipid levels via
301 additive and non-additive GI effects. Among others, we identify GIs for several prominent
302 cardiovascular risk genes that individually are established targets for lipid-lowering drugs. For
303 instance, coRNAi proposed a synergistic, non-additive GI between *HMGCR*, the rate-limiting
304 enzyme during cholesterol biosynthesis and target of statins, and *APOB* encoding apolipoprotein
305 B, a critical constituent of LDL particles. Consistent with the known biological functions of these
306 genes, joint knockdown increased levels of functional LDL-receptor on the cell surface and
307 stimulated internalization of exogenous, fluorescent-labelled LDL. This observation is well in line
308 with results from clinical trials showing that in patients with Familial Hypercholesterolemia and
309 other hyperlipidemias a combination of statins with an antisense inhibitor of apolipoprotein B
310 (mipomersen) efficiently reduces plasma LDL levels more strongly than high-intensity statin
311 treatment alone^{38–41}. Importantly, the additive GI identified from UK Biobank participants carrying
312 PTVs in both, *APOB* and *PCSK9* suggests that similarly beneficial effects can be expected when
313 *APOB* antisense therapies are applied in combination with *PCSK9* inhibitors. Recently, inclisiran,
314 an siRNA targeting *PCSK9* in individuals on maximally tolerated statin doses⁴² led to a
315 persistent, highly significant lowering of LDL in treated individuals relative to placebo in a phase
316 3 study⁴³, introducing siRNAs as an attractive therapeutic modality for lipid-lowering therapies.
317 Our results strongly propose that, on a population level, combination therapies inhibiting both
318 *PCSK9* and *APOB* may lower LDL-C levels and CAD-risk even more substantially than drugs
319 targeting only one of the two genes.

320 APOB PTV-burden was associated not only with LDL and TC, but also HDL and TG, and our
321 PTV-based GI tests propose that joint disruption of *APOB* together with *LPL* reduces TG and
322 increases HDL, most likely in an additive manner. *LPL* encodes for lipoprotein lipase which
323 hydrolyzes TG from apolipoprotein B containing lipoproteins, releasing fatty acids⁴⁴. PTV-burden
324 in *LPL* is dominated by the stop-gain variant p.Ser447Ter (c.1421G>C; rs328) which in our
325 exome-sequenced UK Biobank sub-cohort showed an allele frequency of 9.95%. This variant is
326 known to cause gain-of LPL activity leading to a 0.8-fold reduced risk for ischaemic heart
327 disease⁴⁵, an effect that is likely to be further enhanced by concomitant reduction of
328 apolipoprotein B. The p.Ser447Ter allele was also the main driver behind the only naGI detected
329 between PTV-burden and polygenic risk for plasma lipids and conferred that in *LPL* PTV-carriers
330 polygenic risk for TG is reduced, with presumably non-additive effects being the most
331 pronounced in the upper percentile range of the PRS distribution.

332 A prominent driver of GIs in both our SNP- and coRNAi-based analyses was the 19q13.32 locus
333 which includes *APOE* and apart from plasma lipids and CAD is associated with Alzheimer's
334 disease, longevity and macular degeneration among others¹⁸. Interestingly, our findings indicate
335 that genes other than *APOE* at this locus might contribute to lipid GIs, which is consistent with
336 our earlier findings that knock down of several genes at this locus independently modulate
337 cellular LDL-uptake¹⁷. For instance, both SNP-based GI testing and coRNAi suggested buffering
338 naGIs for *TOMM40* with *SORT1* and *NCAN*, respectively. Variants in *TOMM40* have been
339 hypothesized to modify onset of Alzheimer's disease independently of and in conjunction with
340 *APOE*⁴⁵. Our analyses suggest *TOMM40* might exert similar modifying effects on lipid
341 phenotypes and CAD risk, which will need to be clarified in future studies. Another gene at the
342 19q13.32 locus is *PVRL2*, for which both coRNAi and SNP-PTV analyses proposed GIs with
343 *LDLR*. As a vascular cell adhesion molecule, PVRL2 protein regulates transendothelial migration
344 of leukocytes. PVRL2 levels in the atherosclerotic arterial wall correlate with plasma cholesterol
345 in CAD patients and *Ldlr*-deficient mice and have been linked to the progression of
346 atherosclerosis^{46,47}. It is thus tempting to speculate that the extensive pleiotropy of the 19q13.32
347 locus can at least in part be explained through non-*APOE* related mechanisms⁴⁵.

348 Both, genetic and functional analyses further revealed GIs between *HAVCR1*, *NCAN* and *SIK3*
349 with *HMGCR*, nominating these poorly characterized genes to be explored as potentially
350 attractive new targets for lipid-lowering therapies on top of statins.
351 Consistent with previous assumptions¹⁶, our results show that for regulating plasma lipid levels,
352 additive GIs between gene or variant pairs are common, while non-additive epistasis is rare.
353 Indeed, despite a sample size of over 240,000 exomes, our gene-based PTV-burden GI
354 analyses did not find evidence for pairwise naGIs between lipid genes disrupted by PTVs.
355 Further increasing sample sizes might help uncover non-additive effects, however, at least for
356 lipid traits, their contribution to the overall variance appears to be small. This is consistent with
357 the existence of evolutionary mechanisms that suppress epistatic interactions¹³. Since pairwise
358 naGIs can be expected to be identified the most easily for genes that are disrupted sufficiently
359 frequently in a population by PTVs of large-enough effect size, sequencing of consanguineous or
360 bottlenecked populations might improve the detection rate of naGIs^{22,48}. Interestingly, as
361 observed also here, naGIs seem to be more easily detectable in cell and animal models, for
362 instance through synthetic lethality mapping¹⁴.
363 Integration of population-scale genetics and functional coRNAi screening results yielded a total
364 of twelve aGIs and three naGIs (one of them suggestive) that influence plasma and cellular lipid
365 levels. Such validation via two systematic approaches substantially increases the confidence for
366 committing to time and resource-intense follow-up analyses of such findings, e.g., when
367 exploring the suitability of a gene pair to be jointly targeted in combination therapies.
368 Interestingly, a significant number of GIs identified through genetics and coRNAi in our study do
369 not yet overlap. This may be explained by several reasons: First, our functional analyses were
370 limited to measuring LDL-uptake into cells, which reflects a relevant, yet only a partial aspect of
371 the many possible mechanisms by which a gene can modulate plasma lipid levels. Second,
372 siRNA-based gene knock down captures acute and rather severe functional effects, which may
373 differ from the chronic and often compensated consequences upon lifelong modulation of a
374 gene's function through genetic variation. Third, despite the large number of samples used for
375 genetics-based GI testing, the number of informative high-impact variants in the human germline

376 may still be too discrete to comprehensively identify GIs. Regardless, the availability and rapid
377 development of advanced high-throughput microscopy technology joint with the constantly
378 increasing cohort sizes for genetic analyses will allow up-scaling of the approach taken here in
379 future studies and with a high probability validate further GIs.

380 In conclusion, our study introduces and confirms a strategy to link large-scale genetic data from
381 a population biobank with quantitative, cell-based coRNAi to map GIs that affect blood lipid
382 levels and CAD, an approach that can be applied to other diseases and complex traits. Our
383 unbiased analyses support that mechanisms exist through which multiple genes jointly help
384 maintain blood lipid homeostasis. CAD and ischaemic heart disease remain a substantial global
385 health burden, and doubling-down on lowering atherogenic plasma lipids remains one of the
386 most promising therapeutic approaches. With the encouraging results from recent gene- and
387 antisense-based clinical trials for CAD, our results help prioritize drug target pairs for the
388 development of lipid-lowering combination therapies rooted in human genetics.

389

390 REFERENCES

- 391 1. Willer, C., Arbor, A. & Mohlke, K. Discovery and Refinement of Loci Associated with Lipid
392 Levels Supplementary Information. 1–104 (2013) doi:10.1038/ng.2797.
- 393 2. Liu, D. J. *et al.* Exome-wide association study of plasma lipids in >300,000 individuals.
394 *Nature Genetics* **49**, 1758–1766 (2017).
- 395 3. Klarin, D. *et al.* Genetics of blood lipids among ~300,000 multi-ethnic participants of the
396 Million Veteran Program. *Nature Genetics* **50**, 1514–1523 (2018).
- 397 4. Musunuru, K. & Kathiresan, S. Genetics of Common, Complex Coronary Artery Disease.
398 *Cell* **177**, 132–145 (2019).
- 399 5. Khera, A. v. *et al.* Genome-wide polygenic scores for common diseases identify
400 individuals with risk equivalent to monogenic mutations. *Nature Genetics* **1** (2018)
401 doi:10.1038/s41588-018-0183-z.
- 402 6. Fahed, A. C. *et al.* Polygenic background modifies penetrance of monogenic variants for
403 tier 1 genomic conditions. *Nature Communications* **11**, 1–9 (2020).
- 404 7. Torkamani, A., Wineinger, N. E. & Topol, E. J. The personal and clinical utility of
405 polygenic risk scores. *Nature Reviews Genetics* **19**, 581–590 (2018).
- 406 8. Mars, N. *et al.* Polygenic and clinical risk scores and their impact on age at onset and
407 prediction of cardiometabolic diseases and common cancers. *Nature Medicine* **26**, 549–
408 557 (2020).
- 409 9. Ray, K. K. *et al.* Effect of 1 or 2 Doses of Inclisiran on Low-Density Lipoprotein
410 Cholesterol Levels: One-Year Follow-up of the ORION-1 Randomized Clinical Trial. *JAMA
411 Cardiology* **4**, 1067–1075 (2019).
- 412 10. Michos, E. D., McEvoy, J. W. & Blumenthal, R. S. Lipid management for the prevention
413 of atherosclerotic cardiovascular disease. *New England Journal of Medicine* **381**, 1557–
414 1567 (2019).
- 415 11. Nelson, M. R. *et al.* The support of human genetic evidence for approved drug
416 indications. *Nature Genetics* **47**, 856–860 (2015).
- 417 12. Zheng, J. *et al.* Phenome-wide Mendelian randomization mapping the influence of the
418 plasma proteome on complex diseases. *Nature Genetics* (2020) doi:10.1038/s41588-020-
419 0682-6.
- 420 13. Mackay, T. F. C. Epistasis and quantitative traits: Using model organisms to study gene-
421 gene interactions. *Nature Reviews Genetics* **15**, 22–33 (2014).
- 422 14. Horlbeck, M. A. *et al.* Mapping the Genetic Landscape of Human Cells. *Cell* **174**, 953–
423 967.e22 (2018).

- 424 15. Riordan, J. D. & Nadeau, J. H. From Peas to Disease: Modifier Genes, Network
425 Resilience, and the Genetics of Health. *American Journal of Human Genetics* **101**, 177–
426 191 (2017).
- 427 16. Hill, W. G., Goddard, M. E. & Visscher, P. M. Data and theory point to mainly additive
428 genetic variance for complex traits. *PLoS Genetics* **4**, (2008).
- 429 17. Blattmann, P., Schuberth, C., Pepperkok, R. & Runz, H. RNAi-Based Functional Profiling
430 of Loci from Blood Lipid Genome-Wide Association Studies Identifies Genes with
431 Cholesterol-Regulatory Function. *PLoS Genetics* **9**, (2013).
- 432 18. Sun, B. B. *et al.* Genomic atlas of the human plasma proteome. *Nature* **558**, 73–79
433 (2018).
- 434 19. Axelsson, E. *et al.* Extracting quantitative genetic interaction phenotypes from matrix
435 combinatorial RNAi. *BMC bioinformatics* **12**, 342 (2011).
- 436 20. Bycroft, C. *et al.* The UK Biobank resource with deep phenotyping and genomic data.
437 *Nature* **562**, 203–209 (2018).
- 438 21. Karczewski, K. J. *et al.* The mutational constraint spectrum quantified from variation in
439 141,456 humans. *Nature* **581**, 434–443 (2020).
- 440 22. Narasimhan, V. M. *et al.* Humans With Related Parents. *Science* **352**, 474–477 (2016).
- 441 23. McLaren, W. *et al.* The Ensembl Variant Effect Predictor. *Genome Biology* **17**, 1–14
442 (2016).
- 443 24. Landrum, M. J. *et al.* ClinVar: Improving access to variant interpretations and supporting
444 evidence. *Nucleic Acids Research* **46**, D1062–D1067 (2018).
- 445 25. Zuk, O., Hechter, E., Sunyaev, S. R. & Lander, E. S. The mystery of missing heritability:
446 Genetic interactions create phantom heritability. *Proceedings of the National Academy of
447 Sciences* **109**, 1193–1198 (2012).
- 448 26. Ge, T., Chen, C. Y., Ni, Y., Feng, Y. C. A. & Smoller, J. W. Polygenic prediction via
449 Bayesian regression and continuous shrinkage priors. *Nature Communications* **10**, 1–10
450 (2019).
- 451 27. Bartz, F. *et al.* Identification of Cholesterol-Regulating Genes by Targeted RNAi
452 Screening. *Cell Metabolism* **10**, 63–75 (2009).
- 453 28. Thormaehlen, A. S. *et al.* Systematic Cell-Based Phenotyping of Missense Alleles
454 Empowers Rare Variant Association Studies: A Case for LDLR and Myocardial Infarction.
455 *PLOS Genetics* **11**, e1004855 (2015).
- 456 29. Ichimura, T. *et al.* Kidney injury molecule-1 is a phosphatidylserine receptor that confers
457 a phagocytic phenotype on epithelial cells. *Journal of Clinical Investigation* **118**, 1657–
458 1668 (2008).
- 459 30. Phillips, P. C. Epistasis - The essential role of gene interactions in the structure and
460 evolution of genetic systems. *Nature Reviews Genetics* **9**, 855–867 (2008).

- 461 31. Cholesterol Treatment Trialists' (CTT) Collaborators. Efficacy and safety of cholesterol-
462 lowering treatment: prospective meta-analysis of data from 90 056 participants in 14
463 randomised trials of statins. *The Lancet* **366**, 1267–1278 (2005).
- 464 32. Rauh, G. et al. Identification of a heterozygous compound individual with familial
465 hypercholesterolemia and familial defective apolipoprotein B-100. *Klinische Wochenschrift*
466 **69**, 320–324 (1991).
- 467 33. Benlian, P. et al. Phenotypic Expression in Double Heterozygotes for Familial
468 Hypercholesterolemia and Familial Defective Apolipoprotein B-100. *340345*, (1996).
- 469 34. Tada, H. et al. A novel type of familial hypercholesterolemia: Double heterozygous
470 mutations in LDL receptor and LDL receptor adaptor protein 1 gene. *Atherosclerosis* **219**,
471 663–666 (2011).
- 472 35. Soufi, M., Rust, S., Walter, M. & Schaefer, J. R. A combined LDL receptor/LDL receptor
473 adaptor protein 1 mutation as the cause for severe familial hypercholesterolemia. *Gene*
474 **521**, 200–203 (2013).
- 475 36. Elbitar, S. et al. New Sequencing technologies help revealing unexpected mutations in
476 Autosomal Dominant Hypercholesterolemia. *Scientific Reports* **8**, 1–10 (2018).
- 477 37. De, R. et al. Identifying gene–gene interactions that are highly associated with four
478 quantitative lipid traits across multiple cohorts. *Human Genetics* **136**, 165–178 (2017).
- 479 38. Akdim, F. et al. Effect of Mipomersen, an Apolipoprotein B Synthesis Inhibitor, on Low-
480 Density Lipoprotein Cholesterol in Patients With Familial Hypercholesterolemia. *American*
481 *Journal of Cardiology* **105**, 1413–1419 (2010).
- 482 39. Akdim, F. et al. Efficacy of apolipoprotein B synthesis inhibition in subjects with mild-to-
483 moderate hyperlipidaemia. *European Heart Journal* **32**, 2650–2659 (2011).
- 484 40. McGowan, M. P. et al. Randomized, Placebo-Controlled Trial of Mipomersen in Patients
485 with Severe Hypercholesterolemia Receiving Maximally Tolerated Lipid-Lowering
486 Therapy. *PLoS ONE* **7**, 1–10 (2012).
- 487 41. Fogacci, F. et al. Efficacy and Safety of Mipomersen: A Systematic Review and Meta-
488 Analysis of Randomized Clinical Trials. *Drugs* vol. 79 751–766 (2019).
- 489 42. Fitzgerald, K. et al. A highly durable RNAi therapeutic inhibitor of PCSK9. *New England*
490 *Journal of Medicine* **376**, 41–51 (2017).
- 491 43. Raal, F. J. et al. Inclisiran for the treatment of heterozygous familial
492 hypercholesterolemia. *New England Journal of Medicine* **382**, 1520–1530 (2020).
- 493 44. Merkel, M., Eckel, R. H. & Goldberg, I. J. Lipoprotein lipase: Genetics, lipid uptake, and
494 regulation. *Journal of Lipid Research* **43**, 1997–2006 (2002).
- 495 45. Roses, A. et al. Understanding the genetics of APOE and TOMM40 and role of
496 mitochondrial structure and function in clinical pharmacology of Alzheimer's disease.
497 *Alzheimer's and Dementia* **12**, 687–694 (2016).

- 498 46. Björkegren, J. L. M. *et al.* Plasma Cholesterol-Induced Lesion Networks Activated before
499 Regression of Early, Mature, and Advanced Atherosclerosis. *PLoS Genetics* **10**, (2014).
- 500 47. Rossignoli, A. *et al.* Poliovirus Receptor-Related 2: A Cholesterol-Responsive Gene
501 Affecting Atherosclerosis Development by Modulating Leukocyte Migration.
502 *Arteriosclerosis, Thrombosis, and Vascular Biology* **37**, 534–542 (2017).
- 503 48. Locke, A. E. *et al.* Exome sequencing of Finnish isolates enhances rare-variant
504 association power. *Nature* **572**, 323–328 (2019).
- 505 49. van der Harst, P. & Verweij, N. Identification of 64 Novel Genetic Loci Provides an
506 Expanded View on the Genetic Architecture of Coronary Artery Disease. *Circulation
507 Research* **122**, 433–443 (2018).
- 508 50. Kamat, M. A. *et al.* PhenoScanner V2: An expanded tool for searching human genotype-
509 phenotype associations. *Bioinformatics* **35**, 4851–4853 (2019).
- 510 51. Burton, P. R. *et al.* The Wellcome Trust Case Control Consortium. Genome-wide
511 association study of 14,000 cases of seven common diseases and 3,000 shared controls.
512 *Nature* **447**, 661–678 (2007).
- 513 52. Lonsdale, J. *et al.* The Genotype-Tissue Expression (GTEx) project. *Nature Genetics* **45**,
514 580–585 (2013).
- 515 53. Giambartolomei, C. *et al.* Bayesian Test for Colocalisation between Pairs of Genetic
516 Association Studies Using Summary Statistics. *PLoS Genetics* **10**, (2014).
- 517 54. Hout, C. V. van *et al.* Whole exome sequencing and characterization of coding variation
518 in 49,960 individuals in the UK Biobank. *bioRxiv* 572347 (2019) doi:10.1101/572347.
- 519 55. Schwarz, G. Estimating the Dimension of a Model. *Ann. Statist.* **6**, 461–464 (1978).
- 520 56. Erfle, H. *et al.* Reverse transfection on cell arrays for high content screening microscopy.
521 *Nature protocols* **2**, 392–399 (2007).
- 522 57. Erfle, H. *et al.* Work flow for multiplexing siRNA assays by solid-phase reverse
523 transfection in multiwell plates. *Journal of biomolecular screening* **13**: the official journal of
524 the Society for Biomolecular Screening **13**, 575–580 (2008).
- 525 58. Carpenter, A. E. *et al.* CellProfiler: Image analysis software for identifying and
526 quantifying cell phenotypes. *Genome Biology* **7**, (2006).
- 527 59. Gilbert, D. F., Meinhof, T., Pepperkok, R. & Runz, H. DetecTiff[©]: A novel image analysis
528 routine for high-content screening microscopy. *Journal of Biomolecular Screening* **14**,
529 944–955 (2009).
- 530 60. Malo, N., Hanley, J. A., Cerqueiro, S., Pelletier, J. & Nadon, R. Statistical practice in
531 high-throughput screening data analysis. *Nature Biotechnology* **24**, 167–175 (2006).
- 532 61. Birmingham, A. *et al.* Interference Screens. *Nature Methods* **6**, 569–575 (2010).
- 533 62. Raftery, A. E. Bayesian Model Selection in Social Research. *Sociological Methodology*
534 **25**, 111 (1995).

- 535 63. Yoav Benjamini & Yosef Hochberg. Controlling the False Discovery Rate: A Practical
536 and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series*
537 *B (Methodological)* **57**, 289–300 (1995).
- 538

539 **ACKNOWLEDGEMENTS**

540 This research has been conducted using the UK Biobank resource under application number
541 26041. We thank all the participants and researchers of UK Biobank for making these data open
542 and accessible to the research community. AbbVie, Anylam Pharmaceuticals, AstraZeneca,
543 Biogen, Bristol-Myers Squibb, Pfizer, Regeneron and Takeda are acknowledged for generation
544 and initial quality control of the whole-exome sequencing data. We thank Eric Marshall,
545 Yongsheng Huang and Frank Nothaft for infrastructure support for genetic data analyses. The
546 EMBL Advanced Light Microscopy Facility is acknowledged for supporting high-content
547 microscopic-based screening analyses. We are grateful to Brigitte Joggerst, Susanne Theiss and
548 Miriam Reiss for excellent technical assistance. Support to the study came in part from the
549 Transatlantic Networks of Excellence Program 10CVD03 from Fondation Leducq to HR and RP.
550 MZ was supported by the EMBL EIPOD programme, AT and PB by the EMBL PhD programme.

551

552 **AUTHOR CONTRIBUTIONS**

553 Conceptualization, H.R. and R.P.; Methodology and Investigation, M.Z., Y.H., A.T., C.-YC., R.P.
554 and H.R.; Formal Analysis and Validation, M.Z., Y.H., A.T., C.-YC., J.L., A.H., B.K., E.T. and
555 H.R.; Resources and Data Curation, J.L., P.B., C.W., D.S., S.J., E.T., R.P. and H.R.; Writing –
556 Original Draft, M.Z., A.T. and H.R; Writing – Review and Editing, M.Z., Y.H., P.B., E.T., R.P. and
557 H.R.; Supervision, S.J., W.H., E.T., R.P. and H.R.; Project Administration and Funding
558 Acquisition, S.J., R.P. and H.R.

559

560 **DECLARATION OF INTERESTS**

561 Y.H., C.-YC., J.L., C.W., D.S., S.J., and H.R. are full-time employees at Biogen, Inc. The funders
562 had no role in study design, data collection and analysis, decision to publish, or preparation of
563 the manuscript.

564

565 Further information and requests for resources and reagents should be directed to and will be
566 fulfilled by the Lead Contact, Heiko Runz (heiko.runz@gmail.com)

567 **METHODS**

568 **Gene Selection**

569 We chose to study 30 candidate genes from 18 loci reported as associated through common-
570 variant genome-wide association studies (GWAS) as associated with plasma lipid levels and the
571 risk for CAD. Twenty-eight of these genes had been identified and validated as functional
572 regulators of LDL-uptake and/or cholesterol levels into cells in a previous RNAi-screen analysing
573 a total of 133 genes in 56 lipid and CAD GWAS loci¹⁷ (**Table S1**). Common-variant association
574 signals and published biological evidence for potential roles in lipid regulation were updated for
575 all 30 candidate genes based on the recent literature (e.g., ^{1-3,49}) and queries using the
576 PhenoScanner platform⁵⁰ (<http://www.phenoscaner.medschl.cam.ac.uk/>). Twenty eight genes
577 were validated to reside within loci that are associated at genome-wide significance (p<5e-8)
578 with plasma lipid levels or CAD. SNPs near *FAM174A* (rs383830) and *SEZ6L* (rs688034) had
579 originally been reported as associated with CAD⁵¹, but failed to replicate at genome-wide
580 significance in more recent meta-GWAS. However, since knockdown of both genes had scored
581 as significantly impacting lipid parameters in cells¹⁷ the two genes were maintained for this
582 current study.

583

584 **Colocalization Analysis**

585 Colocalization analysis was performed between the 28 GWAS lead SNPs using summary
586 statistics from the 2013 Global Lipid Genetics Consortium GWAS¹
587 (<http://csg.sph.umich.edu/willer/public/lipids2013/>) and the GTEx liver *cis*-eQTL dataset
588 (N=153)⁵². When a respective locus was associated with multiple lipid phenotypes, the SNP with
589 the lowest reported p-value association with LDL was chosen to be the lead SNP. There was no
590 GTEx liver expression data for four genes (*APOE*, *MYBPHL*, *NCAN*, *SEZ6L*), therefore there
591 were no *cis*-eQTL for these genes to colocalize with. Colocalization analysis was conducted
592 following the methods in Giambartolomei et al., 2014⁵³ using the R ‘coloc’ package on a +/-500kb
593 window around each lead SNP against SNP-to-expression data of all neighbouring genes within
594 that locus. Positive colocalization between liver *cis*-eQTL and GWAS signal was defined as

595 showing a posterior probability of sharing the same SNP (PP4) if larger than 0.8. A lead SNP at
596 the *SORT1/CELSR2* locus (rs629301) showed a positive colocalization signal, but the *cis*-eQTL
597 co-localized with both genes, so SNP-based GIs for these genes could not be analysed
598 separately.

599

600 **UK Biobank lipid and CAD phenotypes**

601 The UK Biobank is a prospective study of over 500,000 participants recruited at an age of 40-69
602 years from 2006-2010 in the United Kingdom. Participant data include health records, medication
603 history and self-reported survey information, together with imputed genome-wide genotypes and
604 biochemical measures²⁰. Baseline biochemical measures including LDL cholesterol (LDL), HDL
605 cholesterol (HDL), triglycerides (TG), and serum total cholesterol (TC) had been obtained in UK
606 Biobank's purpose-built facility in Stockport as described in the UK Biobank online data
607 showcase and protocol (www.ukbiobank.ac.uk). Demographic and other relevant phenotypic
608 information was obtained from standard questionnaire data. Individual lipid phenotypes (LDL,
609 HDL, TG and TC) were modelled as dependent variables using linear regression models against
610 covariates including age, sex, smoking, alcohol drinking status, and BMI. Lipid medication use
611 was obtained from self-reported questionnaire data (UK Biobank fields 6153 and 6177). CAD
612 cases were recognized based on both self-reported diagnosis and Hospital Episode Statistics
613 data in the UK Biobank with a code-based CAD definition as presented in the most recent CAD
614 GWAS that included UK Biobank⁴⁹. In total, 30,125 CAD cases were identified and the cohort
615 was adjusted for age, sex, smoking status, alcohol drinking status, BMI and lipid medication use.
616 All phenotype data were derived from UK Biobank basket "ukb27390" released on March 11,
617 2019.

618

619 **Pairwise gene-based PTV-burden interaction testing**

620 High-impact protein-truncating variants (PTVs) expected to disrupt protein functions were
621 identified from 200,654 whole-exome sequencing (WES) data of UK Biobank participants to
622 conduct pairwise interaction analyses. WES data was generated and quality controlled (QC-ed)

623 as described in Van Hout et al. at the Regeneron Genetics Center as part of a collaboration
624 between AbbVie, Alnylam Pharmaceuticals, AstraZeneca, Biogen, Bristol-Myers Squibb, Pfizer,
625 Regeneron and Takeda and the UK Biobank consortium⁵⁴. PTVs were called from a Regeneron
626 QC-passing “Goldilocks” set of genetic variants using Variant Effect Predictor v96²³ (McLaren et
627 al., 2016) and the LOFTEE plugin²¹. We identified 462,762 high-confidence PTVs with a minor
628 allele frequency of <1% in the canonical transcripts of 18,869 genes. This set included 755 rare
629 PTVs in the 30 lipid genes analysed in this study. PTVs per gene were enumerated, and a PTV-
630 burden association analysis was conducted in 161,508 unrelated (>2nd degree relatedness) UK
631 Biobank participants of European ancestry, as defined by principle components analysis of the
632 genotyping data²⁰. Replication analysis was conducted from an additional 101,827 samples,
633 bringing the total sample size used for calling PTVs from UK Biobank exome sequencing data to
634 302,634. Of these 101,827 samples, 79,462 fulfilled the criteria applied to the discovery cohort,
635 so that an overall sample size of 240,970 exomes was available for replicating findings from the
636 initial PTV-based GI analyses.

637 For pairwise PTV-based interaction testing, QC-ed UK Biobank lipid phenotypes (HDL, LDL, TG
638 and TC) were modelled as dependent variables using the following four linear regression models
639 in R:

640 Model 1 for gene1 PTV-burden only: lipids ~ PTV₁

641 Model 2 for gene2 PTV-burden only: lipids ~ PTV₂

642 Model 3 for gene1 PTV-burden and gene2 PTV burden (*additive GI*): lipids ~ PTV₁ + PTV₂

643 Model 4 for gene1 PTV-burden and gene2 PTV burden (*non-additive GI*): lipids ~ PTV₁ + PTV₂ +
644 PTV₁ * PTV₂

645 Schwarz's Bayesian Information Criterion (BIC)⁵⁵ scoring was used to determine the best model
646 to explain the data and goodness of fit, with the lowest BIC value indicating the best-fitting model
647 describing each gene pair. Model 3 reflected *additive* genetic interactions (aGIs), Model 4 *non-*
648 *additive* gene interactions (naGIs). The model with the lowest BIC was chosen as describing
649 most adequately the type of interaction between each corresponding gene pair.

650

651 **Pairwise SNP interaction testing**

652 To assess whether GWAS lead SNPs modulate plasma lipid levels through joint effects within
653 and across GWAS loci, we conducted pairwise SNP-SNP interaction analysis using genome-
654 wide genotyping data and biochemical measures of lipid species from the UK Biobank. Twenty-
655 eight lead SNPs mapped to the 30 lipid GWAS genes were extracted from genotyping data of
656 378,033 unrelated (removed up to 2nd degree relatedness) participants of European ancestry. A
657 total of 378 pairwise modifier effects were tested by conducting Robust Linear Model Fitting
658 using R, running the following four linear regression models:

659 Model 1 for SNP1 only: lipids ~ SNP₁

660 Model 2 for SNP2 only: lipids ~ SNP₂

661 Model 3 for SNP1 and SNP2 (*additive GI*): lipids ~ SNP₁ + SNP₂

662 Model 4 for SNP1 and SNP2 (*non-additive GI*): lipids ~ SNP₁ + SNP₂ + SNP₁ * SNP₂

663 Schwarz's Bayesian Information Criterion (BIC) scoring was used to determine the best model to
664 explain the data and goodness of fit, with the lowest BIC value indicating the best-fitting model
665 describing each SNP pair. If Model 3 had the lowest BIC value, it reflected an aGI, and if Model 4
666 had the lowest BIC value, it reflected a naGI.

667 A similar strategy was applied for pair-wise interaction testing to explore potential joint effects
668 between the 30 genes on CAD risk by running the following four logistic regression models
669 adjusted for age, sex, smoking status, alcohol drinking status, BMI and lipid medication use:

670 Model 1 for SNP1 only: CAD ~ SNP₁

671 Model 2 for SNP2 only: CAD ~ SNP₂

672 Model 3 for SNP1 and SNP2 (*additive GI*): CAD ~ SNP₁ + SNP₂

673 Model 4 for SNP1 and SNP2 (*non-additive GI*): CAD ~ SNP₁ + SNP₂ + SNP₁ * SNP₂

674 As above, the model with the lowest BIC was chosen as describing most adequately the type of
675 interaction between each corresponding SNP pair.

676

677 **PTV-SNP interaction testing**

678 In order to conduct pairwise interaction analyses between GWAS lead SNPs and PTVs, we
679 assessed the interaction of the 28 lead SNPs with rare PTV burden for each of the 30 genes. For
680 SNP-PTV interaction testing, UK Biobank lipid phenotypes (HDL, LDL, TG and TC) were
681 modelled as dependent variables using the following four linear regression models:

682 Model 1 for gene1 lead SNP only: lipids ~ SNP₁

683 Model 2 for gene2 PTV-burden only: lipids ~ PTV₂

684 Model 3 for gene1 lead SNP and gene2 PTV burden (*additive GI*): lipids ~ SNP₁ + PTV₂

685 Model 4 for gene1 lead SNP and gene2 PTV burden (*non-additive GI*): lipids ~ SNP₁ + PTV₂ +
686 SNP₁*PTV₂

687 As above, the model with the lowest BIC was chosen as describing most adequately the type of
688 interaction between each corresponding SNP-gene pair.

689

690 **PTV-PRS interaction testing**

691 We assessed the interaction effects between polygenic risk score (PRS) and PTVs for each of
692 the four lipid phenotypes. To construct PRS for UK Biobank samples, we first derived the PRS
693 weights for each SNP across the genome using PRS-CS²⁶, which is a Bayesian regression-
694 based algorithm, and publicly available summary statistics from lipid GWAS¹. We applied derived
695 PRS weights to imputed genotypes (with minor allele frequency >0.01 and imputation quality
696 INFO >0.8) of UK Biobank samples and calculated PRS for each lipid, based on the
697 corresponding PRS weights. Note that all SNPs in the gene of interest were excluded from the
698 PRS when testing for PRS-PTV gene interaction. GIs were tested between PRS and PTV-
699 burden for each of the 30 genes by fitting the four linear regression models:

700 Model 1 for PRS only: lipids ~ PRS

701 Model 2 for gene PTV-burden only: lipids ~ PTV

702 Model 3 for PRS and gene PTV burden (*additive GI*): lipids ~ PRS + PTV

703 Model 4 for PRS and gene PTV burden (*non-additive GI*): lipids ~ PRS + PTV + PRS * PTV

704 As above, the model with the lowest BIC was chosen as describing most adequately the type of
705 interaction between each corresponding PRS-gene pair.

706

707 **RNAi interaction testing**

708 **Cells and reagents**

709 HeLa-Kyoto cells are a strongly adherent Hela isolate (gift from S. Narumiya, Kyoto University
710 Japan) that, as we demonstrated earlier, enable reliable measurements of LDL-cholesterol
711 uptake dynamics and show lipid homeostatic mechanisms similar to those described for liver-
712 derived cell models^{17,27,28}. Dil-LDL (Life Technologies), DRAQ5 (Biostatus), Dapi (Molecular
713 Probes), 2-hydroxy-propyl-beta-cyclodextrin (HPCD) (Sigma), Lipofectamine 2000 (Invitrogen)
714 and Benzonase (Novagen) were purchased from the respective suppliers.

715

716 **siRNA selection and production of siRNA microarrays**

717 RNA-interference (RNAi) screening was conducted in glass-bottomed single-well chambered cell
718 culture (Lab-Tek) slides with solid-phase reverse siRNA-transfection of cultured cells (“cell
719 microarrays”) as described previously^{27,56}. Each gene under study was targeted with a single
720 siRNA (Silencer Select, Invitrogen) that had been selected with the EMBL-generated software
721 tool bluegecko (J.K. Hériche, unpublished) based on the alignment to the reference genome, a
722 maximal number of protein-coding transcripts per gene targeted and expected specificity for the
723 target gene. The 28 siRNAs in this study had been validated earlier to significantly enhance or
724 reduce cellular uptake of fluorescent-labelled LDL (Dil-LDL) or free cellular cholesterol levels¹⁷
725 and were shown to efficiently downregulate mRNA or protein levels of their respective target
726 genes (**Table S2**). siRNA sequences are provided in Blattmann et al., 2013 Supplementary
727 Table 4. For the two genes not analysed in our earlier study (*MYLIP*, *PAFAH1B2*), siRNAs used
728 in the current study were prioritized from 3 and 5 siRNAs per gene based on bluegecko *in silico*
729 analyses, knockdown efficiency on target mRNA/protein levels (up to less than 10% residual
730 levels) and/or efficiency to modulate cellular Dil-LDL uptake in preparatory individual single gene
731 knock-down experiments (not shown). The 75% (12/16) of siRNAs that had scored as
732 individually modulating cellular Dil-LDL uptake in our earlier study¹⁷ also met the more stringent
733 criteria of our current study to score as LDL-uptake modulator when used either alone or

734 together with non-silencing control siRNA Neg9 (**Figure 3b**, CTRL column), thereby replicating
735 our earlier results and validating experimental settings for this current study.

736 To cover the total of 435 pairwise siRNA combinations including controls and replicas, five
737 different cell microarrays with 384 spots/array were produced. Per array, the following negative
738 controls were added: eight spots containing *INCENP*-siRNA (s7424) to control for transfection
739 efficiency¹⁷; eight spots containing non-silencing control siRNA Neg1 (s229174), and eight spots
740 containing non-silencing control siRNA (denoted as CTRL throughout the text) Neg9 (s444246).
741 Furthermore, eight spots were added with siRNA targeting *LDLR* (s224006) as a positive control
742 for LDL uptake, as well as eight spots with siRNA targeting *NPC1* (s237198) knockdown of
743 which increases free cellular cholesterol signals²⁷. For pairwise combinatorial RNAi-screening,
744 siRNAs against two genes were printed simultaneously on a respective siRNA-spot, with equal
745 amounts (15 pmol/siRNA) of siRNA per gene. As positive controls, eight spots containing both,
746 non-silencing control siRNA Neg9 (CTRL) (s444246) and siRNA targeting *LDLR* (s224006), and
747 eight spots containing both, non-silencing control siRNA (CTRL) Neg9 (s444246) and siRNA
748 targeting *NPC1* (s237197) were included per array. For all genes, “single-gene knockdown”
749 scenarios [siRNA_{geneA}+Neg9] were added on two spots per array. Each pairwise “combinatorial
750 knockdown” scenario [siRNA_{geneA}+siRNA_{geneB}] was analyzed on one spot per array, with a single
751 spot covering 50-100 informative cells^{28,57} (**Figure S1**).

752 In order to confirm GIs identified with the coRNAi screen, we replicated our analyses with
753 forward transfection in a liquid-phase format with Lipofectamine 2000 reagent in 12-well plates,
754 according to the manufacturer’s instructions. Concentrations of the siRNAs were adjusted to
755 mimic the single knockdown phenotypes from the screen (**Table S2**). 1µl of Lipofectamine 2000
756 was used per each transfection. GIs that showed a statistically significant interaction effects
757 ($p_{\text{fdr}} < 10^{-2}$) in replication analyses and acted in the same direction (same directionality of
758 interaction value) as in the coRNAi screen, were considered as validated (**Table S12**).

759

760 **Cell culture, transfection and LDL-uptake assay**

761 HeLa Kyoto cells were grown in DMEM medium (Gibco) supplemented with 10 % (w/v) fetal calf

762 serum (FCS)(PAA) and 2 mM L-glutamine (Sigma) at 37 °C with 5 % CO₂ and saturated
763 humidity. Cells were plated at a density of 6x10⁴ per plate on the cell microarrays for solid-phase
764 siRNA transfection⁵⁶ and cultivated for 48 hours before performing the LDL-uptake assay. For
765 liquid phase transfection-based validation experiments, cells were plated in 12-well plates the
766 day prior to transfection, and siRNA-transfected cells were cultivated for 48 hours. The assays to
767 monitor cellular uptake of fluorescently-labelled LDL (Dil-LDL) were performed as described in
768 more detail in previous publications^{17,27,28}. In brief, cells cultured in serum-free medium
769 (DMEM/2mM L-glutamine/0.2 % (w/v) BSA) and exposed to 1% 2-hydroxy-propyl-beta-
770 cyclodextrin for 45min were labelled with 50 µg/ml Dil-LDL (Invitrogen) for 30 min at 4 °C. Dil-
771 LDL uptake was stimulated for 20 min at 37.0 °C before washing off non-internalized dye for 1
772 min in acidic (pH 3.5) medium at 4 °C, fixation, and counterstaining for nuclei (Dapi) and cell
773 outlines (DRAQ5). For RNAi-based gene interaction screening, each of the five cell microarrays
774 was assayed in 7-10 biological replicas.

775

776 **Image acquisition and quality control**

777 Image acquisition was performed using an Olympus IX81 automated microscope with Scan^R
778 software and an UPlanSApo 20x/NA 0.40 air objective as described^{17,28}. Images from a total of
779 42 cell microarrays were visually quality controlled. Arrays with insufficient knockdown efficiency
780 where *INCENP* siRNA treated cells did not show the expected multinucleated phenotype in the
781 DAPI channel were excluded. Also arrays with plate effects as evaluated through diagnostic
782 plots with the *spot* function in R, and arrays where knockdown of *LDLR*, or *LDLR* together with
783 negative control siRNA Neg9, did not show a significant difference from controls, were discarded
784 as well. Following these QC criteria, 29 cell microarrays with a total of 11,047 image frames per
785 channel were further analysed. The in-house developed tool HTM Explorer (Ch. Tischer;
786 <https://github.com/embl-cba/shinyHTM>) was then used to select images fulfilling pre-defined
787 criteria for cell number, image sharpness quality, and image background intensity, resulting in a
788 total number of 9,539 (86.35%) QC-ed image frames that were used for subsequent analyses.

789

790 **Image analysis**

791 Automated image analysis was performed using a specifically developed pipeline (available
792 upon request) in the open source software CellProfiler⁵⁸ <http://www.cellprofiler.org> as
793 described^{17,28}. In brief, areas of individual cells were approximated by stepwise dilation of masks
794 on the DAPI (nuclei) and DRAQ5 (cell outlines) channels⁵⁹. For each individual cell, Dil-LDL
795 signal was determined from masks representing intracellular endosome-like vesicular areas that
796 were determined by local adaptive thresholding according to predefined criteria for size and
797 shape (**Figure S1**). Total fluorescence intensity of Dil-signal above local background per cell
798 mask was quantified, and means were calculated from all cells per image. Then, for each siRNA,
799 or siRNA combination (“*treated*”), signals from different images from the same biological replicate
800 were averaged and a robust Z-score was calculated using the median fluorescence signal of all
801 the negative control siRNAs per array (“*median(controls)*”) and by the median absolute deviation
802 of these controls (“*mad(controls)*”) as follows:
$$\text{robust Z-score} = \frac{I_F \text{treated} - \text{median} I_F \text{(controls)}}{\text{mad} I_F \text{(controls)}}$$
^{60,61}. A
803 median robust Z-score was calculated per treatment across all biological replicates and is
804 represented on the plots.

805

806 **RNAi gene interaction testing**

807 To identify pairs of genes for which simultaneous knock-down results in an additive or non-
808 additive gene interaction effects on LDL uptake we conducted a Robust Linear Model fitting in R.
809 RobustZScore values calculated from different biological replicates in the presence of single
810 ([siRNA_{geneA}+Neg9] and [siRNA_{geneB}+Neg9]) or double knock-downs ([siRNA_{geneA}+ siRNA_{geneB}])
811 were considered to be response variable value. Negative control values [Neg9] were included in
812 each fitted dataset to correctly account for baseline LDL uptake. The full regression model
813 considered in the study was

814
$$y = \beta_0 + \beta_A * x_A + \beta_B * x_B + \beta_{AB} * x_A * x_B + \square$$

815 which is equivalent to the short form of the statistical formula:

816
$$y \sim x_A + x_B + x_A * x_B$$

817 In both formulas y corresponds to the robustZscore values of measured LDL uptake; x_A , x_B are
818 encoded predictor variables, which are equal to 1 in case of presence of siRNA_{geneA}, siRNA_{geneB},
819 or both siRNAs accordingly and equal 0 otherwise. The ϵ is a noise term, which is minimised
820 during the fitting process. Model fitting provides estimates of β_0 , β_A , β_B and β_{AB} values. β_0 defines
821 the effect of the negative control on robustZscore values and can be also denoted as an
822 intercept of the linear fit. For our data β_0 is always close to 0 because of the robustZscore
823 definition. The β_A and β_B define individual effects of siRNA_{geneA} and siRNA_{geneB} accordingly. The
824 β_{AB} defines the interaction effect between genes A and B and represents the difference between
825 the observed robustZscore values in case of double knockdown y_{AB} and the expected additive
826 effect of geneA and geneB knockdown ($\beta_{AB} = y_{AB} - \beta_0 - \beta_A - \beta_B$).

827 Subsequently, two strategies were used to evaluate functional interactions for each gene pair
828 using defined statistical model:

829 First, to determine gene pairs for which genetic interactions and additive effects observed upon
830 combinatorial knockdowns, we compared fitting of the whole model to the fitting of reduced
831 model versions. Following models were compared:

832 Model 0 - (only baseline effect β_0 in case of either single or double knockdown): $y \sim 1$

833 Model 1 - effect of siRNA_{geneA} only: $y \sim x_A$

834 Model 2 - effect of siRNA_{geneB} only: $y \sim x_B$

835 Model 3 for additive effect of both siRNAs (additive GI): $y \sim x_A + x_B$

836 Model 4 – full model including genetic interaction (non-additive GI): $y \sim x_A + x_B + x_A * x_B$.

837

838 To determine the best model explaining the data for each gene pair we used Schwarz's
839 Bayesian Information Criterion (BIC). BIC score was calculated for each model fitted to the data,
840 then the model with the lowest BIC value (BIC*) was selected as the best-fitting model. Co-
841 knockdown effects of each gene pair were classified as aGIs or naGIs when model 3 or model 4
842 accordingly were defined to fit data best. Additionally, for the RNAi screen, we used the method
843 published by Raftery, 1995 to define the strength of evidence for the respective model to be
844 selected⁶². Namely, if the difference (ΔBIC) between the BIC value of the best fitting model (the

845 model with the lowest BIC value) and the BIC value of any other model is bigger than 2, then it
846 would indicate a significant evidence for this model (with BIC*) to truly represent the data. In
847 other words, if $\Delta\text{BIC} > 2$ then the model with lowest BIC value (BIC*) was considered as the one
848 most correctly describing the data in comparison to other tested models. If the $\Delta\text{BIC} < 2$, then two
849 models were considered as possible alternatives for representing the dataset.

850 Secondly, to estimate statistical significance of gene interaction effect for each siRNA gene
851 combination, we calculated a p-value from the t-value of the linear regression model term,
852 describing genetic interaction (β_{AB}) as $p_{\text{Val}} = 2 - 2 * p_{\text{norm}}(\text{abs}(t_{\text{Val}}))$. To correct for multiple
853 comparisons, the p-values were adjusted using the false discovery rate (fdr) method⁶³, and the
854 fdr-corrected p-values $< 10^{-2}$ were considered to correspond to significant GIs.

855

856 RT-qPCR analysis

857 Cell lysis and total RNA extraction was done using the RNease Mini Kit (Qiagen). Reverse-
858 transcription was performed with the SuperScript™ III First-Strand Synthesis SuperMix for RT-
859 qPCR (Invitrogen). RT-qPCR data was obtained from three biological replicates/siRNA
860 treatment. Primers can be provided upon request. For each siRNA treatment target mRNA was
861 normalized to that of GAPDH and compared to CTRL siRNA and the log2 of fold change ($2^{\Delta\Delta\text{CT}}$)
862 was calculated (see **Figure S4**). Results were considered as significant if p-values were below
863 0.05 in a two-tailed Student's t-test.

864

865 Immunocytochemistry and confocal microscopy

866 Cells were transfected and cultured as described above then fixed with 3% paraformaldehyde
867 (PFA) at room temperature for 20 min, washed with PBS, incubated with 30mM glycine for 5 min
868 and washed again with PBS. For LDLR staining cells were permeabilized with 0.05% Filipin III
869 (Sigma #F4767) in 10% FCS for 30 min at room temperature. Primary antibody: rabbit
870 monoclonal anti-LDLR (Fitzgerald #20R-LR002) was diluted in 5% FCS overnight at 4 °C.
871 Secondary antibody: goat polyclonal goat anti-rabbit IgG Alexa 568 (Invitrogen #A11011) was
872 diluted 1:400 in 5% FCS. Fixed cells were imaged using a Zeiss LSM 780 confocal microscope

873 using a 63x/NA 1.4 oil objective.

874

875 **FIGURE LEGENDS**

876 **Figure 1**

877 **PTV-burden tests in UK Biobank establish additive GIs for PCSK9-APOB and LPL-APOB**

878 **(a)** Workflow of the study. 30 high-confidence candidate genes for GI testing were chosen from
879 18 GWAS regions associated with blood lipid traits or CAD risk based on colocalization analyses
880 with eQTL/pQTL signals and previously reported lipid-regulatory functions (see Methods).
881 Pairwise GI analyses were conducted from three complementary datasets: protein-truncating
882 variants (PTVs) from exome sequencing in the UK Biobank; lipid/CAD GWAS lead SNPs; and
883 combinatorial RNAi (coRNAi) experiments in cells. Robust linear model fitting was used to
884 identify additive (aGI) and non-additive (naGI) GIs, and genetic and functional data were
885 integrated. **(b)** Gene-based PTV-burden analyses from 161,508 exomes identified an additive GI
886 (aGI) for LDL (and TC; not shown) between *PSCK9* and *APOB*. **(c)** A suggestive non-additive GI
887 (naGI) for HDL (and TG; not shown) between PTVs in *LPL* and *APOB* was validated as aGI for
888 HDL, TG, and also LDL by replication analyses in an additional 79,462 UK Biobank exomes
889 (Table S6). n, number of carriers. (-), predicted loss-of-function due to PTVs.

890

891 **Figure 2**

892 **Pairwise GIs between lipid and CAD GWAS lead SNPs in 387,033 UK Biobank participants**

893 **(a-e)** Circos plots showing aGIs (grey) and naGIs (colored) between GWAS lead SNPs (blue) at
894 the 28 selected lipid/CAD loci (red) for the four tested lipid species and CAD. **(f)** Tests for GIs
895 between polygenic risk scores for the four lipid species and PTV-burden for each of the 30 lipid
896 genes identified a naGI between PTV-burden in *LPL* and the PRS for TG. PRS distribution
897 (mean±SD) for *LPL*-PTV carriers (pink) and non-carriers (blue) are plotted against mean
898 normalized residual TG levels. Each dot reflects mean TG levels at a respective percentile.

899

900 **Figure 3**

901 **Combinatorial RNAi identifies pairwise GIs modulating cellular LDL-uptake**

902 **(a)** coRNAi screen workflow. Customized cell microarrays were generated by pairwise spotting of
903 siRNAs against two different candidate genes on 384 spots/array for solid-phase reverse siRNA
904 transfection of cultured HeLa cells. Cells challenged to internalize fluorescent-labelled LDL (Dil-
905 LDL) over a period of 20 min were imaged on a high-content microscope. Integrated
906 fluorescence intensities for each cell individually were quantified by automated image analysis.
907 Averaged signal intensities per gene pair were tested for GIs in multiple replica experiments per
908 array. GIs suggested in the coRNAi screen as potentially non-additive were subsequently
909 validated in customized experiments using fluid-phase transfection. **(b)** Heatmap visualizing
910 median robust Z-score distribution upon coRNAi of 435 gene pairs assessed for their impact on
911 cellular LDL-uptake. Red, increase. Blue, decrease. CTRL (top row and first column) reflects the
912 relative impact on LDL-uptake when candidate genes were silenced individually (siRNA_{geneAorB} +
913 negative control siRNA). **(c)** 20 gene pairs validated as either buffering or synergistic naGIs on
914 cellular LDL-uptake in independent replica experiments, sorted according to effect size.
915 Interaction Value (right graph) depicts the directionality and difference of the combined effect
916 versus single knockdown effects. **(d-f)** Selected examples of single gene (siRNA_{geneA} + negative
917 control siRNA) and gene pair (siRNA_{geneA+geneB}) siRNA knockdown effects on relative
918 fluorescently-labelled LDL (Dil-LDL) cellular uptake. CTRL, control siRNA. Boxplots represent
919 values between 25th and 75th percentile, whiskers indicate largest value within 1.5 times
920 interquartile range above 75th percentile. Median value is highlighted in the boxplot as a
921 horizontal line. Dots represent robust Z-score values calculated for integrated Dil fluorescence
922 intensities per cell (see Methods). Scale bar=10 µm.

923

924 **Figure 4**

925 **Integrative analysis identifies pairwise GIs supported by both, genetic and functional data**
926 Overlap of GIs identified through genetic analyses and coRNAi. Highlighted are gene pairs
927 identified through either PTV-SNP **(a, b)** or SNP-SNP **(c,d)** GI testing for which pairwise siRNA-
928 knockdown showed corresponding effects on cellular LDL-uptake, validating these GIs as either
929 aGI (a,c) or naGI (b,d). **(e)** TOMM40 as an example for which, consistent with SNP-SNP

930 analyses, siRNA knockdown revealed buffering naGIs when jointly silenced with *SORT1* (left
931 panel) or *NCAN* (right panel). Values on the graphs reflect robust Z-scores values calculated for
932 total intensity of Dil-LDL per cell averaged per image (see Methods). Boxplots represent values
933 between 25th and 75th percentile, whiskers indicate largest value within 1.5 times interquartile
934 range above 75th percentile. Median value is highlighted in the boxplot as a horizontal line. Dots
935 represent robust Z-score values calculated for integrated Dil fluorescence intensities per cell
936 (see Methods). Scale bar= 10 μ m.

937 **TABLES**

938 **Table 1.**
 939 **Non-additive GIs from pairwise PTV-burden and GWAS lead SNP-based GI testing in UK Biobank**
 940

Trait	Gene 1	Gene 2	BIC_Best Model	Lowest ΔBIC	incl. 19q13.32	cis GI	trans GI	MAF (SNP1) N_PTV1 carriers	MAF (SNP2) N_PTV2 carriers
LDL SNP-SNP GIs:									
LDL	NCAN (rs2228603)	TM6SF2 (rs58542926)	4	85.34		+		0.076	0.076
LDL	TM6SF2 (rs58542926)	APOE (rs4420638)	4	63.34	+	(+)		0.076	0.191
LDL	TM6SF2 (rs58542926)	TOMM40 (rs2075650)	4	41.95	+	(+)		0.076	0.147
LDL	BCAM (rs118147862)	APOE (rs4420638)	4	34.25	+	+		0.046	0.191
LDL	NCAN (rs2228603)	APOE (rs4420638)	4	32.61	+	(+)		0.076	0.191
LDL	NCAN (rs2228603)	TOMM40 (rs2075650)	4	30.02	+	(+)		0.076	0.147
LDL	BCAM (rs118147862)	TOMM40 (rs2075650)	4	28.81	+	+		0.046	0.147
LDL	ZNF259 (rs2075290)	APOE (rs4420638)	4	4.68	+		+	0.068	0.191
LDL	CBLC (rs3208856)	APOE (rs4420638)	4	3.34	+	+		0.036	0.191
LDL	SORT1/CELSR2 (rs629301)	TOMM40 (rs2075650)	4	0.57	+		+	0.222	0.147
HDL SNP-SNP GIs:									
HDL	BCAM (rs118147862)	PVRL2 (rs7254892)	4	1.47		+		0.046	0.031
TG SNP-SNP GIs:									
TG	ZNF259 (rs2075290)	SIK3 (rs6589574)	4	31.81		+		0.068	0.084
TG	BCAM (rs118147862)	PVRL2 (rs7254892)	4	21.41	+	+		0.046	0.031
TG	CBLC (rs3208856)	BCAM (rs118147862)	4	20.45	+	+		0.036	0.046
TG	ZNF259 (rs2075290)	PAFAH1B2 (rs4936367)	4	17.82		+		0.068	0.1
TG	LPL (rs12678919)	ZNF259 (rs2075290)	4	13.81			+	0.098	0.068
TG	LPL (rs12678919)	SIK3 (rs6589574)	4	3.22			+	0.098	0.084
TC SNP-SNP GIs:									
TC	NCAN (rs2228603)	TM6SF2 (rs58542926)	4	74.81		+		0.076	0.076
TC	TM6SF2 (rs58542926)	APOE (rs4420638)	4	53.33	+	(+)		0.076	0.191
TC	TM6SF2 (rs58542926)	TOMM40 (rs2075650)	4	38.17	+	(+)		0.076	0.147
TC	NCAN (rs2228603)	APOE (rs4420638)	4	30.93	+	+		0.076	0.191
TC	NCAN (rs2228603)	TOMM40 (rs2075650)	4	28.59	+	+		0.076	0.147
TC	BCAM (rs118147862)	TOMM40 (rs2075650)	4	11.24	+	+		0.046	0.147
TC	BCAM (rs118147862)	APOE (rs4420638)	4	9.05	+	+		0.046	0.191
TC	ZNF259 (rs2075290)	APOE (rs4420638)	4	2.40	+		+	0.147	0.191
TC	SORT1/CELSR2 (rs629301)	TOMM40 (rs2075650)	4	0.31	+		+	0.222	0.147
LDL PTV-SNP GIs:									
LDL	LDLR	PVRL2 (rs7254892)	4	8.318	+		+	33	0.031
HDL PTV-SNP GIs:									
HDL	APOB	LPL (rs12678919)	4	1.601			+	222	0.098
TC PTV-SNP GIs:									
TC	LDLR	PVRL2 (rs7254892)	4	18.014	+		+	33	0.031
TC	LDLR	SIK3 (rs6589574)	4	2.961			+	33	0.084
TC	LDLR	PAFAH1B2 (rs4936367)	4	0.405			+	33	0.1
TG PTV-SNP GIs									
TG	LPL	SIK3 (rs6589574)	4	8.4			+	31011	0.084
TG	BAZ1B	PAFAH1B2 (rs4936367)	4	5.825			+	25	0.1
TG	LPL	ZNF259 (rs2075290)	4	3.894			+	31011	0.147
TG	BAZ1B	NCAN (rs2228603)	4	2.845			+	25	0.076
TG	LPL	PAFAH1B2 (rs4936367)	4	1.648			+	31011	0.1
TG	BAZ1B	TM6SF2 (rs58542926)	4	1.522			+	25	0.076

941
 942 Non-additive genetic interactions (naGIs) identified through GWAS lead SNP- and PTV-SNP-based GI analyses in the
 943 UK Biobank as described in Methods. BIC, Bayesian Information Criterion. A lowest ΔBIC of "4" indicates interaction
 944 model is most compatible with a non-additive interaction effect. MAF (minor allele frequency) estimates and numbers
 945 of rare protein-truncating variant (PTV) carriers are based on genotypes from 387,033 and exomes, respectively, from
 946 161,508 unrelated UK Biobank participants of European ancestry. (+) indicates possible cis-effects of rs4420638 in
 947 APOE on neighbouring genes on Chr.19q13.32. Trans GI indicates genes contributing to pairwise naGIs are located
 948 on different chromosomes. * represents data based on the replication study in additional 79,462 UK Biobank
 949 participants.
 950

951

952

953 **Table 2.**

GI Gene Pair		Interaction type	BIC-based GI testing		RLMF-based GI testing			Validation RNAi GI testing		
Gene1	Gene2		Lowest ΔBIC	Best Model	Robust Zscore	Interaction Value	pVal(fdr)	Robust Zscore	Interaction Value	pVal (fdr)
<i>APOB</i>	<i>HMGCR</i>	Synergistic	4.62	2	2.8	2.18	4.05E-03	3.33	0.97	1.53E-03
<i>HAVCR1</i>	<i>LDLR</i>		4.31	4	-2.18	-1.32	1.68-02	-2.24	-1.24	1.81E-10
<i>LDLR</i>	<i>NCAN</i>		2.36	4	-1.86	-1.28	7.64E-03	-1.23	-2.22	2.20E-11
<i>MYBPHL</i>	<i>SIK3</i>		2.27	4	2.17	1.59	6.59E-03	2.3	0.96	7.91E-03
<i>PAFAH1B1</i>	<i>SIK3</i>		4.49	4	1.62	1.79	2.52E-03	3.3	2.44	8.37E-12
<i>PCSK9</i>	<i>TMEM57</i>		3.82	4	1.46	1.76	2.37E-03	3.21	2.44	3.74E-11
<i>BCAM</i>	<i>LDLRAP1</i>	Buffering	4.70	4	-0.4	-1.83	3.58E-04	-0.05	-0.7	5.20E-03
<i>CELSR2</i>	<i>LPL</i>		0.04	0/2	-1.4	-1.43	7.53E-03	-0.11	-0.86	8.70E-05
<i>CXCL12</i>	<i>PAFAH1B1</i>		5.47	4	-2.17	-1.78	3.69E-04	1.32	-1.92	9.84E-13
<i>HAVCR1</i>	<i>LDLRAP1</i>		13.70	4	-0.56	-2.16	1.03E-05	0.39	-0.73	5.40E-03
<i>HAVCR1</i>	<i>MLXIPL</i>		13.55	4	-0.63	-2.31	1.03E-05	-0.12	-1.2	2.06E-05
<i>HAVCR1</i>	<i>SEZ6L</i>		8.54	4	-1.19	-1.91	3.39E-04	-0.67	-0.77	2.90E-04
<i>HAVCR1</i>	<i>SORT1</i>		11.66	4	-2.09	-2.24	9.34E-05	-0.63	-0.58	1.98E-03
<i>LDLR</i>	<i>LDLRAP1</i>		10.06	4	-2.44	-1.86	9.34E-05	-2.49	-0.92	2.36E-10
<i>LDLR</i>	<i>MLXIPL</i>		5.79	4	-2.03	-1.49	2.97E-03	-2.26	-0.78	4.19E-11
<i>LDLRAP1</i>	<i>SORT1</i>		8.13	4	-1.11	-1.65	2.57E-03	-1.09	-0.95	3.89E-06
<i>MLXIPL</i>	<i>TOMM40</i>		18.73	4	-2.14	-2.67	1.03E-05	-1.94	-0.85	3.82E-08
<i>NCAN</i>	<i>SEZ6L</i>		5.08	4	-0.59	-1.6	5.54E-03	-0.08	-1.56	7.94E-12
<i>NCAN*</i>	<i>TOMM40*</i>		2.76	4	-1.4	-1.56	7.17E-03	-1.49	-1.34	1.75E-08
<i>SORT1*</i>	<i>TOMM40*</i>		5.80	0	0.77	1.84	3.90E-03	-2.48	1.77	0.00E+00

954 **Pairwise GIs identified and validated through coRNAi to impact LDL-uptake into cells**

955

956 Gene pairs identified and independently validated by combinatorial RNAi (coRNAi) as impacting the uptake of
 957 fluorescent-labelled LDL into cells in a non-additive manner. Both, *BIC*, Bayesian Information Criterion and *RLMF*,
 958 Robust Linear Model Fitting were applied for analysis of coRNAi-based GI-testing as described in Methods. *, gene
 959 pairs that both, genetic and coRNAi GI-testing identify as naGIs.

960

961

962

963

964

965

966

967

968

969

970

971

972

973

974

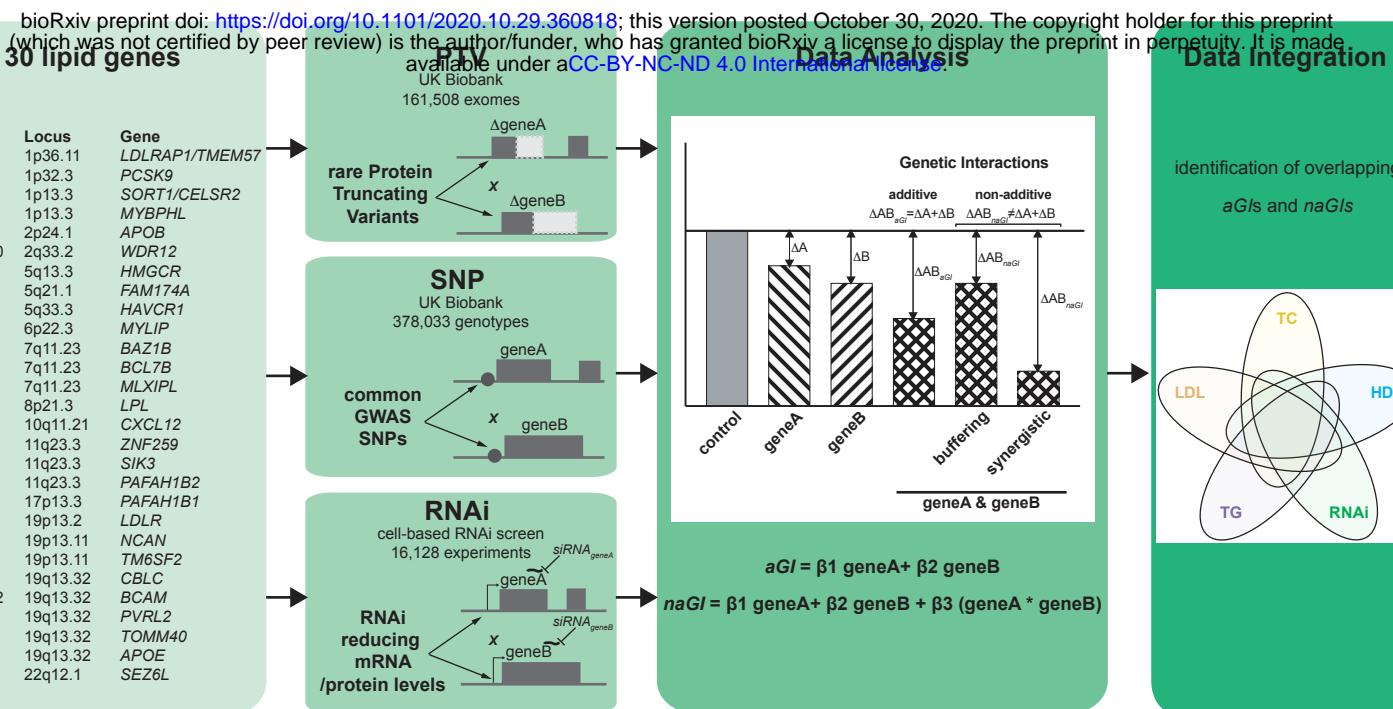
975

976

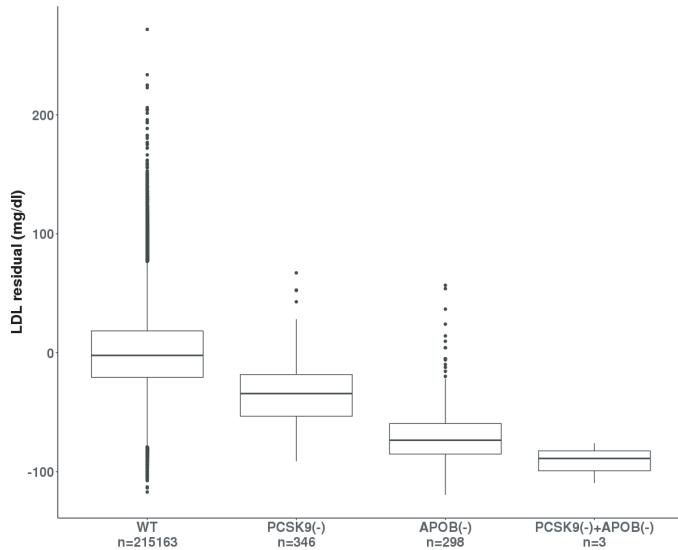
977
978
979
980
981
982
983
984
985
986
987
988
989

Figure 1

a



b



c

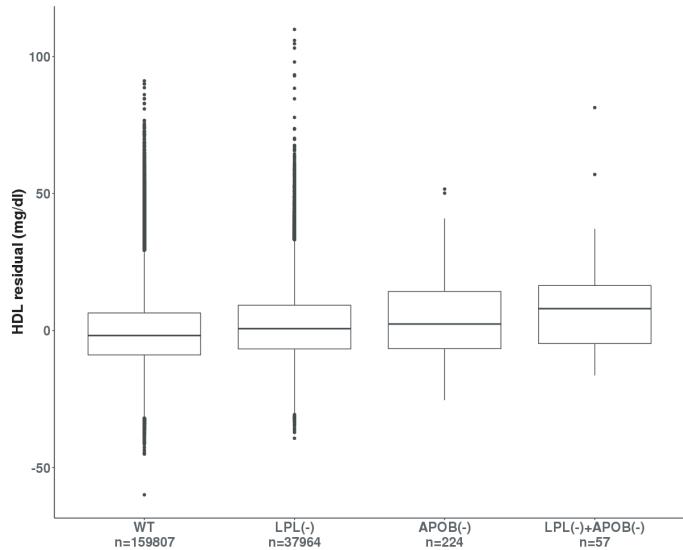


Figure 2

a

LDL

bioRxiv preprint doi: <https://doi.org/10.1101/2020.10.29.360818>; this version posted October 30, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

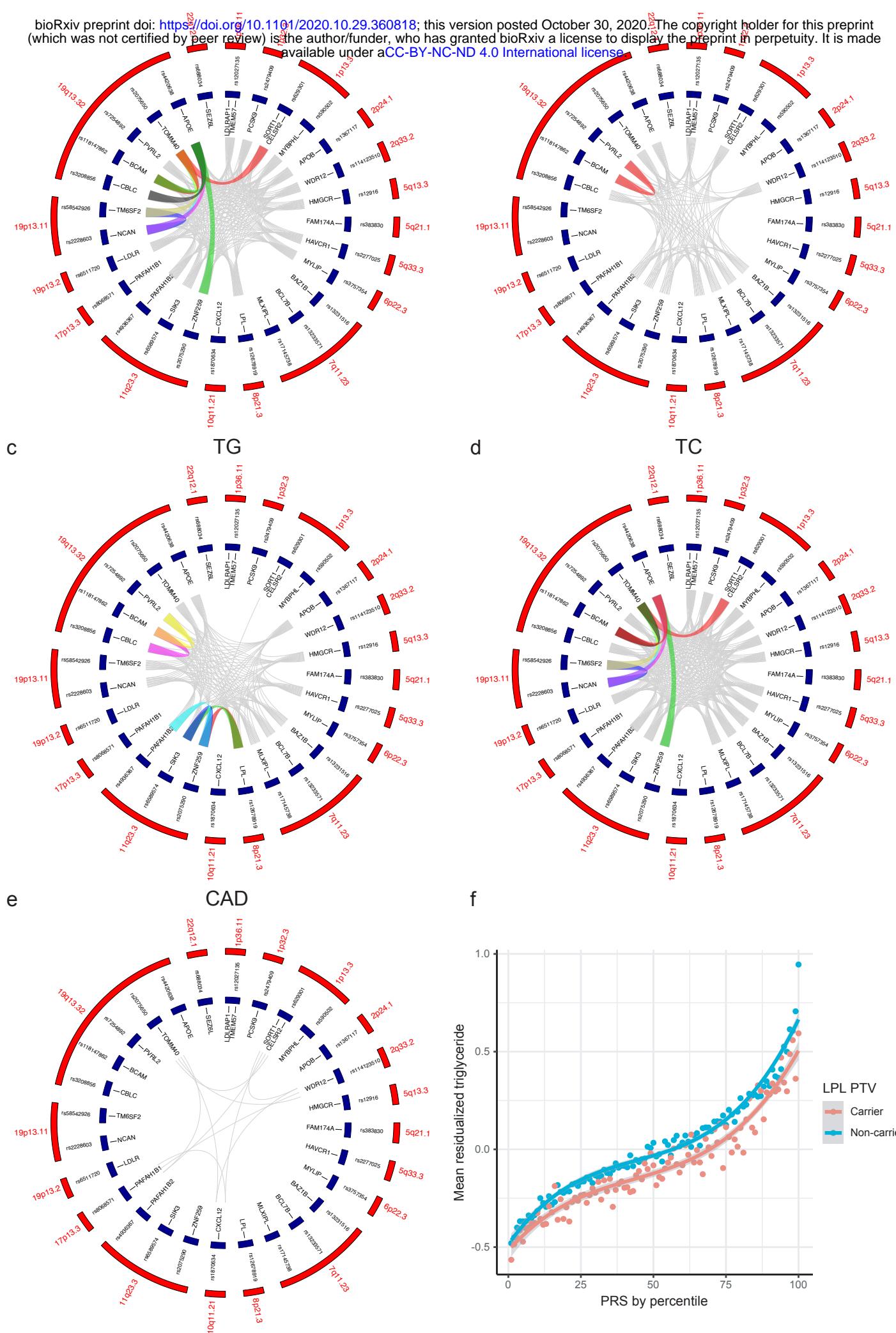
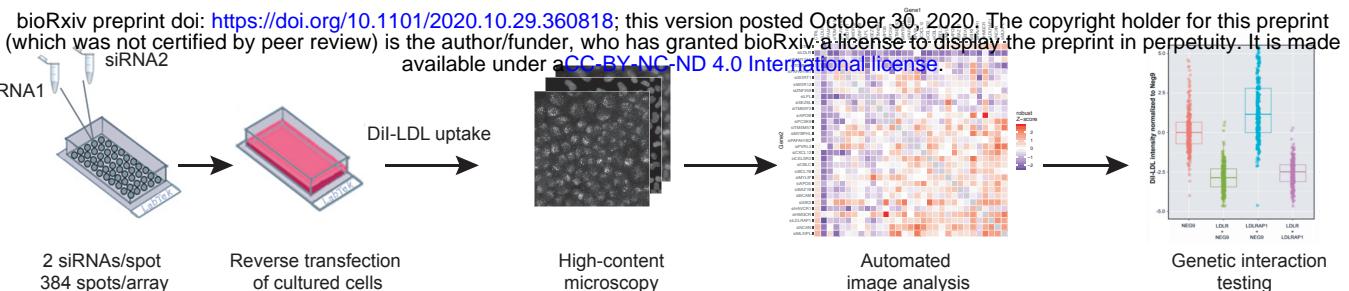
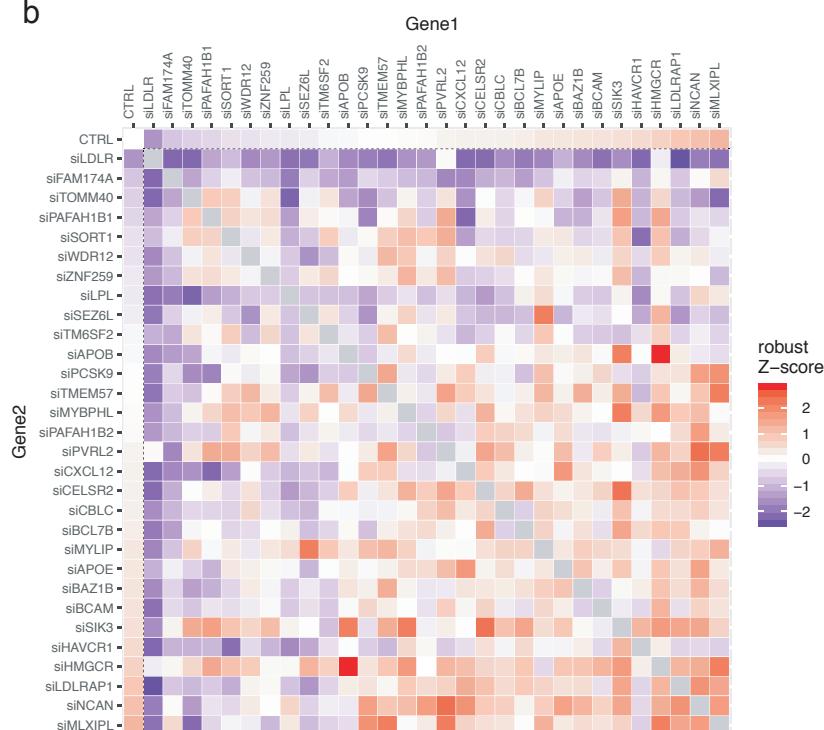


Figure 3

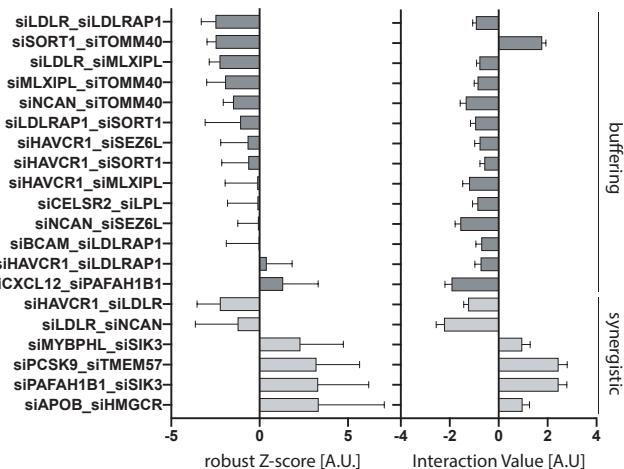
a



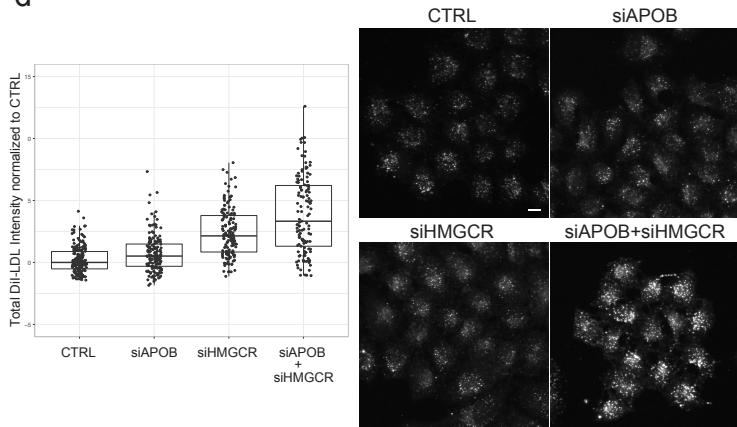
b



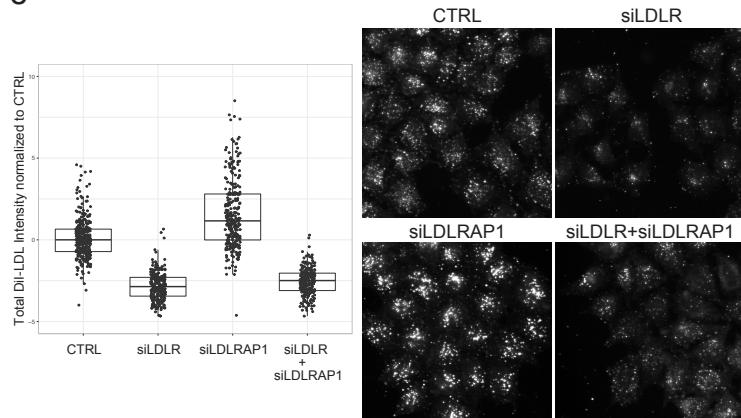
c



d



e



f

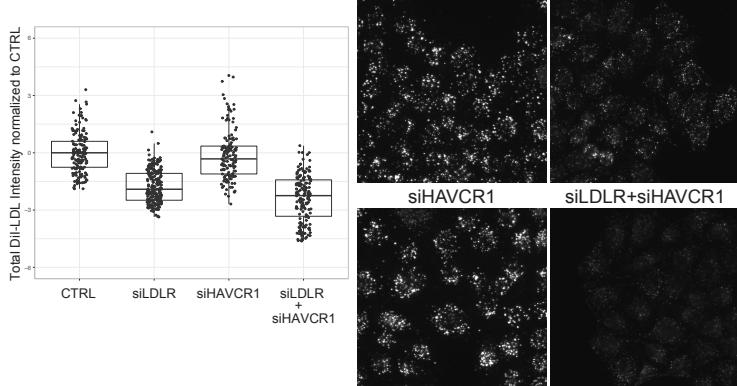


Figure 4

