

Loss-of-function of *MFGF8* and protection against coronary atherosclerosis

Sanni E. Ruotsalainen¹, Ida Surakka², Nina Mars¹, Juha Karjalainen³, Mitja Kurki³, Masahiro Kanai³⁻⁵, Kristi Krebs⁶, Pashupati P. Mishra⁷⁻⁹, Binisha H. Mishra⁷⁻⁹, Juha Sinisalo¹⁰, Priit Palta^{1,6}, Terho Lehtimäki⁷⁻⁹, Olli Raitakari¹¹⁻¹³, *Estonian Biobank research team*, Lili Milani⁶, *The Biobank Japan Project*¹⁴, Yukinori Okada^{5,15}, *FinnGen*, Aarno Palotie^{1,3}, Elisabeth Widen¹, Mark J. Daly^{1,3,4}, Samuli Ripatti^{1,3,16*}

- 1) *Institute for Molecular Medicine Finland (FIMM), HiLIFE, University of Helsinki, Helsinki, Finland*
- 2) *Department of Internal Medicine, University of Michigan, Ann Arbor, Michigan, USA*
- 3) *The Broad Institute of MIT and Harvard, Cambridge, Massachusetts, USA*
- 4) *Analytic and Translational Genetics Unit, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts, USA*
- 5) *Department of Statistical Genetics, Osaka University Graduate School of Medicine, Suita, Japan*
- 6) *Estonian Genome Centre, Institute of Genomics, University of Tartu, Tartu, Estonia*
- 7) *Department of Clinical Chemistry, Faculty of Medicine and Health Technology, Tampere University, Tampere, Finland*
- 8) *Finnish Cardiovascular Research Centre, Faculty of Medicine and Health Technology, Tampere University, Tampere, Finland*
- 9) *Department of Clinical Chemistry, Fimlab Laboratories, Tampere, Finland*
- 10) *Heart and Lung Center, Helsinki University Hospital and Helsinki University, Helsinki, Finland*
- 11) *Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku, Turku, Finland*
- 12) *Centre for Population Health Research, University of Turku, Turku University Hospital, Turku, Finland*
- 13) *Department of Clinical Physiology and Nuclear Medicine, University of Turku, Turku, Finland*
- 14) *Institute of Medical Science, The University of Tokyo*
- 15) *Laboratory for Statistical Analysis, RIKEN Center for Integrative Medical Sciences, Yokohama, Japan*
- 16) *Department of Public Health, Clinicum, Faculty of Medicine, University of Helsinki, Helsinki, Finland*

*Corresponding author

Tel: +358 40 567 0826

E-mail: samuli.ripatti@helsinki.fi

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

Abstract

Cardiovascular diseases are the leading cause of premature death and disability worldwide, with both genetic and environmental determinants. While genome-wide association studies have identified multiple genetic loci associated with cardiovascular diseases, exact genes driving these associations remain mostly uncovered. Due to Finland's population history, many deleterious and high-impact variants are enriched in the Finnish population giving a possibility to find genetic associations for protein-truncating variants that likely tie the association to a gene and that would not be detected elsewhere.

In FinnGen, a large Finnish biobank study, we identified an inframe insertion rs534125149 in *MFGE8* to have protective effect against coronary atherosclerosis (OR = 0.75, $p = 2.63 \times 10^{-16}$) and related endpoints. This variant is highly enriched in Finland (70-fold compared to Non-Finnish Europeans) with allele frequency of 3% in Finland. The protective association was replicated in meta-analysis of biobanks of Japan and Estonian (OR = 0.75, $p = 5.41 \times 10^{-7}$).

Additionally, we identified a splice acceptor variant rs201988637 in *MFGE8*, independent of the rs534125149 and similarly protective in relation to coronary atherosclerosis (OR = 0.72, $p = 7.94 \times 10^{-06}$) and related endpoints, with no significant risk-increasing associations. The protein-truncating variant was also associated with lower pulse pressure, pointing towards a function of *MFGE8* in arterial stiffness and aging also in humans in addition to previous evidence in mice. In conclusion, our results show that inhibiting the production of lactadherin could lower the risk for coronary heart disease substantially.

Keywords: coronary atherosclerosis, myocardial infarction, loss-of-function, *MFGE8*

25 Introduction

26 Cardiovascular disease (CVD) is the leading cause of premature death and disability worldwide, with
 27 both genetic and environmental determinants^{1, 2}. The most common cardiovascular disease is
 28 coronary heart disease (CHD), including coronary atherosclerosis and myocardial infarction, among
 29 others. While genome-wide association studies (GWAS) have identified multiple genetic loci
 30 associated with cardiovascular diseases, exact genes driving these associations remain mostly
 31 uncovered³.

32
 33 Due to Finland's population history, many deleterious and high-impact variants are enriched in the
 34 Finnish population giving a possibility to find genetic associations that would not be detected
 35 elsewhere⁴. Many studies have reported high-impact loss-of-function (LoF) variants associated with
 36 risk factors for CVD, such as blood lipid levels, thus impacting on the CVD risk remarkably. For
 37 example, high-impact LoF variants in genes *LPA*⁴, *PCSK9*⁵, *APOC3*⁶ and *ANGPTL4*⁷ have been
 38 shown to be associated with Lipoprotein(a), LDL-cholesterol (LDL-C) or triglyceride levels and
 39 lowering the CVD risk.

40
 41 Besides blood lipids, other risk factors for CVD include hypertension, smoking and the metabolic
 42 syndrome cluster components. The mechanism that links these risk factors to atherogenesis, however,
 43 remains incompletely elucidated. Many, if not all, of these risk factors, however, also participate in
 44 the activation of inflammatory pathways, and inflammation in turn can alter the function of artery
 45 wall cells in a manner that drives atherosclerosis⁸.

46
 47 Using data from a sizeable Finnish biobank study FinnGen (n = 260 405), we identified an inframe
 48 insertion rs534125149 in *MFG8* to protect against coronary atherosclerosis and related endpoints,
 49 such as myocardial infarction (MI). This variant is highly enriched in Finland, 70- fold compared to

Non-Finnish Europeans (NFE) in the gnomAD genome reference database⁹ with AF of 3% in Finland. This association was also replicated in BioBank Japan (BBJ) and Estonian Biobank (EstBB). We also identified a splice acceptor variant rs201988637 in the same gene, which is also protective against coronary atherosclerosis-related endpoints, indicating that rs534125149 act as a loss-of-function variant in *MFGE8*. Associations of functional variants in *MFGE8* were specific to coronary atherosclerosis-related endpoints, and they did not significantly ($p < 1.75 \times 10^{-5}$) increase risk for any other disease, highlighting *MFGE8* as a potential drug target candidate.

Material and methods

Study cohort and data

We studied total of 2 861 disease endpoints in Finnish biobank study FinnGen (n = 260 405) (**Table 1**). FinnGen (<https://www.finngen.fi/en>) is a large biobank study that aims to genotype 500 000 Finns and combine this data with longitudinal registry data, including national hospital discharge, death, and medication reimbursement registries, using unique national personal identification numbers. FinnGen includes prospective epidemiological and disease-based cohorts as well as hospital biobank samples.

Definition of disease endpoints

All the 2 861 disease-endpoint analysed in FinnGen have been defined based on registry linkage to national hospital discharge, death, and medication reimbursement registries. Diagnoses are based on International Classification of Diseases (ICD) codes and have been harmonized over ICD codes 8, 9 and 10. More detailed lists of the ICD codes used for the disease-endpoints myocardial infarction and coronary atherosclerosis, which are discussed more in this study, are in Supplementary Material. A

complete list of endpoints analysed, and their definitions is available at <https://www.finnngen.fi/en/researchers/clinical-endpoints>.

Table 1: Basic characteristics of the study cohort.

	All	Females	Males
N (%)	260 405	147 061 (56.47 %)	113 344 (43.53 %)
Age (mean (sd))	53.15 (17.55)	51.84 (17.71)	54.85 (17.19)
BMI (mean (sd)) ^a	27.29 (5.36)	27.21 (5.83)	27.38 (4.76)
Statin use (N (%))	86 466 (33.2 %)	40 422 (27.48 %)	46 044 (40.62 %)
Hypertension (N (%))	68 005 (26.11 %)	33 420 (22.72 %)	34 585 (30.51 %)
Smoking (N (%)) ^b	1 733 (1.07 %)	901 (0.96 %)	832 (1.22 %)
Coronary atherosclerosis	28 598 (11.38 %)	9 252 (6.87 %)	19 346 (17.86 %)
Myocardial infarction	14 305 (6.04 %)	3 958 (2.87 %)	10 347 (10.42 %)

^a BMI is available only from 178 966 individuals

^b Smoking information is available only from 98 654 individuals

Genotyping and imputation

FinnGen samples were genotyped with multiple Illumina and Affymetrix arrays (Thermo Fisher Scientific, Santa Clara, CA, USA). Genotype calls were made with GenCall and zCall algorithms for Illumina and AxiomGT1 algorithm for Affymetrix chip genotyping data batchwise. Genotyping data produced with previous chip platforms were lifted over to build version 38 (GRCh38/hg38) following the protocol described here: dx.doi.org/10.17504/protocols.io.nqtdown. Samples with sex discrepancies, high genotype missingness (> 5%), excess heterozygosity ($\pm 4SD$) and non-Finnish ancestry were removed. Variants with high missingness (> 2%), deviation from Hardy–Weinberg equilibrium ($P < 1 \times 10^{-6}$) and low minor allele count ($MAC < 3$) were removed.

90

91 Pre-phasing of genotyped data was performed with Eagle 2.3.5
 92 (<https://data.broadinstitute.org/alkesgroup/Eagle/>) with the default parameters, except the number of
 93 conditioning haplotypes was set to 20,000. Imputation of the genotypes was carried out by using the
 94 population-specific Sequencing Initiative Suomi (SISu) v3 imputation reference panel with Beagle
 95 4.1 (version 08Jun17.d8b, https://faculty.washington.edu/browning/beagle/b4_1.html) as described
 96 in the following protocol: [dx.doi.org/10.17504/protocols.io.nmndc5e](https://doi.org/10.17504/protocols.io.nmndc5e). SISu v3 imputation reference
 97 panel was developed using the high-coverage (25–30x) whole-genome sequencing data generated at
 98 the Broad Institute of MIT and Harvard and at the McDonnell Genome Institute at Washington
 99 University, USA; and jointly processed at the Broad Institute. Variant callset was produced with
 100 Genomic Analysis Toolkit (GATK) HaplotypeCaller algorithm by following GATK best practices
 101 for variant calling. Genotype-, sample- and variant-wise quality control was applied in an iterative
 102 manner by using the Hail framework v0.2. The resulting high-quality WGS data for 3 775 individuals
 103 were phased with Eagle 2.3.5 as described above. As a post-imputation quality control, variants with
 104 INFO score < 0.7 were excluded.

105

106 Association testing

107 A total of 260 405 samples from FinnGen Data Freeze 6 with 2 861 disease endpoints were analyzed
 108 using Scalable and Accurate Implementation of Generalized mixed model (SAIGE), which uses
 109 saddlepoint approximation (SPA) to calibrate unbalanced case-control ratios¹⁰. Models were adjusted
 110 for age, sex, genotyping batch and first ten principal components. All variants reaching genome-wide
 111 significance p-value threshold of 5×10^{-8} are considered as genome-wide significant (GWS), and all
 112 disease- endpoints reaching multiple testing corrected (for the number of endpoints tested = 2 861)
 113 p-value threshold of $0.05/2\,861 = 1.75 \times 10^{-5}$ were considered as phenome-wide significant (PWS).

114

115 Independent GWS loci for atherosclerosis were determined as adding $\pm 0.5\text{Mb}$ around each variant
 116 that reach the genome-wide significance threshold, overlapping regions were merged. The NHGRI-
 117 EBI GWAS Catalog¹¹ was used for assessing the novelty of the independent loci with any CVD-
 118 related endpoint or traditional risk factor for CVD, such as blood lipids, BMI and blood pressure. All
 119 novel loci for CVD were fine-mapped using FINEMAP¹² to determine the credible sets in each signal.
 120

121 In Corogene¹³ (n = 5 300), a sub-cohort of FinnGen where participants have been collected as patients
 122 with coronary artery disease CAD) and other related heart diseases, we tested the association of
 123 rs534125149 with sub-types of coronary heart disease (acute coronary syndrome, stable coronary
 124 heart disease (CHD) and other heart attacks). The acute coronary syndrome was further divided into
 125 unstable Angina pectoris, non-ST segment elevation myocardial infarction (NSTEMI) and ST-
 126 segment elevation myocardial infarction (STEMI). Associations were tested by calculating risk ratios
 127 (RR) for carriers vs. non-carriers of rs534125149 using non-CHD group always as controls and
 128 excluding the other tested groups from the analysis. P-values were calculated using χ^2 - test, and p-
 129 values < 0.05 were considered significant.

130

131 Survival analysis

132
 133

134 Survival analysis for coronary atherosclerosis and myocardial was performed using GATE¹⁴, which
 135 accounts for both population structure and sample relatedness and controls type I error rates even for
 136 phenotypes with extremely heavy censoring. GATE transforms the likelihood of a multivariate
 137 Gaussian frailty model to a modified Poisson generalized linear mixed model (GLMM^{15, 16}), and to
 138 obtain well-calibrated p-values for heavily censored phenotypes, GATE uses the SPA to estimate the
 139 null distribution of the score statistic. For coronary atherosclerosis and myocardial infarction, survival
 140 time from birth to first diagnose was analyzed for both rs534125149 and rs201988637. Models were

adjusted for age, sex, genotyping batch and first ten principal components, similarly to original GWAS analyses.

Replication and biomarker analyses

We tested the association of the two *MFGE8* variants (rs534125149 and rs201988637) with quantitative measurements of cardiometabolic relevance or known risk factors for CVD in two sub-cohorts of FinnGen, the population-based national FINRISK study¹⁷ (n = 26 717) and GeneRISK¹⁸ (n = 7 239). The associations were tested across 66 quantitative measurements of cardiometabolic relevance in FINRISK, and for 158 sub-lipid species in GeneRISK. In Young Finns Study (YFS)¹⁹ cohort (n = 1 934), we tested the association of the two variants with three measurements of arterial relevance (carotid artery distensibility, pulse wave velocity and pulse pressure).

In addition to Finnish cohorts described above, we tested the association of the two variants in Estonian Biobank data (EstBB)^{20,21}, BioBank Japan (BBJ)^{22,23} and UK Biobank (UKBB)²⁴. In EstBB (n = 51 388-137 722) we tested the associations of both variants with body mass index (BMI), systolic and diastolic blood pressure (SBP and DBP) and pulse pressure (PP), in BBJ in we tested the association of rs534125149 with 17 known quantitative risk factors for CVD and lastly, in the UKBB we tested the association of rs201988637 with 79 measurements of cardiometabolic relevance. In all of these biomarker analyses, a linear regression model adjusted for age and sex was used and for all quantitative risk factors rank-based inverse normal transformation was applied prior to analysis. Bonferroni corrected p-value threshold for the number of phenotypes tested was used to assess the significance of resulting associations in each cohort.

For biomarkers that showed significant association in any of the cohorts, we performed a meta-analysis across all cohorts the measurement was available. Meta-analysis was performed using

inverse-variance weighted fixed-effects meta-analysis method^{25, 26}. Bonferroni corrected p-value for number of traits tested ($n = 2$) was used to assess the significance of resulting associations in meta-analysis.

Identifying causal variants

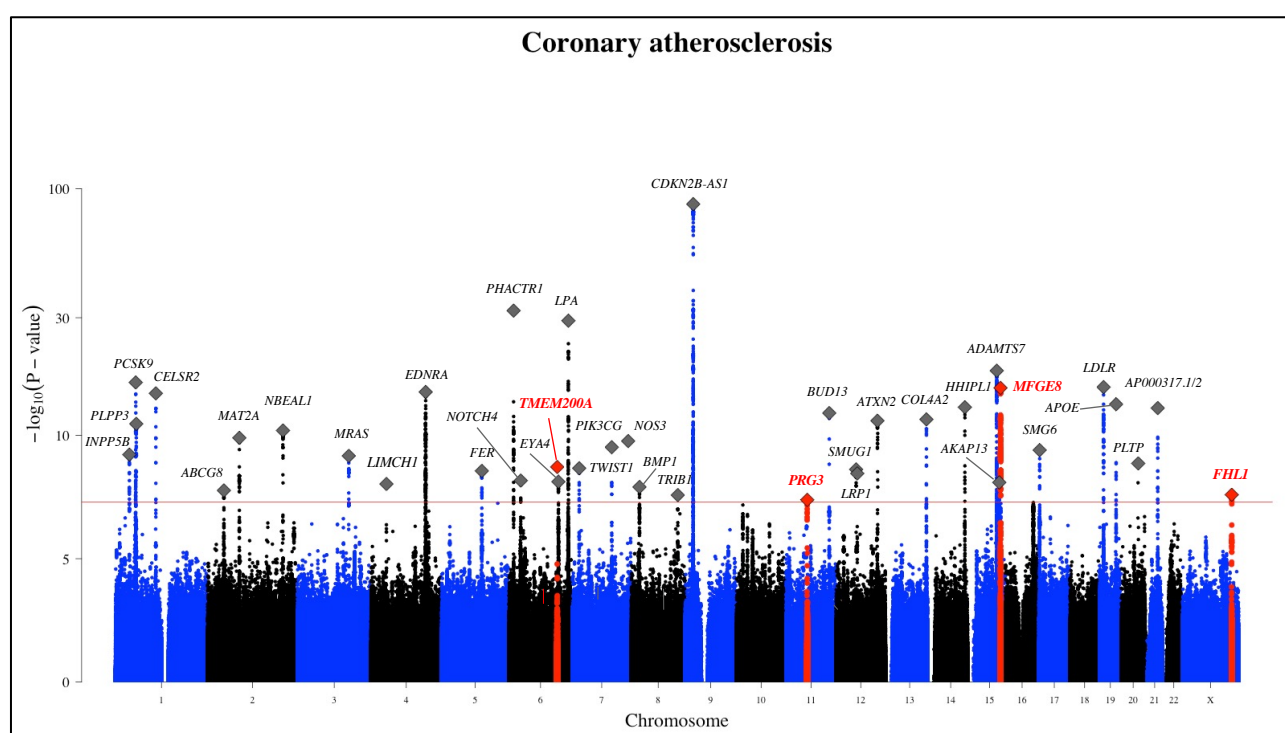
We used FINEMAP¹² on the GWAS summary statistics to identify causal variants underlying the associations for MI (strict definition, i.e., only primary diagnoses accepted) and coronary atherosclerosis. FINEMAP analyses were restricted to a ± 1.5 Mb region around the rs534125149. We assessed variants in the top 95% credible sets, i.e., the sets of variants encompassing at least 95% of the probability of being causal (causal probability) within each causal signal in the genomic region. Credible sets were filtered if minimum linkage disequilibrium (LD, r^2) between the variants in the credible set was < 0.1 , i.e., not clearly representing one signal.

Results

GWAS results for Coronary Atherosclerosis

We identified a total of 2 302 variants associated (GWS, $p < 5 \times 10^{-8}$) with coronary atherosclerosis (detailed description of the definition of the endpoint is in Supplementary Material). These variants were located in 38 distinct genetic loci (a minimum of 0.5 Mb distance from each other; **Figure 1**). Out of the 38 GWS loci, four (within or near genes *MFGE8*, *TMEM200A*, *PRG3* and *FHL1*) were novel for any CVD-related endpoints or risk factor for CVD compared to the GWAS Catalog¹¹ [<https://www.ebi.ac.uk/gwas/>] (**Supplementary Table 1**). Lead variants in these novel loci and their characteristics are listed in **Table 2** and locus zoom plots for each of the loci are in **Supplementary Figure 1**.

191 Among the three novel loci for coronary atherosclerosis, the locus near *MFGE8* had the strongest
 192 association (p -value = 2.63×10^{-16} for top variant rs534125149). The lead variant is an inframe
 193 insertion located in the sixth exon in the *MFGE8* gene (**Supplementary Figure 2**) and it is highly
 194 enriched in the Finnish population compared to NFSEEs (Non-Finnish, Estonian or Swedish
 195 Europeans). Interestingly, *MFGE8* is mainly expressed in coronary and tibial arteries according to
 196 data from GTEx v8 (**Supplementary Figure 3**).
 197



198
 199 **Figure 1:** GWAS results for coronary atherosclerosis in FinnGen. Total number of independent
 200 genome-wide significant associations (GWS; $p < 5 \times 10^{-8}$) is 38, the lead variant in each marked with
 201 diamonds. Four novel associations for CVD-related phenotypes are highlighted with ± 750 Mb
 202 around the lead variant in the region as red and the lead variant marked with red diamond.

203
 204 In addition to *MFGE8*, we identified three additional novel loci to be associated with coronary
 205 atherosclerosis, *TMT200A*, *PRG3* and *FHL1* being the nearest genes of the lead variants. *TMT200A*
 206 and *PRG3* loci had one non-coding low-frequency variant reaching the genome-wide significance
 207 threshold, and *FHL1* had 11. All variants in the credible sets of all these associations were either

intergenic or intronic variants and had no reported significant GWAS associations with any trait in the GWAS Catalog or significant eQTL associations in GTEx. The one variant (rs118042209) in the credible set of *TMEM200A* locus was associated with multiple coronary atherosclerosis related endpoints in FinnGen, including coronary atherosclerosis, ischemic heart disease and angina pectoris, whereas the lead variant in the *PRG3* locus was associated with several cardiomyopathy related endpoints. All variants in the credible set of *FHL1* were associated with multiple coronary atherosclerosis and related endpoints in FinnGen, including angina pectoris and ischemic heart disease. *TMEM200A* have been reported to be associated with ten traits (including height and trauma exposure) and *PRG3* with two traits (eosinophil count and eosinophil percentage of white cells) in the GWAS Catalog. *FHL1* gene had no reported associations in GWAS Catalog.

Table 2: Lead variants in novel loci associated with coronary atherosclerosis.

Lead variant chrom:pos:ref_alt (rsid)	Most severe consequence	Nearest gene	FIN enrichment (NFE)	AF	OR (P-value)	Info	# cs (Post- pr)	# coding in cs(s)
chr15:88901702:C_CTGT (rs534125149)	Inframe insertion	<i>MFGE8</i>	70.59	0.029	0.75 (2.60×10 ⁻¹⁶)	0.99	2 (0.705)	1
chr6:130483492:A_G (rs118042209)	Intergenic variant	<i>TMEM200A</i>	0.87	0.010	0.7 (1.90×10 ⁻⁹)	0.91	1 (0.904)	0
chr11:57380633:A:G (rs764568652)	Intron variant	<i>PRG3</i>	- ^a	0.0003	7.72 (4.10×10 ⁻⁸)	0.89	1 (0.583)	0
chrX:136194941:C_G (rs5974585)	Intron variant	<i>FHL1</i>	1.25	0.49	0.95 (2.55×10 ⁻⁸)	0.99	1 (0.692)	0

^a Variant not present in NFE in gnomAD

In the *PRG3* locus the lead variant is 40.9 kb away from a variant (chr11:57340143:G:A (rs11603691), the p-value for coronary atherosclerosis risk = 0.0269) which has previously been reported to be associated with HDL- cholesterol (HDL-C). However, our association is independent

226 of the previously reported variant rs11603691 (p-value for rs764568652 in a conditional analysis =
 227 4.36×10^{-08})

228

229 **Replication and phenome-wide association results for rs534125149**

230 We observed a highly protective association for the Finnish enriched inframe insertion rs534125149
 231 in the *MFGE8* gene and coronary atherosclerosis related endpoints, including coronary
 232 atherosclerosis (OR = 0.72 [0.63-0.83], $p = 7.94 \times 10^{-06}$) and myocardial infarction (MI) (OR = 0.69
 233 [0.58-0.83], $p = 9.62 \times 10^{-05}$) (**Figure 2**). In total, this variant had PWS significant protective effect
 234 on 14 disease endpoints, all related to coronary atherosclerosis. The majority of these disease
 235 endpoints are highly overlapping and representing major coronary heart disease (CHD) event. For
 236 this variant, we did not detect other phenome-wide significant associations among the 2 861
 237 endpoints in our data. Risk-lowering effect of rs534125149 on CHD was replicated in Biobank
 238 Japan ^{22,23} (BBJ) and the Estonian Biobank (EstBB) ²¹ (35 644 cases and 328 461 controls total: OR
 239 = 0.754 [0.67-0.84], $p = 5.41 \times 10^{-7}$). Association results for rs534125149 with CHD and MI across
 240 different cohorts are shown in **Supplementary Figure 4**.

241

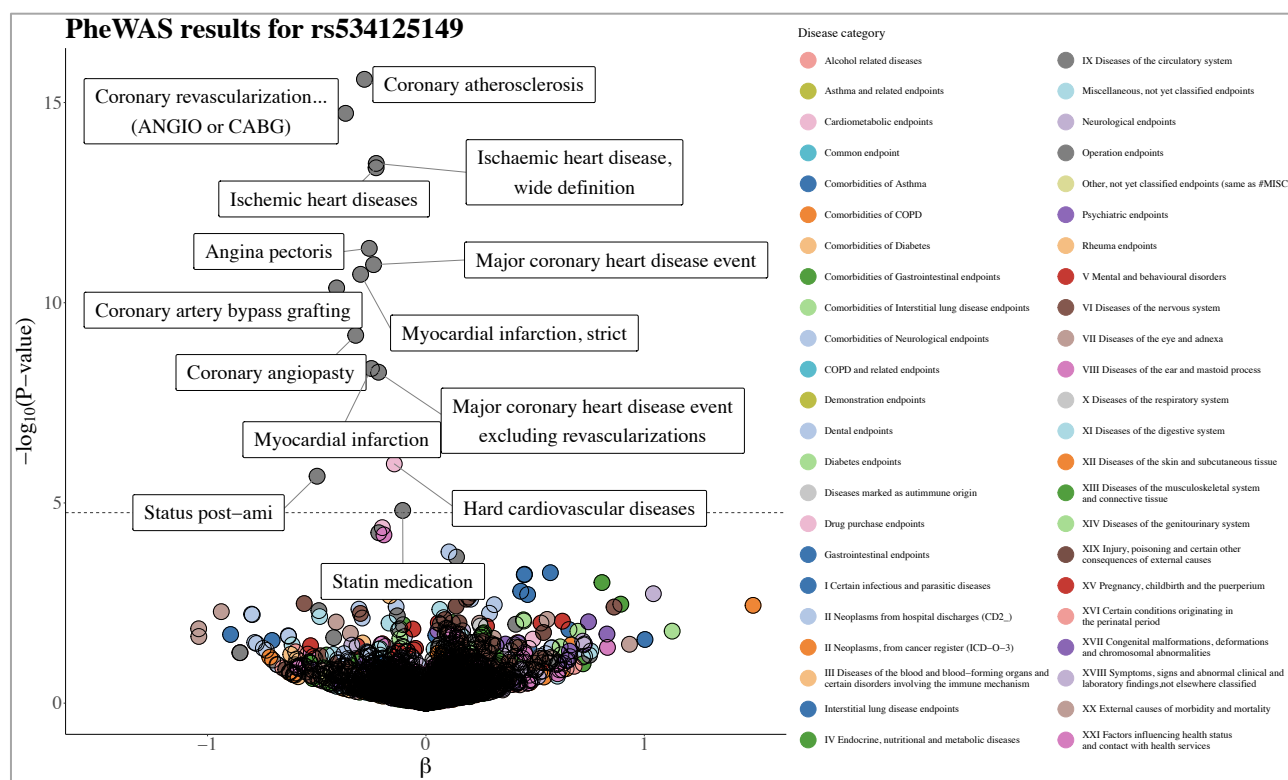


Figure 2: Phenome-wide association study (PheWAS) results for rs534125149. Total number of tested endpoints is 2 861 (A complete list of endpoints analyzed and their definitions is available at <https://www.finngen.fi/en/researchers/clinical-endpoints>). The dashed line represents the phenome-wide significance threshold, multiple testing corrected by the number of endpoints = $0.05/2\,861 = 1.75 \times 10^{-5}$. All endpoints reaching that threshold are labelled in the figure.

Splice acceptor variant rs201988637 in *MFGE8*

In addition to inframe insertion rs53412514, we also identified a splice acceptor variant (rs201988637) in *MFGE8* to have a protective effect on coronary atherosclerosis (OR = 0.72 [0.63-0.83], $p = 7.94 \times 10^{-6}$) and related endpoints. The splice acceptor variant had very similar PheWAS profile as the indel (Supplementary Figure 5) and furthermore the two variants had very similar protective effect sizes for the endpoints (Figure 3 and Supplementary Table 1). Similar to rs534125149, this variant is also highly enriched in Finland (37- fold compared to NFE), allele

frequency in Finland being 0.6 %. Moreover, both the splice acceptor and the inframe insertion variants were enriched to Eastern Finland (**Supplementary Figure 6**).

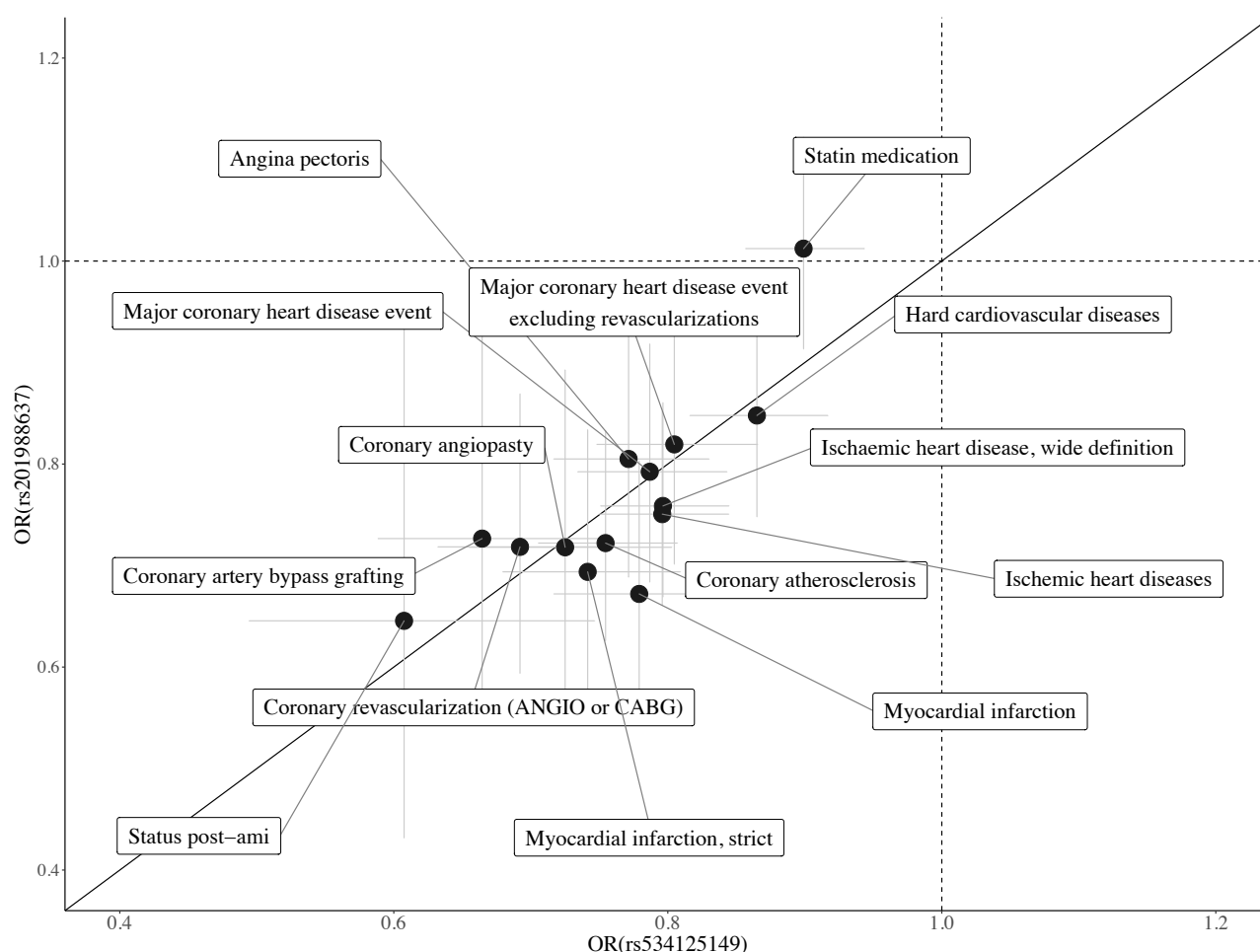


Figure 3: Comparison of the effects (OR) of rs534125149 and rs201988637 for 14 endpoints with $p\text{-value} < 1.75 \times 10^{-5}$ (PWS) for rs534125149. 95% confidence intervals represented as grey lines.

These two variants (rs534125149 and rs201988637) are in low linkage disequilibrium (LD, $r^2 = 0.00015$) and did not have any effect on the other variant's associations with coronary atherosclerosis or MI (**Table 3** and **Supplementary Figure 7**). This indicates that they both have their own, independent protective effect against these endpoints.

Table 3: Results of the conditional analysis on MI and coronary atherosclerosis.

Phenotype	SNPID [chr:position:ref:alt] (rsid)	Most severe consequence	Original GWAS results		Conditional results	
			OR [CI]	P-value	OR [CI]	P-value
Coronary atherosclerosis	chr15:88901702:C:CTGT (rs534125149)	Inframe insertion	0.75 [0.71-0.81]	2.63×10^{-16}	0.75 [0.70-0.80] ^a	7.68×10^{-15} ^a
	chr15:88899813:T:G (rs201988637)	Splice acceptor variant	0.72 [0.63-0.83]	7.94×10^{-6}	0.73 [0.64-0.85] ^b	1.99×10^{-5} ^b
Myocardial infarction, strict	chr15:88901702:C:CTGT (rs534125149)	Inframe insertion	0.74 [0.68-0.81]	1.95×10^{-11}	0.79 [0.73-0.85] ^a	1.92×10^{-10} ^a
	chr15:88899813:T:G (rs201988637)	Splice acceptor variant	0.69 [0.58-0.83]	9.62×10^{-5}	0.71 [0.59-0.85] ^b	4.03×10^{-4} ^b

This table present the conditional analysis results for coronary atherosclerosis and MI (strict definition, only primary diagnoses accepted) where the association has been conditioned on rs534125149 and rs201988637 separately

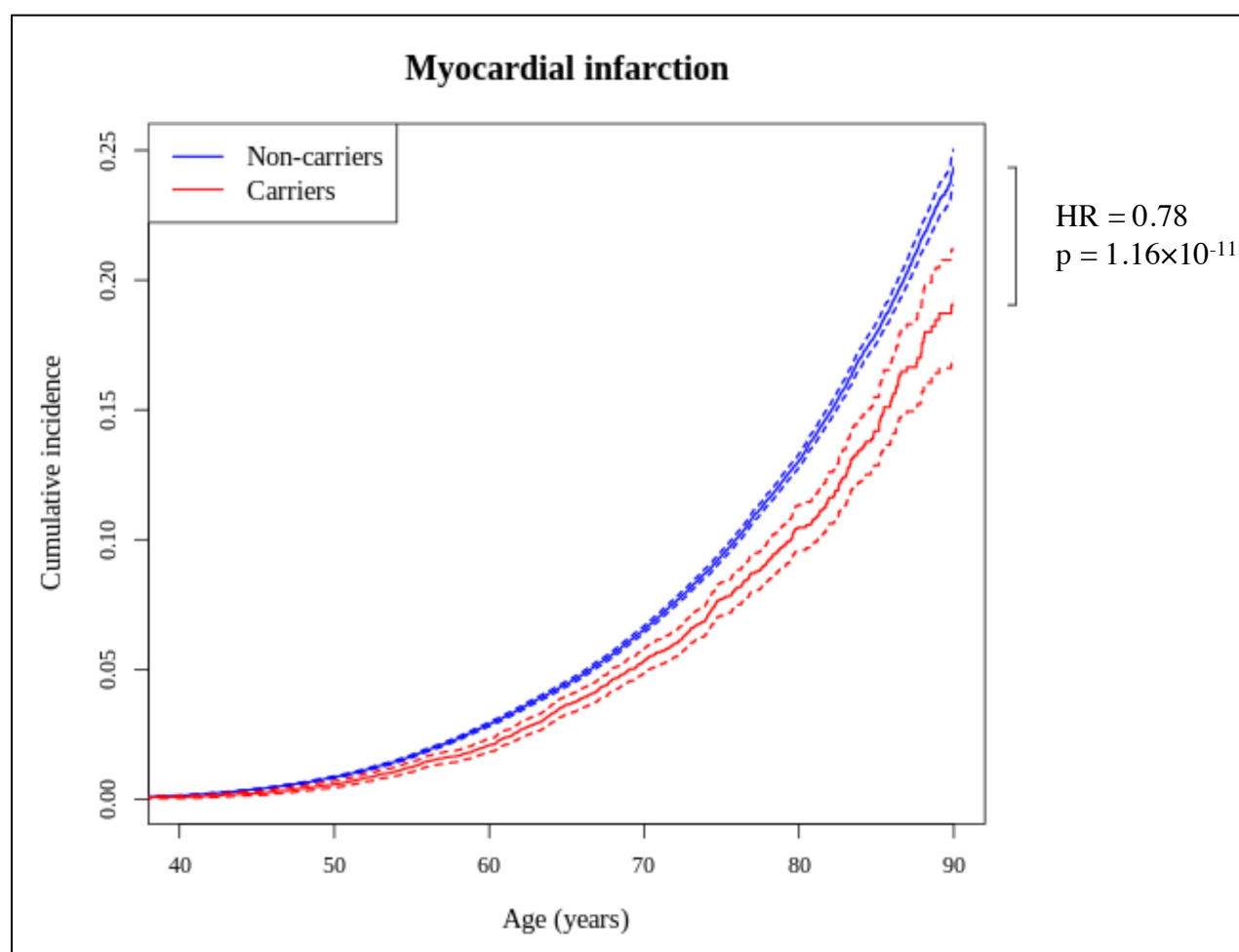
^a Conditional on rs201988637

^b Conditional on rs534125149

Survival analysis

In addition to protection against coronary atherosclerosis and myocardial infarction, both the inframe insertion rs534125149 and splice acceptor variant rs201988637 showed also significant association in survival analysis when analyzing survival time from birth to first diagnose of coronary atherosclerosis (HR = 0.78 [0.74-0.93]), $p = 1.67 \times 10^{-17}$ and HR = 0.77 [0.69-0.88], $p = 5.08 \times 10^{-5}$, respectively) and myocardial infarction (HR = 0.86 [0.80-0.93], $p = 2.63 \times 10^{-10}$ and HR = 0.72 [0.61-0.85], $p = 8.16 \times 10^{-5}$). In addition, when combining the heterozygous and homozygous carriers of both rs534125149 and rs201988637 together, carriers get the first diagnose significantly later than non-carriers (HR = 0.81 [0.77-0.85], $p = 6.4 \times 10^{-16}$ for coronary atherosclerosis and HR = 0.78 [0.72-0.85], $p = 1.16 \times 10^{-11}$ for MI) (**Figure 4**).

291



292

293 **Figure 4:** Cumulative incidence plots for myocardial infarction. Red line represents carriers (homo-
 294 or heterozygous) for either rs534125149 or rs201988637, and blue line represent non-carriers.
 295 Dashed lines represent 95%- confidence intervals.

296

297 Associations with risk factors for CVD and coronary heart disease sub-types

298 We then tested for possible associations between the *MFGE8* variants and risk factors for CVD.

299 The splice acceptor variant rs201988637 was associated with pulse pressure in analyses across four
 300 cohorts with pulse pressure measurements and variant rs201988637 available, with the risk
 301 lowering allele associated with lower pulse pressure ($p = 1.08 \times 10^{-03}$, $\beta = -1.52 [-2.43 - -0.61]$)
 302 mmHg (**Figure 5**). In addition, rs534125149 was significantly associated with height, but further
 303 analysis pointed this signal to be reflecting the association of a known association of *ACAN* with

304 height, located near *MFGE8* (Supplementary material, **Supplementary Figure 8**). No associations
305 with other risk factors were observed.
306
307 In the Corogene cohort (n = 4 896), rs534125149 significantly ($p < 0.05$) lowers the risk for acute
308 coronary syndrome and stable coronary heart disease ($RR = 0.87$ and 0.83 , respectively) compared
309 to healthy controls, but not against other types of heart attack, further pointing to a specific risk-
310 lowering effect of *MFGE8* on ischemic heart diseases (**Supplementary Figure 9**).
311
312 Previously reported common variants near *MFGE8*
313 Previously, common intergenic variant (rs8042271) near *MFGE8* has been reported to associate with
314 coronary heart disease (CHD) risk^{3, 27}. We replicate this association ($OR = 0.90$, $p = 3.69 \times 10^{-10}$ for
315 coronary atherosclerosis) in FinnGen. LD between the common variant rs8042271 and the inframe
316 insertion rs534125149 is 0.154. The LD characteristics for all 3 variants in *MFGE8* (rs534125149,
317 rs201988637 and rs8042271) in FinnGen are in **Supplementary Table 2**. Common variant
318 rs8042271 was in the 95% credible set for MI with the causal probability of 0.003 but was not
319 included in the 95% credible sets for coronary atherosclerosis (**Supplementary Table 3** and
320 **Supplementary Table 4**). The conditional analyses of all three *MFGE8* variants showed that the
321 association of the previously reported common variant rs8042271 can be explained by the inframe
322 insertion variant rs534125149, but not vice versa, and that the association of the splice acceptor
323 variant rs201988637 is independent of both rs534125149 and rs8042271 (**Supplementary Table 5**).
324

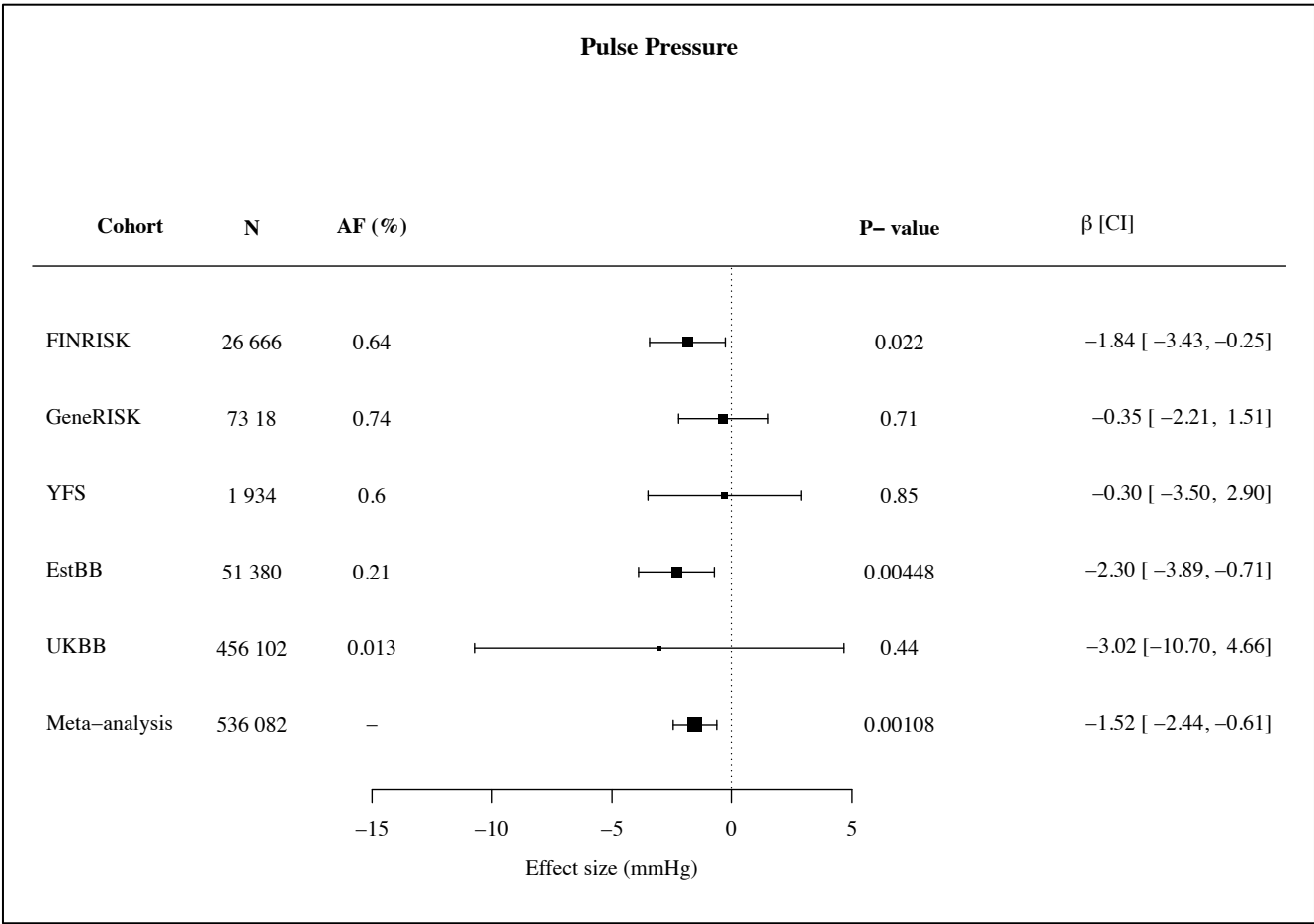


Figure 5: Results for pulse pressure association across all cohorts with splice acceptor variant *rs201988637* present.

Fine-mapping of the *MFGE8* locus

In our fine-mapping analyses, MI had most probably one credible set (set of causal variants) of 32 variants with the highest posterior probability (posterior probability = 0.62), and coronary atherosclerosis had two credible sets of 6 and 45 variants, respectively, with the highest posterior probability (posterior probability = 0.74). For both MI and coronary atherosclerosis, *rs534125149* had the highest probability of being causal (probability of being causal = 0.250 and 0.318, respectively) and was included in the first credible set (Supplementary Table 3, **Supplementary Table 4** and **Supplementary Figure 10**). Splice acceptor variant *rs201988637* was not included in the credible sets for either MI or coronary atherosclerosis, whereas previously reported common

variant rs8042271 was included in the credible set for MI with the probability of being causal = 0.003 (Supplementary Table 3).

Discussion

Here we show that a loss of function for *MFGE8* leads to 28% lower risk of myocardial infarction and reduces risk for other coronary atherosclerosis-related diseases. The effects are specific to coronary heart disease events and no significant association was observed to other diseases in a phenome-wide search, even if this can be due to lower statistical power in rare disease endpoints. Loss-of-function of *MFGE8* was also associated with lower pulse pressure, but not with blood lipids, blood pressure or other known coronary heart disease risk factors. In addition to *MFGE8*, we report three other novel loci associated with coronary atherosclerosis risk, and to our knowledge we report the first CHD associated locus, *FHL1*, in chromosome X.

Our findings allow us to draw several conclusions. First, *MFGE8* is a potential intervention target with specific effects on coronary heart disease. Specific protective effect with the LoF in *MFGE8* shows potential for efficacy of a treatment targeting MFGE8 protein or downstream products. Second, the lack of risk elevation in other diseases provide evidence on the potential safety of the intervention. Previously, the protective effect of loss-of-function variants have been reported for example for *PCSK9*⁵ and *APOC3*⁶, and in phase I, II and III trials, inhibition of PCSK9 have led to significantly decreased LDL-C levels, and in short-term trials, PCSK9 inhibitors have been well-tolerated and have had a low incidence of adverse effects.²⁸ Based on the phenome-wide association profile for the loss-of-function variant, we hypothesize that inhibiting *MFGE8* could lower the CHD risk.

363 An association of loss-of-function of *MFGE8* with lower pulse pressure, a potential biomarker for
364 arterial stiffness ²⁹, are very much in line with previous studies on MFGE8 and the inflammatory
365 ageing process of the arteries, highlighting the possible role of *MFGE8* in arterial ageing and
366 stiffness. The *MFGE8* gene encodes Milk-fat globule-EGF 8 (MFGE8), or lactadherin, which is an
367 integrin-binding glycoprotein implicated in vascular smooth muscle cell (VSMC) proliferation and
368 invasion, and the secretion of pro-inflammatory molecules ^{30,31}. Lactadherin is known to play
369 important roles in several other biological processes, including apoptotic cell clearance and adaptive
370 immunity ³², which are known to contribute to the pathogenesis of ischemic stroke. Initially
371 lactadherin was identified as a bridging molecule between apoptotic cells and phagocytic
372 macrophages ³³⁻³⁵, but a growing evidence has indicated that it is a secreted inflammatory mediator
373 that orchestrates diverse cellular interactions involved in the pathogenesis of various diseases,
374 including vascular metabolic disorders and some tumors ³⁶⁻⁴⁰ and cancers, such as breast ^{38,41},
375 bladder⁴², esophageal⁴³ and colorectal cancer ⁴⁴. Recently, not only has MFG-E8 expression
376 emerged as a molecular hallmark of adverse cardiovascular remodeling with age ⁴⁵⁻⁴⁸, but MFG-E8
377 signaling has also been found to mediate the vascular outcomes of cellular and matrix responses to
378 the hostile stresses associated with hypertension, diabetes, and atherosclerosis ⁴⁹⁻⁵³.

379

380 Arterial inflammation and remodeling are linked to the pathogenesis of age-associated arterial
381 diseases, such as atherosclerosis. Recently, lactadherin has been identified as a novel local
382 biomarker for ageing arterial walls by high-throughput proteomic screening, and it has been shown
383 to also be an element of the arterial inflammatory signaling network ⁵⁴. The transcription,
384 translation, and signaling levels of MFG-E8 are increased in aged, atherosclerotic, hypertensive,
385 and diabetic arterial walls in vivo, as well as activated VSMCs and a subset of macrophages in
386 vitro. During aging, both MFG-E8 transcription and translation increase within the arterial walls
387 and hearts of various species, including rats, humans, and monkeys^{48,55-57}, and MFG-E8 is markedly

up-regulated in rat aortic walls with ageing⁴⁸. High levels of MFG-E8 have also been detected within endothelial cells, SMC, and macrophages of atherosclerotic aortae in both mice and humans^{53, 58}. Furthermore, in the advanced atherosclerotic plaques found in murine models, decreased macrophage MFG-E8 levels are associated with an inhibition of apoptotic cell engulfment, leading to the accumulation of cellular debris during the pathogenesis of atherosclerosis. Lastly, lactadherin has shown tissue protection in various models of organ injury, including suppression of inflammation and apoptosis in intestinal ischemia in mice ⁵⁹, as well as inducing recovery from ischemia by facilitating angiogenesis ⁶⁰.

Our study does, however, have a few limitations. First, our primary association results come from Finnish population with considerable elevation in allele frequency in *MFGE8* variants among Finns. Therefore, the replication of the association in other populations has reduced statistical power. However, there were enough carriers combined in Japanese, Estonian and UK samples to replicate robustly both the protective association with coronary artery disease and for pulse pressure. Secondly, although our data shows association with pulse pressure which has previously been linked to arterial stiffness, the direct effect of the genetic variants on arterial stiffness and arterial aging needs further evidence.

In conclusion, our results show that inhibiting production of lactadherin could reduce the risk for coronary atherosclerosis substantially and thus present *MFGE8* as a potential therapeutic target for atherosclerotic cardiovascular disease. Our study also highlights the potential of FinnGen, as a large-scale biobank study in isolated population to identify high-impact variants either very rare or absent in other populations.

References

- 413 1. Kessler, T., Erdmann, J. & Schunkert, H. Genetics of coronary artery disease and myocardial
414 infarction-2013. *Curr. Cardiol. Rep.* **15**, 368 (2013).
- 415 2. O'Donnell, C. J. & Nabel, E. G. Genomics of cardiovascular disease. *N. Engl. J. Med.* **365**, 2098-
416 2109 (2011).
- 417 3. Nikpay, M. *et al.* A comprehensive 1000 Genomes-based genome-wide association meta-
418 analysis of coronary artery disease. *Nat. Genet.* **47**, 1121 (2015).
- 419 4. Lim, E. T. *et al.* Distribution and medical impact of loss-of-function variants in the Finnish
420 founder population. *PLoS genetics* **10** (2014).
- 421 5. Cohen, J. C., Boerwinkle, E., Mosley Jr, T. H. & Hobbs, H. H. Sequence variations in PCSK9,
422 low LDL, and protection against coronary heart disease. *N. Engl. J. Med.* **354**, 1264-1272 (2006).
- 423 6. TG and HDL Working Group of the Exome Sequencing Project, National Heart, Lung, and
424 Blood Institute. Loss-of-function mutations in APOC3, triglycerides, and coronary disease. *N. Engl.*
425 *J. Med.* **371**, 22-31 (2014).
- 426 7. Dewey, F. E. *et al.* Inactivating variants in ANGPTL4 and risk of coronary artery disease. *N.*
427 *Engl. J. Med.* **374**, 1123-1133 (2016).
- 428 8. Libby, P. *et al.* Lewis EF. *Atherosclerosis. Nature reviews Disease primers* **5**, 56 (2019).
- 429 9. Karczewski, K. J. *et al.* The mutational constraint spectrum quantified from variation in 141,456
430 humans. *Nature* **581**, 434-443 (2020).
- 431 10. Zhou, W. *et al.* Efficiently controlling for case-control imbalance and sample relatedness in
432 large-scale genetic association studies. *Nat. Genet.* **50**, 1335-1341 (2018).
- 433 11. Buniello, A. *et al.* The NHGRI-EBI GWAS Catalog of published genome-wide association
434 studies, targeted arrays and summary statistics 2019. *Nucleic Acids Res.* **47**, D1005-D1012 (2018).
- 435 12. Benner, C. *et al.* FINEMAP: efficient variable selection using summary data from genome-wide
436 association studies. *Bioinformatics* **32**, 1493-1501 (2016).
- 437 13. Vaara, S. *et al.* Cohort profile: the Corogene study. *Int. J. Epidemiol.* **41**, 1265-1271 (2012).
- 438 14. Dey, R. *et al.* An efficient and accurate frailty model approach for genome-wide survival
439 association analysis controlling for population structure and relatedness in large-scale biobanks.
440 *bioRxiv* (2020).
- 441 15. Breslow, N. E. & Clayton, D. G. Approximate inference in generalized linear mixed models.
442 *Journal of the American statistical Association* **88**, 9-25 (1993).
- 443 16. Chen, H. *et al.* Control for population structure and relatedness for binary traits in genetic
444 association studies via logistic mixed models. *The American Journal of Human Genetics* **98**, 653-
445 666 (2016).

- 446 17. Borodulin, K. *et al.* Forty-year trends in cardiovascular risk factors in Finland. *The European*
447 *Journal of Public Health* **25**, 539-546 (2014).
- 448 18. Widen, E. *et al.* Communicating polygenic and non-genetic risk for atherosclerotic
449 cardiovascular disease-An observational follow-up study. *medRxiv* (2020).
- 450 19. Raitakari, O. T. *et al.* Cohort profile: the cardiovascular risk in Young Finns Study. *Int. J.*
451 *Epidemiol.* **37**, 1220-1226 (2008).
- 452 20. Mitt, M. *et al.* Improved imputation accuracy of rare and low-frequency variants using
453 population-specific high-coverage WGS-based imputation reference panel. *European Journal of*
454 *Human Genetics* **25**, 869 (2017).
- 455 21. Leitsalu, L. *et al.* Cohort profile: Estonian biobank of the Estonian genome center, university of
456 Tartu. *Int. J. Epidemiol.* **44**, 1137-1147 (2015).
- 457 22. Nagai, A. *et al.* Overview of the BioBank Japan Project: study design and profile. *Journal of*
458 *epidemiology* **27**, S2-S8 (2017).
- 459 23. Sakaue, S. *et al.* A global atlas of genetic associations of 220 deep phenotypes. *medRxiv*,
460 2020.10. 23.20213652 (2021).
- 461 24. Bycroft, C. *et al.* The UK Biobank resource with deep phenotyping and genomic data. *Nature*
462 **562**, 203 (2018).
- 463 25. Cooper, H., Hedges, L. V. & Valentine, J. C. in *The handbook of research synthesis and meta-*
464 *analysis* (Russell Sage Foundation, 2019).
- 465 26. Evangelou, E. & Ioannidis, J. P. Meta-analysis methods for genome-wide association studies
466 and beyond. *Nature Reviews Genetics* **14**, 379-389 (2013).
- 467 27. Soubeyrand, S. *et al.* Regulation of MFGE8 by the intergenic coronary artery disease locus on
468 15q26. 1. *Atherosclerosis* **284**, 11-17 (2019).
- 469 28. Dadu, R. T. & Ballantyne, C. M. Lipid lowering with PCSK9 inhibitors. *Nature Reviews*
470 *Cardiology* **11**, 563 (2014).
- 471 29. Benetos, A. *et al.* Mortality and cardiovascular events are best predicted by low
472 central/peripheral pulse pressure amplification but not by high blood pressure levels in elderly
473 nursing home subjects: the PARTAGE (Predictive Values of Blood Pressure and Arterial Stiffness
474 in Institutionalized Very Aged Population) study. *J. Am. Coll. Cardiol.* **60**, 1503-1511 (2012).
- 475 30. Oshima, K. *et al.* Lactation-Dependent Expression of an mRNA Splice Variant with an Exon for
476 a Multiply O-Glycosylated Domain of Mouse Milk Fat Globule Glycoprotein MFG-E8. *Biochem.*
477 *Biophys. Res. Commun.* **254**, 522-528 (1999).
- 478 31. Aoki, N. *et al.* Immunologically cross-reactive 57 kDa and 53 kDa glycoprotein antigens of
479 bovine milk fat globule membrane: isoforms with different N-linked sugar chains and differential
480 glycosylation at early stages of lactation. *Biochimica et Biophysica Acta (BBA)-General Subjects*
481 **1200**, 227-234 (1994).

482 32. Deroide, N. *et al.* MFG-E8 inhibits inflammasome-induced IL-1 β production and limits
483 postischemic cerebral injury. *J. Clin. Invest.* **123**, 1176-1181 (2013).

484 33. Hanayama, R. *et al.* Identification of a factor that links apoptotic cells to phagocytes. *Nature*
485 **417**, 182-187 (2002).

486 34. Hanayama, R. *et al.* Autoimmune disease and impaired uptake of apoptotic cells in MFG-E8-
487 deficient mice. *Science* **304**, 1147-1150 (2004).

488 35. Yoshida, H. *et al.* Phosphatidylserine-dependent engulfment by macrophages of nuclei from
489 erythroid precursor cells. *Nature* **437**, 754-758 (2005).

490 36. Tahara, H. *et al.* Emerging concepts in biomarker discovery; the US-Japan Workshop on
491 Immunological Molecular Markers in Oncology. *Journal of Translational Medicine* **7**, 1-25 (2009).

492 37. Neutzner, M. *et al.* MFG-E8/lactadherin promotes tumor growth in an angiogenesis-dependent
493 transgenic mouse model of multistage carcinogenesis. *Cancer Res.* **67**, 6777-6785 (2007).

494 38. TAYLOR, M. R., COUTO, J. R., SCALLAN, C. D., CERIANI, R. L. & PETERSON, J. A.
495 Lactadherin (formerly BA46), a membrane-associated glycoprotein expressed in human milk and
496 breast carcinomas, promotes Arg-Gly-Asp (RGD)-dependent cell adhesion. *DNA Cell Biol.* **16**, 861-
497 869 (1997).

498 39. Jinushi, M. *et al.* Milk fat globule EGF-8 promotes melanoma progression through coordinated
499 Akt and twist signaling in the tumor microenvironment. *Cancer Res.* **68**, 8889-8898 (2008).

500 40. Raymond, A., Ensslin, M. A. & Shur, B. D. SED1/MFG-E8: a bi-motif protein that orchestrates
501 diverse cellular interactions. *J. Cell. Biochem.* **106**, 957-966 (2009).

502 41. Yu, L. *et al.* MFG-E8 overexpression is associated with poor prognosis in breast cancer
503 patients. *Pathology-Research and Practice* **215**, 490-498 (2019).

504 42. Sugano, G. *et al.* Milk fat globule—epidermal growth factor—factor VIII (MFG-E8)/lactadherin
505 promotes bladder tumor development. *Oncogene* **30**, 642-653 (2011).

506 43. Kanemura, T. *et al.* Immunoregulatory influence of abundant MFG-E8 expression by
507 esophageal cancer treated with chemotherapy. *Cancer science* **109**, 3393-3402 (2018).

508 44. Jia, M. *et al.* Prognostic correlation between MFG-E8 expression level and colorectal Cancer.
509 *Arch. Med. Res.* **48**, 270-275 (2017).

510 45. Wang, M., Khazan, B. & G Lakatta, E. Central arterial aging and angiotensin II signaling.
511 *Current hypertension reviews* **6**, 266-281 (2010).

512 46. Wang, M., Monticone, R. E. & Lakatta, E. G. Arterial aging: a journey into subclinical arterial
513 disease. *Curr. Opin. Nephrol. Hypertens.* **19**, 201 (2010).

514 47. Ortiz, A. *et al.* Clinical usefulness of novel prognostic biomarkers in patients on hemodialysis.
515 *Nature reviews Nephrology* **8**, 141 (2012).

- 516 48. Fu, Z. *et al.* Milk fat globule protein epidermal growth factor-8: a pivotal relay element within
517 the angiotensin II and monocyte chemoattractant protein-1 signaling cascade mediating vascular
518 smooth muscle cells invasion. *Circ. Res.* **104**, 1337-1346 (2009).
- 519 49. Bagnato, C. *et al.* Proteomics analysis of human coronary atherosclerotic plaque: a feasibility
520 study of direct tissue proteomics by liquid chromatography and tandem mass spectrometry.
521 *Molecular & Cellular Proteomics* **6**, 1088-1102 (2007).
- 522 50. Li, X. *et al.* Proteomics approach to study the mechanism of action of grape seed
523 proanthocyanidin extracts on arterial remodeling in diabetic rats. *Int. J. Mol. Med.* **25**, 237-248
524 (2010).
- 525 51. Strøm, C. C. *et al.* Identification of a core set of genes that signifies pathways underlying
526 cardiac hypertrophy. *Comp. Funct. Genomics* **5**, 459-470 (2004).
- 527 52. Lin, Y. *et al.* Comparative proteomic analysis of rat aorta in a subtotal nephrectomy model.
528 *Proteomics* **10**, 2429-2443 (2010).
- 529 53. Ait-Oufella, H. *et al.* CLINICAL PERSPECTIVE. *Circulation* **115**, 2168-2177 (2007).
- 530 54. Wang, M., H Wang, H. & G Lakatta, E. Milk fat globule epidermal growth factor VIII signaling
531 in arterial wall remodeling. *Current vascular pharmacology* **11**, 768-776 (2013).
- 532 55. Peng, S., Glennert, J. & Westermark, P. Medin-amyloid: a recently characterized age-associated
533 arterial amyloid form affects mainly arteries in the upper part of the body. *Amyloid* **12**, 96-102
534 (2005).
- 535 56. Peng, S. *et al.* Medin and medin-amyloid in ageing inflamed and non-inflamed temporal
536 arteries. *The Journal of Pathology: A Journal of the Pathological Society of Great Britain and*
537 *Ireland* **196**, 91-96 (2002).
- 538 57. Grant, J. E. *et al.* Quantification of protein expression changes in the aging left ventricle of
539 *Rattus norvegicus*. *Journal of proteome research* **8**, 4252-4263 (2009).
- 540 58. Bagnato, C. *et al.* Proteomics analysis of human coronary atherosclerotic plaque: a feasibility
541 study of direct tissue proteomics by liquid chromatography and tandem mass spectrometry.
542 *Molecular & Cellular Proteomics* **6**, 1088-1102 (2007).
- 543 59. Cheyuo, C. *et al.* Recombinant human MFG-E8 attenuates cerebral ischemic injury: its role in
544 anti-inflammation and anti-apoptosis. *Neuropharmacology* **62**, 890-900 (2012).
- 545 60. Silvestre, J. *et al.* Lactadherin promotes VEGF-dependent neovascularization. *Nat. Med.* **11**,
546 499-506 (2005).

547

548 Acknowledgements

549

We would like to thank all participants of all study cohorts for their generous participation. We also want to thank Dr. Kaoru Ito at RIKEN Center for Integrative Medical Sciences for supporting the study. This work was supported by the Academy of Finland Center of Excellence in Complex Disease Genetics (Grant No 312062 and 336820 to S.R.), the Finnish Foundation for Cardiovascular Research, the Sigrid Juselius Foundation, University of Helsinki HiLIFE Fellow, Grand Challenge grants and Horizon 2020 Research and Innovation Programme (grant number 101016775 “INTERVENE” to S.R.), Academy of Finland grant number 331671 to N.M., the European Union through the European Regional Development Fund (Project No. 2014-2020.4.01.16-0125 to L.M.), the Estonian Research Council grant PRG184 [to L.M.] and the Doctoral Programme in Population Health, University of Helsinki [to S.E.R.].

The FinnGen project is funded by two grants from Business Finland (HUS 4685/31/2016 and UH 4386/31/2016) and the following industry partners: AbbVie Inc., AstraZeneca UK Ltd, Biogen MA Inc., Celgene Corporation, Celgene International II Sàrl, Genentech Inc., Merck Sharp & Dohme Corp, Pfizer Inc., GlaxoSmithKline Intellectual Property Development Ltd., Sanofi US Services Inc., Maze Therapeutics Inc., Janssen Biotech Inc, and Novartis AG. Following biobanks are acknowledged for delivering biobank samples to FinnGen: Auria Biobank (www.auria.fi/biopankki), THL Biobank (www.thl.fi/biobank), Helsinki Biobank (www.helsinginbiopankki.fi), Biobank Borealis of Northern Finland (<https://www.ppshep.fi/Tutkimus-ja-opetus/Biopankki/Pages/Biopankki-Borealis-briefly-in-English.aspx>), Finnish Clinical Biobank Tampere (www.tays.fi/en-US/Research_and_development/Finnish_Clinical_Biobank_Tampere), Biobank of Eastern Finland (www.ita-suomenbiopankki.fi/en), Central Finland Biobank (www.ksshp.fi/fi-FI/Potilaille/Biopankki), Finnish Red Cross Blood Service Biobank (www.veripalvelu.fi/verenluovutus/biopankkitoiminta) and Terveystalo Biobank

574 (www.terveystalo.com/fi/Yritystietoa/Terveystalo-Biopankki/Biopankki/). All Finnish Biobanks are
575 members of BBMRI.fi infrastructure (www.bbmri.fi) and FinBB (<https://finbb.fi/>).
576

577 Estonian Biobank research team

578 Tõnu Esko,
579 Andres Metspalu,
580 Reedik Mägi,
581 Mari Nelis
582

583 Contributors of FinnGen

584 Steering Committee

585 Aarno Palotie Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland
586 Mark Daly Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland
587

588 Pharmaceutical companies

589 Bridget Riley-Gills Abbvie, Chicago, IL, United States
590 Howard Jacob Abbvie, Chicago, IL, United States
591 Dirk Paul Astra Zeneca, Cambridge, United Kingdom
592 Heiko Runz Biogen, Cambridge, MA, United States
593 Sally John Biogen, Cambridge, MA, United States
594 Robert Plenge Celgene, Summit, NJ, United States/Bristol Myers Squibb, New York, NY, United States
595 Mark McCarthy Genentech, San Francisco, CA, United States
596 Julie Hunkapiller Genentech, San Francisco, CA, United States
597 Meg Ehm GlaxoSmithKline, Brentford, United Kingdom
598 Kirsi Auro GlaxoSmithKline, Brentford, United Kingdom
599 Caroline Fox Merck, Kenilworth, NJ, United States
600 Anders Mälarstig Pfizer, New York, NY, United States
601 Katherine Klinger Sanofi, Paris, France
602 Deepak Raipal Sanofi, Paris, France

603	Tim Behrens	Maze Therapeutics, San Francisco, CA, United States
604	Robert Yang	Janssen Biotech, Beerse, Belgium
605	Richard Siegel	Novartis, Basel, Switzerland
606	University of Helsinki & Biobanks	
607	Tomi Mäkelä	HiLIFE, University of Helsinki, Finland, Finland
608	Jaakko Kaprio	Institute for Molecular Medicine Finland, HiLIFE, Helsinki, Finland, Finland
609	Petri Virolainen	Auria Biobank / University of Turku / Hospital District of Southwest Finland, Turku,
610	Finland	
611	Antti Hakanen	Auria Biobank / University of Turku / Hospital District of Southwest Finland, Turku,
612	Finland	
613	Terhi Kilpi	THL Biobank / The National Institute of Health and Welfare Helsinki, Finland
614	Markus Perola	THL Biobank / The National Institute of Health and Welfare Helsinki, Finland
615	Jukka Partanen	Finnish Red Cross Blood Service / Finnish Hematology Registry and Clinical Biobank,
616	Helsinki, Finland	
617	Anne Pitkäranta	Helsinki Biobank / Helsinki University and Hospital District of Helsinki and Uusimaa,
618	Helsinki	
619	Juhani Junttila	Northern Finland Biobank Borealis / University of Oulu / Northern Ostrobothnia Hospital
620	District, Oulu, Finland	
621	Raisa Serpi	Northern Finland Biobank Borealis / University of Oulu / Northern Ostrobothnia Hospital
622	District, Oulu, Finland	
623	Tarja Laitinen	Finnish Clinical Biobank Tampere / University of Tampere / Pirkanmaa Hospital District,
624	Tampere, Finland	
625	Johanna Mäkelä	Finnish Clinical Biobank Tampere / University of Tampere / Pirkanmaa Hospital District,
626	Tampere, Finland	
627	Veli-Matti Kosma	Biobank of Eastern Finland / University of Eastern Finland / Northern Savo Hospital District,
628	Kuopio, Finland	
629	Urho Kujala	Central Finland Biobank / University of Jyväskylä / Central Finland Health Care District,
630	Jyväskylä, Finland	
631		
632	Other Experts/ Non-Voting Members	

633	Outi Tuovila	Business Finland, Helsinki, Finland
634	Raimo Pakkanen	Business Finland, Helsinki, Finland
635		
636	Scientific Committee	
637	Pharmaceutical companies	
638	Jeffrey Waring	Abbvie, Chicago, IL, United States
639	Ali Abbasi	Abbvie, Chicago, IL, United States
640	Mengzhen Liu	Abbvie, Chicago, IL, United States
641	Ioanna Tachmazidou	Astra Zeneca, Cambridge, United Kingdom
642	Chia-Yen Chen	Biogen, Cambridge, MA, United States
643	Heiko Runz	Biogen, Cambridge, MA, United States
644	Shameek Biswas	Celgene, Summit, NJ, United States/Bristol Myers Squibb, New York, NY, United States
645	Julie Hunkapiller	Genentech, San Francisco, CA, United States
646	Meg Ehm	GlaxoSmithKline, Brentford, United Kingdom
647	Neha Raghavan	Merck, Kenilworth, NJ, United States
648	Adriana Huertas-Vazquez	Merck, Kenilworth, NJ, United States
649	Anders Mälarstig	Pfizer, New York, NY, United States
650	Xinli Hu	Pfizer, New York, NY, United States
651	Katherine Klinger	Sanofi, Paris, France
652	Matthias Gossel	Sanofi, Paris, France
653	Robert Graham	Maze Therapeutics, San Francisco, CA, United States
654	Tim Behrens	Maze Therapeutics, San Francisco, CA, United States
655	Beryl Cummings	Maze Therapeutics, San Francisco, CA, United States
656	Wilco Fleuren	Janssen Biotech, Beerse, Belgium
657	Dawn Waterworth	Janssen Biotech, Beerse, Belgium
658	Nicole Renaud	Novartis, Basel, Switzerland
659	Ma'en Obeidat	Novartis, Basel, Switzerland
660		
661	University of Helsinki & Biobanks	
662	Samuli Ripatti	Institute for Molecular Medicine Finland, HiLIFE, Helsinki, Finland

663	Johanna Schleutker	Auria Biobank / Univ. of Turku / Hospital District of Southwest Finland, Turku, Finland
664	Markus Perola	THL Biobank / The National Institute of Health and Welfare Helsinki, Finland
665	Mikko Arvas	Finnish Red Cross Blood Service / Finnish Hematology Registry and Clinical Biobank,
666	Helsinki, Finland	
667	Olli Carpén	Helsinki Biobank / Helsinki University and Hospital District of Helsinki and Uusimaa,
668	Helsinki	
669	Reetta Hinttala	Northern Finland Biobank Borealis / University of Oulu / Northern Ostrobothnia Hospital
670	District, Oulu, Finland	
671	Johannes Kettunen	Northern Finland Biobank Borealis / University of Oulu / Northern Ostrobothnia Hospital
672	District, Oulu, Finland	
673	Johanna Mäkelä	Finnish Clinical Biobank Tampere / University of Tampere / Pirkanmaa Hospital District,
674	Tampere, Finland	
675	Arto Mannermaa	Biobank of Eastern Finland / University of Eastern Finland / Northern Savo Hospital
676	District, Kuopio, Finland	
677	Jari Laukkanen	Central Finland Biobank / University of Jyväskylä / Central Finland Health Care District,
678	Jyväskylä, Finland	
679	Urho Kujala	Central Finland Biobank / University of Jyväskylä / Central Finland Health Care District,
680	Jyväskylä, Finland	
681		
682	Clinical Groups	
683	Neurology Group	
684	Reetta Kälviäinen	Northern Savo Hospital District, Kuopio, Finland
685	Valtteri Julkunen	Northern Savo Hospital District, Kuopio, Finland
686	Hilkka Soininen	Northern Savo Hospital District, Kuopio, Finland
687	Anne Remes	Northern Ostrobothnia Hospital District, Oulu, Finland
688	Mikko Hiltunen	Northern Savo Hospital District, Kuopio, Finland
689	Jukka Peltola	Pirkanmaa Hospital District, Tampere, Finland
690	Pentti Tienari	Hospital District of Helsinki and Uusimaa, Helsinki, Finland
691	Juha Rinne	Hospital District of Southwest Finland, Turku, Finland
692	Roosa Kallionpää	Hospital District of Southwest Finland, Turku, Finland

693	Ali Abbasi	Abbvie, Chicago, IL, United States
694	Adam Ziemann	Abbvie, Chicago, IL, United States
695	Jeffrey Waring	Abbvie, Chicago, IL, United States
696	Sahar Esmaeeli	Abbvie, Chicago, IL, United States
697	Nizar Smaoui	Abbvie, Chicago, IL, United States
698	Anne Lehtonen	Abbvie, Chicago, IL, United States
699	Susan Eaton	Biogen, Cambridge, MA, United States
700	Heiko Runz	Biogen, Cambridge, MA, United States
701	Sanni Lahdenperä	Biogen, Cambridge, MA, United States
702	Janet van Adelsberg	Celgene, Summit, NJ, United States/ Bristol Myers Squibb, New York, NY, United States
703	Shameek Biswas	Celgene, Summit, NJ, United States/ Bristol Myers Squibb, New York, NY, United States
704	Julie Hunkapiller	Genentech, San Francisco, CA, United States
705	Natalie Bowers	Genentech, San Francisco, CA, United States
706	Edmond Teng	Genentech, San Francisco, CA, United States
707	Sarah Pendergrass	Genentech, San Francisco, CA, United States
708	Onuralp Soylemez	Merck, Kenilworth, NJ, United States
709	Kari Linden	Pfizer, New York, NY, United States
710	Fanli Xu	GlaxoSmithKline, Brentford, United Kingdom
711	David Pulford	GlaxoSmithKline, Brentford, United Kingdom
712	Kirsi Auro	GlaxoSmithKline, Brentford, United Kingdom
713	Laura Addis	GlaxoSmithKline, Brentford, United Kingdom
714	John Eicher	GlaxoSmithKline, Brentford, United Kingdom
715	Minna Raivio	Hospital District of Helsinki and Uusimaa, Helsinki, Finland
716	Sarah Pendergrass	Genentech, San Francisco, CA, United States
717	Beryl Cummings	Maze Therapeutics, San Francisco, CA, United States
718	Juulia Partanen	Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland
719		
720	Gastroenterology Group	
721	Martti Färkkilä	Hospital District of Helsinki and Uusimaa, Helsinki, Finland
722	Jukka Koskela	Hospital District of Helsinki and Uusimaa, Helsinki, Finland

723	Sampsa Pikkarainen	Hospital District of Helsinki and Uusimaa, Helsinki, Finland
724	Airi Jussila	Pirkanmaa Hospital District, Tampere, Finland
725	Katri Kaukinen	Pirkanmaa Hospital District, Tampere, Finland
726	Timo Blomster	Northern Ostrobothnia Hospital District, Oulu, Finland
727	Mikko Kiviniemi	Northern Savo Hospital District, Kuopio, Finland
728	Markku Voutilainen	Hospital District of Southwest Finland, Turku, Finland
729	Ali Abbasi	Abbvie, Chicago, IL, United States
730	Graham Heap	Abbvie, Chicago, IL, United States
731	Jeffrey Waring	Abbvie, Chicago, IL, United States
732	Nizar Smaoui	Abbvie, Chicago, IL, United States
733	Fedik Rahimov	Abbvie, Chicago, IL, United States
734	Anne Lehtonen	Abbvie, Chicago, IL, United States
735	Keith Usiskin	Celgene, Summit, NJ, United States/ Bristol Myers Squibb, New York, NY, United States
736	Tim Lu	Genentech, San Francisco, CA, United States
737	Natalie Bowers	Genentech, San Francisco, CA, United States
738	Danny Oh	Genentech, San Francisco, CA, United States
739	Sarah Pendergrass	Genentech, San Francisco, CA, United States
740	Kirsi Kalpala	Pfizer, New York, NY, United States
741	Melissa Miller	Pfizer, New York, NY, United States
742	Xinli Hu	Pfizer, New York, NY, United States
743	Linda McCarthy	GlaxoSmithKline, Brentford, United Kingdom
744	Onuralp Soylemez	Merck, Kenilworth, NJ, United States
745	Mark Daly	Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland
746		
747	Rheumatology Group	
748	Kari Eklund	Hospital District of Helsinki and Uusimaa, Helsinki, Finland
749	Antti Palomäki	Hospital District of Southwest Finland, Turku, Finland
750	Pia Isomäki	Pirkanmaa Hospital District, Tampere, Finland
751	Laura Pirilä	Hospital District of Southwest Finland, Turku, Finland
752	Oili Kaipainen-Seppänen	Northern Savo Hospital District, Kuopio, Finland

753	Johanna Huhtakangas	Northern Ostrobothnia Hospital District, Oulu, Finland
754	Ali Abbasi	Abbvie, Chicago, IL, United States
755	Jeffrey Waring	Abbvie, Chicago, IL, United States
756	Fedik Rahimov	Abbvie, Chicago, IL, United States
757	Apinya Lertratanakul	Abbvie, Chicago, IL, United States
758	Nizar Smaoui	Abbvie, Chicago, IL, United States
759	Anne Lehtonen	Abbvie, Chicago, IL, United States
760	David Close	Astra Zeneca, Cambridge, United Kingdom
761	Marla Hochfeld	Celgene, Summit, NJ, United States/ Bristol Myers Squibb, New York, NY, United States
762	Natalie Bowers	Genentech, San Francisco, CA, United States
763	Sarah Pendergrass	Genentech, San Francisco, CA, United States
764	Onuralp Soylemez	Merck, Kenilworth, NJ, United States
765	Kirsi Kalpala	Pfizer, New York, NY, United States
766	Nan Bing	Pfizer, New York, NY, United States
767	Xinli Hu	Pfizer, New York, NY, United States
768	Jorge Esparza Gordillo	GlaxoSmithKline, Brentford, United Kingdom
769	Kirsi Auro	GlaxoSmithKline, Brentford, United Kingdom
770	Dawn Waterworth	Janssen Biotech, Beerse, Belgium
771	Nina Mars	Institute for Molecular Medicine Finland, HiLIFE, Helsinki, Finland
772		
773	Pulmonology Group	
774	Tarja Laitinen	Pirkanmaa Hospital District, Tampere, Finland
775	Margit Pelkonen	Northern Savo Hospital District, Kuopio, Finland
776	Paula Kauppi	Hospital District of Helsinki and Uusimaa, Helsinki, Finland
777	Hannu Kankaanranta	Pirkanmaa Hospital District, Tampere, Finland
778	Terttu Harju	Northern Ostrobothnia Hospital District, Oulu, Finland
779	Riitta Lahesmaa	Hospital District of Southwest Finland, Turku, Finland
780	Nizar Smaoui	Abbvie, Chicago, IL, United States
781	Alex Mackay	Astra Zeneca, Cambridge, United Kingdom
782	Glenda Lassi	Astra Zeneca, Cambridge, United Kingdom

783	Susan Eaton	Biogen, Cambridge, MA, United States
784	Steven Greenberg	Celgene, Summit, NJ, United States/ Bristol Myers Squibb, New York, NY, United States
785	Hubert Chen	Genentech, San Francisco, CA, United States
786	Sarah Pendergrass	Genentech, San Francisco, CA, United States
787	Natalie Bowers	Genentech, San Francisco, CA, United States
788	Joanna Betts	GlaxoSmithKline, Brentford, United Kingdom
789	Soumitra Ghosh	GlaxoSmithKline, Brentford, United Kingdom
790	Kirsi Auro	GlaxoSmithKline, Brentford, United Kingdom
791	Rajashree Mishra	GlaxoSmithKline, Brentford, United Kingdom
792	Sina Rüeger	Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland
793		
794	Cardiometabolic Diseases Group	
795	Teemu Niiranen	The National Institute of Health and Welfare Helsinki, Finland
796	Felix Vaura	The National Institute of Health and Welfare Helsinki, Finland
797	Veikko Salomaa	The National Institute of Health and Welfare Helsinki, Finland
798	Markus Juonala	Hospital District of Southwest Finland, Turku, Finland
799	Kaj Metsärinne	Hospital District of Southwest Finland, Turku, Finland
800	Mika Kähönen	Pirkanmaa Hospital District, Tampere, Finland
801	Juhani Junttila	Northern Ostrobothnia Hospital District, Oulu, Finland
802	Markku Laakso	Northern Savo Hospital District, Kuopio, Finland
803	Jussi Pihlajamäki	Northern Savo Hospital District, Kuopio, Finland
804	Daniel Gordin	Hospital District of Helsinki and Uusimaa, Helsinki, Finland
805	Juha Sinisalo	Hospital District of Helsinki and Uusimaa, Helsinki, Finland
806	Marja-Riitta Taskinen	Hospital District of Helsinki and Uusimaa, Helsinki, Finland
807	Tiinamaija Tuomi	Hospital District of Helsinki and Uusimaa, Helsinki, Finland
808	Jari Laukkanen	Central Finland Health Care District, Jyväskylä, Finland
809	Benjamin Challis	Astra Zeneca, Cambridge, United Kingdom
810	Dirk Paul	Astra Zeneca, Cambridge, United Kingdom
811	Julie Hunkapiller	Genentech, San Francisco, CA, United States
812	Natalie Bowers	Genentech, San Francisco, CA, United States

813	Sarah Pendergrass	Genentech, San Francisco, CA, United States
814	Onuralp Soylemez	Merck, Kenilworth, NJ, United States
815	Jaakko Parkkinen	Pfizer, New York, NY, United States
816	Melissa Miller	Pfizer, New York, NY, United States
817	Russell Miller	Pfizer, New York, NY, United States
818	Audrey Chu	GlaxoSmithKline, Brentford, United Kingdom
819	Kirsi Auro	GlaxoSmithKline, Brentford, United Kingdom
820	Keith Usiskin	Celgene, Summit, NJ, United States/ Bristol Myers Squibb, New York, NY, United States
821	Amanda Elliott	Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland / Broad
822	Institute, Cambridge, MA, United States	
823	Joel Rämö	Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland
824	Samuli Ripatti	Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland
825	Mary Pat Reeve	Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland
826	Sanni Ruotsalainen	Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland
827		
828	Oncology Group	
829	Tuomo Meretoja	Hospital District of Helsinki and Uusimaa, Helsinki, Finland
830	Heikki Joensuu	Hospital District of Helsinki and Uusimaa, Helsinki, Finland
831	Olli Carpén	Hospital District of Helsinki and Uusimaa, Helsinki, Finland
832	Lauri Aaltonen	Hospital District of Helsinki and Uusimaa, Helsinki, Finland
833	Johanna Mattson	Hospital District of Helsinki and Uusimaa, Helsinki, Finland
834	Annika Auranen	Pirkanmaa Hospital District, Tampere, Finland
835	Peeter Karihtala	Northern Ostrobothnia Hospital District, Oulu, Finland
836	Saila Kauppila	Northern Ostrobothnia Hospital District, Oulu, Finland
837	Päivi Auvinen	Northern Savo Hospital District, Kuopio, Finland
838	Klaus Elenius	Hospital District of Southwest Finland, Turku, Finland
839	Johanna Schleutker	Hospital District of Southwest Finland, Turku, Finland
840	Relja Popovic	Abbvie, Chicago, IL, United States
841	Jeffrey Waring	Abbvie, Chicago, IL, United States
842	Bridget Riley-Gillis	Abbvie, Chicago, IL, United States

843	Anne Lehtonen	Abbvie, Chicago, IL, United States
844	Jennifer Schutzman	Genentech, San Francisco, CA, United States
845	Julie Hunkapiller	Genentech, San Francisco, CA, United States
846	Natalie Bowers	Genentech, San Francisco, CA, United States
847	Sarah Pendergrass	Genentech, San Francisco, CA, United States
848	Andrey Loboda	Merck, Kenilworth, NJ, United States
849	Aparna Chhibber	Merck, Kenilworth, NJ, United States
850	Heli Lehtonen	Pfizer, New York, NY, United States
851	Stefan McDonough	Pfizer, New York, NY, United States
852	Marika Crohns	Sanofi, Paris, France
853	Sauli Vuoti	Sanofi, Paris, France
854	Diptee Kulkarni	GlaxoSmithKline, Brentford, United Kingdom
855	Kirsi Auro	GlaxoSmithKline, Brentford, United Kingdom
856	Esa Pitkänen	Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland
857	Nina Mars	Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland
858	Mark Daly	Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland
859		
860	Ophthalmology Group	
861	Kai Kaarniranta	Northern Savo Hospital District, Kuopio, Finland
862	Joni A Turunen	Hospital District of Helsinki and Uusimaa, Helsinki, Finland
863	Terhi Ollila	Hospital District of Helsinki and Uusimaa, Helsinki, Finland
864	Sanna Seitsonen	Hospital District of Helsinki and Uusimaa, Helsinki, Finland
865	Hannu Uusitalo	Pirkanmaa Hospital District, Tampere, Finland
866	Vesa Aaltonen	Hospital District of Southwest Finland, Turku, Finland
867	Hannele Uusitalo-Järvinen	Pirkanmaa Hospital District, Tampere, Finland
868	Marja Luodonpää	Northern Ostrobothnia Hospital District, Oulu, Finland
869	Nina Hautala	Northern Ostrobothnia Hospital District, Oulu, Finland
870	Mengzhen Liu	Abbvie, Chicago, IL, United States
871	Heiko Runz	Biogen, Cambridge, MA, United States
872	Stephanie Loomis	Biogen, Cambridge, MA, United States

873	Erich Strauss	Genentech, San Francisco, CA, United States
874	Natalie Bowers	Genentech, San Francisco, CA, United States
875	Hao Chen	Genentech, San Francisco, CA, United States
876	Sarah Pendergrass	Genentech, San Francisco, CA, United States
877	Anna Podgornaia	Merck, Kenilworth, NJ, United States
878	Juha Karjalainen	Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland / Broad
879	Institute, Cambridge, MA,	United States
880	Esa Pitkänen	Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland
881		
882	Dermatology Group	
883	Kaisa Tasanen	Northern Ostrobothnia Hospital District, Oulu, Finland
884	Laura Huilaja	Northern Ostrobothnia Hospital District, Oulu, Finland
885	Katariina Hannula-Jouppi	Hospital District of Helsinki and Uusimaa, Helsinki, Finland
886	Teea Salmi	Pirkanmaa Hospital District, Tampere, Finland
887	Sirkku Peltonen	Hospital District of Southwest Finland, Turku, Finland
888	Leena Koulu	Hospital District of Southwest Finland, Turku, Finland
889	Kirsi Kalpala	Pfizer, New York, NY, United States
890	Ying Wu	Pfizer, New York, NY, United States
891	David Choy	Genentech, San Francisco, CA, United States
892	Sarah Pendergrass	Genentech, San Francisco, CA, United States
893	Nizar Smaoui	Abbvie, Chicago, IL, United States
894	Fedik Rahimov	Abbvie, Chicago, IL, United States
895	Anne Lehtonen	Abbvie, Chicago, IL, United States
896	Dawn Waterworth	Janssen Biotech, Beerse, Belgium
897		
898	Odontology Group	
899	Pirkko Pussinen	Hospital District of Helsinki and Uusimaa, Helsinki, Finland
900	Aino Salminen	Hospital District of Helsinki and Uusimaa, Helsinki, Finland
901	Tuula Salo	Hospital District of Helsinki and Uusimaa, Helsinki, Finland
902	David Rice	Hospital District of Helsinki and Uusimaa, Helsinki, Finland

903	Pekka Nieminen	Hospital District of Helsinki and Uusimaa, Helsinki, Finland
904	Ulla Palotie	Hospital District of Helsinki and Uusimaa, Helsinki, Finland
905	Juha Sinisalo	Hospital District of Helsinki and Uusimaa, Helsinki, Finland
906	Maria Siponen	Northern Savo Hospital District, Kuopio, Finland
907	Liisa Suominen	Northern Savo Hospital District, Kuopio, Finland
908	Päivi Mäntylä	Northern Savo Hospital District, Kuopio, Finland
909	Ulvi Gursoy	Hospital District of Southwest Finland, Turku, Finland
910	Vuokko Anttonen	Northern Ostrobothnia Hospital District, Oulu, Finland
911	Kirsi Sipilä	Northern Ostrobothnia Hospital District, Oulu, Finland
912	Sarah Pendergrass	Genentech, San Francisco, CA, United States
913		
914	Women's Health and Reproduction Group	
915	Hannele Laivuori	Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland
916	Venla Kurra	Pirkanmaa Hospital District, Tampere, Finland
917	Oskari Heikinheimo	Hospital District of Helsinki and Uusimaa, Helsinki, Finland
918	Ilkka Kalliala	Hospital District of Helsinki and Uusimaa, Helsinki, Finland
919	Laura Kotaniemi-Talonen	Pirkanmaa Hospital District, Tampere, Finland
920	Kari Nieminen	Pirkanmaa Hospital District, Tampere, Finland
921	Päivi Polo	Hospital District of Southwest Finland, Turku, Finland
922	Kaarin Mäkikallio	Hospital District of Southwest Finland, Turku, Finland
923	Eeva Ekholm	Hospital District of Southwest Finland, Turku, Finland
924	Marja Väärasmäki	Northern Ostrobothnia Hospital District, Oulu, Finland
925	Outi Uimari	Northern Ostrobothnia Hospital District, Oulu, Finland
926	Laure Morin-Papunen	Northern Ostrobothnia Hospital District, Oulu, Finland
927	Marjo Tuppurainen	Northern Savo Hospital District, Kuopio, Finland
928	Katja Kivinen	Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland
929	Elisabeth Widen	Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland
930	Taru Tukiainen	Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland
931	Mary Pat Reeve	Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland
932	Mark Daly	Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland

933	Liu Aoxing	Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland
934	Eija Laakkonen	University of Jyväskylä, Jyväskylä, Finland
935	Niko Välimäki	University of Helsinki, Helsinki, Finland
936	Lauri Aaltonen	Hospital District of Helsinki and Uusimaa, Helsinki, Finland
937	Johannes Kettunen	Northern Ostrobothnia Hospital District, Oulu, Finland
938	Mikko Arvas	Finnish Red Cross Blood Service, Helsinki, Finland
939	Jeffrey Waring	Abbvie, Chicago, IL, United States
940	Bridget Riley-Gillis	Abbvie, Chicago, IL, United States
941	Mengzhen Liu	Abbvie, Chicago, IL, United States
942	Janet Kumar	GlaxoSmithKline, Brentford, United Kingdom
943	Kirsi Auro	GlaxoSmithKline, Brentford, United Kingdom
944	Andrea Ganna	Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland
945	Sarah Pendergrass	Genentech, San Francisco, CA, United States
946		
947	FinnGen Analysis working group	
948	Justin Wade Davis	Abbvie, Chicago, IL, United States
949	Bridget Riley-Gillis	Abbvie, Chicago, IL, United States
950	Danjuma Quarless	Abbvie, Chicago, IL, United States
951	Fedik Rahimov	Abbvie, Chicago, IL, United States
952	Sahar Esmaeeli	Abbvie, Chicago, IL, United States
953	Slavé Petrovski	Astra Zeneca, Cambridge, United Kingdom
954	Eleonor Wigmore	Astra Zeneca, Cambridge, United Kingdom
955	Adele Mitchell	Biogen, Cambridge, MA, United States
956	Benjamin Sun	Biogen, Cambridge, MA, United States
957	Ellen Tsai	Biogen, Cambridge, MA, United States
958	Denis Baird	Biogen, Cambridge, MA, United States
959	Paola Bronson	Biogen, Cambridge, MA, United States
960	Ruoyu Tian	Biogen, Cambridge, MA, United States
961	Stephanie Loomis	Biogen, Cambridge, MA, United States
962	Yunfeng Huang	Biogen, Cambridge, MA, United States
963	Joseph Maranville	Celgene, Summit, NJ, United States/ Bristol Myers Squibb, New York, NY, United States

964	Shameek Biswas	Celgene, Summit, NJ, United States/ Bristol Myers Squibb, New York, NY, United States
965	Elmutaz Mohammed	Celgene, Summit, NJ, United States/ Bristol Myers Squibb, New York, NY, United States
966	Samir Wadhawan	Celgene, Summit, NJ, United States/ Bristol Myers Squibb, New York, NY, United States
967	Erika Kvikstad	Celgene, Summit, NJ, United States/ Bristol Myers Squibb, New York, NY, United States
968	Minal Caliskan	Celgene, Summit, NJ, United States/ Bristol Myers Squibb, New York, NY, United States
969	Diana Chang	Genentech, San Francisco, CA, United States
970	Julie Hunkapiller	Genentech, San Francisco, CA, United States
971	Tushar Bhangale	Genentech, San Francisco, CA, United States
972	Natalie Bowers	Genentech, San Francisco, CA, United States
973	Sarah Pendergrass	Genentech, San Francisco, CA, United States
974	Kirill Shkura	Merck, Kenilworth, NJ, United States
975	Victor Neduva	Merck, Kenilworth, NJ, United States
976	Xing Chen	Pfizer, New York, NY, United States
977	Åsa Hedman	Pfizer, New York, NY, United States
978	Karen S King	GlaxoSmithKline, Brentford, United Kingdom
979	Padhraig Gormley	GlaxoSmithKline, Brentford, United Kingdom
980	Jimmy Liu	GlaxoSmithKline, Brentford, United Kingdom
981	Clarence Wang	Sanofi, Paris, France
982	Ethan Xu	Sanofi, Paris, France
983	Franck Auge	Sanofi, Paris, France
984	Clement Chatelain	Sanofi, Paris, France
985	Deepak Rajpal	Sanofi, Paris, France
986	Dongyu Liu	Sanofi, Paris, France
987	Katherine Call	Sanofi, Paris, France
988	Tai-He Xia	Sanofi, Paris, France
989	Beryl Cummings	Maze Therapeutics, San Francisco, CA, United States
990	Matt Brauer	Maze Therapeutics, San Francisco, CA, United States
991	Huilei Xu	Novartis, Basel, Switzerland
992	Amy Cole	Novartis, Basel, Switzerland
993	Jonathan Chung	Novartis, Basel, Switzerland

994	Jaison Jacob	Novartis, Basel, Switzerland
995	Katrina de Lange	Novartis, Basel, Switzerland
996	Jonas Zierer	Novartis, Basel, Switzerland
997	Mitja Kurki	Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland / Broad
998	Institute, Cambridge, MA, United States	
999	Samuli Ripatti	Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland
1000	Mark Daly	Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland
1001	Juha Karjalainen	Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland / Broad
1002	Institute, Cambridge, MA, United States	
1003	Aki Havulinna	Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland
1004	Juha Mehtonen	Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland
1005	Priit Palta	Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland
1006	Shabbeer Hassan	Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland
1007	Pietro Della Briotta Parolo	Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland
1008	Wei Zhou	Broad Institute, Cambridge, MA, United States
1009	Mutaamba Maasha	Broad Institute, Cambridge, MA, United States
1010	Shabbeer Hassan	Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland
1011	Susanna Lemmelä	Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland
1012	Manuel Rivas	University of Stanford, Stanford, CA, United States
1013	Aarno Palotie	Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland
1014	Arto Lehisto	Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland
1015	Andrea Ganna	Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland
1016	Vincent Llorens	Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland
1017	Hannele Laivuori	Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland
1018	Mari E Niemi	Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland
1019	Taru Tukiainen	Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland
1020	Mary Pat Reeve	Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland
1021	Henrike Heyne	Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland
1022	Nina Mars	Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland
1023	Kimmo Palin	University of Helsinki, Helsinki, Finland

1024	Javier Garcia-Tabuenca	University of Tampere, Tampere, Finland
1025	Harri Siirtola	University of Tampere, Tampere, Finland
1026	Tuomo Kiiskinen	Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland
1027	Jiwoo Lee	Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland / Broad
1028	Institute, Cambridge, MA, United States	
1029	Kristin Tsuo	Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland / Broad
1030	Institute, Cambridge, MA, United States	
1031	Amanda Elliott	Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland / Broad
1032	Institute, Cambridge, MA, United States	
1033	Kati Kristiansson	THL Biobank / The National Institute of Health and Welfare Helsinki, Finland
1034	Mikko Arvas	Finnish Red Cross Blood Service / Finnish Hematology Registry and Clinical Biobank,
1035	Helsinki, Finland	
1036	Kati Hyvärinen	Finnish Red Cross Blood Service, Helsinki, Finland
1037	Jarmo Ritari	Finnish Red Cross Blood Service, Helsinki, Finland
1038	Miika Koskinen	Helsinki Biobank / Helsinki University and Hospital District of Helsinki and Uusimaa,
1039	Helsinki	
1040	Olli Carpén	Helsinki Biobank / Helsinki University and Hospital District of Helsinki and Uusimaa,
1041	Helsinki	
1042	Johannes Kettunen	Northern Finland Biobank Borealis / University of Oulu / Northern Ostrobothnia Hospital
1043	District, Oulu, Finland	
1044	Katri Pylkäs	University of Oulu, Oulu, Finland
1045	Marita Kalaoja	University of Oulu, Oulu, Finland
1046	Minna Karjalainen	University of Oulu, Oulu, Finland
1047	Tuomo Mantere	Northern Finland Biobank Borealis / University of Oulu / Northern Ostrobothnia Hospital
1048	District, Oulu, Finland	
1049	Eeva Kangasniemi	Finnish Clinical Biobank Tampere / University of Tampere / Pirkanmaa Hospital District,
1050	Tampere, Finland	
1051	Sami Heikkinen	University of Eastern Finland, Kuopio, Finland
1052	Arto Mannermaa	Biobank of Eastern Finland / University of Eastern Finland / Northern Savo Hospital District,
1053	Kuopio, Finland	

1054	Eija Laakkonen	University of Jyväskylä, Jyväskylä, Finland
1055	Samuel Heron	University of Turku, Turku, Finland
1056	Dhanaprakash Jambulingam	University of Turku, Turku, Finland
1057	Venkat Subramaniam Rathinakannan	University of Turku, Turku, Finland
1058	Nina Pitkänen	Auria Biobank / University of Turku / Hospital District of Southwest Finland, Turku,
1059	Finland	
1060		
1061	Biobank directors	
1062	Lila Kallio	Auria Biobank / University of Turku / Hospital District of Southwest Finland, Turku,
1063	Finland	
1064	Sirpa Soini	THL Biobank / The National Institute of Health and Welfare Helsinki, Finland
1065	Jukka Partanen	Finnish Red Cross Blood Service / Finnish Hematology Registry and Clinical Biobank,
1066	Helsinki, Finland	
1067	Eero Punkka	Helsinki Biobank / Helsinki University and Hospital District of Helsinki and Uusimaa,
1068	Helsinki	
1069	Raisa Serpi	Northern Finland Biobank Borealis / University of Oulu / Northern Ostrobothnia Hospital
1070	District, Oulu, Finland	
1071	Johanna Mäkelä	Finnish Clinical Biobank Tampere / University of Tampere / Pirkanmaa Hospital District,
1072	Tampere, Finland	
1073	Veli-Matti Kosma	Biobank of Eastern Finland / University of Eastern Finland / Northern Savo Hospital District,
1074	Kuopio, Finland	
1075	Teijo Kuopio	Central Finland Biobank / University of Jyväskylä / Central Finland Health Care District,
1076	Jyväskylä, Finland	
1077		
1078	FinnGen Teams	
1079	Administration	
1080	Anu Jalanko	Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland
1081	Huei-Yi Shen	Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland
1082	Risto Kajanne	Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland
1083	Mervi Aavikko	Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland

1084

1085

1086 **Analysis**

1087 Mitja Kurki Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland / Broad

1088 Institute, Cambridge, MA, United States

1089 Juha Karjalainen Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland / Broad

1090 Institute, Cambridge, MA, United States

1091 Pietro Della Briotta Parolo Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland

1092 Arto Lehisto Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland

1093 Juha Mehtonen Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland

1094 Wei Zhou Broad Institute, Cambridge, MA, United States

1095 Masahiro Kanai Broad Institute, Cambridge, MA, United States

1096 Mutaamba Maasha Broad Institute, Cambridge, MA, United States

1097

1098 **Clinical Endpoint Development**

1099 Hannele Laivuori Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland

1100 Aki Havulinna Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland

1101 Susanna Lemmelä Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland

1102 Tuomo Kiiskinen Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland

1103 L. Elisa Lahtela Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland

1104

1105 **Communication**

1106 Mari Kaunisto Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland

1107

1108 **E-Science**

1109 Elina Kilpeläinen Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland

1110 Timo P. Sipilä Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland

1111 Georg Brein Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland

1112 Oluwaseun Alexander Dada Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland

1113	Awaisa Ghazal	Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland
1114	Anastasia Shcherban	Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland
1115		
1116	Genotyping	
1117	Kati Donner	Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland
1118	Timo P. Sipilä	Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland
1119		
1120	Sample Collection Coordination	
1121	Anu Loukola	Helsinki Biobank / Helsinki University and Hospital District of Helsinki and Uusimaa,
1122	Helsinki	
1123		
1124	Sample Logistics	
1125	Päivi Laiho	THL Biobank / The National Institute of Health and Welfare Helsinki, Finland
1126	Tuuli Sistonen	THL Biobank / The National Institute of Health and Welfare Helsinki, Finland
1127	Essi Kaiharju	THL Biobank / The National Institute of Health and Welfare Helsinki, Finland
1128	Markku Laukkanen	THL Biobank / The National Institute of Health and Welfare Helsinki, Finland
1129	Elina Järvensivu	THL Biobank / The National Institute of Health and Welfare Helsinki, Finland
1130	Sini Lähteenmäki	THL Biobank / The National Institute of Health and Welfare Helsinki, Finland
1131	Lotta Männikkö	THL Biobank / The National Institute of Health and Welfare Helsinki, Finland
1132	Regis Wong	THL Biobank / The National Institute of Health and Welfare Helsinki, Finland
1133		
1134	Registry Data Operations	
1135	Hannele Mattsson	THL Biobank / The National Institute of Health and Welfare Helsinki, Finland
1136	Kati Kristiansson	THL Biobank / The National Institute of Health and Welfare Helsinki, Finland
1137	Susanna Lemmelä	Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland
1138	Sami Koskelainen	THL Biobank / The National Institute of Health and Welfare Helsinki, Finland
1139	Tero Hiekkalinna	THL Biobank / The National Institute of Health and Welfare Helsinki, Finland
1140	Teemu Paaanen	THL Biobank / The National Institute of Health and Welfare Helsinki, Finland
1141		
1142		

1143 Sequencing Informatics

1144 Priit Palta Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland

1145 Kalle Pärn Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland

1146 Shuang Luo Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland

1147 Vishal Sinha Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland

1148

1149 Trajectory

1150 Tarja Laitinen Pirkanmaa Hospital District, Tampere, Finland

1151 Mary Pat Reeve Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland

1152 Harri Siirtola University of Tampere, Tampere, Finland

1153 Javier Gracia-Tabuenca University of Tampere, Tampere, Finland

1154 Mika Helminen University of Tampere, Tampere, Finland

1155 Tiina Luukkaala University of Tampere, Tampere, Finland

1156 Iida Vähätalo University of Tampere, Tampere, Finland

1157

1158 Data protection officer

1159 Tero Jyrhämä Institute for Molecular Medicine Finland, HiLIFE, University of Helsinki, Finland

1160

1161 FinBB - Finnish biobank cooperative

1162 Marco Hautalahti

1163 Laura Mustaniemi

1164 Mirkka Koivusalo

1165 Sarah Smith

1166 Tom Southerington

1167

1168 Declaration of Interests

1169 The authors declare no competing interests.

1170

1171 Ethics statement

1172 Patients and control subjects in FinnGen provided informed consent for biobank research, based on
 1173 the Finnish Biobank Act. Alternatively, separate research cohorts, collected prior the Finnish Biobank
 1174 Act came into effect (in September 2013) and start of FinnGen (August 2017), were collected based
 1175 on study-specific consents and later transferred to the Finnish biobanks after approval by Fimea, the
 1176 National Supervisory Authority for Welfare and Health. Recruitment protocols followed the biobank
 1177 protocols approved by Fimea. The Coordinating Ethics Committee of the Hospital District of Helsinki
 1178 and Uusimaa (HUS) approved the FinnGen study protocol Nr HUS/990/2017.

1179
 1180 The FinnGen study is approved by Finnish Institute for Health and Welfare (permit numbers:
 1181 THL/2031/6.02.00/2017, THL/1101/5.05.00/2017, THL/341/6.02.00/2018,
 1182 THL/2222/6.02.00/2018, THL/283/6.02.00/2019, THL/1721/5.05.00/2019,
 1183 THL/1524/5.05.00/2020, and THL/2364/14.02/2020), Digital and population data service agency
 1184 (permit numbers: VRK43431/2017-3, VRK/6909/2018-3, VRK/4415/2019-3), the Social Insurance
 1185 Institution (permit numbers: KELA 58/522/2017, KELA 131/522/2018, KELA 70/522/2019, KELA
 1186 98/522/2019, KELA 138/522/2019, KELA 2/522/2020, KELA 16/522/2020 and Statistics Finland
 1187 (permit numbers: TK-53-1041-17 and TK-53-90-20).

1188
 1189 The Biobank Access Decisions for FinnGen samples and data utilized in FinnGen Data Freeze 6
 1190 include: THL Biobank BB2017_55, BB2017_111, BB2018_19, BB_2018_34, BB_2018_67,
 1191 BB2018_71, BB2019_7, BB2019_8, BB2019_26, BB2020_1, Finnish Red Cross Blood Service
 1192 Biobank 7.12.2017, Helsinki Biobank HUS/359/2017, Auria Biobank AB17-5154, Biobank Borealis
 1193 of Northern Finland_2017_1013, Biobank of Eastern Finland 1186/2018, Finnish Clinical Biobank
 1194 Tampere MH0004, Central Finland Biobank 1-2017, and Terveystalo Biobank STB 2018001.

1195

1196 The FINRISK data used for the study were obtained from THL Biobank with application number
1197 BB2015_55.1 and UKBB data using the UK Biobank Resource with application number 22627.
1198 The study was approved by the Estonian Committee on Bioethics and Human Research (approval
1199 number 1.1-12/624).

1200 Figures and tables

1201 **Figure 1:** GWAS results for coronary atherosclerosis in FinnGen. Total number of independent
1202 genome-wide significant associations (GWS; $p < 5 \times 10^{-8}$) is 38, the lead variant in each marked with
1203 diamonds. Four novel associations for CVD- related phenotypes are highlighted with ± 750 Mb
1204 around the lead variant in the region as red and the lead variant marked with red diamond.....10

1205 **Figure 2:** Phenome-wide association study (PheWAS) results for rs534125149. Total number of
1206 tested endpoints is 2 861 (A complete list of endpoints analyzed and their definitions is available at
1207 <https://www.finnngen.fi/en/researchers/clinical-endpoints>). The dashed line represents the phenome-
1208 wide significance threshold, multiple testing corrected by the number of endpoints = $0.05 / 2\,861 =$
1209 1.75×10^{-5} . All endpoints reaching that threshold are labelled in the figure.....13

1210 **Figure 3:** Comparison of the effects (OR) of rs534125149 and rs201988637 for 14 endpoints with
1211 p-value $< 1.75 \times 10^{-5}$ (PWS) for rs534125149. 95% confidence intervals represented as grey lines.14

1212 **Figure 4:** Cumulative incidence plots for myocardial infarction. Red line represents carriers (homo-
1213 or heterozygous) for either rs534125149 or rs201988637, and blue line represent non-carriers.
1214 Dashed lines represent 95%- confidence intervals.16

1215 **Figure 5:** Results for pulse pressure association across all cohorts with splice acceptor variant
1216 rs201988637 present.18

1217
1218 **Table 1:** Basic characteristics of the study cohort.....5

1219 **Table 2:** Lead variants in novel loci associated with coronary atherosclerosis.11

1220 **Table 3:** Results of the conditional analysis on MI and coronary atherosclerosis.15

