THE LANCET
Diabetes & Endocrinology

Supplementary appendix

This appendix formed part of the original submission and has been peer reviewed. We post it as supplied by the authors.

Supplement to: Saleheen D, Haycock PC, Zhao W, et al. Apolipoprotein(a) isoform size, lipoprotein(a) concentration, and coronary artery disease: a mendelian randomisation analysis. *Lancet Diabetes Endocrinol* 2017; published online April 10. http://dx.doi.org/10.1016/S2213-8587(17)30088-8.

Supplementary Material

Supplementary Methods

Biochemical Analyses

In PROMIS, non-fasting blood samples (with the time since last meal recorded) have been drawn by phlebotomists from each participant and centrifuged within 45 minutes of venepuncture. Serum, plasma and whole blood samples were stored at -70°C within 45 minutes of venepuncture[1]. All samples were transported on dry-ice to the central laboratory at the Center for Non-Communicable Diseases (CNCD), Pakistan, where serum and plasma samples were aliquoted in different storage vials. Samples from cases and controls were randomly assorted in storage boxes (each box containing 81 samples) and stored at -70°C for any subsequent laboratory analyses. All biochemical assays were conducted in automated auto-analysers. At CNCD Pakistan, measurements for total-cholesterol, HDL-C, LDL-C, triglycerides, creatinine were conducted in serum samples using enzymatic assays; whereas levels of HbA1c were measured using a turbidemetric assay in whole-blood samples (all using kits manufactured by Roche Diagnostics). For further measurements, aliquots of serum and plasma samples were transported on dry ice to the Translational Research Center (TRC), University of Pennsylvania, USA, where following biochemical assays were conducted: apolipoproteins (apoA-I, apoA-II, apoB, apoC-III, apoE) and non-esterified fatty acids were measured through immunoturbidometric assays using kits by Roche Diagnostics or Kamiya; lipoprotein (a) levels were determined through a turbidimetric assay using reagents and calibrators from Denka Seiken (Niigata, Japan); LpPLA2 mass and activity levels were determined using immunoassays manufactured by diaDexus; measurements for insulin, leptin and adiponectin were made using radio-immunoassays by LINCO; levels of adhesion molecules (ICAM-1, VCAM-1, P- and E-Selectin) were determined through enzymatic assays by R&D; and measurements for CRP, ALT, AST, cystatin-C, ferritin, ceruloplasmin, thyroid stimulating hormone, alkaline phosphatase, sodium, potassium, choloride, phosphate, sex-harmone binding globulin were made using enzymatic assays manufactured by Abott Diagnostics. Glomerular filtration rate (eGFR) was estimated from serum creatinine levels using the MDRD equation[2].

Statistical Methods

Genome-wide analyses Participants were genotyped on the Illumina 660W-Quad or the Illumina OmniExpress Bead chip. Quality control (QC) criteria on the genotyped data included, removal of participants or single nucleotide polymorphisms (SNP's) that had a missing rate >5%. SNPs with a MAF <1% and a P-value of $<10^{-7}$ for the Hardy-Weinberg equilibrium test were also excluded from the analyses. Further QC included removal of participants with discrepancy between their reported sex and genetic sex determined from the X chromosome. To identify sample duplications, unintentional use of related samples (cryptic relatedness) and sample contamination (individuals who seem to be related to nearly everyone in the sample), identity-by-descent (IBD) analyses were

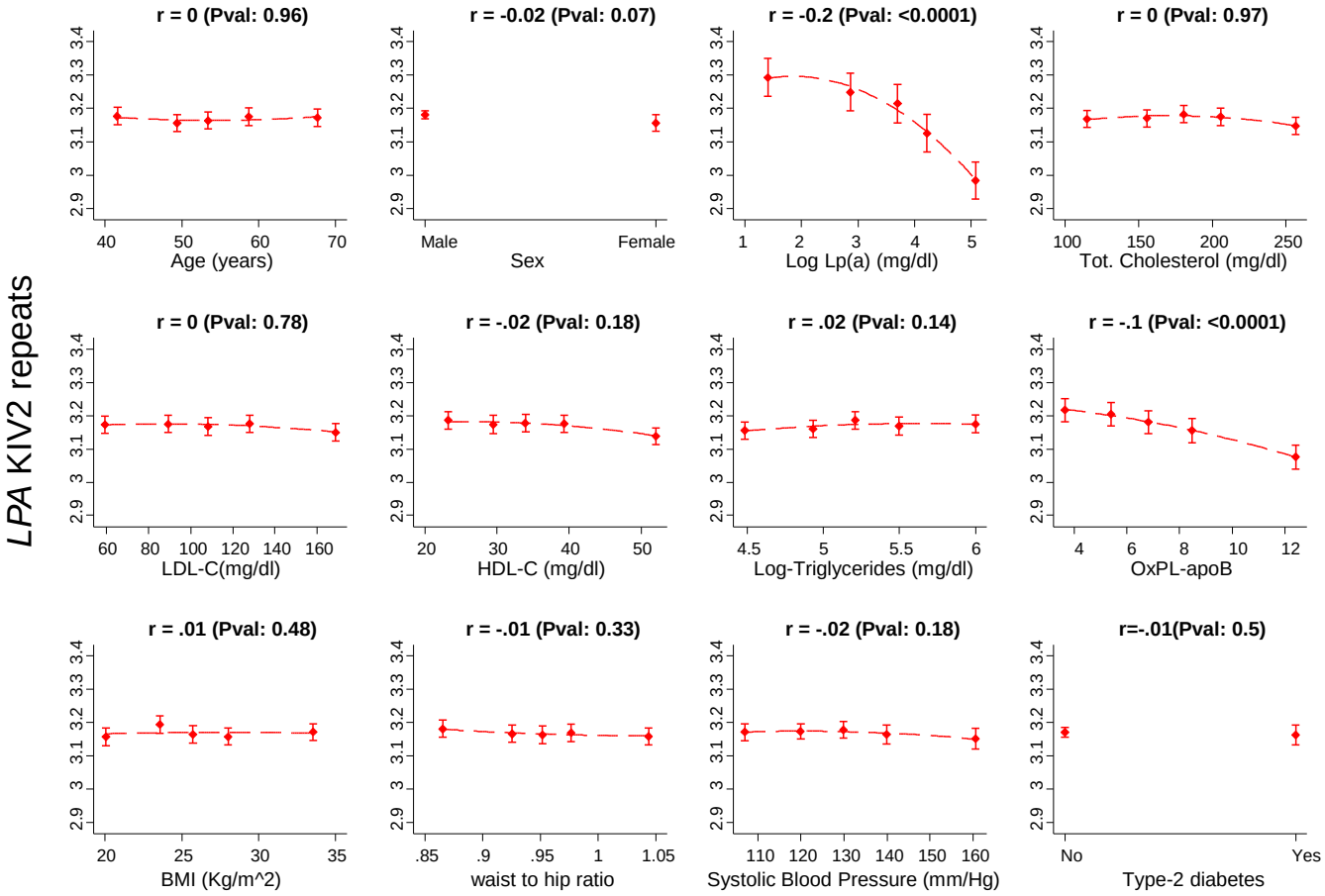
conducted in PLINK[4]. To account for population substructure in Pakistanis, principal component analyses were conducted on the pruned dataset as described previously[1]. After implementing QC, imputations were conducted using the global reference panel in the 1000-Genomes project (March 2012 (v3) release) using IMPUTE-2. All analyses were conducted in SNPTEST, adjusted for the first 10 principal components, age and sex.

Genetic Risk Score Analysis We utilized a two-sample MR method[5] to estimate effects for a multi-SNP genetic instrument by using summary statistics from the CARDIoGRAMplusC4D consortium. This method has been previously validated to infer causal effects (odds ratio) and associated standard error[6].

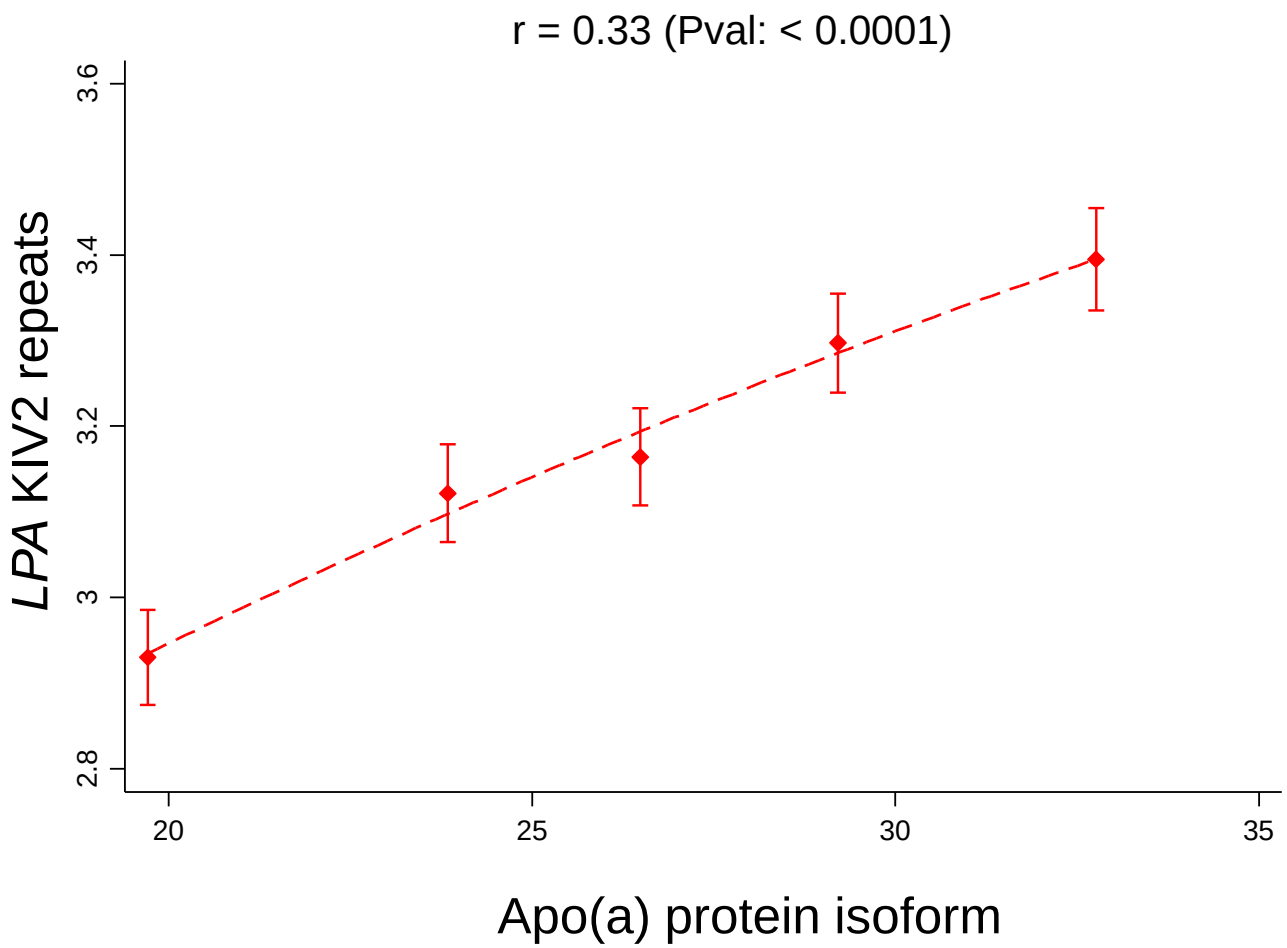
Supplementary References

1. Saleheen D, Zaidi M, Rasheed A et al. The Pakistan Risk of Myocardial Infarction Study: a resource for the study of genetic, lifestyle and other determinants of myocardial infarction in South Asia. Eur J Epidemiol. 2009;24:329-338.
2. Levey AS, Coresh J, Greene T, Stevens LA, Zhang YL, Hendriksen S, Kusek JW, Van Lente F; Chronic Kidney Disease Epidemiology Collaboration. Using standardized serum creatinine values in the modification of diet in renal disease study equation for estimating glomerular filtration rate. Ann Intern Med. 2006 Aug 15;145(4):247-54.
3. Purcell S, Neale B, Todd-Brown K et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007;81:559-575.
4. Byun YS, Lee JH, Arsenault BJ, Yang X, Bao W, DeMicco D, Laskey R, Witztum JL, Tsimikas S and Investigators TNTT. Relationship of Oxidized Phospholipids on Apolipoprotein B-100 to Cardiovascular Outcomes in Patients Treated With Intensive Versus Moderate Atorvastatin Therapy: The TNT Trial. J Am Coll Cardiol. 2015;65:1286-95.
5. Dastani Z, Hivert MF, Timpson N, et al. Novel loci for adiponectin levels and their influence on type 2 diabetes and metabolic traits: a multi-ethnic meta-analysis of 45,891 individuals. PLoS Genet. 2012;8(3):e1002607.
6. Evans, D.M. & Davey Smith, G. Mendelian Randomization: New Applications in the Coming Age of Hypothesis-Free Causality. Annu Rev Genomics Hum Genet. 2015;16:327-50.

eFigure 1a. Cross-sectional associations of *LPA* KIV2 repeats in PROMIS controls participants

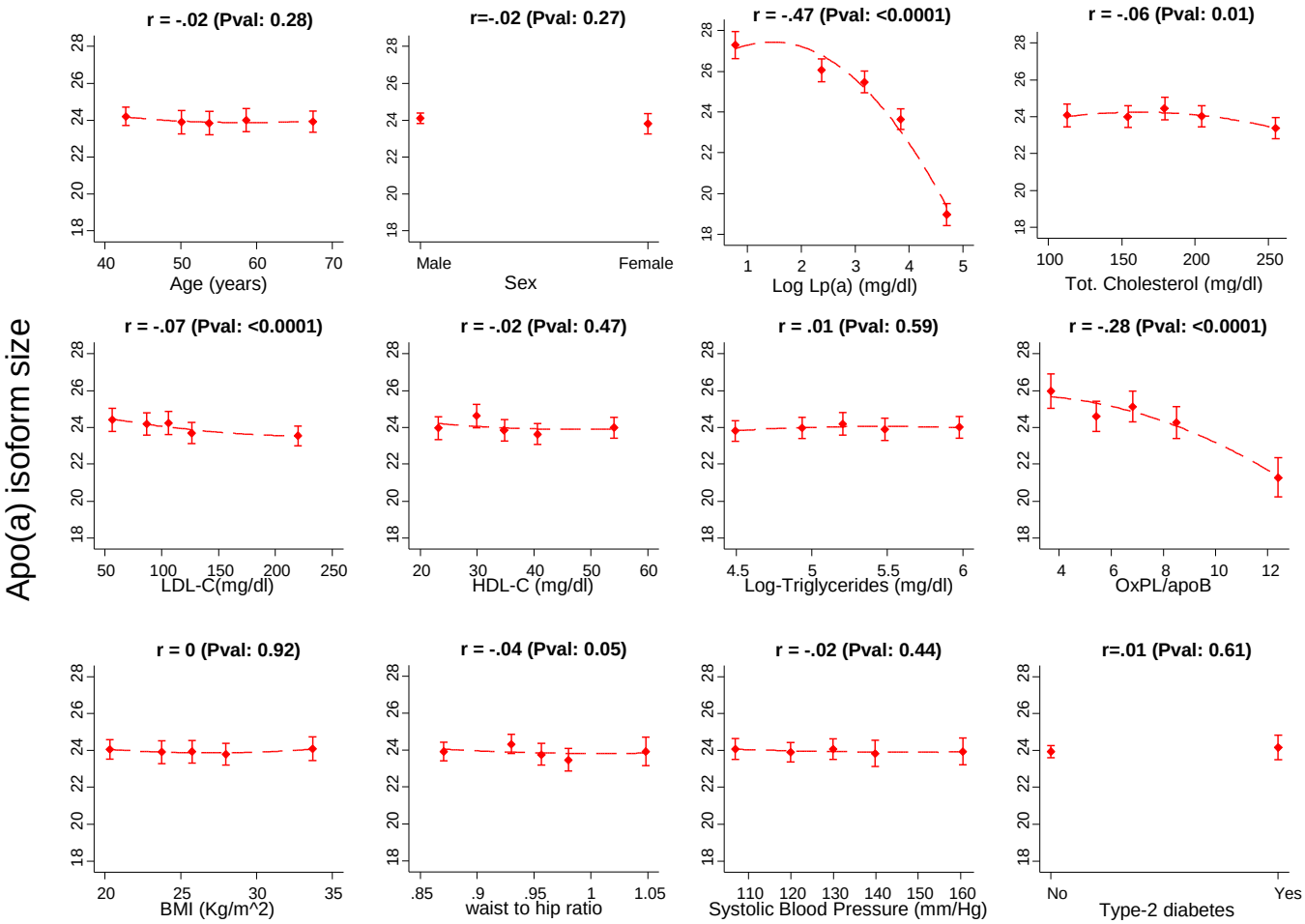


eFigure 1b. Cross-sectional association of *LPA* KIV2 repeats with apo(a) protein isoforms in PROMIS controls participants

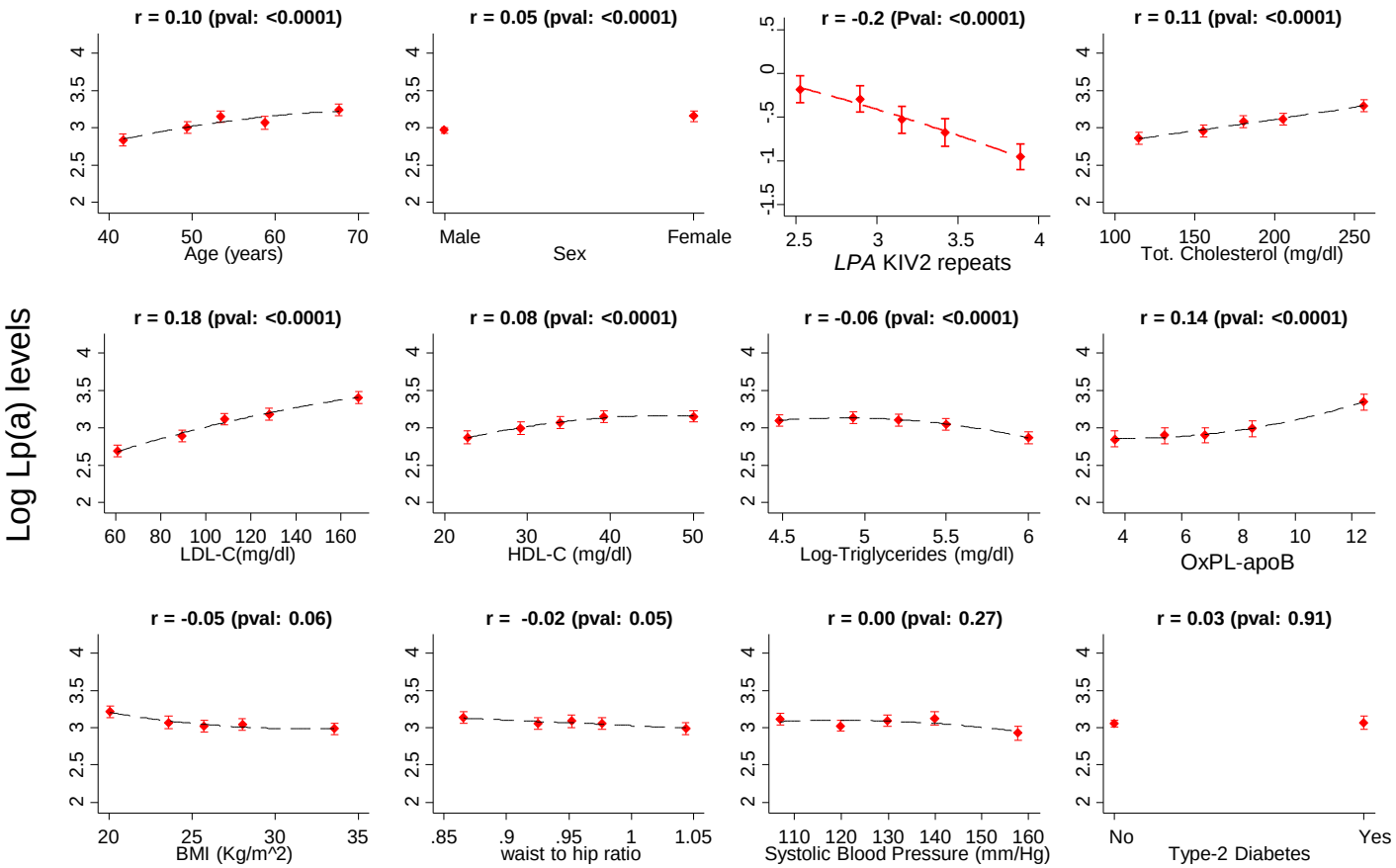


Partial correlation coefficients (r) adjusted for age and sex have been calculated using all data from the study control participants.

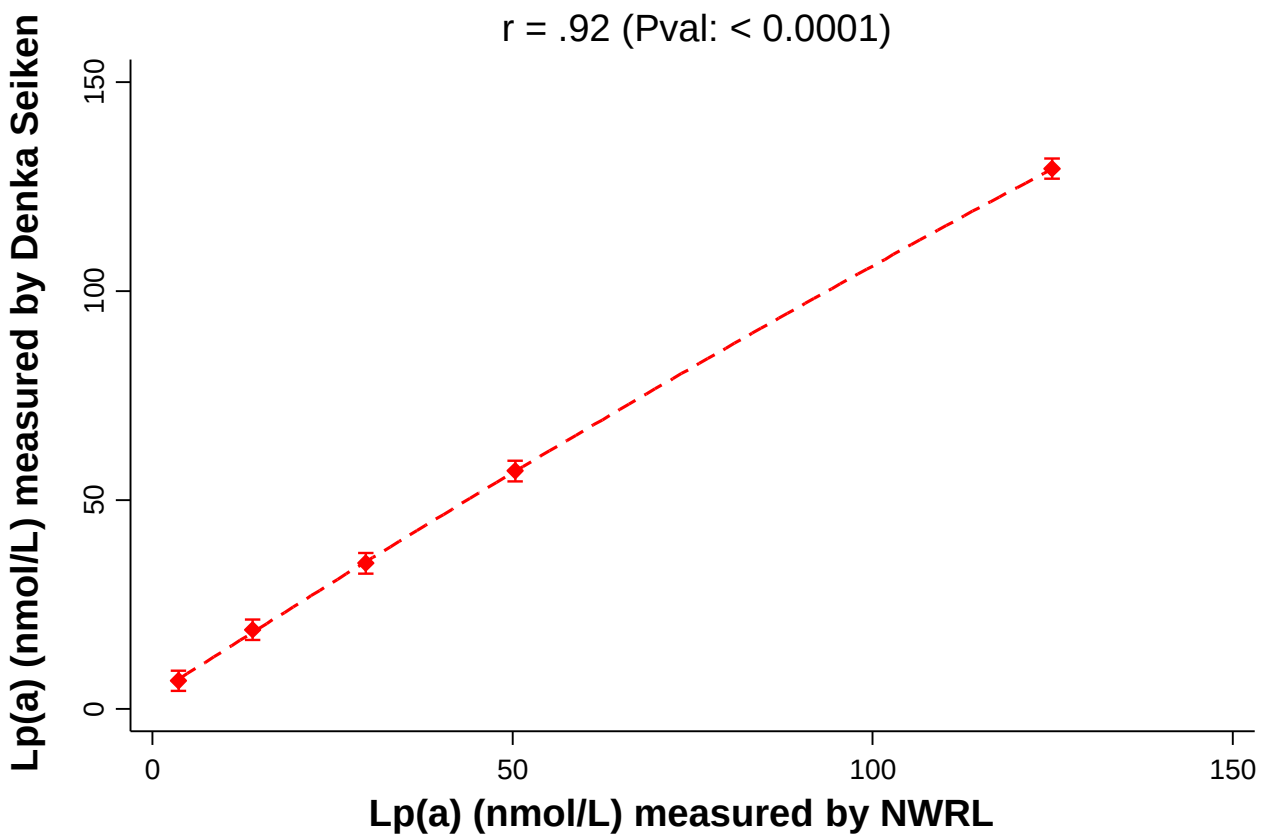
eFigure 2. Cross-sectional associations of apo(a) protein isoform in PROMIS controls participants



eFigure 3. Cross-sectional associations of Lp(a) concentration in PROMIS control participants

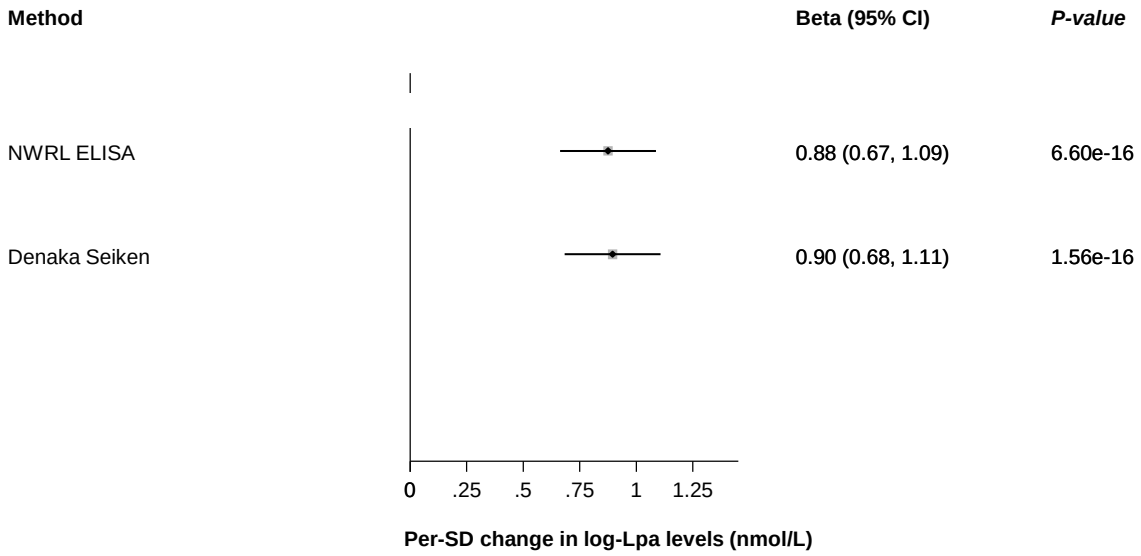


eFigure 4. Correlation of Lp(a) levels measured by the Denka Seiken immunoturbidemetric method and the Northwest Lipid Metabolism and Diabetes Research Laboratories (NWRL) ELISA method (n = 2617)

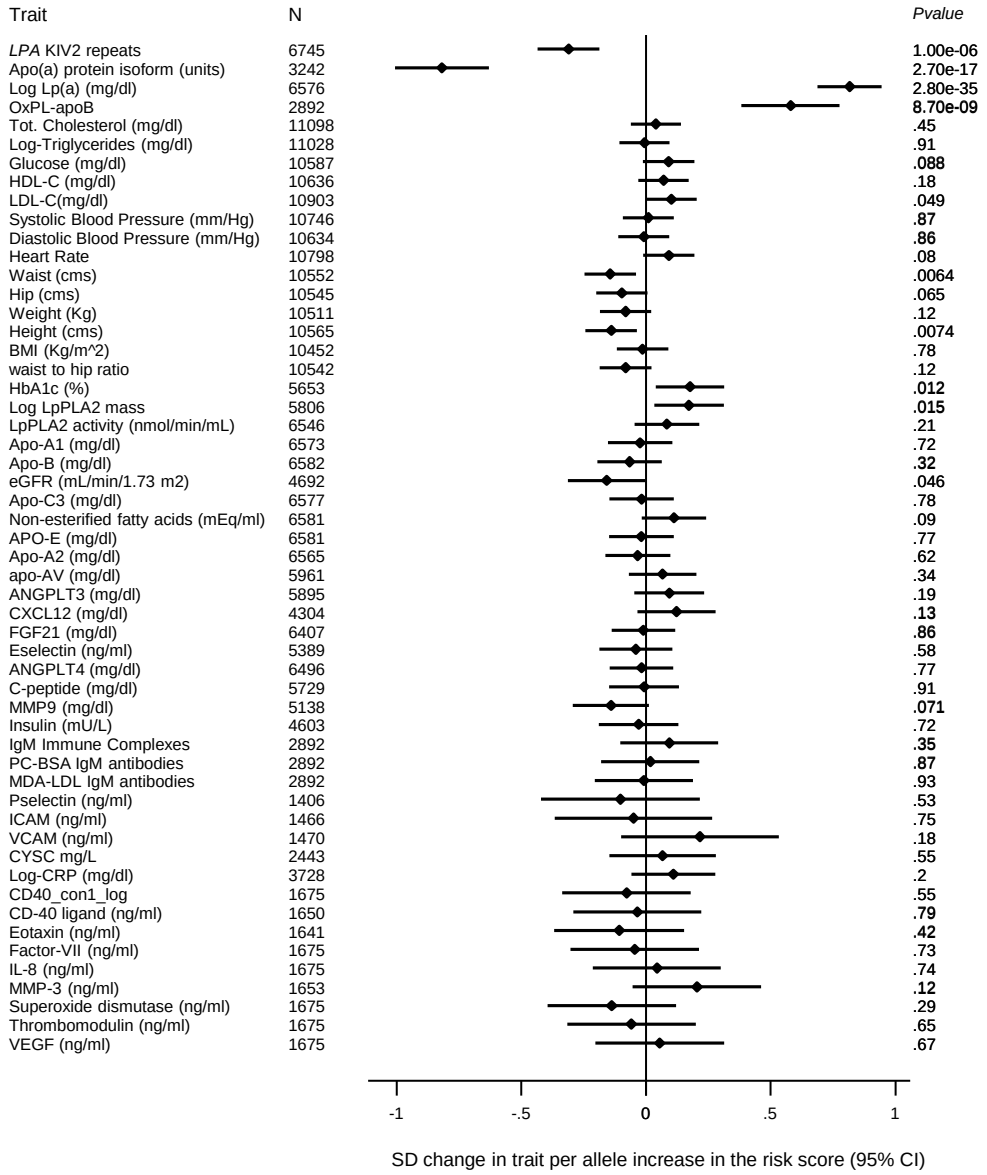


Partial correlation coefficients adjusted for age and sex were calculated to assess correlations between the two traits. Association of the two traits was further characterized by using linear regression analyses adjusted for age and sex. From each fitted regression model, overall adjusted mean values and 95 percent confidence intervals for Lp(a) (nmol/L) measured by NWRL within quintiles of Lp(a) (nmol/L) measured by Denka Seiken, were obtained.

eFigure 5. Association of rs10455872 and rs3798220 variants with Lp(a) levels measured by the Denka Seiken immunoturbidemetric method and the Northwest Lipid Metabolism and Diabetes Research Laboratories (NWRL) ELISA method (n = 2617)

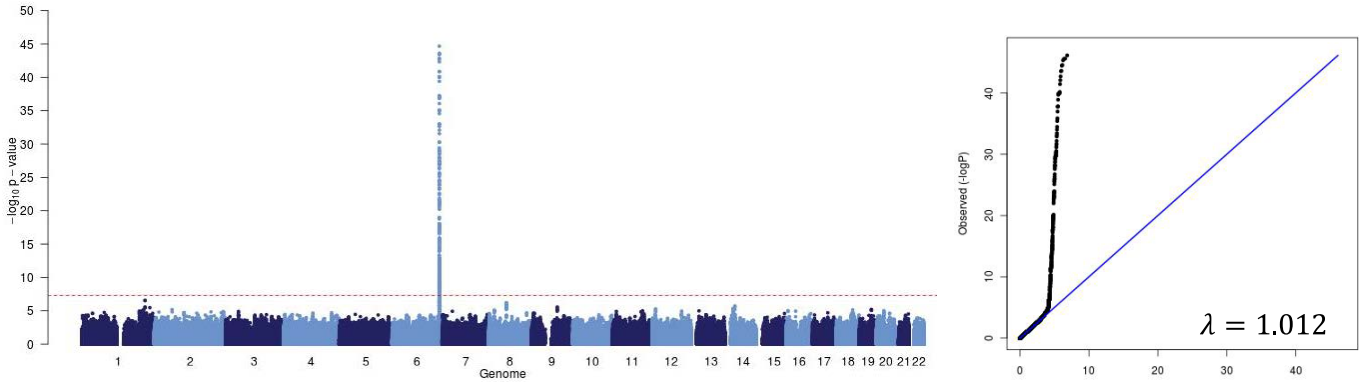


eFigure 6: Association of the genetic risk score, comprising of rs10455872 and rs3798220 variants, with *LPA* KIV2 repeats, Lp(a) levels, apo(a) isoform size and other traits

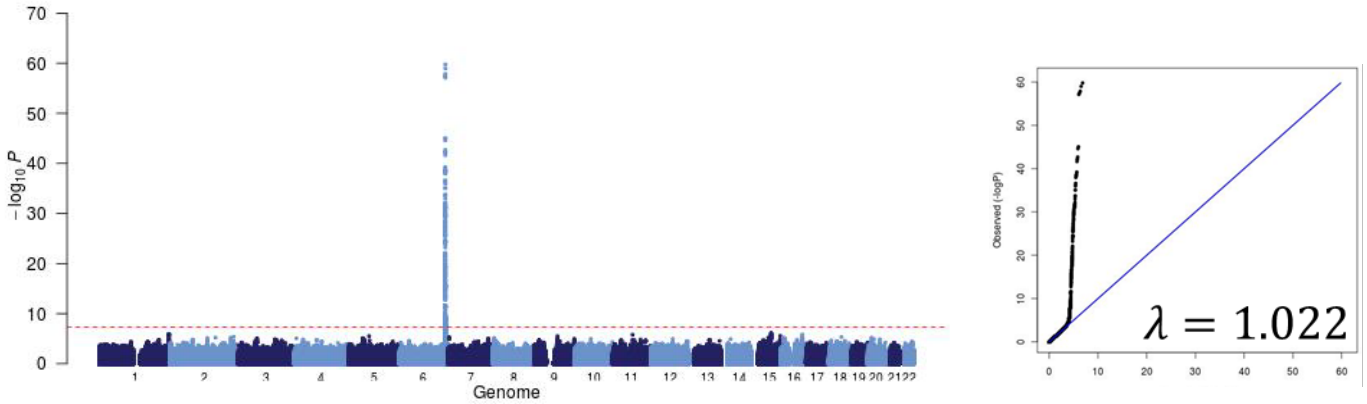


eFigure 7a-b. Genetic determinants of apo(a) isoform size and KIV2 repeats

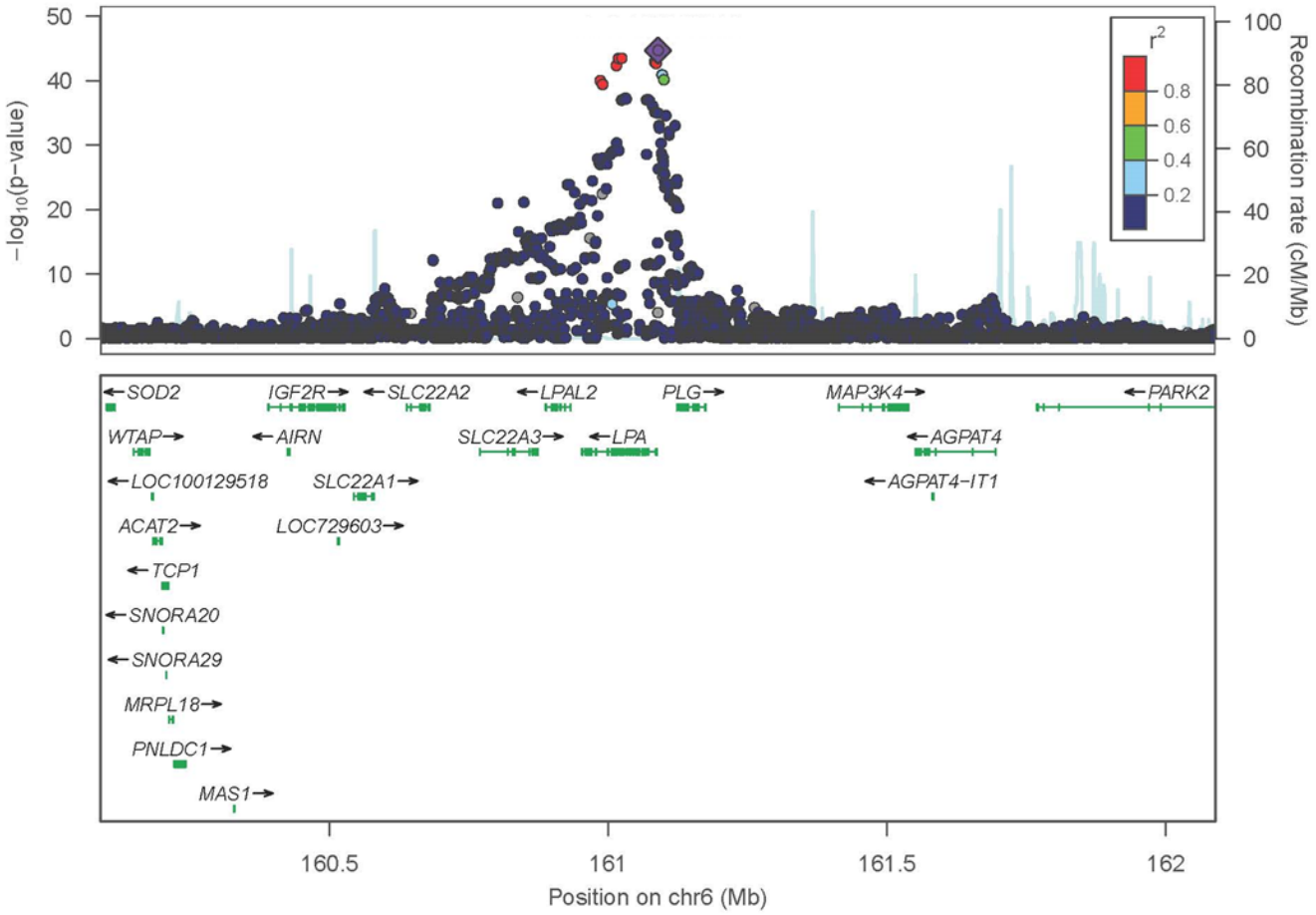
(a) Apo(a) isoform size



(b) *LPA* KIV2 repeats

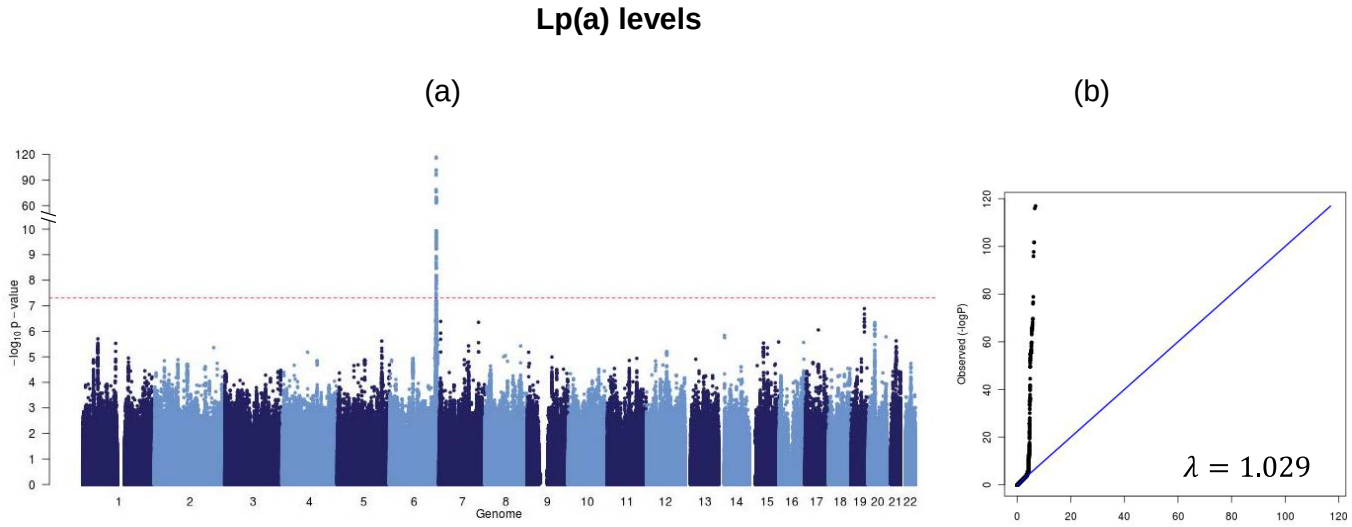


eFigure 8. Analyses of the *LPA* locus in association with apo(a) protein isoform size

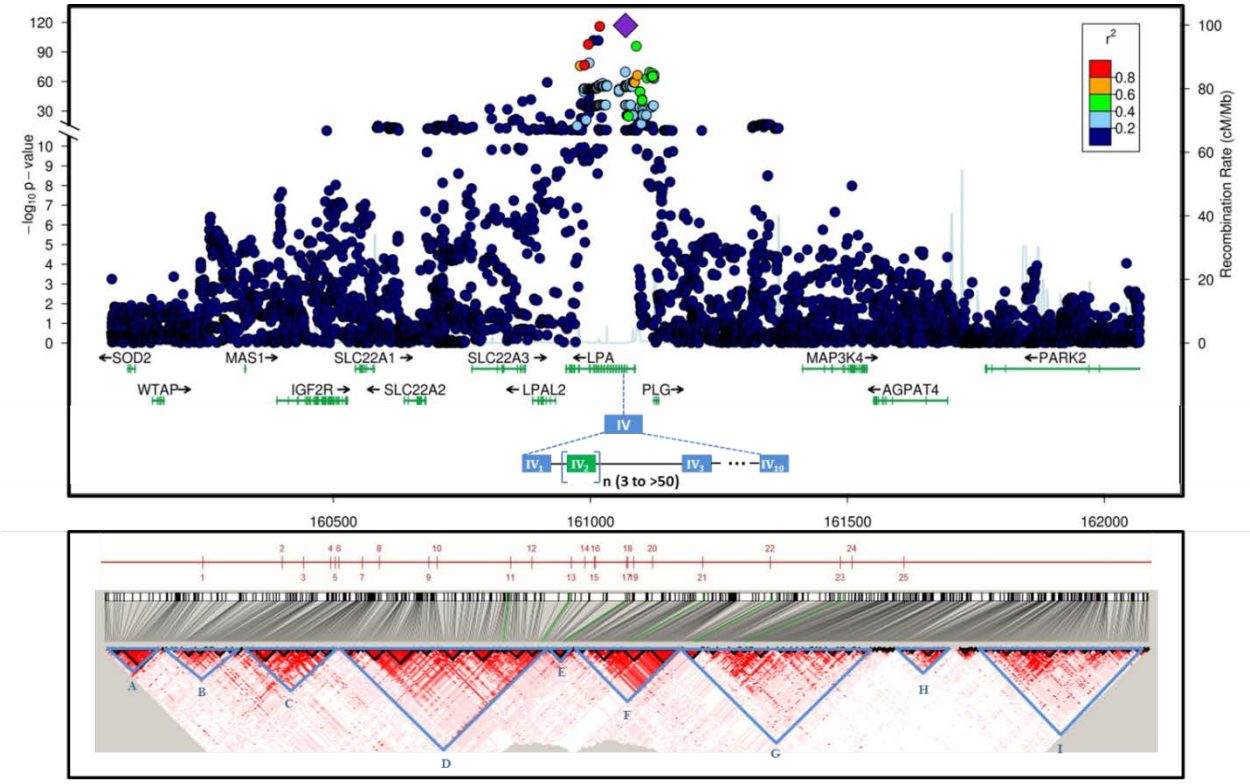


The following SNPs were independently associated with apo(a) protein isoform size on step-wise conditional analyses in PROMIS: rs12525588 rs7761293 rs7770628 rs1367211 rs13213129 rs2665356 rs3798950 rs142231215. These SNPs were further used to create a genetic risk score to assess association with CHD risk.

eFigure 9. Genetic determinants of Lp(a) levels



eFigure 10. Analyses of the *LPA* locus in association with Lp(a) levels



The following SNPs were independently associated with lp(a) levels on step-wise conditional analyses in PROMIS: rs73028918, rs9456488, rs77153348, rs8191871, rs2297362, rs73023700, rs3737088, rs148868952, rs316024, rs17589858, rs4708867, rs58432601, rs6919346, rs62441903, rs41272114, rs7765803, rs4708876, rs41269876, rs1367210, rs783144, rs2064712, rs7756550, rs12174748, rs35533223, rs79067311. These SNPs were further used to create a genetic risk score to assess association with CHD risk.

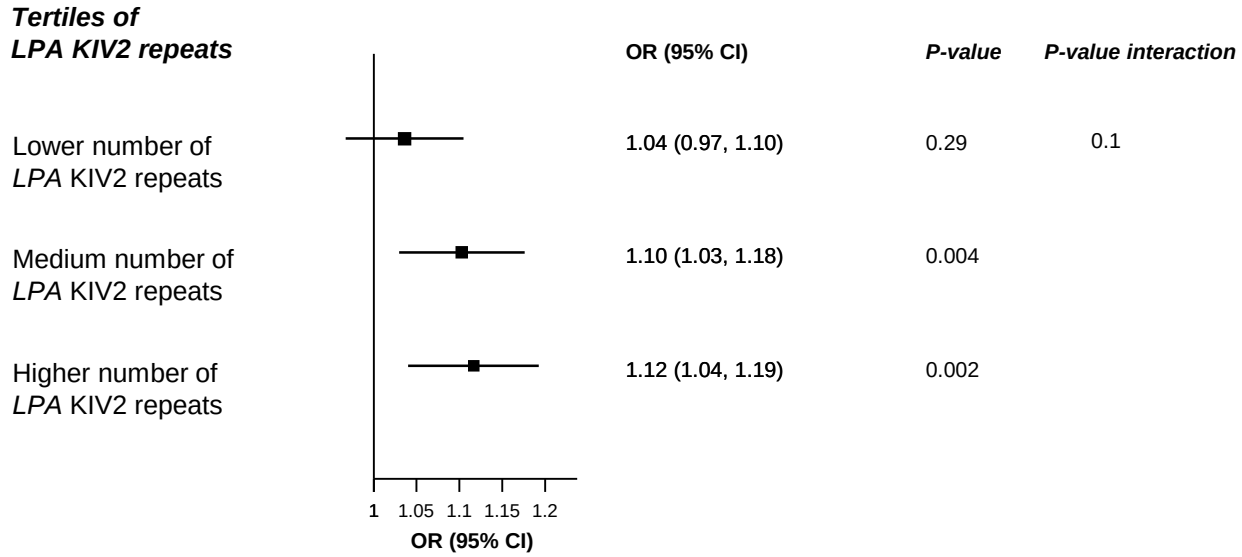
eTable 1. Analyses of rs3777392 and rs2457564 variants with various traits in international consortia

	P-values	
	rs3777392	**rs2457564**
ICBP (n = 134,433)		
Systolic blood pressure	0.78	0.74
Diastolic blood pressure	0.12	0.02
Pulse Pressure	0.46	0.09
Mean arterial pressure	0.80	0.41
GLGC (n = 134,433)		
Total Cholesterol	0.12	0.02
LDL-C	0.13	0.13
HDL-C	0.92	0.02
Log-triglycerides	0.25	0.92
MAGIC (n = 46,368)		
Post-load glucose	0.84	0.33
Fasting glucose	0.93	0.66
Hba1c	0.43	0.45
Fasting insulin	0.24	0.28
Fasting pro-insulin	0.7	0.73
HOMA-B	0.28	0.43
HOMA-IR	0.17	0.70
GIANT (n=183,727)		
Waist-to-hip ratio	0.14	0.03
Height	0.41	0.91
BMI	0.64	0.12

eTable 2. Odds ratios for myocardial infarction in PROMIS

	1-SD higher LPA KIV2 repeats (qPCR) *(6020 MI cases and 6686 controls)*		***1-SD higher log-Lp(a) levels*** *(5878 MI cases and 6886 controls)*	
Model	**OR (95% CI)**	***P-value***	**OR (95% CI)**	***P-value***
Age, gender, recruitment center and ethnicity	0.93 (0.90 – 0.97)	<0.0001	1.15 (1.11 - 1.20)	<0.0001
plus tobacco use, history of diabetes, and hypertension and waist-to-hip ratio	0.92 (0.89 – 0.96)	<0.0001	1.16 (1.12 - 1.21)	<0.0001
plus apo(a1) and apo(b)	0.92 (0.88 – 0.96)	<0.0001	1.11 (1.06 - 1.15)	<0.0001
plus log-Lp(a) levels / *LPA* KIV2 repeats (qPCR)	0.93 (0.90 -0.97)	0.001	1.10 (1.05 - 1.14)	<0.0001

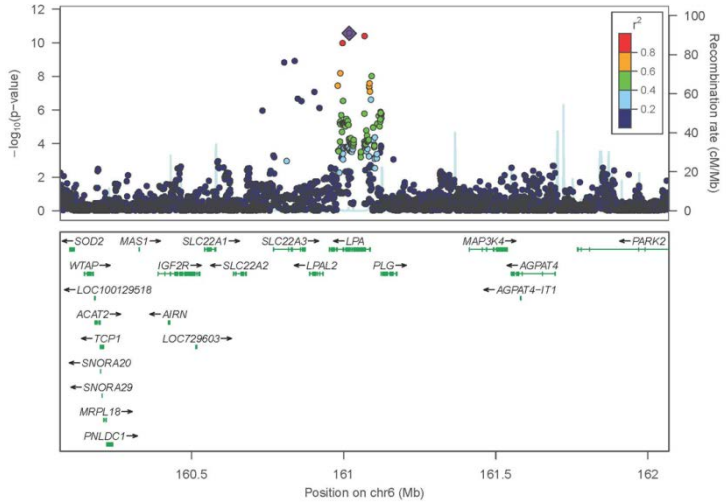
eFigure 11. Analyses of per-SD increase in Lp(a) levels with MI risk across tertiles of *LPA* KIV2 repeats



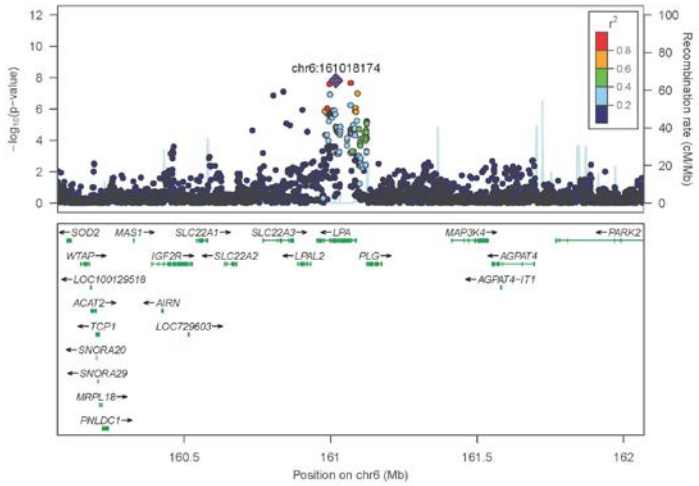
Interaction *P-value* was calculated by fitting an interaction term (x) in the logistic regression model for MI as an outcome while treating *LPA* KIV2 repeats on a continuous scale

eFigure 12. Association of the *LPA* locus with OxPL-apoB (a) unadjusted, (b) adjusted for *LPA* KIV2 repeats, (c) adjusted for Lp(a) levels and (d) adjusted for KIV2 repeats and Lp(a) levels

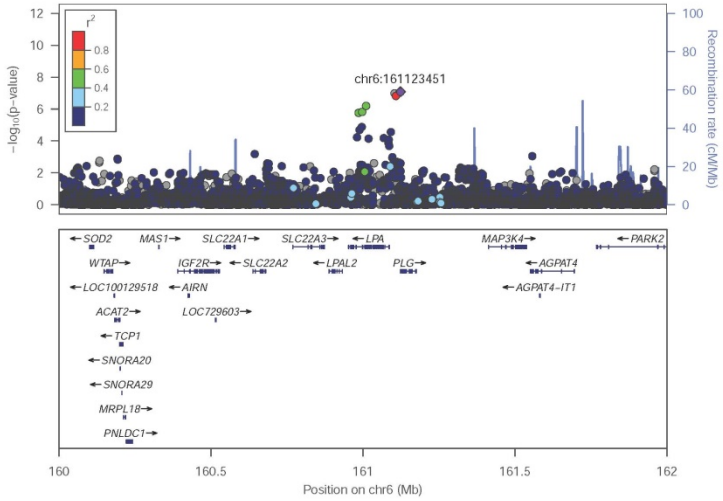
(a) Unadjusted



(b) Adjusted for *LPA* KIV2 repeats



(c) Adjusted for log-Lp(a) levels



(d) Adjusted for *LPA* KIV2 repeats and log-Lp(a) levels