

Clinical and Population Studies

OPEN

Genetic Factors Explain a Major Fraction of the 50% Lower Lipoprotein(a) Concentrations in Finns

Gertraud Erhart, Claudia Lamina, Terho Lehtimäki, Pedro Marques-Vidal, Mika Kähönen, Peter Vollenweider, Olli T. Raitakari, Gérard Waeber, Barbara Thorand, Konstantin Strauch, Christian Gieger, Thomas Meitinger, Annette Peters, Florian Kronenberg, Stefan Coassin

Objective—Lp(a) (lipoprotein(a)) concentrations are widely genetically determined by the *LPA* isoforms and show 5-fold interpopulation differences. Two- to 3-fold differences have been reported even within Europe. Finns represent a distinctive population isolate within Europe and have been repeatedly reported to present lower Lp(a) concentrations than Central Europeans. The significance of this finding was unclear for a long time because of the difficult comparability of Lp(a) assays. Recently, a large standardized study in >50 000 individuals from 7 European populations confirmed this observation but could not provide insights into the causes.

Approach and Results—We investigated Lp(a) concentrations, *LPA* isoforms, and genotypes of established genetic variants affecting Lp(a) concentrations (*LPA* variants, *APOE* isoforms, and *PCSK9* R46L) in the Finnish YFS (Cardiovascular Risk in Young Finns Study) population (n=2281) and 3 Non-Finnish Central European populations (n=10 003). We observed ≈50% lower Lp(a) concentrations in Finns. The isoform distribution was shifted toward longer isoforms, and the percentage of low-molecular-weight isoform carriers was reduced. Most interestingly, however, Lp(a) was reduced in each single-isoform group. In contrast to the known inverse relationship between *LPA* isoforms and Lp(a) concentrations, especially very short isoforms presented unexpectedly low Lp(a) concentrations in Finns. The investigated genetic variants, as well as age, sex, and renal function, explained 71.8% of the observed population differences.

Conclusions—The population differences in Lp(a) concentrations between Finnish and Central European populations originate not only from a different *LPA* isoform distribution but suggest the existence of novel functional variation in the small-isoform range.

Visual Overview—An online [visual overview](#) is available for this article. (*Arterioscler Thromb Vasc Biol.* 2018;38:1230-1241. DOI: 10.1161/ATVBAHA.118.310865.)

Key Words: atherosclerosis ■ genetics, population ■ lipoprotein(a) ■ risk factors

Lp(a) (lipoprotein(a)) is a major genetically determined cardiovascular risk factor. High Lp(a) plasma concentrations are associated with coronary heart disease, myocardial infarction, aortic valve calcification and stenosis, carotid atherosclerosis, stroke, and venous thromboembolism.^{1,2} Lp(a) concentrations are controlled mostly genetically by the *LPA* locus, which encodes apo(a) (apolipoprotein(a)) as the primary structural protein of Lp(a) and explains 70% to 90% of the Lp(a) concentrations.¹

The *LPA* gene consists of 10 different KIV (kringle IV) domains (KIV-1 to KIV-10), 1 KV domain (kringle V), and an

inactive protease domain.³ The KIV-2 domain is encoded by a 5.6-kb large, coding copy number variation, which is present in 1 to >40 repeats³ and thus generates >40 isoforms in the population.³ The KIV-2 number is inversely correlated to the Lp(a) plasma concentration and explains 30% to 70% of the concentrations.¹ Low-molecular-weight (LMW) isoforms (≤22 KIV repeats) are associated with ≈4- to 5-fold higher concentrations than high-molecular-weight (HMW) isoforms (>22 KIV).¹ The smaller allele thus determines the LMW/HMW group assignment.⁴ Although >95% of the population is heterozygous on DNA, only 30% to 70% present both

Received on: December 27, 2017; final version accepted on: February 26, 2018.

From the Division of Genetic Epidemiology, Department of Medical Genetics, Molecular and Clinical Pharmacology, Medical University of Innsbruck, Austria (G.E., C.L., F.K., S.C.); Department of Clinical Chemistry, Fimlab Laboratories (T.L.), Finnish Cardiovascular Research Center (T.L., M.K.), and Department of Clinical Physiology, Tampere University Hospital (M.K.), University of Tampere, Finland; Department of Medicine, Internal Medicine, Lausanne University Hospital, Switzerland (P.M.-V., P.V., G.W.); Department of Clinical Physiology and Nuclear Medicine, Turku University Hospital, Finland (O.T.R.); Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku, Finland (O.T.R.); Institute of Epidemiology II (B.T., C.G., A.P.), Institute of Genetic Epidemiology (K.S., C.G.), and Research Unit of Molecular Epidemiology (C.G.), Helmholtz Zentrum München—German Research Center for Environmental Health, Neuherberg, Germany; German Center for Diabetes Research, Neuherberg, Germany (B.T., A.P.); Institute of Medical Informatics, Biometry, and Epidemiology, Ludwig-Maximilians-Universität, Munich, Germany (K.S.); Institute of Human Genetics, Technische Universität München, Germany (T.M.); Institute of Human Genetics, Helmholtz Zentrum München, Neuherberg, Germany (T.M.); Munich Cluster for Systems Neurology, Germany (T.M.); and German Centre for Cardiovascular Research, Partner Site Munich Heart Alliance (A.P.).

The online-only Data Supplement is available with this article at <http://atvb.ahajournals.org/lookup/suppl/doi:10.1161/ATVBAHA.118.310865/-/DC1>.

Correspondence to Stefan Coassin, PhD, Division of Genetic Epidemiology, Department of Medical Genetics, Molecular and Clinical Pharmacology, Medical University of Innsbruck, Schöpfstraße 41, A-6020 Innsbruck, Austria. E-mail stefan.coassin@i-med.ac.at

© 2018 The Authors. *Arteriosclerosis, Thrombosis, and Vascular Biology* is published on behalf of the American Heart Association, Inc., by Wolters Kluwer Health, Inc. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution, and reproduction in any medium, provided that the original work is properly cited.

Arterioscler Thromb Vasc Biol is available at <http://atvb.ahajournals.org>

DOI: 10.1161/ATVBAHA.118.310865

Nonstandard Abbreviations and Acronyms

apo(a)	apolipoprotein(a)
eGFR	estimated glomerular filtration rate
HMW	high molecular weight
KIV	kringle IV
KORA	Cooperative Health Research in the Region of Augsburg, Kooperative Gesundheitsforschung in der Region Augsburg
KV	kringle V
LMW	low-molecular weight
LOF	loss of function
Lp(a)	lipoprotein(a)
LPA	apolipoprotein(a) gene
MAF	minor allele frequency
NFE	Non-Finnish Central European
SISu	Sequencing Initiative in Suomi
YFS	Cardiovascular Risk in Young Finns Study

isoforms also in plasma.^{3,5} Null alleles can be caused by large alleles, which tend to be nonexpressed or by 3 known loss-of-function (LOF) mutations.^{6–8}

To date, most studies did not identify Lp(a)-modifying genes outside the *LPA* locus. Only recently, 2 additional genes were implicated into Lp(a) metabolism: *APOE*^{9–12} and *PCSK9*.^{13,14} The isoform APOE2 is associated with 15% lower Lp(a) concentrations,^{9,10} whereas the Lp(a)-increasing effect of APOE4⁹ is controversial^{9,15} and vanishes after apo(a) isoform adjustment.¹⁰ In line with the effects of *PCSK9* inhibitors,¹⁶ the *PCSK9* LOF mutation R46L (rs11591147) with a minor allele frequency (MAF) of 3% was associated with ≈10% lower Lp(a) concentrations in a Danish population.¹³

Lp(a) concentrations show pronounced interethnic differences. In whites, Malay, Chinese, and Inuit, the distribution of Lp(a) concentration is extremely right skewed¹⁷ (median Lp(a), ≈10–12 mg/dL), whereas Africans present a rather Gaussian-like distribution and 2- to 3-fold higher Lp(a) levels than whites.^{3,5,17} Indians present intermediary concentrations.¹⁷ However, lacking standardization of Lp(a) quantification¹⁸ and the difficult conversion between mass-based (mg/dL) and molar-based (nmol/L) quantification¹⁸ hampers precise interpopulation comparisons.

The distribution and impact of the apo(a) isoforms differ between ethnicities.^{3,5} The isoform explains 28% of the Lp(a) concentration in Sudanese but 70% in Chinese.¹⁹ Chinese have a low number of null alleles, whereas Germans, Ghanaians, and San have more 1-isoform carriers, which tend to show lower Lp(a) concentrations than 2-isoform carriers.²⁰ In whites, the nonexpressed isoform is mostly the larger allele, whereas in blacks, the distribution of nonexpressed alleles is more U shaped, with an increased proportion of nonexpressed isoforms occurring both among very short and very long isoforms.⁵ Accordingly, differences in the frequency of some Lp(a) concentration-modifying sequence variations (SNPs [single nucleotide polymorphisms] and STRs [short tandem repeats]) have been shown between populations.^{3,5}

Within Europeans, Finns represent a genetically distinct population isolate.^{21,22} Lp(a) concentrations have been reported to be lower in Finns than in Central European populations,^{7,23–27} but the significance of this finding was unclear

because of the lacking standardization in Lp(a) quantification and small sample numbers in early studies. Only recently, this was confirmed by Waldeyer et al²⁸ in a standardized study of 7 European cohorts (n=56 804). The authors confirmed a clear reduction of Lp(a) in Finns and assumed a potential difference in isoform distributions as possible cause.

Indeed, several putative Lp(a)-regulating SNPs present marked differences in MAF in Finns compared with Non-Finnish Central Europeans (NFE). The donor splice-site mutation rs41272114,⁷ which is associated with an Lp(a) reduction by 5.25 mg/dL,⁷ is more frequent in Finns (MAF=0.0635) than in Tyroleans (MAF=0.053)⁷ or in British (PROCARDIS study [Precocious Coronary Artery Disease Study], MAF=0.03²⁹; UK Biobank, MAF=0.02).³⁰ Also, a second rare splice-site variant (rs143431368)⁶ shows frequency differences between Finns and NFEs (MAF, ≈0.03 in Finns versus 0.003 in NFE) as reported by the Exome Aggregation Consortium (ExAC)³¹ and UK Biobank.³⁰ Finally, also differences in the allele frequencies for *PCSK9* R46L and *APOE2* have been reported in Finns. Both are associated with reduced Lp(a),^{9,10,13} but, although *APOE2* frequency is reduced in Finns by ≈50³² and does, therefore, not contribute to the lower Lp(a) concentrations, the *PCSK9* R46L LOF variant is reported to be markedly more frequent (MAF=0.179 versus 0.035 in the ExAC data³¹).

To investigate the factors contributing to the population differences in Lp(a) concentrations, we report here a comprehensive investigation in a Finnish (n=2281) and 3 large Central European studies (n_{total}=10003). By combining SNP and isoform data for all samples, we assess (1) the distribution of isoforms in Finns compared with NFE and (2) dissect the relative contribution of isoforms and established Lp(a)-regulating SNPs to the low Lp(a) levels observed in Finns. Most importantly, all measurements of Lp(a) concentrations and apo(a) isoforms have been done centrally in 1 laboratory to keep methodological influences low.

Materials and Methods

Populations

Details on the study populations are reported in Table 1. The CoLaus³³ study (Cohorte Lausannoise) is a single-center, prospective study, including 6182 randomly selected white subjects aged 35 to 75 years from the city of Lausanne in Switzerland. Only individuals with 4 grandparents of European origin were included in the genetic study. The major aims of the CoLaus study are the investigation of prevalence and determinants of cardiovascular disease and cardiovascular risk factors in the Lausanne population. The analysis data set comprised of 3998 participants with all genotypes and Lp(a) values available.

The KORA cohorts (Cooperative Health Research in the Region of Augsburg, Kooperative Gesundheitsforschung in der Region Augsburg) are several population-based cohorts representative of the general population in Augsburg, Germany, and 2 surrounding counties. Ten-year age–sex strata have been sampled from the 25- to 74-year-old population with a stratum size of 640 subjects. The KORA F3³⁴ study was conducted in the years 2004/2005 as a follow-up study of the KORA S3 survey (1994/1995). The KORA F4³⁴ study was conducted in the years 2006/2008 and is a follow-up study of the KORA S4 survey (1999/2001). Both studies contain an independent nonoverlapping sample drawn from the same study region. The analyses data set comprised of 3089 participants in KORA F3 and 2916 participants in KORA F4, from which genotypes and Lp(a) values were available.

The YFS (Cardiovascular Risk in Young Finns Study)³⁵ is a prospective multicenter study from Finland initiated in 1980 (baseline

Table 1. Descriptive Statistics Continuous Variables Shown as Mean±SD and Percentiles (25%/50%/75%)

	YFS (n=2281)*	NFE (n=10 003)*	CoLaus (n=3998)*	KORA F3 (n=3089)*	KORA F4 (n=2916)*
Age, y	31.70±4.99 (27/33/36)	55.60±12.25 (46/55/65)	58.54±10.52 (49/58/66)	57.32±12.90 (46/57/67)	56.16±13.30 (44/56/67)
Women, n (%)	1255 (55.0%)	5230 (52.3%)	2141 (53.6%)	1584 (51.3%)	1505 (51.6%)
Lp(a), mg/dL; all subjects	12.45±5.29 (2.78/6.32/15.19)	22.47±26.09 (5.07/11.59/30.43)	23.37±26.82 (5.10/11.94/31.51)	21.93±26.00 (4.89/11.07/28.40)	21.80±24.63 (5.23/11.73/30.60)
Lp(a), mg/dL; LMW isoforms	34.23±22.31 (15.62/33.54/46.56)	52.64±33.24 (28.28/50.75/71.93)	51.59±33.27 (37.67/56.54/78.59)	50.68±34.22 (18.34/49.39/72.48)	47.57±31.12 (22.27/46.59/66.08)
Lp(a), mg/dL; HMW isoforms	8.36±8.86 (2.36/5.28/11.18)	13.60±14.50 (4.02/8.91/17.30)	13.92±14.71 (4.00/9.20/18.16)	13.39±14.56 (4.00/8.64/16.55)	13.36±14.14 (4.00/8.83/17.25)
LMW isoform carriers, n (%)	361 (15.8%)	2272 (22.7%)	846 (21.2%)	707 (22.9%)	719 (24.7%)
Isoform 1 (n KIV)†	30.55±7.04 (24/29/34)	27.62±7.04 (23/27/32)	28.25±7.07 (23/28/33)	27.37±6.40 (23/27/32)	27.03±6.52 (23/26/31)
LDL cholesterol, mg/dL	126.8±32.6 (104.4/123.7/147.0)	132.6±34.9 (108.3/131.0/154.7)	133.6±36.3 (108.3/131.5/158.6)	128.0±32.6 (105.0/126.0/148.0)	136.1±34.9 (111.0/134.0/158.0)
LDL-C corrected, mg/dL‡	123.0±32.3 (100.11/119.7/142.3)	125.8±34.6 (101.7/124.1/147.7)	126.5±35.9 (101.5/125.3/149.8)	121.4±32.4 (98.7/119.9/141.6)	129.5±34.5 (104.9/127.6/151.7)
HDL cholesterol, mg/dL	49.9±12.3 (41.0/49.1/57.6)	59.9±17.1 (47.0/58.0/69.6)	63.7±18.1 (50.3/61.9/73.5)	58.8±17.1 (46.0/56.0/69.0)	55.9±14.4 (45.0/54.0/65.0)
Total cholesterol, mg/dL	199.7±38.0 (174.0/197.2/224.3)	218.6±40.1 (192.0/216.6/243.6)	220.9±40.3 (193.4/220.4/247.5)	218.2±40.0 (191.0/216.0/243.0)	216.0±40.0 (188.0/214.0/240.0)
Triglycerides, mg/dL§	119±76 (71/97/142)	135±100 (77/109/162)	120±78 (71/97/142)	165±126 (88/135/201)	125±89 (72/105/151)
Type-2 diabetes mellitus, n (%)	78/2266 (3.4%)	912/9994 (9.1%)	436/3996 (10.9%)	240/3082 (7.8%)	236/2916 (8.1%)
eGFR, mL/min per 1.73 m ²	109.7±11.3 (102.6/111.8/117.6)	85.4±16.6 (74.7/86.9/97.2)	81.7±14.9 (71.9/82.2/92.3)	87.9±17.6 (78.2/90.7/100.3)	87.7±16.9 (77.3/89.1/100.0)
SNP MAF/APOE haplotype frequency					
rs143431368 (LPA)	0.027	0.002	0.002	0.002	0.002
rs41272114 (LPA)	0.049	0.031	0.034	0.029	0.028
rs3798220 (LPA)	0.012	0.015	0.015	0.016	0.014
rs10455872 (LPA)	0.029	0.056	0.056	0.053	0.060
rs11591147 (PCSK9)	0.039	0.013	0.010	0.015	0.016
rs7412 (APOE)	0.043	0.068	0.052	0.084	0.073
rs429358 (APOE)	0.195	0.131	0.128	0.136	0.132
APOE2/E2	0.001	0.003	0.001	0.006	0.004
APOE2/E3	0.062	0.111	0.089	0.130	0.121
APOE2/E4	0.020	0.018	0.013	0.026	0.016
APOE3/E3	0.580	0.640	0.669	0.611	0.632
APOE3/E4	0.302	0.210	0.213	0.209	0.207
APOE4/E4	0.034	0.017	0.015	0.018	0.020

eGFR indicates estimated glomerular filtration rate; HDL, high-density lipoprotein; HMW, high molecular weight; KIV, kringle IV; KORA, Cooperative Health Research in the Region of Augsburg, Kooperative Gesundheitsforschung in der Region Augsburg; LDL, low-density lipoproteins; LDL-C, low-density lipoprotein cholesterol; LMW, low-molecular weight; Lp(a), lipoprotein(a); MAF, minor allele frequency; NFE, Non-Finnish Central European; SNP, single nucleotide polymorphisms; and YFS, Cardiovascular Risk in Young Finns Study.

*Based on samples with available Lp(a) concentration.

†Shorter isoform present.

‡LDL cholesterol levels corrected for Lp(a) concentrations.⁵⁵

§KORA F3 participants were nonfasting.

age, 3–18 years). Children (4320) aged 3, 6, 9, 12, 15, and 18 years were randomly chosen from the population register from the 5 Finnish university cities with medical schools (Helsinki, Kuopio, Oulu, Tampere, and Turku). Several follow-ups during 30 years to

investigate childhood risk factors for cardiometabolic outcomes have been performed. Lp(a) phenotype information was available for 2281 participants. The measurements included in this work have been done on materials from the follow-up in the year 2001, where 2283

individuals (63.5% of the original cohort) participated in clinical examinations, and 2620 (72.9%) returned questionnaires.

Informed consent was obtained from each participant, and the studies were approved by the respective institutional review boards.

Lp(a) Quantification Using ELISA

Lp(a) plasma concentrations were measured using a sandwich ELISA method described by Kronenberg et al³⁶ with minor modifications. All measurements were done in the same laboratory (Division of Genetic Epidemiology, Medical University of Innsbruck, Austria).

In brief, the ELISA plates (Nunc-Immuno MicroWell Maxi Sorp flat bottom design, MaxiSorp surface treatment; Thermo Fisher Scientific, Waltham) were coated using an affinity-purified polyclonal rabbit anti-human apo(a) antibody in a final concentration of 5 µg/mL in 1× PBS containing 1 mg/mL Na₂CO₃. The plates were incubated with 100 µL antibody dilution (3 hours, 37°C), washed 3× (1× PBS+0.05% v/v Tween-20), and blocked with 200-µL 0.1% wt/vol casein in 1× PBS pH 7.3 (30 minutes, 37°C). To ensure measuring each sample within the linear range of optical density, all samples were diluted into the ELISA plate twice (1:150 and 1:1500 in Assay Buffer [Microcoat, Bernried, DE] of 1:30 and 1:1000 predilutions in 1× PBS, pH 7.3). A 7-point standard curve ranging from 0.32 mg/dL to 5 µg/dL was created (with an additional blank representing the zero point). Duplicate determinations of 4 reference samples were used as longitudinal interassay controls.

The coated plates were incubated with the analyte for 1 hour at 37°C. Detection was performed using a horseradish peroxidase-conjugated monoclonal antibody (1A2; in 0.1% wt/vol casein, 1× PBS, pH 7.3) directed against the KIV domain³⁷ and not cross-reacting with plasminogen (1 hour, 37°C). After 3 washing steps as described above, 100 µL Blue Star TMB substrate (Adaltis, Guidonia Montecelio, IT) was added (30 minutes, room temperature). Reaction was stopped by adding 50 µL 0.5 mol/L sulfuric acid. Measurement of the absorption (dual wavelength, analyte: 450 nm, reference: 690 nm) was done using a Microplate Reader (BioRad Benchmark Plus; Bio-Rad Laboratories, Hercules), and concentrations were calculated based on the standard curve (expressed as mg/dL). All dilution and pipetting steps were done using liquid handling robotics (Tecan, Männedorf, CH).

apo(a) Isoform Determination by Western Blot

apo(a) isoform determination has been done as previously described in Kronenberg et al³⁸ with minor modifications. All measurements were done in the same laboratory (Division of Genetic Epidemiology, Medical University of Innsbruck, Austria) and evaluated by the same experienced scientist.

In brief, Lp(a) concentrations of each sample were determined by ELISA as described above, and all plasma samples were then diluted with PBS to a standardized Lp(a) concentration of 30 ng/µL. Ten microliters of this dilution were mixed with 20 µL of reducing sample buffer (15.38% v/v glycerol, 7.69% 2-mercaptoethanol, 3.85% v/v SDS, 0.177% v/v 4-ethylmorpholine, and 0.077% wt/vol bromophenol blue) resulting in an Lp(a) concentration of 10 ng/µL. Protein was denatured at 98°C for 5 minutes. Subsequently, 150 ng (15 µL of the 10-ng/µL solution) were applied on a 1.46% agarose gel supplied with 0.08% SDS and separated for 18 hours (0.04 A, constant current). Samples with an Lp(a) concentration <15 ng/µL (ie, 1.5 mg/dL) were not diluted with PBS, but a higher volume of sample was applied on the gel to obtain an equal apo(a) mass. A size standard consisting of a mixture of 5 plasma samples expressing only 1 apo(a) isoform each (13, 19, 23, 27, and 35 KIV repeats, determined by pulsed-field gel electrophoresis³⁹) was applied every seventh sample on the gel to detect uneven gel running and minimize distance between samples and standards. After electrophoresis, the proteins were transferred to a polyvinylidene difluoride membrane (Immobilion-P; Millipore, Darmstadt, DE) using semidry blotting (Perfect Blue Semi Dry Blotter; VWR, Vienna, Austria). The membrane was first equilibrated in methanol for 30 seconds, and then rinsed in water and afterward equilibrated in blotting buffer (20% v/v ethanol, 16 mmol/L Tris, and 120 mmol/L glycine). The transfer was done for

45 minutes and 250 mA having the blotting stack semidry in Blotting Buffer. The membrane was then blocked for at least 30 minutes at 37°C using buffer C (85 mmol/L NaCl, 10 mmol/L Tris, 0.2% Triton X-100, and 1% BSA) and incubated in the first antibody (1A2³⁷ in a concentration of 168 ng/mL in buffer C) for 2 hours at room temperature on a shaker. After 3 wash steps in TTBS (Tris-buffered saline buffer with Tween-20; 20 mmol/L Tris-HCl pH 7.4, 0.5 mol/L NaCl, and 0.05% Tween-20) of 15 minutes each, the membrane was incubated with horseradish peroxidase-conjugated goat anti-mouse antibody (401253, Millipore, Darmstadt, DE; diluted 1:13333 in buffer C) for 45 minutes on a shaker at room temperature. Residual secondary antibodies were removed by 3 washing steps (15 minutes in TTBS each); enhanced chemiluminescence substrate (WesternBright Chemiluminescence Spray; Biozym, Hessisch Oldendorf, DE) was added, and the signals were detected on an Amersham Hyperfilm for enhanced chemiluminescence (GE Healthcare, Vienna, Austria).

Genotyping

1000 Genomes⁴⁰-imputed genome-wide genotyping data were available for CoLauS and YFS, whereas genotypes imputed using the haplotype reference consortium⁴¹ panel were available for KORA F3 and KORA F4. These data were used for evaluation of LPA SNPs rs3798220, rs10455872, rs143431368, and rs41272114, as well as PCSK9 rs11591147 (PCSK9 R46L) and the APOE isoforms E2, E3, and E4 (rs7412 and rs429358). APOE haplotypes are defined as follows: APOE2/E2: rs7412=TT and rs429358=TT, APOE2/E3: rs7412=CT and rs429358=TT; APOE2/E4: rs7412=CT and rs429358=CT, APOE3/E3: rs7412=CC and rs429358=TT; APOE3/E4: rs7412=CC and rs429358=CT, APOE4/E4: rs7412=CC and rs429358=CC. Imputation quality is given in Table VII in the [online-only Data Supplement](#) and in Mack et al 2017.¹⁰

Statistical Methods

In samples heterozygous for apo(a) isoforms, the smaller of the 2 isoforms present is commonly used for assignment to the LMW (≤22 KIV) or the HMW (>22 KIV) groups.⁴ Because genotyping using pulsed-field gel electrophoresis is not feasible on population scale, we refer to the Western blot result throughout the article, when assigning zygosity. For statistical analysis and data presentation of heterozygous samples, the smaller isoform present was termed isoform 1, independent of its size, and the larger isoform was named isoform 2. Accordingly, all isoform-based adjustments and stratifications were done based on the smaller isoform (unless otherwise stated).

Because of the skewed distribution of the Lp(a) trait, nonparametric tests were used to compare median values of Lp(a) between the studies (Mann-Whitney-Wilcoxon test). Dunn post hoc test for multiple comparisons was used to test for significant median differences between all possible pairwise tests between studies. Proportions were tested by χ^2 test. In regression and mediation models, inverse-normal transformation was used for Lp(a) to adhere to the normal distribution assumption for statistical tests. Equal distribution of isoforms within carriers of rs10455872 and rs3798220 in YFS and NFE was tested using the Kolmogorov-Smirnov test.

We used a multivariable regression model to evaluate the joint influence of the SNPs rs10455872, rs3798220, rs14341368, rs41272114, and rs11591147, APOE2 and APOE4-carriers, the isoforms (isoform 1), age, sex, and estimated glomerular filtration rate (eGFR) on Lp(a). Imputed genotype scores were used on a continuous scale, which corresponds to an additive coding. The eGFR was included to account for renal function, which is one of the few established nongenetic modifiers of Lp(a) concentrations.^{42,43} eGFR was calculated according to Levey et al,⁴⁴ (CDK EPI [Chronic Kidney Disease Epidemiology Collaboration] creatinine equation). For the isoforms, a nonlinear spline was used. The multivariable regression models were performed in each of the 4 populations individually and for all Non-Finnish European studies combined (NFE) using a mixed model accounting for the study as a random effect (package mgcv in R).

To investigate whether the lower Lp(a) values in Finns can be explained by the evaluated variables, we performed a formal

mediation analysis. This means that the effect on Lp(a) of being Finnish is decomposed into the effect portions, which are mediated by the isoforms, by the SNPs, and by other factors, which differ between YFS and the Non-Finnish European studies (see the Results section). Because the participants in the YFS are younger than the participants in the other Non-Finnish studies, age and, in consequence, better kidney function could also possibly explain the difference in Lp(a) distribution. Therefore, eGFR and age were also included as potential mediating variables. Technically, a binary variable (Finnish versus Non-Finnish) was created and regressed onto inverse-normal transformed Lp(a) values. This effect is denoted as the total effect. Then, this simple regression model was adjusted for the potential mediating variables (age, sex, eGFR, isoform 1, and the investigated SNPs). The percentage by which the total effect is attenuated is the relative mediated effect, which can be decomposed into their single parts (indirect effects, herein after referred to as explained effects). The remaining effect is the direct effect of the variable being Finnish. This is actually the remainder, which cannot be explained by the included variables (herein after referred to as unexplained effects). Assumptions of the mediation model are (1) association of the mediators with the outcome (this was set to a P value <0.1) and (2) association of the exposure (the Finnish versus Non-Finnish effect) with the outcome (P value, <0.1). Only those variables were included in the mediation analysis that fulfilled these assumptions. Model assumptions, decomposition into single indirect effects (time of resampling, $n=100$), and bootstrap confidence intervals (number of times of bootstrap resampling, $n=500$) were performed using the package *mma* in R. All statistical analyses were done with R (<http://www.r-project.org>).

Results

YFS Present Reduced Lp(a) Concentrations Compared With Central Europeans

We determined Lp(a) concentrations and apo(a) isoforms in the YFS ($n=2281$) and in the 3 Central European population CoLaus ($n=3998$), KORA F3 ($n=3089$), and KORA F4 ($n=2916$). All measurements were performed in the same laboratory using standardized ELISA and Western blot assays. Median Lp(a) concentrations were similar in all Central European populations ($P>0.59$ for all pairwise comparisons between Central European populations) but significantly different between YFS and all Central European studies (Figure 1A; Table I in the [online-only Data Supplement](#); all $P \leq 1.38 \times 10^{-58}$ for all pairwise comparisons using Dunn post hoc test). Therefore, for all subsequent analyses, CoLaus, KORA F3, and KORA F4 were combined to the Non-Finnish Europeans (NFE) population. The median Lp(a) concentration in YFS was significantly lower than in NFE (6.3 versus 11.6 mg/dL; $P=8.15 \times 10^{-91}$; Figure 1A). This difference was present both in men and women and when restricting the analysis only to the age range ≥ 32 to ≤ 39 years allowing a similar age distribution of all studies (Table I in the [online-only Data Supplement](#)). When stratifying Lp(a) concentrations into 5-mg/dL strata, concentrations of 0 to 5 mg/dL were overrepresented in YFS compared with NFE (42.3% versus 24.7%; $P=5.22 \times 10^{-64}$; Figure 1B). Accordingly, less individuals showed concentrations >30 mg/dL and >50 mg/dL, which are commonly deemed as thresholds for increased cardiovascular risk^{45,46} (>30 mg/dL, 12.1% versus 25.3%; >50 mg/dL, 3.9% versus 14.1%; $P=2.32 \times 10^{-41}$ and $P=1.17 \times 10^{-41}$; Figure III in the [online-only Data Supplement](#)).

Differences in Isoform Distributions Between YFS and NFE

The Lp(a) concentration is highly dependent on the number of KIV repeats. We, therefore, investigated whether the apo(a) isoform distribution is shifted toward larger isoforms in YFS compared with NFE, which could potentially explain the lower Lp(a) levels observed in YFS.

Isoforms ≤ 25 KIV repeats were more frequent in NFE than in YFS, whereas YFS presented higher frequencies of isoforms with >28 KIV domains (Figure 2A). Accordingly, LMW frequency was significantly reduced in YFS compared with the NFE populations (361/2282 [15.8%] versus 2272/10003 [22.7%]; $P=5.83 \times 10^{-13}$; Figure 2B). The median size of the smaller isoform (isoform 1) was larger in YFS than in NFE (29 versus 27 KIV repeats; $P=8.21 \times 10^{-37}$; Figure 2C; Table II in the [online-only Data Supplement](#)). In heterozygous individuals, the differences in the distribution of the larger isoform (isoform 2) were less pronounced but still significant (36 versus 35 KIV repeats; $P=3.54 \times 10^{-12}$; Figure 2D; Table II in the [online-only Data Supplement](#)).

Interestingly, the Lp(a) concentrations were lower in YFS than in NFE even within the LMW and HMW groups (Figure 3A). Because this could arise from a shift of the isoform distribution toward higher isoforms within the groups (thus reducing the group mean without changing group assignment), we stratified the Lp(a) isoforms into more granular subgroups as done in earlier studies.⁴⁷ YFS showed reduced Lp(a) concentrations in all subgroups (Figure 3B; Table IV in the [online-only Data Supplement](#)). Also, the percentage of homozygous null allele carriers (ie, samples that do not present any of the 2 *LPA* isoforms in plasma as detectable by Western blotting) was higher in YFS than in NFE ($n=36$ [1.6%] versus $n=70$ [0.7%]; $P=7.25 \times 10^{-5}$). However, the median Lp(a) concentration of the entire YFS cohort increased only marginally from 6.32 to 6.46 mg/dL after exclusion of the homozygous null allele carriers.

Frequency of Established Lp(a)-Regulating SNPs

The *APOE* isoforms and PCSK9 R46L have been recently established as novel genetic regulators of Lp(a).^{9,10,13} As described in detail in the Introduction section, both genetic variants show differences in allele frequencies between Finns and Central Europeans. The same applies to 2 well-known *LPA* splice-site variants^{6,7} (Table 1). We, therefore, evaluated to which extent these variants contribute to the differences in Lp(a) concentrations between YFS and NFE. YFS presented a higher frequency of E3/E4 and E4/E4 ($P=3.72 \times 10^{-20}$ and $P=1.04 \times 10^{-6}$, respectively) and lower frequencies of E2/E3 ($P=1.41 \times 10^{-11}$; Table 1). Conversely, rs11591147, rs143431368, and rs41272114 were more frequent in YFS than in NFE (Table 1; $P=4.79 \times 10^{-29}$, $P=6.77 \times 10^{-83}$, and $P=1.44 \times 10^{-8}$, respectively). They represent, therefore, valid putative determinants for the observed differences in Lp(a).

Moreover, the *LPA* variants rs3798220⁴⁸ and rs10455872⁴⁸, which have been reported to tag small isoforms (19–21 and 17–20 KIV repeats, respectively),⁴⁸ were included. Both showed the same isoform distribution in YFS as in NFE ($P=0.61$ and

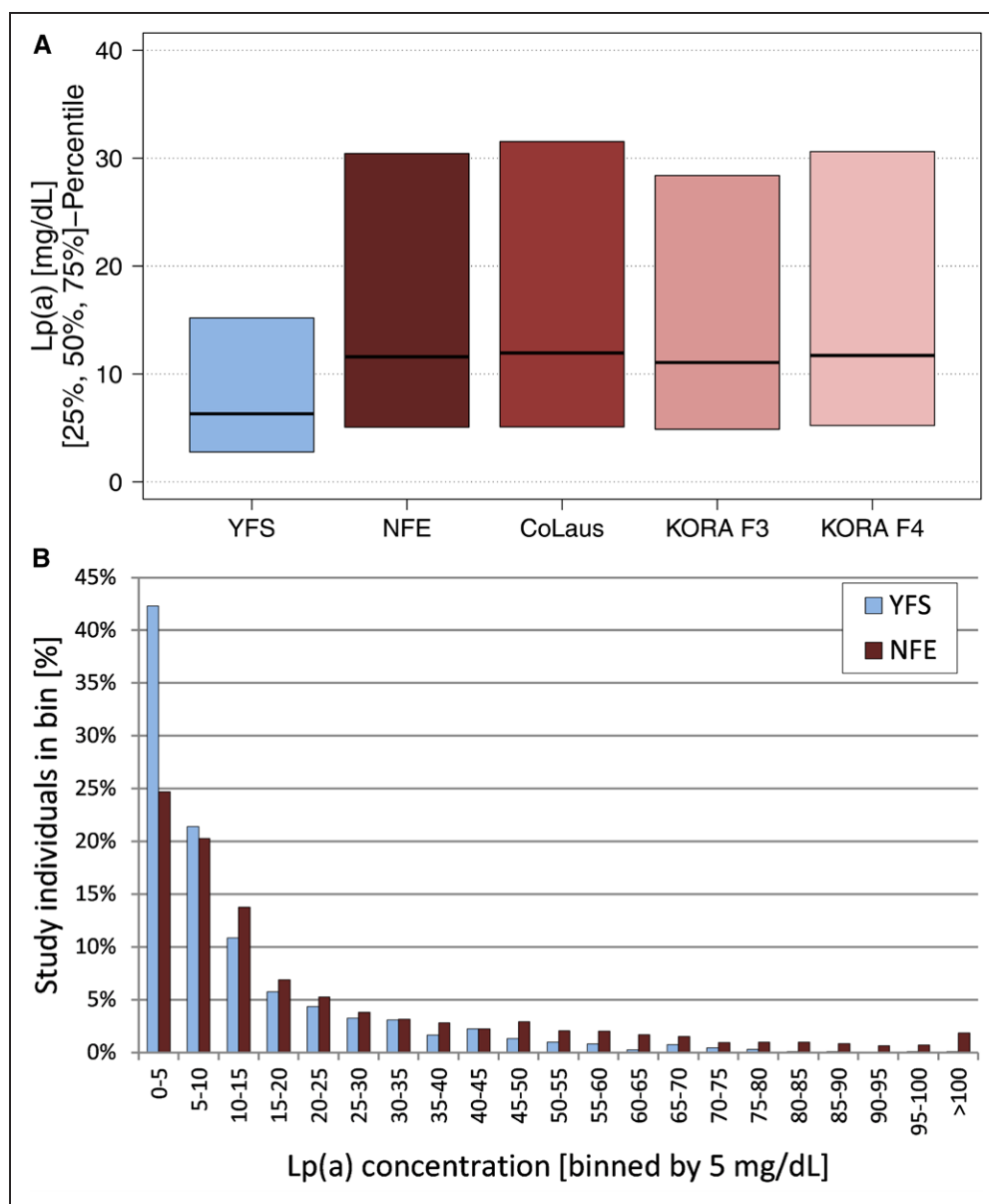


Figure 1. Distribution of Lp(a) (lipoprotein(a)) concentrations. **A**, Median Lp(a) concentrations for each population and the combined Non-Finnish Europeans (NFE) group. YFS (Cardiovascular Risk in Young Finns Study) shows $\approx 50\%$ lower Lp(a) concentrations. Because of the large variance of Lp(a) concentrations, only medians and 25%/75% percentiles are shown to enhance comprehensibility. Full figure is given in Figure I in the [online-only Data Supplement](#). **B**, Frequencies of Lp(a) concentrations for YFS and NFE stratified by 5 mg/dL strata. The 0 to 5 mg/dL group is markedly larger, and the relative frequencies in almost all strata >30 mg/dL are substantially smaller in YFS than in NFE. Figure II in the [online-only Data Supplement](#) shows nonclustered NFE. KORA, Cooperative Health Research in the Region of Augsburg, Kooperativ Gesundheitsforschung in der Region Augsburg.

$P=0.20$; Figure VII in the [online-only Data Supplement](#)). The rs10455872 frequency was $\approx 50\%$ lower in YFS (which is also in line with the lower frequency of short isoforms), whereas no pronounced frequency difference was observed for the rarer rs3798220 variant.

Association of Potential Mediating Factors With Lp(a) Concentrations

We investigated *LPA* isoform distribution, *LPA* and *PCSK9* LOF frequencies, and *APOE* isoform frequencies, as well as sex, age, and renal function, estimated by eGFR^{1,5} as potential mediators of the observed population differences.

Results from multiple regression models for each factor on Lp(a) are reported in Table 2 (YFS and NFE) and Table V in the [online-only Data Supplement](#) (single studies). Among the genetic factors, isoforms were highly associated with Lp(a) in a nonlinear fashion (Figure 4), whereas *APOE2* and the SNPs rs143411368 and rs41272114, which are more frequent in the YFS, showed an Lp(a)-decreasing effect. Conversely, rs10455872 and rs378220 were associated with increased Lp(a), but only rs10455872 shows a pronounced frequency difference between YFS and NFE and thus potentially contributes to the Lp(a) difference. No association was observed for *PCSK9* R46L (rs11591147) and *APOE4*

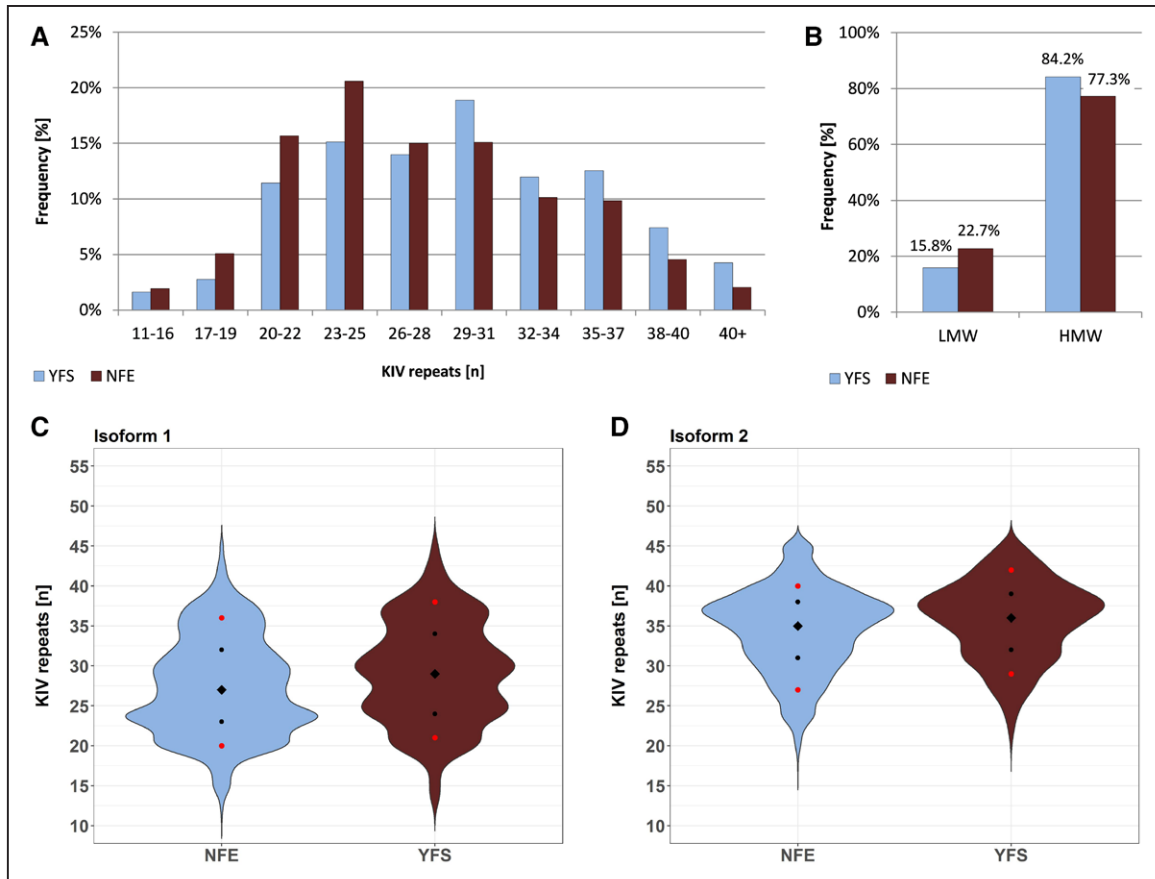


Figure 2. apo(a) (apolipoprotein(a)) isoform distribution in YFS (Cardiovascular Risk in Young Finns Study) and Non-Finnish Central Europeans (NFE). **A**, Isoform 1 in YFS and NFE stratified by groups of 3 KIV (kringle IV) repeats (except the group with shortest isoforms 11–16). YFS shows a higher frequency of carriers >28 KIV repeats, compared with NFE. The number of carriers in each isoform stratum is given in Table IV in the [online-only Data Supplement](#). **B**, Low-molecular-weight (LMW; at least 1 isoform with ≤ 22 KIV repeats) and high-molecular-weight (HMW) carriers (only isoforms with >22 repeats) in YFS and NFE. HMW individuals are more common in YFS. $P=5.83 \times 10^{-13}$. **C** and **D**, apo(a) isoform distribution of the 2 alleles. The violin plots show distributions of the smaller allele (isoform 1) and the larger allele (isoform 2, if detectable). The isoform distribution in YFS is shifted toward larger alleles. Diamonds, median; black dots, 25th and 75th percentiles; red dots, 10th and 90th percentiles. For details and unclustered NFE data, see Figure IV in the [online-only Data Supplement](#) and Table III in the [online-only Data Supplement](#).

carrier status, despite their frequency difference between YFS and NFE.

The multiple adjusted effect of isoform 1 on Lp(a) is shown in Figure 4. In general, Lp(a) decreases with increasing KIV repeats, although not linearly. Because in YFS the distribution of isoforms is shifted to higher repeat numbers, lower Lp(a) values are expected. Both in YFS and NFE, Lp(a) decreases rather sharply at >22 repeats. The nonlinear relationship is consistent in all 3 Non-Finnish studies and is also consistent for HMW isoforms in YFS. Interestingly, the effect of isoform 1 being <20 KIV repeats on Lp(a) concentrations differs between YFS and NFE. Although in NFE the Lp(a) concentrations still rise for isoforms <20, the Lp(a) values decrease in YFS.

Among the study-specific factors, age and female sex were associated with increased Lp(a) values, whereas eGFR, which is higher in the YFS, was associated with a decreased Lp(a) (although not significantly). Because YFS individuals are younger and present better renal function than those in the NFE studies, Lp(a) values are lower in YFS also in consequence of the age effect.

Mediation Analysis Explaining the Differences in Lp(a) Concentrations Between Finns and Non-Finns

The multiple regression model (Table 2) indicates that both genetic and environmental factors, which are associated with reduced Lp(a) and which are more frequent in YFS than in NFE, contribute to lowering Lp(a) in YFS. To identify the relative contribution of these factors, we conducted a formal mediation analysis. All variables of the multiple regression models were included as potential mediators of the “Finns effect” on Lp(a), except SNP rs11591147, which was not significant in the multiple regression model and thus did not fulfill the criteria to be included as a mediator (Table VI in the [online-only Data Supplement](#)), and rs3798220, which was not identified as an independent modifier (Table VII in the [online-only Data Supplement](#); column P value 2). Altogether, 71.8% of the Finns effect is determined by the included mediators, which leaves 28.2% unexplained (Figure 5). The largest contributors to reduced Lp(a) concentrations in YFS are the isoforms (27.3%), followed by age (26.4%), eGFR (10.4%), rs41272114 (4.0%), and rs143411368 (3.8%), whereas *APOE2* frequency contributed inversely (−1.8%).

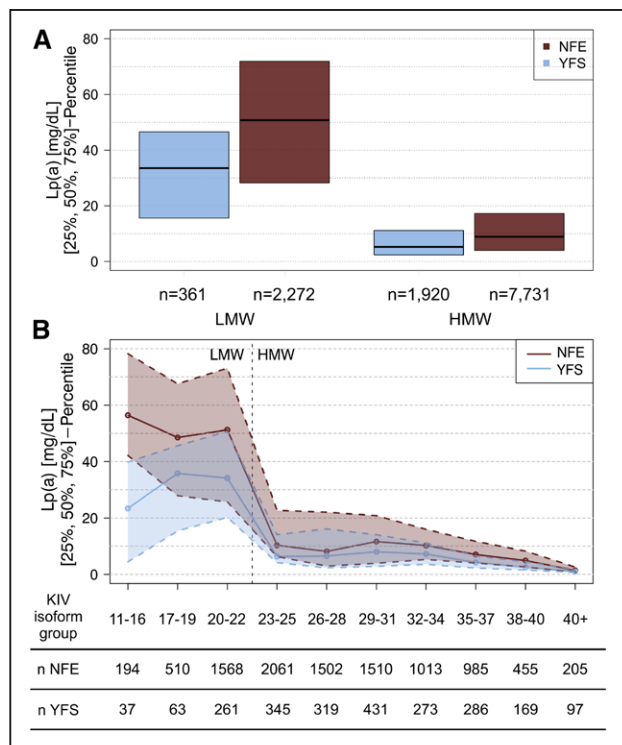


Figure 3. Lp(a) (lipoprotein(a)) concentrations stratified by apo(a) (apolipoprotein(a)) isoform size. **A**, Median Lp(a) concentration stratified by low-molecular-weight (LMW)/high-molecular-weight (HMW) classification. In both groups, YFS (Cardiovascular Risk in Young Finns Study) presents lower median Lp(a) concentrations than Non-Finnish Central Europeans (NFE). The nonclustered populations are shown in Figure V in the [online-only Data Supplement](#). **B**, Median Lp(a) concentrations stratified by isoform size in groups of 3 to 5 KIV (kringle IV) isoforms (based on isoform 1). YFS individuals express lower Lp(a) concentrations than NFE individuals in all isoform groups. In the group 11 to 16 KIV repeats, Finns have an even lower Lp(a) concentration than in the larger 17 to 19 KIV repeats group. Stratified number of carriers for each population and median Lp(a) concentrations are given in Tables IV and V in the [online-only Data Supplement](#) and Figure VI in the [online-only Data Supplement](#).

Discussion

Lp(a) concentrations show pronounced interethnic and interpopulation variability.^{5,17,19} Several reports showed lower Lp(a) concentrations in Finns than in other European populations,^{23–27,49} but the small sample numbers of early studies and the nonstandardized Lp(a) quantification methods hampered a direct comparison. Waldeyer et al²⁸ recently confirmed lower median Lp(a) concentrations ($\approx 58\%$) in the Finnish FINRISK cohort compared with Central European populations. However, the causes could not be addressed. Various causes are conceivable: (1) differences in the isoform distribution, (2) differences in the frequency of Lp(a)-lowering variants, or (3) differences in modifier genes.

To dissect the relative contribution of these factors in a standardized study, we determined apo(a) isoforms and Lp(a) concentrations with the same assays for 10003 Central European whites from 3 population-based studies (NFE) and 1 Finnish study (YFS; $n=2281$).¹⁰ The contributions of isoform distribution, *LPA* LOF mutations and variants, *APOE* genotypes, and *PCSK9* R46L were assessed.

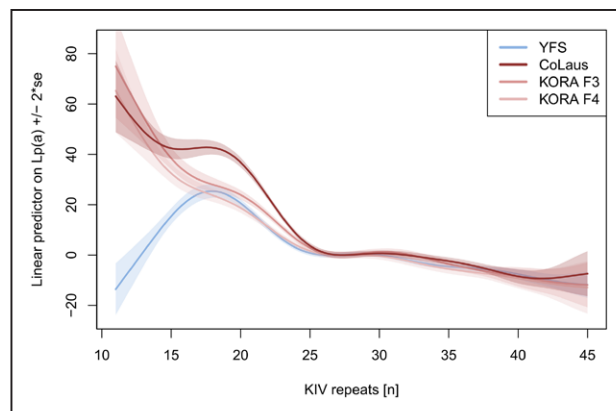


Figure 4. Nonlinear spline, showing the effect of isoform 1 (ie, in heterozygotes, the smaller allele present) on untransformed Lp(a) (lipoprotein(a)) concentrations. The splines are centered at 0 at isoform 1=27 (median on the Non-Finnish Central Europeans studies). The splines are derived from the multiple regression model additionally adjusting for the variables age, sex, estimated glomerular filtration rate, rs143411368, rs41272114, rs11591147, *APOE2*, and *APOE4*-carriers (Table V in the [online-only Data Supplement](#)). The spline for inverse-transformed Lp(a) is shown in Figure VIII in the [online-only Data Supplement](#). YFS, Cardiovascular Risk in Young Finns Study; KORA, Cooperative Health Research in the Region of Augsburg; KIV, kringle IV.

We found that a consistent part of the differences in Lp(a) concentrations can be explained by the common action of (1) isoform distribution, (2) *LPA* variants, and (3) *APOE* genotypes. Additionally, we observed that in Finns, the Lp(a) concentrations were reduced in all apo(a) isoform groups.

The isoform distribution in Finns was shifted toward longer isoforms (Figure 2), which accounted for 27% of the interpopulation differences in Lp(a) concentrations (Figure 5). It thus represents the largest contributor to the interpopulation differences without being, however, the only causal factor. This is in line with previous reports showing that differences in isoform distribution among populations only partially account for interpopulation differences in Lp(a) concentrations.^{1,5}

We found that in YFS, the Lp(a) concentrations were reduced even within the same-isoform group. This mirrors the differences between blacks and whites, where blacks present higher Lp(a) concentrations over the whole range of isoforms.^{17,50} The effect was most striking in the 11 to 16 KIV repeats group, which showed even lower Lp(a) than the 17 to 19 KIV repeats group (Figures 3B and 4). Finns have a higher frequency of LOF mutations and variants in conserved noncoding regions than Central Europeans.^{6,51} Because *LPA* is reported to be LOF tolerant by ExAC (according to the pLI/pRec scores on the ExAC website³¹) and by phenome-wide studies,³⁰ and it has been shown that Lp(a)-modifying SNPs can cluster with particular isoforms^{3,52}; it is tempting to speculate that novel functional variants may exist in the 11 to 16 KIV repeat carriers in YFS. Indeed, especially short isoforms according to the Western blot might represent truncated protein forms caused by novel nonsense or splice-site variants. Such truncated forms have been observed previously in Tyroleans with low Lp(a).⁸

Table 2. Results From Multiple Regression Analyses in the YFS and in NFE Combined (Using a Random Effects Model)

Variables	Frequencies/Mean Values in YFS Compared With NFE	Results From Multiple Regression Model on Lp(a)*				Expected Consequence of the Differences in Variable Distribution on Lp(a) Levels in YFS Compared With NFE
		YFS		NFE (Mixed Model)		
		β	P Value	β	P Value	
Age	↓	0.07	0.3962	0.13	8.19×10^{-11}	↓
Sex (F)	↔	0.62	0.0368	2.77	4.84×10^{-20}	↔
Isoforms†	↑	NA	2.60×10^{-124}	NA	$<1 \times 10^{-300}$	↓
eGFR	↑	-0.03	0.1899	-0.03	0.0053	↓
rs143411368‡	↑	-2.6	5.33×10^{-9}	-3.22	0.0334	↓
rs41272114‡	↑	-2.45	6.38×10^{-19}	-6.05	7.67×10^{-58}	↓
rs11591147‡	↑	0.53	0.9545	-0.58	0.3769	↔
rs10455872‡	↓	11.68	1.33×10^{-13}	18.69	1.50×10^{-74}	↓
rs3798220	↔	28.03	3.37×10^{-14}	43.54	1.35×10^{-89}	↔
APOE2	↓	-2.34	0.0071	-3.60	1.60×10^{-16}	↑
APOE4	↑	-0.21	0.6147	-0.34	0.23570	↔

eGFR indicates estimated glomerular filtration rate; Lp(a), lipoprotein(a); NFE, Non-Finnish Central Europeans; and YFS, Cardiovascular Risk in Young Finns Study. * β estimate is derived from model on Lp(a). P value is derived from model on inverse-normal transformed Lp(a) values to ensure normal distribution assumption. †P values for isoforms refer to the significance of a nonlinear spline of isoform 1 on inverse-normal transformed Lp(a); for effect size, see nonlinear spline in Figure 5. ‡For SNPs (single nucleotide polymorphisms), an additive inheritance model was assumed. NA, effect is given as spline in Figure 4.

Several studies proposed that (even frequent) SNPs contribute to interethnic differences in Lp(a) levels.^{53,54} The *APOE* isoforms vary between Finns and Central Europeans, with *APOE2* being rarer and *APOE4* being more frequent in Finns. Although the individual impact of these genetic variants on Lp(a) has been described previously,^{6,7,9,10,12} their cumulative contribution in establishing low Lp(a) levels in Finns has not been evaluated to date. We showed that the frequency distribution of 2 *LPA* LOF variants and rs10455872 contributed for a total of $\approx 10\%$ to the population difference. Conversely, *APOE2* counteracts this by increasing Lp(a) and so contributes to closing the gap between the populations (Figure 5; Table VI in the [online-only Data Supplement](#)). Among the *APOE* genotypes, only *APOE2* showed an appreciable contribution in the mediation analysis in this sample set, which is in line with previous results¹⁰ (Figure 5; Table VI in the [online-only Data Supplement](#)). Of note, rs41272114—a major contributor in our mediation analysis—has been proposed also as a cause for the Lp(a) differences between whites and Africans.⁵⁴ No effect of rs11591147 was found, which is likely because of its small absolute effects on Lp(a) concentrations (1 mg/dL).¹³ We could also not confirm the high MAF (17.9%) of rs11591147 reported in Finns by ExAC (Table 1), but we rather support the MAF determined in the SISu (Sequencing Initiative in Suomi) project (4.2%).⁶

Importantly, the focus of the work at hands was to show that the observed differences in Lp(a) are not just because of a shifted isoform distribution and evaluate the contribution of known functional factors affecting Lp(a) concentrations. This left $\approx 28\%$ of variance unexplained and suggests the existence of additional functional variation. Various types of genetic variants may contribute to this observation. The KIV-2 region plays a special role because it encompasses a considerable part of the *LPA* gene but is currently not captured by common

genetic data set. However, also differences in the effect magnitude or frequency of known variation outside the KIV-2 may play role. Comparative ultradeep sequencing studies and a

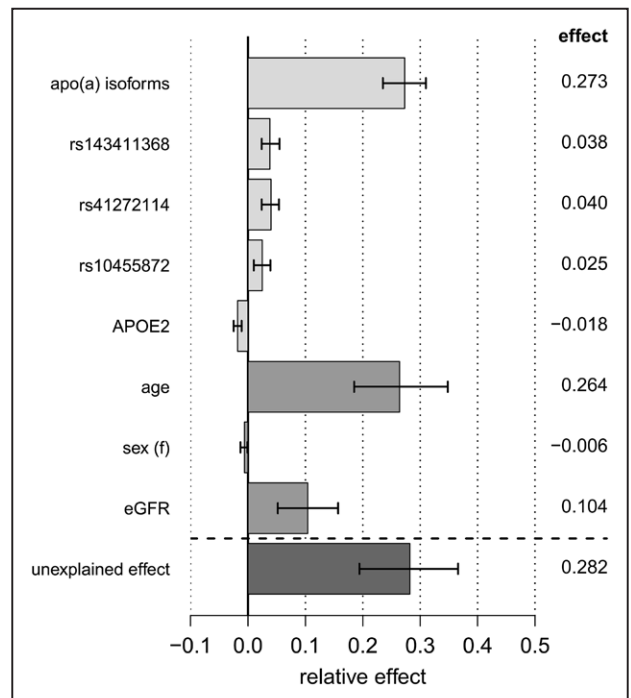


Figure 5. Mediation analysis explaining the differences in Lp(a) (lipoprotein(a)) concentrations between YFS (Cardiovascular Risk in Young Finns Study) and Non-Finnish Central Europeans: decomposition of the total effect on inverse-normal transformed Lp(a) values into explained and unexplained relative effects (including 95% confidence interval). Exact values are reported in Table VII in the [online-only Data Supplement](#). eGFR, estimated glomerular filtration rate; apo(a), apolipoprotein(a).

dedicated study design will be required to pinpoint the genetic determinants of the population differences reported here.

Strengths and Limitations

Our study presents strengths and limitations. To the best of our knowledge, here we present the largest investigation of Lp(a) concentrations in European populations accounting for Lp(a) isoform data to date. All measurements and Western blots were done in a single laboratory using the same methods in all populations. The YFS individuals are younger and present better renal function—a major determinant of Lp(a) levels¹—than the NFE. Accordingly, environmental differences account for 34% of the population differences, and age and eGFR were the second largest contributors to the population differences. This underscores the importance of matching these factors, despite the often alleged robustness Lp(a) to nongenetic influences.

The pronounced effect of age is surprising and could reflect kidney function parameters, which are not reflected by eGFR, other unmeasured age-related confounders (eg, hormonal status), or undetermined study-specific confounders, which are specific for the YFS study. The difference between YFS and NFE is present in both sexes (Table I in the [online-only Data Supplement](#)) and also when looking only at the age range ≥ 32 to ≤ 39 , where all studies overlap. This indicates that the observed Lp(a) reduction is not only because of different proportion of postmenopausal individuals in NFE. Finally, it has to be noted that the mediation analysis assumes a linear and homogeneous effect of the factors. Therefore, the differing effect of low isoforms on Lp(a) in YFS cannot be fully accounted for and may contribute to the unexplained effect.

Importantly, no other study with combined information on Lp(a) concentrations and isoform distribution in Finns was available until now, except 1 small study in 181 individuals.²⁷ We could successfully identify and quantify determinants of the Lp(a) differences between Finns and NFE and show that a limited number of genetic elements (isoforms and SNPs) already explain remarkable $\approx 37\%$ of the population difference. This may indicate that a small number of factors can already explain a consistent part of differences of Lp(a) between Finns and NFE populations. Finally, the availability of Western blot data allowed identifying a surprising Lp(a)-lowering effect in short isoforms in YFS, which warrants further research.

Conclusion

We demonstrate that the lower Lp(a) concentrations in Finns are not caused solely by differences in isoform distribution but are lowered across the whole range of isoforms. Frequency differences of common LOF mutations and *APOE* isoforms significantly contribute to these findings but do not fully explain it. Of particular note, especially very small isoforms with 11 to 16 KIV repeats present massively lowered Lp(a) concentrations, which might hint toward the existence of novel LOF mutations. Therefore, future sequencing studies in *LPA* might be fruitful to identify novel Lp(a)-lowering variants.

Acknowledgments

This work was part of the bachelor thesis in Biomedical Sciences of G. Erhart at the Health University of Applied Sciences Tyrol.

Sources of Funding

This study was supported by the Austrian Science Fund (FWF) Project Number P266600-B13 to C. Lamina. Other funding sources are given in the [online-only Data Supplement](#).

Disclosures

None.

References

- Kronenberg F, Utermann G. Lipoprotein(a): resurrected by genetics. *J Intern Med*. 2013;273:6–30. doi: 10.1111/j.1365-2796.2012.02592.x.
- Nordestgaard BG, Langsted A. Lipoprotein (a) as a cause of cardiovascular disease: insights from epidemiology, genetics, and biology. *J Lipid Res*. 2016;57:1953–1975. doi: 10.1194/jlr.R071233.
- Schmidt K, Noureen A, Kronenberg F, Utermann G. Structure, function, and genetics of lipoprotein (a). *J Lipid Res*. 2016;57:1339–1359. doi: 10.1194/jlr.R067314.
- Erqou S, Thompson A, Di Angelantonio E, Saleheen D, Kaptoge S, Marcovina S, Danesh J. Apolipoprotein(a) isoforms and the risk of vascular disease: systematic review of 40 studies involving 58,000 participants. *J Am Coll Cardiol*. 2010;55:2160–2167. doi: 10.1016/j.jacc.2009.10.080.
- Enkmaa B, Anuurad E, Berglund L. Lipoprotein (a): impact by ethnicity and environmental and medical conditions. *J Lipid Res*. 2016;57:1111–1125. doi: 10.1194/jlr.R051904.
- Lim ET, Würtz P, Havulinna AS, et al; Sequencing Initiative Suomi (SISu) Project. Distribution and medical impact of loss-of-function variants in the Finnish founder population. *PLoS Genet*. 2014;10:e1004494. doi: 10.1371/journal.pgen.1004494.
- Ogorelkova M, Gruber A, Utermann G. Molecular basis of congenital Lp(a) deficiency: a frequent apo(a) ‘null’ mutation in caucasians. *Hum Mol Genet*. 1999;8:2087–2096. doi: 10.1093/hmg/8.11.2087.
- Parson W, Kraft HG, Niederstätter H, Lingenhel AW, Köchl S, Fresser F, Utermann G. A common nonsense mutation in the repetitive kringle IV-2 domain of human apolipoprotein(a) results in a truncated protein and low plasma Lp(a). *Hum Mutat*. 2004;24:474–480. doi: 10.1002/humu.20101.
- Moriarty PM, Varvel SA, Gordts PL, McConnell JP, Tsimikas S. Lipoprotein(a) mass levels increase significantly according to APOE genotype: an analysis of 431 239 patients. *Arterioscler Thromb Vasc Biol*. 2017;37:580–588. doi: 10.1161/ATVBAHA.116.308704.
- Mack S, Coassin S, Rueedi R, et al; KORA-Study Group. A genome-wide association meta-analysis on lipoprotein (a) concentrations adjusted for apolipoprotein (a) isoforms. *J Lipid Res*. 2017;58:1834–1844. doi: 10.1194/jlr.M076232.
- Anuurad E, Lu G, Rubin J, Pearson TA, Berglund L. ApoE genotype affects allele-specific apo[a] levels for large apo[a] sizes in African Americans: the Harlem-Basset Study. *J Lipid Res*. 2007;48:693–698. doi: 10.1194/jlr.M600431-JLR200.
- Li J, Lange LA, Sabourin J, Duan Q, Valdar W, Willis MS, Li Y, Wilson JG, Lange EM. Genome- and exome-wide association study of serum lipoprotein (a) in the Jackson Heart Study. *J Hum Genet*. 2015;60:755–761. doi: 10.1038/jhg.2015.107.
- Langsted A, Nordestgaard BG, Benn M, Tybjaerg-Hansen A, Kamstrup PR. PCSK9 R46L loss-of-function mutation reduces lipoprotein(a), LDL cholesterol, and risk of aortic valve stenosis. *J Clin Endocrinol Metab*. 2016;101:3281–3287. doi: 10.1210/jc.2016-1206.
- Romagnuolo R, Scipione CA, Boffa MB, Marcovina SM, Seidah NG, Koschinsky ML. Lipoprotein(a) catabolism is regulated by proprotein convertase subtilisin/kexin type 9 through the low density lipoprotein receptor. *J Biol Chem*. 2015;290:11649–11662. doi: 10.1074/jbc.M114.611988.
- Tiret L, de Knijff P, Menzel HJ, Ehnholm C, Nicaud V, Havekes LM. ApoE polymorphism and predisposition to coronary heart disease in youths of different European populations. The EARS Study. European

- Atherosclerosis Research Study. *Arterioscler Thromb*. 1994;14:1617–1624. doi: 10.1161/01.ATV.14.10.1617.
16. Reyes-Soffer G, Pavlyha M, Ngai C, et al. Effects of PCSK9 inhibition with alirocumab on lipoprotein metabolism in healthy humans. *Circulation*. 2017;135:352–362. doi: 10.1161/CIRCULATIONAHA.116.025253.
 17. Sandholzer C, Hallman DM, Saha N, Sigurdsson G, Lackner C, Császár A, Boerwinkle E, Utermann G. Effects of the apolipoprotein(a) size polymorphism on the lipoprotein(a) concentration in 7 ethnic groups. *Hum Genet*. 1991;86:607–614.
 18. Marcovina SM, Albers JJ. Lipoprotein (a) measurements for clinical application. *J Lipid Res*. 2016;57:526–537. doi: 10.1194/jlr.R061648.
 19. Scholz M, Kraft HG, Lingenhel A, Delport R, Vorster EH, Bickeböller H, Utermann G. Genetic control of lipoprotein(a) concentrations is different in Africans and Caucasians. *Eur J Hum Genet*. 1999;7:169–178. doi: 10.1038/sj.ejhg.5200290.
 20. Helmholt M, Bigge J, Muche R, Mainoo J, Thiery J, Seidel D, Armstrong VW. Contribution of the apo[a] phenotype to plasma Lp[a] concentrations shows considerable ethnic variation. *J Lipid Res*. 1991;32:1919–1928.
 21. Salmela E, Lappalainen T, Fransson I, Andersen PM, Dahlman-Wright K, Fiebig A, Sistonen P, Savontaus ML, Schreiber S, Kere J, Lahermo P. Genome-wide analysis of single nucleotide polymorphisms uncovers population structure in Northern Europe. *PLoS One*. 2008;3:e3519. doi: 10.1371/journal.pone.0003519.
 22. Lao O, Lu TT, Nothnagel M, et al. Correlation between genetic and geographic structure in Europe. *Curr Biol*. 2008;18:1241–1248. doi: 10.1016/j.cub.2008.07.049.
 23. Salomaa V, Rasi V, Kulathinal S, Vahtera E, Jauhiainen M, Ehnholm C, Pekkanen J. Hemostatic factors as predictors of coronary events and total mortality: the FINRISK '92 Hemostasis Study. *Arterioscler Thromb Vasc Biol*. 2002;22:353–358.
 24. Kivimäki M, Magnussen CG, Juonala M, Kähönen M, Kettunen J, Loo BM, Lehtimäki T, Viikari J, Raitakari OT. Conventional and Mendelian randomization analyses suggest no association between lipoprotein(a) and early atherosclerosis: the Young Finns Study. *Int J Epidemiol*. 2011;40:470–478. doi: 10.1093/ije/dyq205.
 25. Leino A, Impivaara O, Kaitsaari M, Järvisalo J. Serum concentrations of apolipoprotein A-I, apolipoprotein B, and lipoprotein(a) in a population sample. *Clin Chem*. 1995;41:1633–1636.
 26. Kunutsor SK, Khan H, Nyssönen K, Laukkanen JA. Lipoprotein(a) and risk of sudden cardiac death in middle-aged Finnish men: a new prospective cohort study. *Int J Cardiol*. 2016;220:718–725. doi: 10.1016/j.ijcard.2016.06.069.
 27. Klausen IC, Beisiegel U, Menzel HJ, Rosseneu M, Nicaud V, Faergeman O. Apo(a) phenotypes and Lp(a) concentrations in offspring of men with and without myocardial infarction. The EARS Study. European Atherosclerosis Research Study. *Arterioscler Thromb Vasc Biol*. 1995;15:1001–1008. doi: 10.1161/01.ATV.15.8.1001.
 28. Waldeyer C, Makarova N, Zeller T, et al. Lipoprotein(a) and the risk of cardiovascular disease in the European population: results from the BiomarCaRE consortium. *Eur Heart J*. 2017;38:2490–2498. doi: 10.1093/eurheartj/ehx166.
 29. Kyriakou T, Seedorf U, Goel A, Hopewell JC, Clarke R, Watkins H, Farrall M; PROCARDIS Consortium. A common LPA null allele associates with lower lipoprotein(a) levels and coronary artery disease risk. *Arterioscler Thromb Vasc Biol*. 2014;34:2095–2099. doi: 10.1161/ATVBAHA.114.303462.
 30. Emdin CA, Khera AV, Natarajan P, et al; CHARGE–Heart Failure Consortium; CARDIoGRAM Exome Consortium. Phenotypic characterization of genetically lowered human lipoprotein(a) levels. *J Am Coll Cardiol*. 2016;68:2761–2772. doi: 10.1016/j.jacc.2016.10.033.
 31. Lek M, Karczewski KJ, Minikel EV, et al; Exome Aggregation Consortium. Analysis of protein-coding genetic variation in 60,706 humans. *Nature*. 2016;536:285–291. doi: 10.1038/nature19057.
 32. Ehnholm C, Lukka M, Kuusi T, Nikkilä E, Utermann G. Apolipoprotein E polymorphism in the Finnish population: gene frequencies and relation to lipoprotein concentrations. *J Lipid Res*. 1986;27:227–235.
 33. Firmann M, Mayor V, Vidal PM, et al. The CoLaus study: a population-based study to investigate the epidemiology and genetic determinants of cardiovascular risk factors and metabolic syndrome. *BMC Cardiovasc Disord*. 2008;8:6. doi: 10.1186/1471-2261-8-6.
 34. Heid IM, Boes E, Müller M, et al. Genome-wide association analysis of high-density lipoprotein cholesterol in the population-based KORA study sheds new light on intergenic regions. *Circ Cardiovasc Genet*. 2008;1:10–20. doi: 10.1161/CIRCGENETICS.108.776708.
 35. Juonala M, Viikari JS, Raitakari OT. Main findings from the prospective cardiovascular risk in Young Finns Study. *Curr Opin Lipidol*. 2013;24:57–64. doi: 10.1097/MOL.0b013e32835a7ed4.
 36. Kronenberg F, Lobentanz EM, König P, Utermann G, Dieplinger H. Effect of sample storage on the measurement of lipoprotein[a], apolipoproteins B and A-IV, total and high density lipoprotein cholesterol and triglycerides. *J Lipid Res*. 1994;35:1318–1328.
 37. Dieplinger H, Gruber G, Krasznai K, Reschauer S, Seidel C, Burns G, Müller HJ, Császár A, Vogel W, Robenek H. Kringle 4 of human apolipoprotein[a] shares a linear antigenic site with human catalase. *J Lipid Res*. 1995;36:813–822.
 38. Kronenberg F, Kuen E, Ritz E, Junker R, König P, Kraatz G, Lhotta K, Mann JF, Müller GA, Neyer U, Riegel W, Reigler P, Schwenger V, Von Eckardstein A. Lipoprotein(a) serum concentrations and apolipoprotein(a) phenotypes in mild and moderate renal failure. *J Am Soc Nephrol*. 2000;11:105–115.
 39. Kraft HG, Köchl S, Menzel HJ, Sandholzer C, Utermann G. The apolipoprotein (a) gene: a transcribed hypervariable locus controlling plasma lipoprotein (a) concentration. *Hum Genet*. 1992;90:220–230. doi: 10.1007/BF00220066.
 40. Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, Korbel JO, Marchini JL, McCarthy S, McVean GA, Abecasis GR; 1000 Genomes Project Consortium. A global reference for human genetic variation. *Nature*. 2015;526:68–74. doi: 10.1038/nature15393.
 41. Loh PR, Danecek P, Palamara PF, Fuchsberger C, A Reshef Y, K Finucane H, Schoenherr S, Forer L, McCarthy S, Abecasis GR, Durbin R, L Price A. Reference-based phasing using the Haplotype Reference Consortium panel. *Nat Genet*. 2016;48:1443–1448. doi: 10.1038/ng.3679.
 42. Kronenberg F, Trenkwalder E, Lingenhel A, Friedrich G, Lhotta K, Schober M, Moes N, König P, Utermann G, Dieplinger H. Renovascular arteriovenous differences in Lp[a] plasma concentrations suggest removal of Lp[a] from the renal circulation. *J Lipid Res*. 1997;38:1755–1763.
 43. Kronenberg F. Causes and consequences of lipoprotein(a) abnormalities in kidney disease. *Clin Exp Nephrol*. 2014;18:234–237. doi: 10.1007/s10157-013-0875-8.
 44. Levey AS, Stevens LA, Schmid CH, Zhang YL, Castro AF III, Feldman HI, Kusek JW, Eggers P, Van Lente F, Greene T, Coresh J; CKD-EPI (Chronic Kidney Disease Epidemiology Collaboration). A new equation to estimate glomerular filtration rate. *Ann Intern Med*. 2009;150:604–612. doi: 10.7326/0003-4819-150-9-200905050-00006.
 45. Kamstrup PR, Tybjaerg-Hansen A, Steffensen R, Nordestgaard BG. Genetically elevated lipoprotein(a) and increased risk of myocardial infarction. *JAMA*. 2009;301:2331–2339. doi: 10.1001/jama.2009.801.
 46. Nordestgaard BG, Chapman MJ, Ray K, et al; European Atherosclerosis Society Consensus Panel. Lipoprotein(a) as a cardiovascular risk factor: current status. *Eur Heart J*. 2010;31:2844–2853. doi: 10.1093/eurheartj/ehq386.
 47. Laschkolnig A, Kollerits B, Lamina C, Meisinger C, Rantner B, Stadler M, Peters A, Koenig W, Stöckl A, Dähnhardt D, Böger CA, Krämer BK, Fraedrich G, Strauch K, Kronenberg F. Lipoprotein (a) concentrations, apolipoprotein (a) phenotypes, and peripheral arterial disease in three independent cohorts. *Cardiovasc Res*. 2014;103:28–36. doi: 10.1093/cvr/cvu107.
 48. Clarke R, Peden JF, Hopewell JC, et al; PROCARDIS Consortium. Genetic variants associated with Lp(a) lipoprotein level and coronary disease. *N Engl J Med*. 2009;361:2518–2528. doi: 10.1056/NEJMoa0902604.
 49. Ogorelkova M, Kraft HG, Ehnholm C, Utermann G. Single nucleotide polymorphisms in exons of the apo(a) kringle IV types 6 to 10 domain affect Lp(a) plasma concentrations and have different patterns in Africans and Caucasians. *Hum Mol Genet*. 2001;10:815–824. doi: 10.1093/hmg/10.8.815.
 50. Gaw A, Boerwinkle E, Cohen JC, Hobbs HH. Comparative analysis of the apo(a) gene, apo(a) glycoprotein, and plasma concentrations of Lp(a) in three ethnic groups. Evidence for no common “null” allele at the apo(a) locus. *J Clin Invest*. 1994;93:2526–2534. doi: 10.1172/JCI117263.
 51. Chheda H, Palta P, Pirinen M, McCarthy S, Walter K, Koskinen S, Salomaa V, Daly M, Durbin R, Palotie A, Aittokallio T, Ripatti S. Whole-genome view of the consequences of a population bottleneck using 2926 genome sequences from Finland and United Kingdom. *Eur J Hum Genet*. 2017;25:477–484. doi: 10.1038/ejhg.2016.205.
 52. Coassin S, Erhart G, Weissensteiner H, et al. A novel but frequent variant in LPA KIV-2 is associated with a pronounced Lp(a) and cardiovascular risk reduction. *Eur Heart J*. 2017;38:1823–1831. doi: 10.1093/eurheartj/ehx174.

53. Deo RC, Wilson JG, Xing C, Lawson K, Kao WH, Reich D, Tandon A, Akylbekova E, Patterson N, Mosley TH Jr, Boerwinkle E, Taylor HA Jr. Single-nucleotide polymorphisms in LPA explain most of the ancestry-specific variation in Lp(a) levels in African Americans. *PLoS One*. 2011;6:e14581. doi: 10.1371/journal.pone.0014581.
54. Chretien JP, Coresh J, Berthier-Schaad Y, Kao WH, Fink NE, Klag MJ, Marcovina SM, Giaculli F, Smith MW. Three single-nucleotide polymorphisms in LPA account for most of the increase in lipoprotein(a) level elevation in African Americans compared with European Americans. *J Med Genet*. 2006;43:917–923. doi: 10.1136/jmg.2006.042119.
55. Kinpara K, Okada H, Yoneyama A, Okubo M, Murase T. Lipoprotein(a)-cholesterol: a significant component of serum cholesterol. *Clin Chim Acta*. 2011;412:1783–1787. doi: 10.1016/j.cca.2011.05.036.

Highlights

- Lp(a) (lipoprotein(a)) concentrations are massively reduced by $\approx 50\%$ in Finns ($n=2281$) compared with Central Europeans ($n=10\,003$).
- Finns presented a shift in apo(a) isoform distribution toward larger isoforms and frequency differences in established Lp(a)-regulating genetic variants (*LPA* loss-of-functions, *APOE*, and *PCSK9*).
- Lp(a) was reduced over the whole isoform range. Of particular note, very small isoforms showed surprisingly low Lp(a) concentrations.
- 71.8% of the differences between the populations can be explained by a combination of genetic and environmental factors.

Arteriosclerosis, Thrombosis, and Vascular Biology



JOURNAL OF THE AMERICAN HEART ASSOCIATION

Genetic Factors Explain a Major Fraction of the 50% Lower Lipoprotein(a) Concentrations in Finns

Gertraud Erhart, Claudia Lamina, Terho Lehtimäki, Pedro Marques-Vidal, Mika Kähönen, Peter Vollenweider, Olli T. Raitakari, Gérard Waeber, Barbara Thorand, Konstantin Strauch, Christian Gieger, Thomas Meitinger, Annette Peters, Florian Kronenberg and Stefan Coassin

Arterioscler Thromb Vasc Biol. 2018;38:1230-1241; originally published online March 22, 2018;

doi: 10.1161/ATVBAHA.118.310865

Arteriosclerosis, Thrombosis, and Vascular Biology is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231

Copyright © 2018 American Heart Association, Inc. All rights reserved.

Print ISSN: 1079-5642. Online ISSN: 1524-4636

The online version of this article, along with updated information and services, is located on the World Wide Web at:

<http://atvb.ahajournals.org/content/38/5/1230>

Free via Open Access

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in *Arteriosclerosis, Thrombosis, and Vascular Biology* can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the [Permissions and Rights Question and Answer](#) document.

Reprints: Information about reprints can be found online at:

<http://www.lww.com/reprints>

Subscriptions: Information about subscribing to *Arteriosclerosis, Thrombosis, and Vascular Biology* is online at:

<http://atvb.ahajournals.org/subscriptions/>

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in *Arteriosclerosis, Thrombosis, and Vascular Biology* can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the [Permissions and Rights Question and Answer](#) document.

Reprints: Information about reprints can be found online at:
<http://www.lww.com/reprints>

Subscriptions: Information about subscribing to *Arteriosclerosis, Thrombosis, and Vascular Biology* is online at:
<http://atvb.ahajournals.org/subscriptions/>

Major Resources Table

Antibodies

Target antigen	Vendor or Source	Catalog #	Working concentration	Lot # (preferred but not required)
affinity-purified polyclonal rabbit anti-human apo(a) antibody	produced in-house	produced in-house	5 µg/ml in 1x PBS containing 1 mg/mL NaN ₃ for ELISA	-
horseradish-peroxidase-conjugated monoclonal anti-apo(a) antibody 1A2	produced in-house	produced in-house	Conjugate stock stored 1:7,000 in glycerol, then diluted 1:4,800 in Assay Buffer (Microcoat, Bernried, D)	-
1A2 monoclonal anti-apo(a) antibody for Western blotting	produced in-house	produced in-house	168 ng/ml in Buffer C (85 mM NaCl, 10 mM TRIS, 0.2% Triton X-100, 1% BSA)	-
horseradish-peroxidase conjugated goat anti-mouse	Millipore	401253	1:13,333 in Buffer C (85 mM NaCl, 10 mM TRIS, 0.2% Triton X-100, 1% BSA)	-

SUPPLEMENTAL MATERIAL

Genetic factors explain a major fraction of the 50% lower lipoprotein(a) concentrations in Finns

Running title: Genetic factors lower Lp(a) in Finns

Gertraud Erhart¹, Claudia Lamina¹, Terho Lehtimäki², Pedro Marques-Vidal³, Mika Kähönen⁴, Peter Vollenweider³, Olli T Raitakari^{5,6}, Gérard Waeber³, Barbara Thorand^{7,8}, Konstantin Strauch^{9,10}, Christian Gieger^{7,9,11}, Thomas Meitinger^{12,13,14}, Annette Peters^{7,8,15}, Florian Kronenberg¹, Stefan Coassin¹

¹ Division of Genetic Epidemiology, Department of Medical Genetics, Molecular and Clinical Pharmacology, Medical University of Innsbruck, Innsbruck, Austria,

² Department of Clinical Chemistry, Fimlab Laboratories, and Finnish Cardiovascular Research Center - Tampere, Faculty of Medicine and Life Sciences, University of Tampere, Tampere 33520, Finland

³ Internal Medicine, Lausanne University Hospital, 1015 Lausanne, Switzerland

⁴ Department of Clinical Physiology, Tampere University Hospital, and Finnish Cardiovascular Research Center - Tampere, Faculty of Medicine and Life Sciences, University of Tampere, Tampere 33521, Finland

⁵ Department of Clinical Physiology and Nuclear Medicine, Turku University Hospital, Turku, Finland

⁶ Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku, Turku, Finland

⁷ Institute of Epidemiology II, Helmholtz Zentrum München—German Research Center for Environmental Health, 85764 Neuherberg, Germany

⁸ German Center for Diabetes Research (DZD), 85764 Neuherberg, Germany

⁹ Institute of Genetic Epidemiology, Helmholtz Zentrum München—German Research Center for Environmental Health, 85764 Neuherberg, Germany

¹⁰ Institute of Medical Informatics, Biometry and Epidemiology, Chair of Genetic Epidemiology, Ludwig-Maximilians-Universität, 81377 Munich, Germany

¹¹ Research Unit of Molecular Epidemiology, Helmholtz Zentrum München—German Research Center for Environmental Health, 85764 Neuherberg, Germany

¹² Institute of Human Genetics, Technische Universität München, 81675 München, Germany

¹³ Institute of Human Genetics, Helmholtz Zentrum München, 85764 Neuherberg, Germany

¹⁴ Munich Cluster for Systems Neurology (SyNergy), 81377 Munich, Germany

¹⁵ DZHK (German Centre for Cardiovascular Research), partner site Munich Heart Alliance, 80802 Munich, Germany

Address of correspondence:

Stefan Coassin, PhD

Division of Genetic Epidemiology

Department of Medical Genetics, Molecular and Clinical Pharmacology

Medical University of Innsbruck

Schöpfstr. 41, A-6020 Innsbruck, AUSTRIA

Phone: (+43) 512 9003 70576, Fax: (+43) 512 9003 73561

E-mail: Stefan.Coassin@i-med.ac.at

Contents

- Contents 2
- Supplementary Note 3
- Supplementary Figures..... 4
 - Supplementary Figure I: 4
 - Supplementary Figure II: 5
 - Supplementary Figure III: 6
 - Supplementary Figure IV: 7
 - Supplementary Figure V: 8
 - Supplementary Figure VI: 9
 - Supplementary Figure VII: 10
 - Supplementary Figure VIII: 11
- Supplementary Tables..... 12
 - Supplementary Table I: 12
 - Supplementary Table II: 12
 - Supplementary Table III: 13
 - Supplementary Table IV: 14
 - Supplementary Table V: 15
 - Supplementary Table VI: 16
 - Supplementary Table VII: 17
 - Supplementary Table VIII: 18
- References..... 18

Supplementary Note

The study was supported by the Austrian Genome Project 'GOLD' to FK, by the Austrian Science Fund (FWF) Project Number P266600-B13 to CL and by the intramural funding of the Medical University of Innsbruck for young scientists (MUI-START, project 2015-06-007) to SC.

The Young Finns Study has been financially supported by the Academy of Finland: grants 286284 (T.L.), 134309 (Eye), 126925, 121584, 124282, 129378 (Salve), 117787 (Gendi), and 41071 (Skidi); the Social Insurance Institution of Finland; Kuopio, Tampere and Turku University Hospital Medical Funds (grant X51001 for T.L.); Juho Vainio Foundation; Paavo Nurmi Foundation; Finnish Foundation of Cardiovascular Research (T.L.); Finnish Cultural Foundation; Tampere Tuberculosis Foundation (T.L.); Emil Aaltonen Foundation (T.L.); Yrjö Jahnsson Foundation (T.L.) and Signe ja Ane Gyllenberg's Foundation (T.L.) and Diabetes Research Foundation of Finnish Diabetes Association (T.L.). We gratefully acknowledge the THL DNA laboratory for its skillful work to produce the DNA samples used in this study and Irina Lisinen for the expert technical assistance in the statistical analyses.

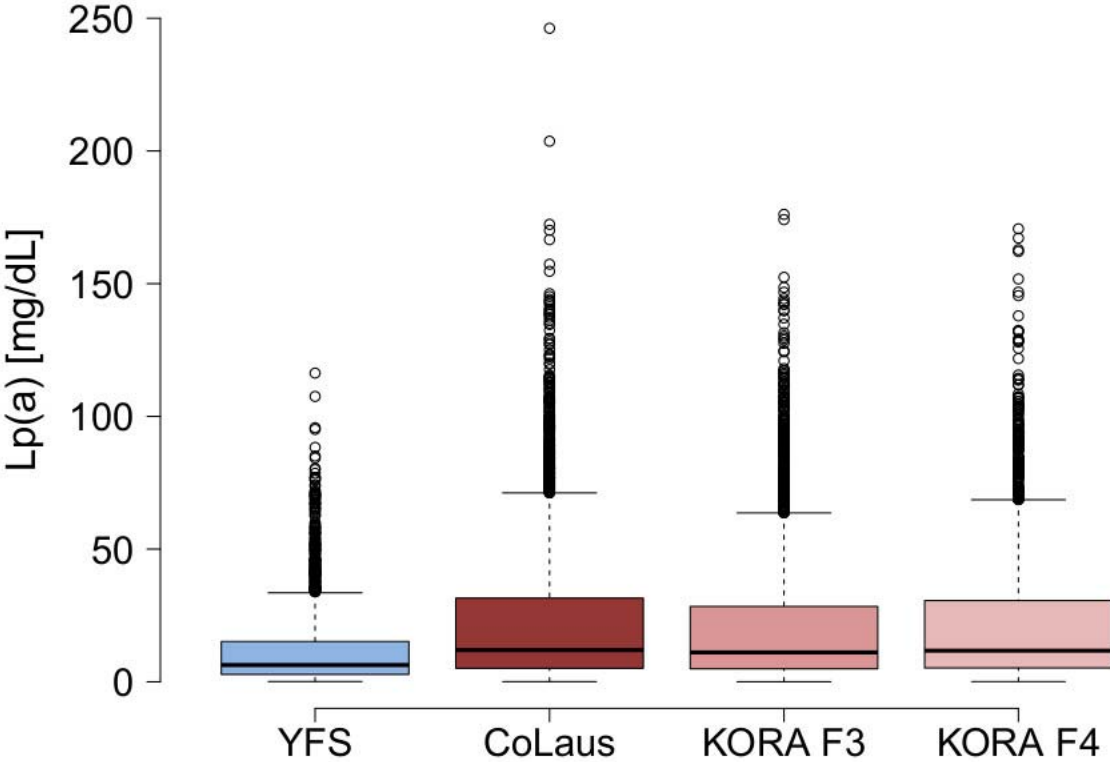
The CoLaus study was and is supported by research grants from GlaxoSmithKline, the Faculty of Biology and Medicine of Lausanne, and the Swiss National Science Foundation (grants 33CSCO-122661, 33CS30-139468 and 33CS30-148401).

The KORA-Study Group consists of A. Peters (speaker), J. Heinrich, R. Holle, R. Leidl, C. Meisinger, K. Strauch, and their co-workers, who are responsible for the design and conduct of the KORA studies. The KORA study was initiated and financed by the Helmholtz Zentrum München – German Research Center for Environmental Health, which is funded by the German Federal Ministry of Education and Research (BMBF) and by the State of Bavaria. Furthermore, KORA research was supported within the Munich Center of Health Sciences (MC-Health), Ludwig-Maximilians-Universität, as part of LMUinnovativ.

Supplementary Figures

Supplementary Figure I:

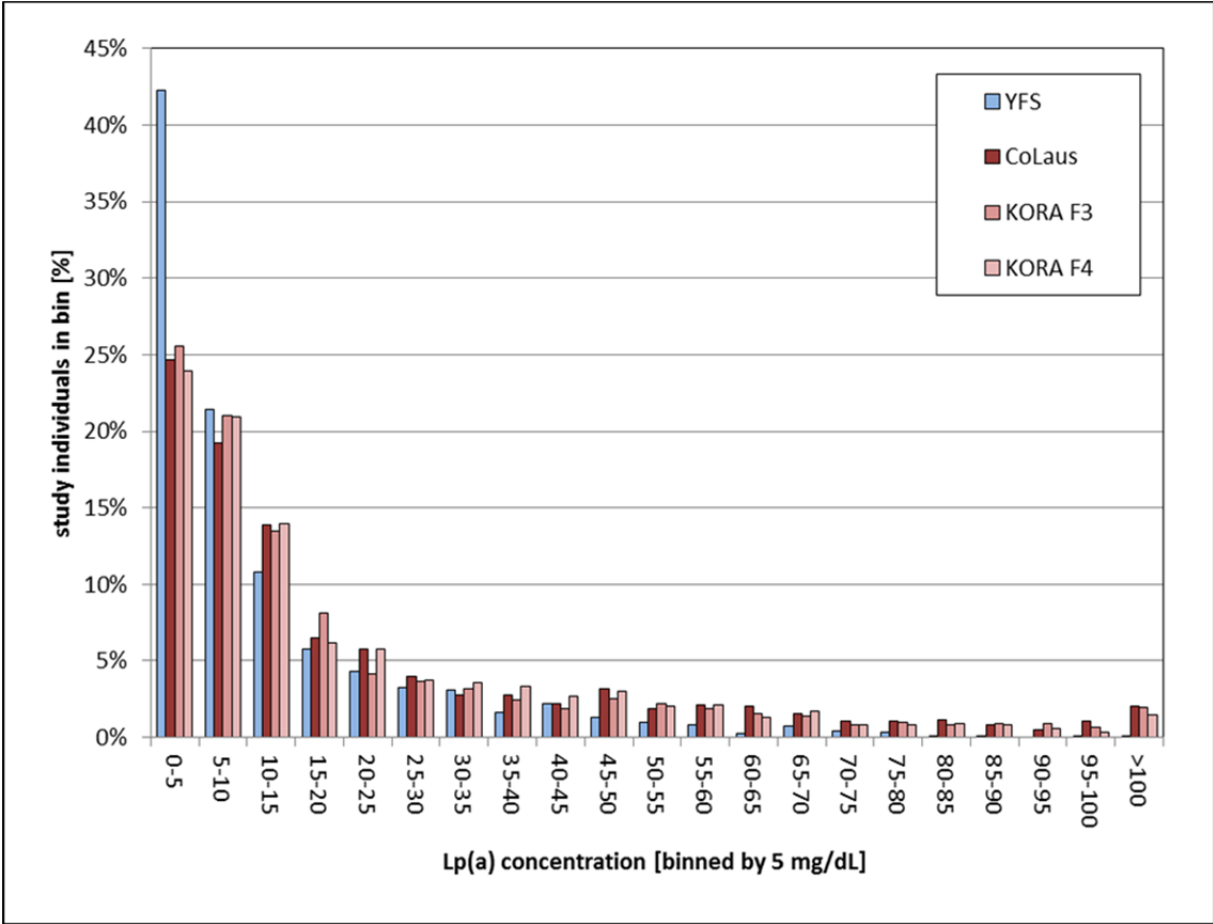
Full boxplots of Lp(a) concentrations for each population.



Supplementary Figure II:

Distribution of Lp(a) concentrations

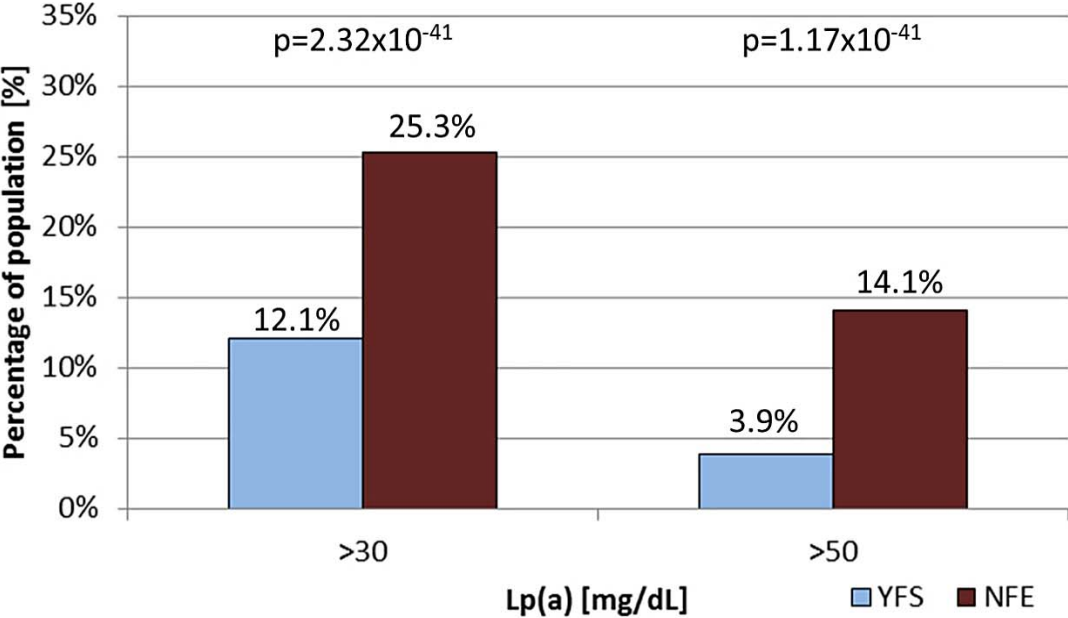
Frequencies of Lp(a) concentrations in studies. Percent of individuals of each study in the bins of 5 mg/dL Lp(a) are shown. An elevated proportion of YFS shows Lp(a) concentrations of 0-5 mg/dL.



Supplementary Figure III:

Percentage of populations with Lp(a) concentrations >30 mg/dL and >50 mg/dL

Percentage of YFS and NFE individuals presenting Lp(a) concentrations increasing CVD risk (>30¹ and >50² mg/dL).



Supplementary Figure IV:

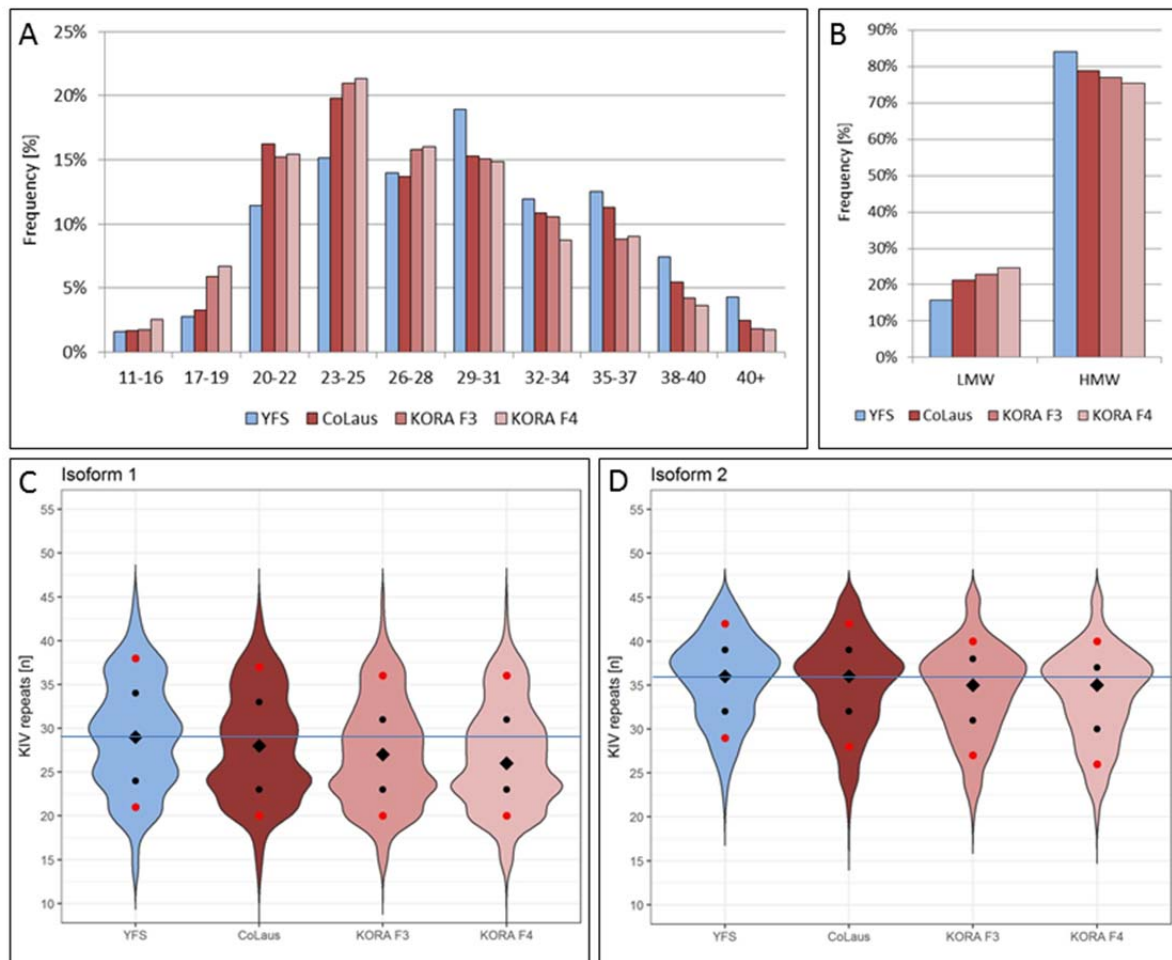
Apo(a) isoform distribution in the populations.

A: Apo(a) isoform 1 distribution in the populations. Panel A shows the apo(a) isoform 1 distribution in the different populations. YFS has markedly higher frequencies of carriers of isoforms containing >28 KIV repeats in isoform 1, so subsequently in YFS is also a higher frequency of probands grouped into HMW, which is shown in panel B. Table with the numbers of probands in each group is given in Figure 3 of the main manuscript.

B: Frequencies LMW vs. HMW. Frequencies of probands in each study grouped into LMW or HMW are shown. YFS show reduced LMW frequency and accordingly a higher HMW frequency compared to the NFE-populations.

C and D: Apo(a) isoform distribution of the two alleles. The Violin plots show distributions of the smaller allele (Isoform 1, panel C) and the larger allele (if detectable on the Western blot, Isoform 2, panel D) in the populations. In both panels the distribution of the apo(a) isoforms is shifted towards larger alleles in YFS compared to the other populations. In the Isoform 2, the samples in the CoLaus study have the same median apo(a) isoform as YFS.

diamonds: median; black dots: 25th, 75th percentile; red dots: 10th and 90th percentiles



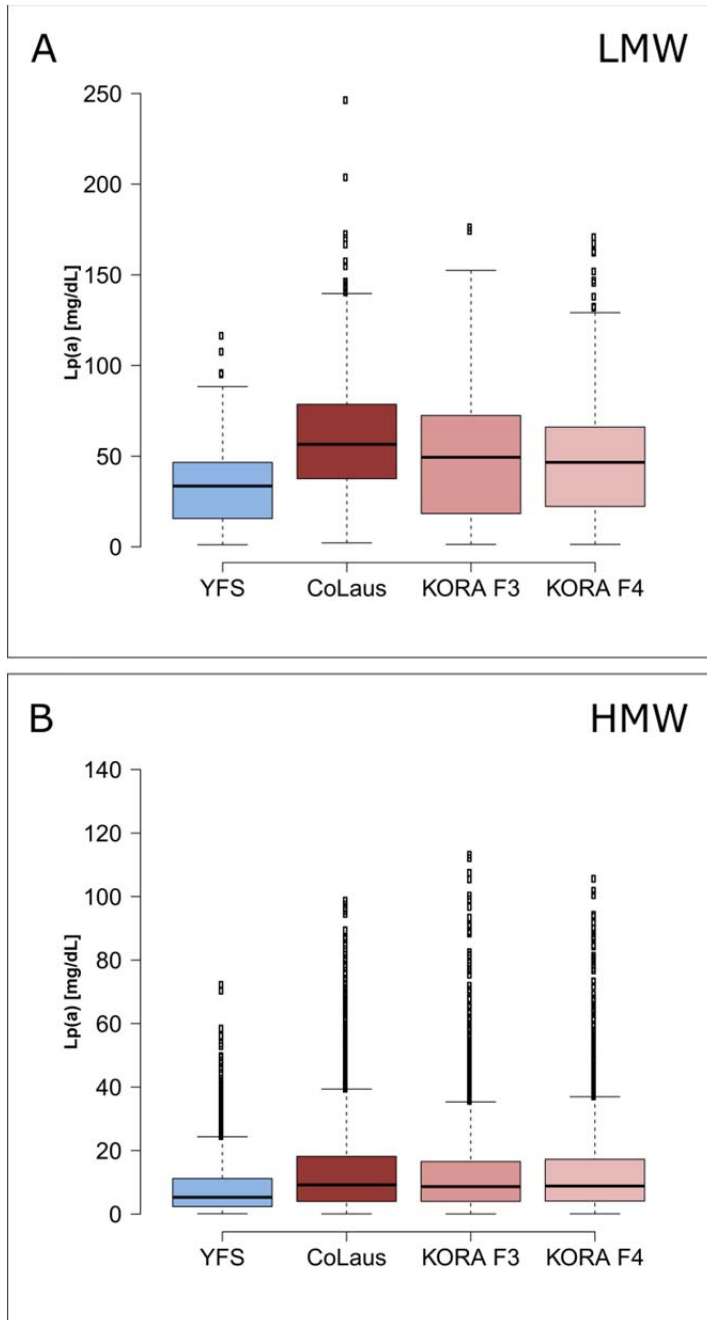
Supplementary Figure V:

Boxplots of Lp(a) concentrations by LMW/HMW status.

A: Lp(a) in LMW carriers. The boxplots of Lp(a) concentrations are shown for all LMW isoform carriers in each population.

B: Lp(a) in HMW carriers. The boxplots of Lp(a) concentrations are shown for all HMW isoform carriers in each population.

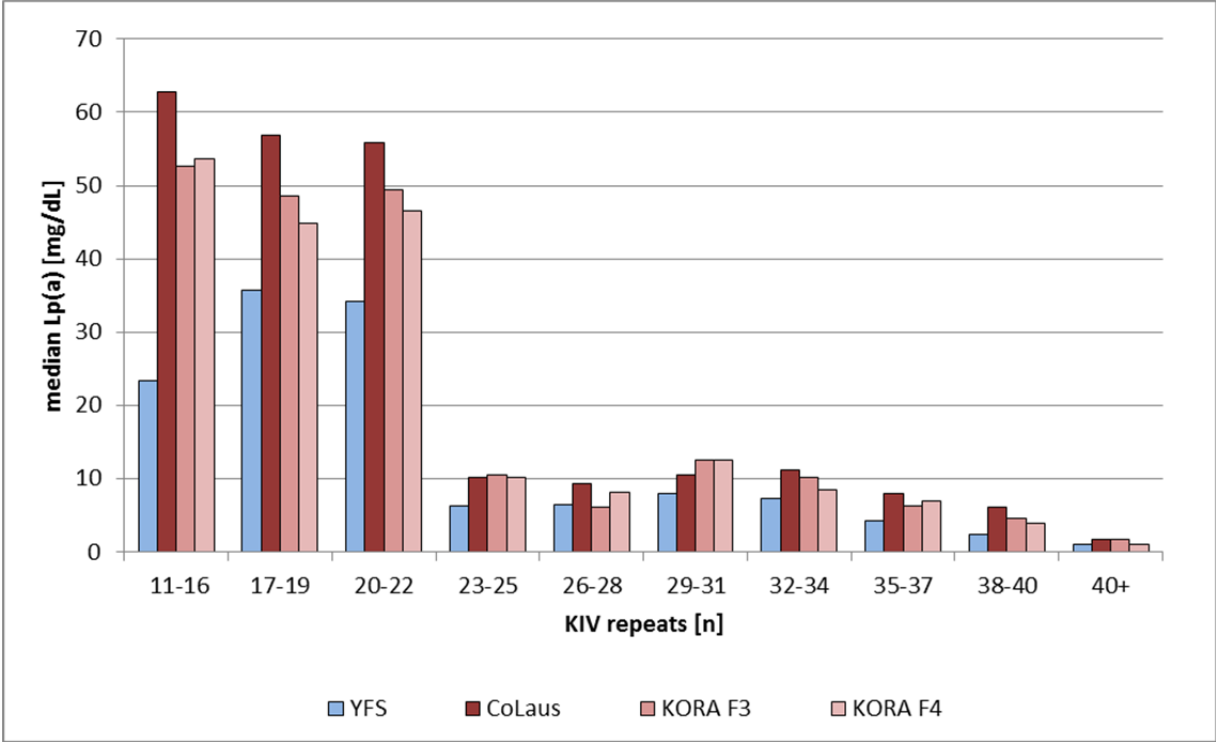
The exact median and quartile values are reported in Table 1 of the main document.



Supplementary Figure VI:

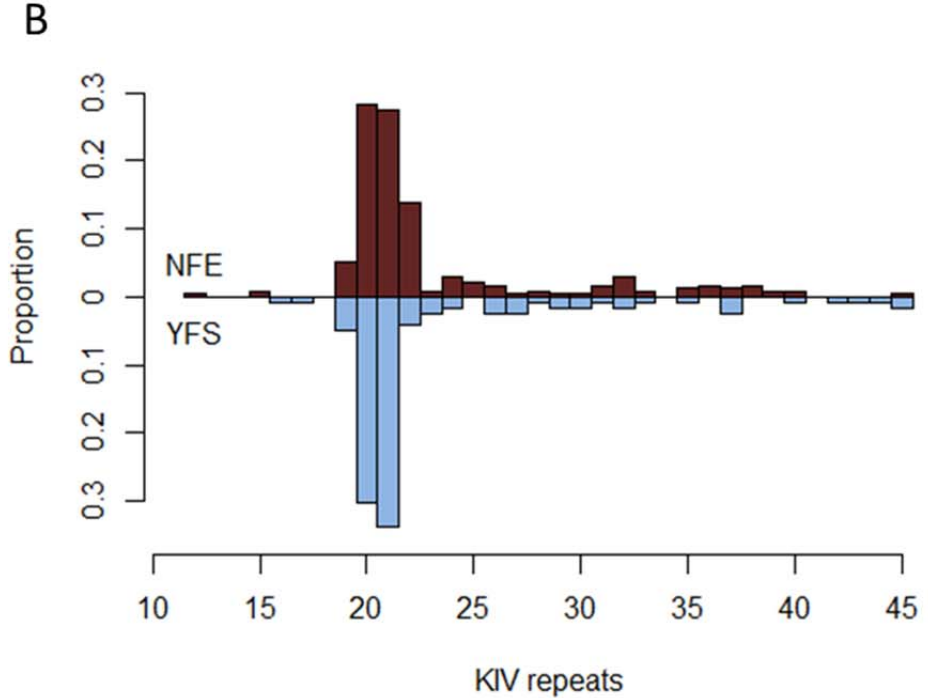
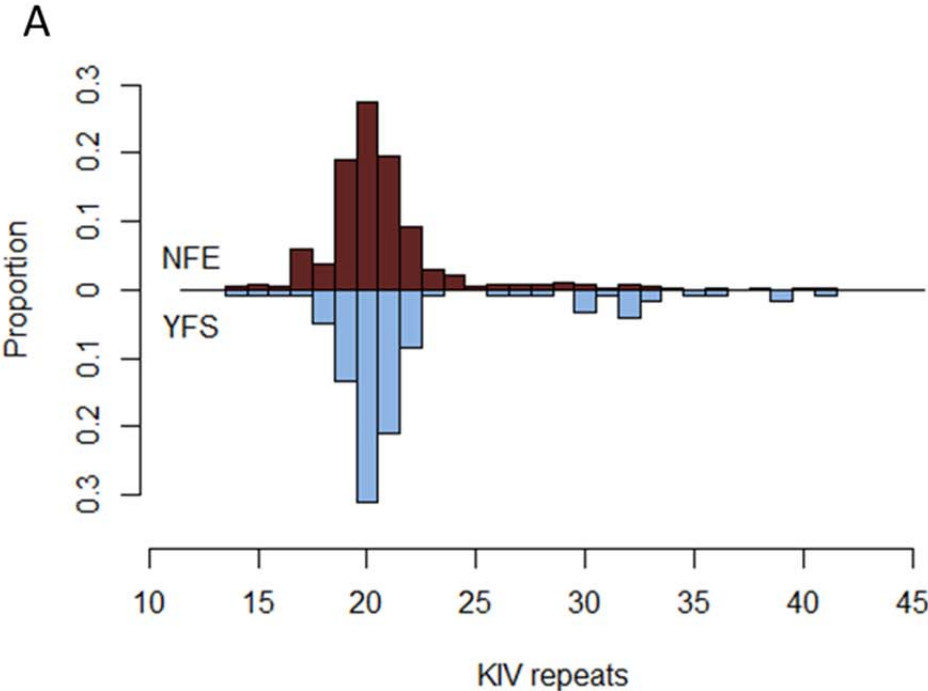
Median Lp(a) concentrations.

Histogram showing the median Lp(a) concentration in each population stratified by KIV repeats. YFS have lower Lp(a) concentrations than the NFE subgroups over the entire isoform range. The number of probands in each group is given in Supplementary Table IV and the exact values are given in Supplementary Table V.



Supplementary Figure VII:

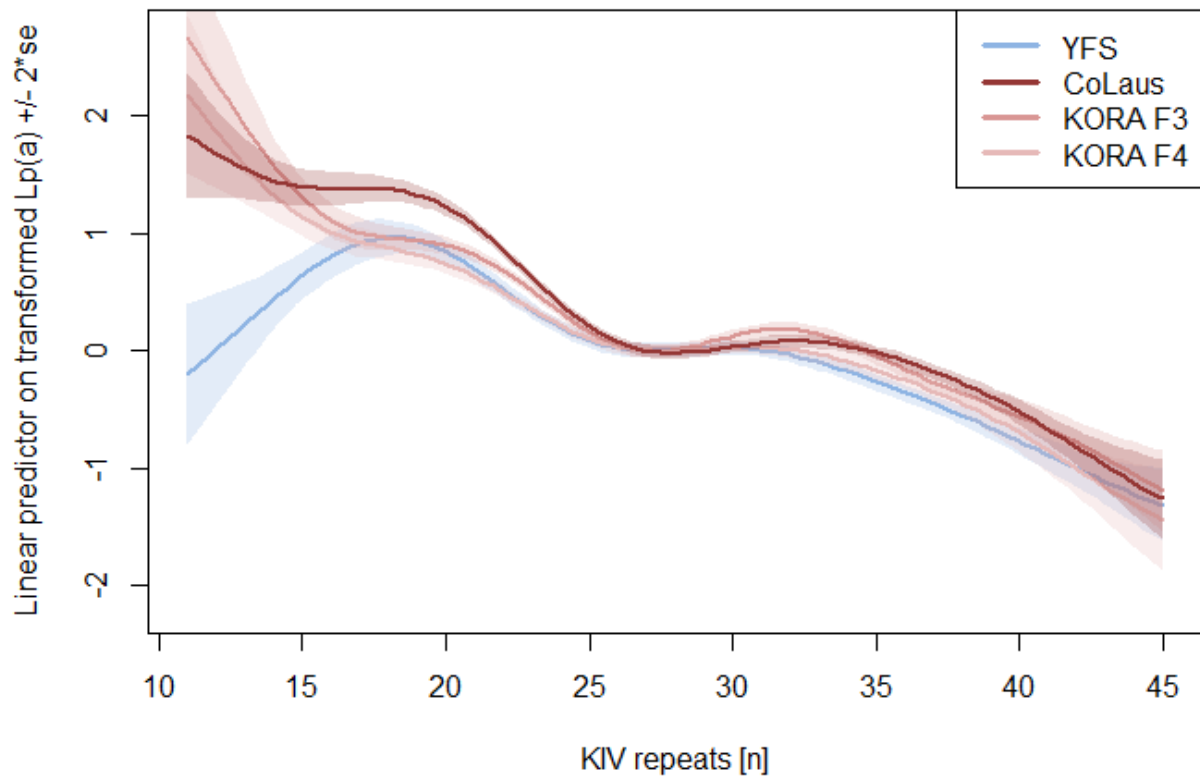
Histograms showing the distribution of rs10455872 (A) and rs3798220 (B) in the different populations. No differences in the distribution were observed ($p=0.20$ and $p=0.61$).The



Supplementary Figure VIII:

Nonlinear Spline

Spline showing the effect of the small isoform (isoform 1) on inverse normal transformed Lp(a) levels. The spline are centered at 0 at isoform 1 = 27 (median on the NFE studies) and is derived from the multiple regression model additionally adjusting for the variables age, sex, eGFR, rs143411368, rs41272114, rs11591147, rs10455872, rs3798220, APOE2 and APOE4-carriers (Supplementary Table VI). se: standard error.



Supplementary Tables

Supplementary Table I:

Lp(a) concentrations [mg/dL] stratified by (a) sex and (b) additionally including only participants in the age range ≥ 32 to ≤ 39 years, where all studies overlap.

Study	25 th percentile	Median	Mean	75 th percentile	Wilcoxon-test p-value
All participants					
Men, YFS (n= 1,026)	2.62	6.09	12.15	14.39	9.66x10 ⁻³⁶
Men, NFE (n= 4,773)	4.46	10.65	20.98	28.16	
Women, YFS (n= 1,255)	2.86	6.56	12.70	15.74	5.61x10 ⁻⁵⁹
Women, NFE (n= 5,230)	5.62	12.64	23.83	32.51	
Age range ≥ 32 to ≤ 39 years.					
Men, YFS (n=528)	2.75	6.42	12.79	15.19	3.083x10 ⁻⁰⁵
Men, NFE (n=311)	4.01	9.58	17.96	21.27	
Women, YFS (n=658)	2.80	6.58	12.67	15.59	1.42x10 ⁻⁰⁷
Women, NFE (n=354)	4.80	10.04	17.89	22.94	

Supplementary Table II:

Lp(a) concentrations in percentiles.

Lp(a) concentration percentile data is shown in YFS and NFE - both grouped and split up in single populations.

Study	Percentiles Lp(a) [mg/dL]						
	5	10	25	50	75	90	95
YFS (n=2,281)	0.96	1.33	2.78	6.32	15.19	33.05	45.7
NFE (n=10,003)	1.57	2.29	5.07	11.59	30.43	60.35	79.84
CoLaus (n=3,998)	1.42	2.17	5.1	11.94	31.51	62.74	81.71
KORA F3 (n=3,089)	1.82	2.52	4.89	11.07	28.4	59.48	81.26
KORA F4 (n=2,916)	1.46	2.13	5.23	11.73	30.6	56.36	73.19

Supplementary Table III:

Apo(a) isoform percentiles of both isoforms

Table shows percentiles of the apo(a) isoform size stratified in the smaller and – if determinable – the larger isoform in YFS and NFE, both grouped and in single populations. In YFS both isoforms are larger than in NFE.

Isoform	Study	Percentiles KIV repeats						
		5	10	25	50	75	90	95
Isoform 1	YFS (n=2,281)	20.0	21.0	24.0	29.0	34.0	38.0	40.0
	NFE (n=10,003)	19.0	20.0	23.0	27.0	32.0	36.0	38.0
	CoLaus (n=3,998)	20.0	20.0	23.0	28.0	33.0	37.0	39.0
	KORA F3 (n=3,089)	19.0	20.0	23.0	27.0	32.0	36.0	38.0
	KORA F4 (n=2,916)	19.0	20.0	23.0	26.0	31.0	36.0	38.0
Isoform 2	YFS (n=2,281)	27.0	29.0	32.0	36.0	39.0	42.0	43.0
	NFE (n=10,003)	25.0	27.0	31.0	35.0	38.0	40.0	42.0
	CoLaus (n=3,998)	25.0	28.0	32.0	36.0	39.0	42.0	43.0
	KORA F3 (n=3,089)	25.0	27.0	31.0	35.0	38.0	40.0	42.0
	KORA F4 (n=2,916)	24.0	26.0	30.0	35.0	37.0	40.0	41.0

Supplementary Table IV:

Numbers of probands in isoform 1 strata of all studies

	Probands in each group [n]									
KIV-repeats (grouped)	11-16	17-19	20-22	23-25	26-28	29-31	32-34	35-37	38-40	40+
YFS (n=2,281)	37	63	261	345	319	431	273	286	169	97
NFE (n=10,003)	194	510	1568	2061	1502	1510	1013	985	455	205
CoLaus	66	131	649	793	548	610	433	450	218	100
KORA F3	54	183	470	647	487	466	325	272	130	55
KORA F4	74	196	449	621	467	434	255	263	107	50

Supplementary Table V:

Median [25%/75% percentiles] Lp(a) concentrations per isoform group

Table shows the median and the 25%/75% percentiles for each apo(a) isoform group.

Lp(a) [mg/dL]	Isoform group									
	11-16	17-19	20-22	23-25	26-28	29-31	32-34	35-37	38-40	40+
YFS (n=2,281)	23.41 [4.45 / 39.80]	35.79 [15.40 / 45.58]	34.15 [20.22 / 50.68]	6.32 [4.18 / 14.08]	6.53 [2.35 / 16.08]	8.02 [2.89 / 14.05]	7.24 [3.59 / 11.10]	4.21 [2.29 / 6.55]	2.49 [1.58 / 4.46]	1.05 [0.69 / 1.83]
NFE (n=10,003)	56.43 [42.17 / 78.23]	48.54 [27.89 / 67.57]	51.23 [25.72 / 73.09]	10.33 [6.30 / 22.68]	8.19 [2.97 / 22.03]	11.62 [3.92 / 20.85]	10.26 [5.30 / 15.96]	7.05 [4.00 / 11.71]	4.91 [2.71 / 8.27]	1.44 [0.71 / 2.82]
CoLaus (n=3,998)	62.76 [48.54 / 88.28]	56.94 [40.93 / 73.15]	55.86 [35.07 / 78.53]	10.20 [6.38 / 23.30]	9.39 [3.17 / 23.56]	10.59 [3.28 / 21.37]	11.27 [5.27 / 18.18]	7.94 [4.19 / 12.61]	6.08 [2.98 / 9.61]	1.68 [0.85 / 3.39]
KORA F3 (n=3,089)	52.62 [41.14 / 70.87]	48.55 [16.42 / 68.00]	49.35 [18.28 / 73.97]	10.50 [6.46 / 22.07]	6.08 [2.94 / 19.62]	12.64 [4.49 / 20.11]	10.19 [5.60 / 15.53]	6.24 [3.78 / 9.70]	4.64 [2.74 / 7.40]	1.75 [0.73 / 2.28]
KORA F4 (n=2,916)	53.66 [34.59 / 70.04]	44.87 [30.76 / 60.85]	46.59 [16.21 / 66.09]	10.16 [6.20 / 23.03]	8.21 [2.79 / 21.88]	12.49 [4.62 / 20.26]	8.54 [4.65 / 14.35]	7.00 [3.95 / 11.69]	4.01 [2.34 / 7.11]	1.00 [0.59 / 1.81]

Supplementary Table VI:

Results of multiple regression analysis in all studies individually. The effect of the isoforms (isoform 1) on Lp(a) and inverse normally transformed Lp(a) levels in all individual studies is shown as nonlinear spline in Figure 5 in the main document.

Variable	Results from multiple regression model on Lp(a)* in							
	YFS		KORA F3		KORA F4		CoLaus	
	β	p	β	p	β	p	β	p
Age	0.07	0.3962	0.13	0.0025	0.1	0.0002	0.14	9.26×10^{-06}
Sex (f)	0.62	0.0368	2.72	2.21×10^{-09}	2.46	2.61×10^{-05}	2.84	5.40×10^{-09}
Isoforms [§]	--	2.60×10^{-124}	--	1.19×10^{-165}	--	1.51×10^{-132}	--	1.10×10^{-242}
eGFR	-0.03	0.1899	-0.04	0.0073	-0.03	0.2565	-0.02	0.4521
rs143411368 [#]	-2.6	5.33×10^{-09}	-3.57	0.3511	-8.19	0.0093	1.29	0.5796
rs41272114 [#]	-2.45	6.38×10^{-19}	-4.4	2.28×10^{-14}	-4.66	5.69×10^{-14}	-6.55	4.17×10^{-32}
rs11591147 [#]	0.53	0.9545	-4.11	0.0056	0.66	0.5601	1.87	0.5118
rs10455872 [#]	11.68	1.33×10^{-13}	24.12	6.68×10^{-39}	20.98	1.17×10^{-28}	13.14	8.59×10^{-17}
rs3798220 [#]	28.03	3.37×10^{-14}	41.13	2.42×10^{-26}	43.83	4.44×10^{-31}	44.3	6.95×10^{-35}
APOE2 carrier	-2.34	0.0071	-5.02	3.80×10^{-11}	-2.19	0.0026	-3.01	8.44×10^{-06}
APOE4 carrier	-0.21	0.6147	-1.96	0.0100	0.48	0.4403	0.06	0.8923

* β estimates are derived from models on Lp(a); p-values are derived from models on inverse normal transformed Lp(a) values to ensure normal distribution assumption

[§] p-values for isoforms refers to the significance of a nonlinear spline of isoform 1 on inverse-normal transformed Lp(a); for effect size: see nonlinear spline in Figure 4 and Supplementary Figure VIII

[#] For SNPs, an additive inheritance model was assumed

Supplementary Table VII:

Mediation analysis results on inverse normal transformed Lp(a) including only those variables, which qualify for mediation analysis

Variable	YFS		Identified as mediator	Relative effects	
	P-value 1	P-value 2		estimate §	95 % CI*
Explained effects[†]				0.718	[0.634, 0.806]
Age	<0.001	<0.001	✓	0.264	[0.185, 0.348]
Sex (f)	<0.001	0.019	✓	-0.006	[-0.013, -0.001]
Apo(a) isoforms	<0.001	<0.001	✓	0.273	[0.235, 0.310]
eGFR	<0.001	<0.001	✓	0.104	[0.052, 0.157]
rs143411368	<0.001	<0.001	✓	0.038	[0.024, 0.055]
rs41272114	<0.001	<0.001	✓	0.040	[0.024, 0.054]
rs11591147	0.601	--	--	--	--
rs10455872	<0.001	0.020	✓	0.025	[0.010, 0.039]
rs3798220	<0.001	0.164	--	--	--
APOE2 carrier	<0.001	<0.001	✓	-0.018	[-0.025, -0.011]
APOE4 carrier	0.103	<0.001	--	--	--
Unexplained effect[‡]				0.282	[0.194,0.366]

P-value 1: Type-3 tests in the full model, testing, whether factors are associated with Lp(a); P-value 2: Tests of relationship of each factor with the predictor (YFS/NFE-difference)

*95%CI were derived with bootstrap samples assuming normal approximation

[†] "indirect effect" of the mediation analysis

[‡] "direct effect" of the mediation analysis

§ estimates rounded, may not add up to exactly to 1

Supplementary Table VIII:

Imputation details

This table shows the imputation details and quality of the populations in this study.

	YFS	CoLaus	KORA F3	KORA F4
Microarray used	Illumina 670k	Affymetrix GeneChip® Human Mapping 500K	Illumina Omni 2.5/ Illumina Omni Express	Affymetrix Axiom
QC before imputation	callrate >95%, p-value HWE > 10^{-6} , MAF>1%	callrate >90%*, p-value HWE > 10^{-7}	callrate >98%, p-value HWE > 10^{-10}	callrate >98%, p-value HWE > 10^{-10}
Program used for imputation	IMPUTE	minimac	IMPUTE	IMPUTE
Reference panel for imputation	1000G phase1v3 ³	1000G phase1v3 ³	HRC panel ⁴	HRC panel ⁴

References

1. Kamstrup PR, Tybjaerg-Hansen A, Steffensen R, Nordestgaard BG. Genetically elevated lipoprotein(a) and increased risk of myocardial infarction. *JAMA*. 2009;301:2331-2339.
2. Nordestgaard BG, Chapman MJ, Ray K, et al. Lipoprotein(a) as a cardiovascular risk factor: current status. *Eur Heart J*. 2010;31:2844-2853.
3. Auton A, Abecasis GR, Altshuler DM, et al. A global reference for human genetic variation. *Nature*. 2015;526:68-74.
4. Loh P-R, Danecek P, Palamara PF, et al. Reference-based phasing using the Haplotype Reference Consortium panel. *Nat Genet*. 2016;48:1443-1448.