Multi-ancestry study of blood lipid levels identifies four loci interacting with physical activity

Tuomas O. Kilpeläinen et al.*

Many genetic loci affect circulating lipid levels, but it remains unknown whether lifestyle factors, such as physical activity, modify these genetic effects. To identify lipid loci interacting with physical activity, we performed genome-wide analyses of circulating HDL cholesterol, LDL cholesterol, and triglyceride levels in up to 120,979 individuals of European, African, Asian, Hispanic, and Brazilian ancestry, with follow-up of suggestive associations in an additional 131,012 individuals. We find four loci, in/near CLASP1, LHX1, SNTA1, and CNTNAP2, that are associated with circulating lipid levels through interaction with physical activity; higher levels of physical activity enhance the HDL cholesterol-increasing effects of the CLASP1, LHX1, and SNTA1 loci and attenuate the LDL cholesterol-increasing effect of the CNTNAP2 locus. The CLASP1, LHX1, and SNTA1 regions harbor genes linked to muscle function and lipid metabolism. Our results elucidate the role of physical activity interactions in the genetic contribution to blood lipid levels.
Circulating levels of blood lipids are strongly linked to the risk of atherosclerotic cardiovascular disease. Regular physical activity (PA) improves blood lipid profile by increasing the levels of high-density lipoprotein cholesterol (HDL-C) and decreasing the levels of low-density lipoprotein cholesterol (LDL-C) and triglycerides (TG)\(^1\). However, there is individual variation in the response of blood lipids to PA, and twin studies suggest that some of this variation may be due to genetic differences\(^2\). The genes responsible for this variability remain unknown.

More than 500 genetic loci have been found to be associated with blood levels of HDL-C, LDL-C, or TG in published genome-wide association studies (GWAS)\(^3\)–\(^12\). At present, it is not known whether any of these main effect associations are modiﬁed by PA. Understanding whether the impact of lipid loci can be modiﬁed by PA is important because it may give additional insight into biological mechanisms and identify subpopulations in whom PA is particularly beneﬁcial.

Here, we report results from a genome-wide meta-analysis of gene–PA interactions on blood lipid levels in up to 120,979 adults of European, African, Asian, Hispanic, or Brazilian ancestry, with follow-up of suggestive associations in an additional 131,012 individuals. We show that four loci, in/near \(\text{CLASP1}, \text{LHX1}, \text{SNTA1}\), and \(\text{CNPNAZ2}\), are associated with circulating lipid levels through interaction with PA. None of these four loci have been identiﬁed in published main effect GWAS of lipid levels. The \(\text{CLASP1}, \text{LHX1}, \text{and SNTA1}\) regions harbor genes linked to muscle function and lipid metabolism. Our results elucidate the role of PA interactions in the genetic contribution to blood lipid levels.

Results

**Genome-wide interaction analyses in up to 250,564 individuals.** We assessed effects of gene–PA interactions on serum HDL-C, LDL-C, and TG levels in 86 cohorts participating in the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Gene-Lifestyle Interactions Working Group\(^13\). PA was harmonized across participating studies by categorizing it into a dichotomous variable. The participants were deﬁned as inactive if their reported weekly energy expenditure in moderate- to vigorous-intensity leisure-time or commuting PA was less than 225 metabolic equivalent (MET) minutes per week (corresponding to approximately 1 h of moderate-intensity PA), while all other participants were deﬁned as physically active (Supplementary Data 1).

The analyses were performed in two stages. Stage 1 consisted of genome-wide meta-analyses of linear regression results from 42 cohorts, including 120,979 individuals of European \([n=84,902]\), African \([n=20,487]\), Asian \([n=6,403]\), Hispanic \([n=47,491]\), or Brazilian \([n=44,438]\) ancestry (Supplementary Tables 1 and 2; Supplementary Data 2; Supplementary Note 1). All variants that reached two-sided \(P<1 \times 10^{-6}\) in the Stage 1 multi-ancestry meta-analyses or ancestry-specific meta-analyses were taken forward to linear regression analyses in Stage 2, which included 44 cohorts and 131,012 individuals of European \([n=107,617]\), African \([n=5384]\), Asian \([n=6590]\), or Hispanic \([n=11,421]\) ancestry (Supplementary Tables 3 and 4; Supplementary Data 3; Supplementary Note 2). The summary statistics from Stage 1 and Stage 2 were subsequently meta-analyzed to identify lipid loci whose effects are modiﬁed by PA.

We identiﬁed lipid loci interacting with PA by three different approaches applied to the meta-analysis of Stage 1 and Stage 2: (i) we screened for genome-wide signiﬁcant SNP × PA-interaction effects \((P_{\text{INT}}<5 \times 10^{-8})\); (ii) we screened for genome-wide signiﬁcant 2 degree of freedom (2df) joint test of SNP main effect and SNP × PA interaction\(^14\) \((P_{\text{JOINT}}<5 \times 10^{-8})\); and (iii) we screened all previously known lipid loci\(^3\)–\(^12\) for signiﬁcant SNP × PA-interaction effects, Bonferroni-correcting for the number of independent variants tested \((r^2<0.1\) within 1 Mb distance; \(P_{\text{INT}}=0.05/501=1.0 \times 10^{-3}\)).

PA modiﬁes the effect of four loci on lipid levels. Three novel loci (>1 Mb distance and \(r^2<0.1\) with any previously identiﬁed lipid locus) were identiﬁed: in \(\text{CLASP1}\) \((rs2862183, P_{\text{INT}}=8 \times 10^{-9})\), near \(\text{LHX1}\) \((rs295849, P_{\text{INT}}=3 \times 10^{-8})\), and in \(\text{SNTA1}\) \((rs141588480, P_{\text{INT}}=2 \times 10^{-6})\), which showed a genome-wide signiﬁcant SNP × PA interaction on HDL-C in all ancestries combined (Table 1, Figs. 1–4). Higher levels of PA enhanced the HDL cholesterol-increasing effects of the \(\text{CLASP1}, \text{LHX1}, \text{and SNTA1}\) loci. A novel locus in \(\text{CNPNAZ2}\) \((rs190748049)\) was genome-wide signiﬁcant in the joint test of SNP main effect and SNP × PA interaction \((P_{\text{JOINT}}=4 \times 10^{-4})\) and showed moderate evidence of SNP × PA interaction \((P_{\text{INT}}=2 \times 10^{-6})\) in the meta-analysis of LDL-C in all ancestries combined (Table 1, Fig. 5). The LDL-C-increasing effect of the \(\text{CNPNAZ2}\) locus was attenuated in the physically active group as compared to the inactive group. None of these four loci have been identiﬁed in previous main effect GWAS of lipid levels.

No interaction between known main effect lipid loci and PA. Of the previously known 260 main effect loci for HDL-C, 202 for LDL-C, and 185 for TG\(^3\)–\(^12\), none reached the Bonferroni-corrected threshold (two-sided \(P_{\text{INT}}=1.0 \times 10^{-4}\) for SNP × PA interaction alone (Supplementary Data 4-6). We also found no signiﬁcant interaction between a combined score of all published European-ancestry loci for HDL-C, LDL-C, or TG with PA (Supplementary Data 7–9) using our European-ancestry summary results (two-sided \(P_{\text{LDL-C}}=0.14, P_{\text{LDL-C}}=0.77,\) and \(P_{\text{TG}}=0.86\), respectively), suggesting that the beneﬁcial effect of PA on lipid levels may be independent of genetic risk\(^15\).

Potential functional roles of the loci interacting with PA. While the mechanisms underlying the beneﬁcial effect of PA on circulating lipid levels are not fully understood, it is thought that the changes in plasma lipid levels are primarily due to an improve-ment in the ability of skeletal muscle to utilize lipids for energy due to enhanced enzymatic activities in the muscle\(^16,17\). Of the four loci we found to interact with PA, three, in \(\text{CLASP1}, \text{near LHX1}, \text{and SNTA1}\), harbor genes that may play a role in muscle function\(^18,19\) and lipid metabolism\(^20,21\).

The lead variant \(rs2862183\) (minor allele frequency (MAF) 22%) in the \(\text{CLASP1}\) locus which interacts with PA on HDL-C levels is an intronic SNP in \(\text{CLASP1}\) that encodes a microtubule-associated protein (Fig. 2). The \(rs2862183\) SNP is associated with \(\text{CLASP1}\) expression in \(\text{esophagus muscularis}\) \((P=3 \times 10^{-7})\) and is in strong linkage disequilibrium \((r^2>0.79)\) with \(rs13403769\) variant that shows the strongest association with \(\text{CLASP1}\) expression in the region \((P=7 \times 10^{-7})\). Another potent causal candidate gene in this locus is the nearby \(\text{GLI2}\) gene which has been found to play a role in skeletal myogenesis\(^18\) and the conversion of glucose to lipids in mouse adipose tissue\(^20\) by inhibiting hedgehog signaling.

The \(rs295849\) (MAF 38%) variant near \(\text{LHX1}\) interacts with PA on HDL-C levels. However, the more likely causal candidate gene in this locus is acetyl-CoA carboxylase (\(\text{ACACA}\)), which plays a crucial role in fatty acid metabolism\(^21\) (Fig. 3). Rare acetyl-CoA carboxylase deficiency has been linked to hypotonic myopathy, severe brain damage, and poor growth\(^22\).

The lead variant in the \(\text{SNTA1}\) locus \((rs141588480)\) interacts with PA on HDL-C and is an insertion only found in individuals
of African (MAF 6%) or Hispanic (MAF 1%) ancestry. The rs141588480 insertion is in the SNTA1 gene that encodes the syntrophin alpha 1 protein, located at the neuromuscular junction and altering intracellular calcium ion levels in muscle tissue (Fig. 4). Snta1-null mice exhibit differences in muscle regeneration after a cardiotoxin injection19. Two weeks following the injection into mouse tibialis anterior, the muscle showed decreased phasic inhibition and a decreased number of interneurons24. Knockout are used as an animal model of autism and show altered synaptosome morphology, dendritic spine density, and brain connectivity25.

The fourth locus interacting with PA is CNTNAP2, with the lead variant (rs190748049) intronic and no other genes nearby (Fig. 5). The rs190748049 variant is most common in African-ancestry (MAF 8%), less frequent in European-ancestry (MAF 2%), and absent in Asian- and Hispanic-ancestry populations. The protein coded by the CNTNAP2 gene, contactin-associated protein like-2, is a member of the neurolin protein family. The protein is located at the juxtaparanodes of myelinated axons where it may have an important role in the differentiation of the axon into specific functional subdomains. Mice with a Cntnap2 knockout are used as an animal model of autism and show altered phasic inhibition and a decreased number of interneurons26. Human CNTNAP2 variants have been associated with risk of autism and related behavioral disorders27.

### Joint test of SNP main effect and SNP × PA interaction

We found 101 additional loci that reached genome-wide significance in the 2df joint test of SNP main effect and SNP × PA interaction on HDL-C, LDL-C, or TG. However, none of these loci showed evidence of SNP × PA interaction (P_{INT} > 0.001) (Supplementary Data 10). All 101 main effect-driven loci have been identified in previous GWAS of lipid levels3–12.

### Discussion

In this genome-wide study of up to 250,564 adults from diverse ancestries, we found evidence of interaction with PA for four loci, in/near CLASP1, LHX1, and SNTA1, and CNTNAP2. Higher levels of PA enhanced the HDL cholesterol-increasing effects of CLASP1, LHX1, and SNTA1 loci and attenuated the LDL cholesterol-increasing effect of the CNTNAP2 locus. None of these four loci have been identified in previous main effect GWAS for lipid levels3–12.

The loci in/near CLASP1, LHX1, and SNTA1 harbor genes linked to muscle function18,19 and lipid metabolism20,21. More specifically, the GLI2 gene within the CLASP1 locus has been found to play a role in myogenesis18 as well as in the conversion of glucose to lipids in adipose tissue20; the ACACA gene within the LHX1 locus plays a crucial role in fatty acid metabolism21 and has been connected to hypotonic myopathy22; and the SNTA1 gene is linked to muscle regeneration19. These functions may relate to differences in the ability of skeletal muscle to use lipids as an energy source, which may modify the benefit of PA on blood lipid levels16,17.

The inclusion of diverse ancestries in the present meta-analyses allowed us to identify two loci that would have been missed in meta-analyses of European-ancestry individuals alone. In particular, the lead variant (rs141588480) in the SNTA1 locus is only polymorphic in African and Hispanic ancestries, and the lead
variant (rs190748049) in the CNTNAP2 locus is four times more frequent in African-ancestry than in European-ancestry. Our findings highlight the importance of multi-ancestry investigations of gene-lifestyle interactions to identify novel loci.

We did not find additional novel loci when jointly testing for SNP main effect and interaction with PA. While 101 loci reached genome-wide significance in the joint test on HDL-C, LDL-C, or TG, all of these loci have been identified in previous GWAS of lipid levels3–12, and none of them showed evidence of SNP × PA interaction. The 2df joint test bolsters the power to detect novel loci when both main and an interaction effect are present14. The lack of novel loci identified by the 2df test suggests that the loci

---

**Fig. 2** Interaction of rs2862183 in CLASP1 with physical activity on HDL cholesterol levels. The beta and 95% confidence intervals in the forest plot (a) is shown for the rs2862183 × physical activity interaction term, i.e., it indicates the increase in logarithmically transformed HDL cholesterol levels in the active group as compared to the inactive group per each T allele of rs2862183. The $-\log_{10}(P_{\text{value}})$ in the association plot (b) is also shown for the rs2862183 × physical activity interaction term. The P values are two-sided and were obtained by a meta-analysis of linear regression model results. The figure was generated using LocusZoom (http://locuszoom.org)

---

**Fig. 3** Interaction of rs295849 near LHX1 with physical activity on HDL cholesterol levels. The beta and 95% confidence intervals in the forest plot (a) is shown for the rs295849 × physical activity interaction term, i.e., it indicates the increase in logarithmically transformed HDL cholesterol levels in the active group as compared to the inactive group per each G allele of rs295849. The $-\log_{10}(P_{\text{value}})$ in the association plot (b) is also shown for the rs295849 × physical activity interaction term. The P values are two-sided and were obtained by a meta-analysis of linear regression model results. The figure was generated using LocusZoom (http://locuszoom.org)
showing the strongest SNP × PA interaction on lipid levels are not the same loci that show a strong main effect on lipid levels.

In summary, we identified four loci containing SNPs that enhance the beneficial effect of PA on lipid levels. The identification of the SNTA1 and CNTNAP2 loci interacting with PA was made possible by the inclusion of diverse ancestries in the analyses. The gene regions that harbor loci interacting with PA involve pathways targeting muscle function and lipid metabolism. Our findings elucidate the role and underlying mechanisms of PA interactions in the genetic regulation of blood lipid levels.
Methods

Study design. The present study collected summary data from 86 participating cohorts and no individual-level data were exchanged. For each of the participating cohorts, the appropriate ethics review board approved the data collection and all participants provided informed consent.

We included men and women 18–80 years of age and of European, African, Asian, Hispanic, or Brazilian ancestry. The meta-analyses were performed in two stages. Stage 1 meta-analyses included 42 studies with a total of 120,979 individuals of European (n = 84,902), African (n = 20,487), Asian (n = 6,603), Hispanic (n = 4,749), or Brazilian ancestry (n = 4,438) [Supplementary Table 1; Supplementary Data 2; Supplementary Note 1]. Stage 2 meta-analyses included 44 studies with a total of 131,012 individuals of European (n = 107,617), African (n = 53,843), Asian (n = 6,603), or Hispanic (n = 11,421) ancestry [Supplementary Table 3; Supplementary Data 3; Supplementary Note 2]. Studies participating in Stage 1 meta-analyses carried out genome-wide analyses, whereas studies participating in Stage 2 only performed analyses for 17,711 variants that reached \( P < 10^{-8} \) in the Stage 1 meta-analyses and were observed in at least two different Stage 1 studies with a pooled sample size \( >4,000 \). The Stage 1 and Stage 2 meta-analyses were performed in all ancestries combined and in each ancestry separately.

Outcome traits: LDL-C, HDL-C, and TG. The levels of LDL-C were either directly assayed or derived using the Friedewald equation (if TG \( \leq 400 \text{mgdl}^{-1} \)). We adjusted LDL-C levels for lipid-lowering drug use if statin use was reported or if unspecified lipid-lowering drug use was listed after 1994, when statin use became common. For directly assayed LDL-C, we divided the LDL-C value by 0.7. If LDL-C was derived using the Friedewald equation, we first adjusted total cholesterol for statin use (total cholesterol divided by 0.8) before the usual calculation. Inactive individuals were defined as individuals who were nonfasting, we did not include either TG or calculated LDL-C in the present analyses. The HDL-C and TG variables were natural log-transformed, while LDL-C was not transformed.

PA variable. The participating studies used a variety of ways to assess and quantify PA (Supplementary Data 1). To harmonize the PA variable across all participating studies, we coded a dichotomous variable, inactive vs. active, that could be applied in a relatively uniform way in all studies, and that would be congruent with previous findings on SNP \( \times \) PA interactions and the relationship between PA and disease outcomes. Inactive individuals were defined as those with \( \leq 225 \text{MET-min} \) per week of moderate-to-vigorous leisure-time or commuting PA (n = 84,495; 34% of all participants) [Supplementary Data 1]. We considered all other participants as physically active. In studies where MET-min per week was not measured, the level of PA was not available, we defined inactive individuals as those engaging in \( \leq 1 \text{h/week} \) of moderate-intensity leisure-time PA or commuting PA. In studies with PA measures that were not comparable to either MET-min or hours/week of PA, we defined the inactive group using a percentage cut-off, where individuals in the lowest 25% of PA levels were defined as inactive and all other individuals as active.

Genotyping and imputation. Genotyping was performed by each participating study using Illumina or Affymetrix arrays. Imputation was conducted on the cosmopolitan reference panel from the 1000 Genomes Project Phase 1 Integrated Release version 3 Haploview (2010–2011 haplotypes). Only autosomal variants were considered. Specific details of each participating study’s genotyping platform and imputation software are described in Supplementary Tables 2 and 4.

Quality control. The participating studies excluded variants with MAF < 1%. We performed QC for all study-specific results using the EasyQC package in R\(^{48}\). For each study, we generated individual QC plots, which included quality control filtering criteria defined in previous publications. We performed individual QC plots. In each study, we identified quality-control (QC) criteria (e.g., genotype calling rate, minor allele frequency, and Hardy–Weinberg equilibrium) and performed the meta-analysis.

Analysis methods. All participating studies used the following model to test for interaction:

\[
E[Y] = \beta_0 + \beta_F \times PA + \beta_L \times G + \beta_{INT} \times G \times PA + \beta_G \times C,
\]

where Y is the HDL-C, LDL-C, or TG value, PA is the PA variable with 0 or 1 coding for active or inactive group, and G is the dosage of the imputed genetic variant coded additively from 0 to 2. The C is the vector of covariates which included age, sex, study center (for multi-center studies), and genome-wide principal components. From this model, the studies provided the estimated genetic main effect (\( \beta_L \)), estimated interaction effect (\( \beta_{INT} \)), and a robust estimate of the covariance between \( \beta_L \) and \( \beta_{INT} \). Using these estimates, we performed inverse variance-weighted meta-analyses for the SNP \( \times \) PA interaction term alone, and 2df joint meta-analyses of the SNP effect and SNP \( \times \) PA interaction combined by the method of Mannling et al.13 using the METAL meta-analysis software. We applied genomic control correction twice in Stage 1, first for study-specific GWAS results and again for meta-analysis results, whereas genomic control correction was not applied to the Stage 2 results as interaction testing was only performed at select variants.

Combined PA-interaction effect of all known lipid loci. To identify all published SNPs associated with HDL-C, LDL-C, or TG, we extended the previous curated list of lipid loci by Davis et al.4 by searching PubMed and Google Scholar databases and screening the GWAS Catalog. After LD pruning by \( r^2 < 0.1 \) in the 1000 Genomes European-ancestry reference panel, 260 independent loci remained associated with HDL cholesterol, 202 with LDL cholesterol, and 185 with TG [Supplementary Data 7–9]. To approximate the combined PA interaction of all known European-ancestry loci associated with HDL-C, LDL-C, or TG, we calculated their combined interaction effect as the weighted sum of the individual SNP coefficients in our genome-wide summary results for European-ancestry. This approach has been described previously in detail by Dastani et al.31 and incorporated in the package “gtx” in R. We did not weigh the loci by their main effect estimates from the discovery GWAS data.

Data availability. The meta-analysis summary results are available for download on the CHARGE dbGaP website under accession phs009390.

References

1. Leon, A. S. & Sanchez, O. A. Response of blood lipids to exercise training alone or combined with dietary intervention. Med. Sci. Sports Exerc. 33, S502–S515 (2001). discussion S528–S529.
2. Lakka, H. M., Tremblay, A., Despres, J. P. & Bouchard, C. Effects of long-term negative energy balance with exercise on plasma lipid and lipoprotein levels in identical twins. Atherosclerosis 172, 127–133 (2004).
3. Below, J. E. et al. Meta-analysis of lipid-trait in Hispanics identifies novel loci, population-specific effects, and tissue-specific enrichment of eQTLs. Sci. Rep. 9, 19429 (2019).
4. Davis, J. P. et al. Common, low-frequency, and rare genetic variants associated with lipoprotein subclasses and triglyceride measures in Finnish men from the METSIM study. Sci. Rep. 7, 40423 (2017).
5. Kanai, M. et al. Genetic analysis of quantitative traits in the Japanese population links cell types to complex human diseases. Nat. Genet. 48, 390–400 (2018).
6. Liu, D. J. et al. Exome-wide association study of plasma lipids in >300,000 individuals. Nat. Genet. 49, 1758–1766 (2017).
7. Lu, X. et al. Genetic susceptibility to lipid levels and lipid change over time and risk of incident hyperlipidemia in Chinese populations. Circ Cardiovasc Genet. 9, 37–44 (2016).
8. Lu, X. et al. Exome chip meta-analysis identifies novel loci and East Asian-specific coding variants that contribute to lipid levels and coronary artery disease. Nat. Genet. 49, 1722–1730 (2017).
9. Nagy, R. et al. Exploration of haplotype research consortium imputation for genome-wide association studies in 20,032 Generation Scotland participants. Genome Med. 9, 23 (2017).

Received: 6 June 2018 Accepted: 7 December 2018
Published online: 22 January 2019
10. Southam, L. et al. Whole genome sequencing and imputation in isolated populations identify genetic associations with medically-relevant complex traits. *Nat. Commun.* 8, 15606 (2017).

11. Spracklen, C. N. et al. Association analyses of East Asian individuals and trans-ancestry analyses with European individuals reveal new loci associated with cholesterol and triglyceride levels. *Hum. Mol. Genet.* 26, 1770–1784 (2017).

12. van Leeuwen, E. M. et al. Meta-analysis of 49,549 individuals imputed with the 1000 Genomes Project reveals an exonic damaging variant in ANGPTL4 determining fasting TG levels. *J. Med. Genet.* 53, 441–449 (2016).

13. Rao, D. C. et al. Multiancestry study of gene-lifestyle interactions for cardiovascular traits in 610 475 individuals from 124 cohorts: design and rationale. *Circ. Cardiovasc. Genet.* 10, e001649 (2017).

14. Manning, A. K. et al. Meta-analysis of gene-environment interaction: joint estimation of SNP and SNP x environment regression coefficients. *Genet. Epidemiol.* 35, 11–18 (2011).

15. Khera, A. V. et al. Genetic risk, adherence to a healthy lifestyle, and coronary disease. *N. Engl. J. Med.* 375, 2349–2358 (2016).

16. Harrison, M. et al. Lipoprotein particle distribution and skeletal muscle lipoprotein lipase activity after acute exercise. *Lipids Health Dis.* 11, 64 (2012).

17. Riedl, I. et al. Regulation of skeletal muscle transcriptome in elderly men after 6 weeks of endurance training at lactate threshold intensity. *Exp. Gerontol.* 45, 896–903 (2010).

18. McDermott, A. et al. Gli2 and Gli3 have redundant and context-dependent functions in skeletal muscle formation. *Development* 132, 345–357 (2005).

19. Hosaka, Y. et al. Alpha1-syntrophin-deficient skeletal muscle exhibits hypertrophy and aberrant formation of neuromuscular junctions during regeneration. *J. Cell Biol.* 158, 1097–1102 (2002).

20. Shi, Y. & Long, F. Hedgehog signaling via Gli2 prevents obesity induced by high-fat diet in adult mice. *Elife* 6, e31649 (2017).

21. Tong, L. Acetyl-CoA carboxylase: crucial metabolic enzyme and target for drug discovery. *Cell. Mol. Life Sci.* 62, 1784–1803 (2005).

22. Blom, W., de Meuinc Keizer, S. M. & Scholte, H. R. Acetyl-CoA carboxylase deficiency: an inborn error of de novo fatty acid synthesis. *N. Engl. J. Med.* 305, 466–473 (1981).

23. Wu, G. et al. Alpha-1-syntrophin mutation and the long-QT syndrome: a review of 221 cases. *Pediatr. Genet.* 18 (2011).

24. Winkler, T. W. et al. Quality control and conduct of genome-wide association studies. *Nat. Genet.* 46, 297–303 (2014).

25. Welter, D. et al. The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. *Nucleic Acids Res.* 42, D1001–D1006 (2014).
Ervin R. Fox, Sharon L.R. Kardia, Ching-Ti Liu, Dennis O. Mook-Kanamori, Michael A. Province, Susan Redline, Cornelia M. van Duijn, Jerome I. Rotter, Charles B. Kooperberg, W. James Gauderman, Bruce M. Psaty, Kenneth Rice, Patricia B. Munroe, Myriam Fornage, L. Adrienne Cupples, Charles N. Rotimi, Alanna C. Morrison, Dabeeru C. Rao & Ruth J.F. Loos.

1Novo Nordisk Foundation Center for Basic Metabolic Research, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen 2200, Denmark. 2Department of Environmental Medicine and Public Health, The Icahn School of Medicine at Mount Sinai, New York 10029 NY, USA. 3Center for Research on Genomics and Global Health, National Institute on Aging, National Institutes of Health, Bethesda 20892 MD, USA. 4Internal Medicine, Gerontology and Geriatrics, Leiden University Medical Center, Leiden 2300 RC, The Netherlands. 5Division of Biostatistics, Washington University School of Medicine, St. Louis 63110 MO, USA. 6Department of Genetic Epidemiology, University of Regensburg, Regensburg 93051, Germany. 7Preventive Medicine, Brigham and Women’s Hospital, Boston 02215 MA, USA. 8Harvard Medical School, Boston 02118 MA, USA. 9Clinical and Translational Epidemiology Unit, Massachusetts General Hospital, Boston 02114 MA, USA. 10Department of Medicine, Harvard Medical School, Boston 02115 MA, USA. 11Clinical Pharmacology, William Harvey Research Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London E1CM 6BQ, UK. 12Department of Epidemiology, Harvard School of Public Health, Boston 02115 MA, USA. 13Centre de Bioinformatique, Biostatistique et Biologie Intégrative (C3BI), Institut Pasteur, Paris 75015, France. 14Human Genetics Center, Department of Epidemiology, Human Genetics, and Environmental Sciences, School of Public Health, The University of Texas Health Science Center at Houston, Houston 77030 TX, USA. 15Cardiovascular Division, Department of Medicine, Washington University, St. Louis 63110 MO, USA. 16Epidemiology, University of North Carolina Gillings School of Global Public Health, Chapel Hill 27514 NC, USA. 17The Institute for Translational Genomics and Population Sciences, Division of Genomic Outcomes, Department of Pediatrics, Los Angeles Biomedical Research Institute at Harbor-UCLA Medical Center, Torrance 900502 CA, USA. 18Department of Epidemiology, Erasmus University Medical Center, Rotterdam 3015 CE, The Netherlands. 19Department of Epidemiology, University of Alabama at Birmingham, Birmingham 35294 AL, USA. 20Division of Statistical Genomics, Department of Genetics, Washington University School of Medicine, St. Louis 63108 MO, USA. 21Department of Epidemiology, School of Public Health, University of Michigan, Ann Arbor 48109 MI, USA. 22Jackson Heart Study, Department of Medicine, University of Mississippi Medical Center, Jackson 39213 MS, USA. 23Institute of Molecular Medicine, McGovern Medical School, University of Texas Health Science Center at Houston, Houston 77030 TX, USA. 24Division of Sleep and Circadian Disorders, Brigham and Women’s Hospital, Boston 02115 MA, USA. 25Cardiovascular Health Research Unit, Biostatistics and Medicine, University of Washington, Seattle 98101 WA, USA. 26Centre for Genomic & Experimental Medicine, Institute of Genetics & Molecular Medicine, University of Edinburgh, Edinburgh EH4 2UX, UK. 27Genome Institute of Singapore, Agency for Science Technology and Research, Singapore 138672, Singapore. 28Biostatistics, Boston University School of Public Health, Boston 02118 MA, USA. 29Postgraduate Program in Epidemiology, Federal University of Pelotas, Pelotas 96020220 RS, Brazil. 30Medical Research Council Integrative Epidemiology Unit, University of Bristol, Bristol BS8 2BN, UK. 31Laboratory of Genomics and Molecular Cardiology, Heart Institute (InCor), University of São Paulo Medical School, São Paulo 01246903 SP, Brazil. 32Epidemiology and Biostatistics, University of Giorgi at Athens College of Public Health, Athens 30602 GA, USA. 33Public Health Sciences, Biostatistical Sciences, Wake Forest University Health Sciences, Winston-Salem 27157 NC, USA. 34Medical Research Council Human Genetics Unit, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh EH4 2XU, UK. 35Saw Swee Hock School of Public Health, National University Health System and National University of Singapore, Singapore 117549, Singapore. 36Icelandic Heart Association, 201 Kopavogur, Iceland. 37Department of Biostatistics, University of Michigan, Ann Arbor 48109 MI, USA. 38Health Disparities Research Section, Laboratory of Epidemiology and Population Sciences, National Institute on Aging, National Institutes of Health, Baltimore 21224 MD, USA. 39Estonian Genome Center, University of Tartu, Tartu 51010, Estonia. 40Department of Epidemiology, University of Groningen, University Medical Center Groningen, Groningen 9700 RB, The Netherlands. 41CNRS UMR 8199, European Genomic Institute for Diabetes (EGRID), Institut Pasteur de Lille, University of Lille, Lille 59000, France. 42Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Stockholm 17177, Sweden. 43Department of Biostatistical Sciences, Wake Forest School of Medicine, Winston-Salem 27157 NC, USA. 44Department of Medical Epidemiology and Biostatistics, University of London, London 821P, UK. 45Department of Health, Bethesda and Epidemiology, University of Ioannina Medical School, Ioannina 45110, Greece. 46Molecular Genetics and Genomics Program, Wake Forest School of Medicine, Winston-Salem 27157 NC, USA. 47Centre for Cognitive Ageing and Cognitive Epidemiology, The University of Edinburgh, Edinburgh EH8 9JZ, UK. 48Department of Occupational and Environmental Health and State Key Laboratory of Environmental Health for Incubating, Tongji Medical College, Huazhong University of Science and Technology, Wuhan 430014, China. 49Department of Biostatistics and Center for Statistical Genetics, University of Michigan, Ann Arbor 48109 MI, USA. 50MRC Epidemiology Unit, University of Cambridge, Cambridge CB2 0QQ, UK. 51Research Unit of Molecular Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg 85764, Germany. 52Department of Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg 85764, Germany. 53Unit of Cardiovascular Epidemiology, Institute of Environmental Medicine, Karolinska Institutet, Stockholm 17177, Sweden. 54Department of Clinical Chemistry, Fimlab Laboratories, Tampere 33014, Finland. 55Department of Clinical Chemistry, Finnish Cardiovascular Research Center—Tampere, Faculty of Medicine and Life Sciences, University of Tampere, Tampere 33014, Finland. 56Foundation for Research in Health Exercise and Nutrition, Kuopio Research Institute of Exercise Medicine, Kuopio 70100, Finland. 57College of Medicine, Biological Sciences and Psychology, Health Sciences, The Infant Mortality and Morbidity Studies (TIMMS), Leicester LE1 7RH, UK. 58Institute for Maternal and Child Health—IRCCS “Burlo Garofolo”, Trieste 34137, Italy. 59Department of Computational Biology, University of Lausanne, Lausanne 1015, Switzerland. 60Swiss Institute of Bioinformatics, 1015 Lausanne, Switzerland. 61Department of Gene Diagnostics and Therapeutics, Research Institute, National Center for Global Health and Medicine, Tokyo 1628655, Japan. 62Department of Clinical Sciences, Genetic and Molecular Epidemiology Unit, Lund University Diabetes Centre, Skåne University Hospital, Malmö 20502, Sweden. 63University of Groningen, University Medical Center Groningen, Department of Cardiology, Groningen 9700 RB, The Netherlands. 64Survey Research Center, Institute for Social Research, University of Michigan, Ann Arbor 48104 MI, USA. 65Division of Epidemiology, Department of Medicine, Vanderbilt University School of Medicine, Nashville 37203 TN, USA. 66Department of Population and Quantitative Health Sciences, Case Western Reserve University, Cleveland 44106 OH, USA. 67Division of General Internal Medicine, Department of Medicine, Johns Hopkins University School of Medicine, Baltimore 21287 MD, USA. 68Dean’s Office, University of Kentucky College of Public Health, Lexington 40536 KY, USA. 69Human Genome Sequencing Center, Baylor College of Medicine, Houston 77030 TX, USA. 70Department of Medical Sciences, University of Trieste, Trieste 34137, Italy. 71Southwings Health & Medical School, University of Dundee, Dundee DD1 9SY Scotland, UK. 72Clinical Epidemiology, Leiden University Medical Center, Leiden 2300 RC, Netherlands. 73Department of Medicine, Faculty of Medicine, University of Kelaniya, Ragama 11600, Sri Lanka. 74Department of Internal Medicine, Section on
Lifelines Cohort Study

Behrooz Z. Alizadeh40, H. Marike Boezen40, Lude Franke147, Gerjan Navis167, Marianne Rots168, Morris Swertz147, Bruce H.R. Wolffenbuttel149 & Cisca Wijmenga147

167Department of Internal Medicine, Division of Nephrology, University of Groningen, University Medical Center Groningen, Groningen 9713 GZ, The Netherlands. 168Department of Medical Biology, University of Groningen, University Medical Center Groningen, Groningen 9713 GZ, The Netherlands