Replication and fine-mapping of genetic predictors of lipid traits in African–Americans

QiPing Feng1, Wei-Qi Wei2, Rebecca T Levinson3, Jonathan D Mosley1 and C Michael Stein1,4

Circulating lipid concentrations are among the strongest modifiable risk factors for coronary artery disease (CAD). Most genetic studies have focused on Caucasian populations with little information available for populations of African ancestry. Using a cohort of ~2800 African–Americans (AAs) from a biobank at Vanderbilt University (BioVU), we sought to trans-ethnically replicate genetic variants reported by the Global Lipids Genetics Consortium to be associated with lipid traits in Caucasians, followed by fine-mapping those loci using all available variants on the MetaboChip. In AAs, we replicated one of 56 SNPs for total cholesterol (TC) (rs6511720 in LDLR, \( P = 2.15 \times 10^{-8} \)), one of 63 SNPs for high-density lipoprotein cholesterol (HDL-C) (rs3764261 in CETP, \( P = 1.13 \times 10^{-5} \)), two of 46 SNPs for low-density lipoprotein cholesterol (LDL-C) (rs629301 in CELSR2/SORT1, \( P = 1.11 \times 10^{-5} \) and rs6511720 in LDLR, \( P = 2.47 \times 10^{-5} \)) and one of 34 SNPs for TG (rs645040 in MSL2L1, \( P = 4.29 \times 10^{-4} \)). Using all available variants on MetaboChip for fine mapping, we identified additional variants associated with TC (APOE), HDL-C (LPL and CETP) and LDL-C (APOE). Furthermore, we identified two loci significantly associated with non-HDL-C: APOE/APOC1/TOMM40 and PCSK9. In conclusion, the genetic architecture of lipid traits in AAs differs substantially from that in Caucasians and it remains poorly characterized.

Journal of Human Genetics (2017) 62, 895–901; doi:10.1038/jhg.2017.55; published online 25 May 2017
understand genetic predictors for non-HDL-C in AAs. Another lipid trait, remnant cholesterol, is a measure of cholesterol content of triglyceride-rich lipoproteins and includes very low-density lipoprotein, intermediate-density lipoprotein and chylomicron remnants. Elevated remnant cholesterol levels have been associated with increased risk of CAD independent of HDL-C. However, little is known about the genetic predictors of remnant cholesterol level in AAs.

In this study of AAs, we set out to trans-ethnically replicate genetic variants reported by the GLGC to be associated with lipid traits in Caucasians, followed by fine-mapping of those loci using all available variants on the MetaboChip, a platform designed for metabolic traits and enriched for rare and low-frequency variants.

MATERIALS AND METHODS

Study population

The study was approved by the Institutional Review Board of Vanderbilt University. The study cohort consisted of third-party-identified AAs older than 18 years of age who had lipid measurements and genotyping information available in the Vanderbilt DNA biobank (BioVU). BioVU accrues DNA samples from blood drawn for routine clinical testing after these samples are scheduled to be discarded. BioVU and sample handling have been previously described. Samples and existing genotypes in BioVU are linked to a de-identified version of each individual’s electronic health record (EHR).

Phenotyping: extraction of lipid measurements from electronic health records

Lipid measurements were extracted from each individual’s EHR, and median LDL-C, HDL-C, TC and TG calculated. Lipid measurements after statin exposure were excluded from the analyses. We manually reviewed charts for individuals having extreme lipid levels to exclude the data entry errors. We further calculated non-HDL-C for each individual by subtracting HDL-C and remnant cholesterol by subtracting both HDL-C and LDL-C from TC.

Genotyping

Genotyping was performed using the Illumina MetaboChip custom BeadChip targeting 196,725 genetic variants. The genotyping data were curated and SNPs that deviated significantly from Hardy–Weinberg equilibrium (HWE) (P ≤ 1.0 × 10−6). We removed samples: (1) per-individual call rate < 95% and SNPs that deviated significantly from Hardy–Weinberg equilibrium (HWE) (P ≤ 1.0 × 10−6). We removed samples: (1) per-individual call rate < 95%; (2) with mismatch between genetic and EHR sex; (3) with a cryptic relationship closer than a third-degree relative. Twenty-nine individuals were removed from the analysis.

Candidate SNPs

The GLGC previously identified genetic variants associated with four lipid traits. We extracted genotype information for these variants, including 63 variants for HDL-C, 46 variants for LDL-C, 56 variants for TC and 34 variants for TG. We also extracted all available genetic variants (from transcription start to transcription end) for GLGC-identified genes, including 3939 variants for HDL-C, 1995 variants for LDL-C, 2430 variants for TC and 3888 variants for TG.

Table 1 Cohort characteristics

| N of individuals included | N of female (%) | Lipid levels (mg dl⁻¹) | Age at measurement | N of statin users (%) |
|--------------------------|----------------|------------------------|--------------------|----------------------|
| Total Cholesterol (TC)   | 2778           | 1740 (62.6)            | 185.3 ± 45.2       | 44.7 ± 15.4          | 1168 (42.0)          |
| High-density lipoprotein cholesterol (HDL-C) | 2550           | 1624 (63.7)            | 53.2 ± 17.5        | 44.9 ± 15.2          | 1091 (42.8)          |
| Low-density lipoprotein cholesterol (LDL-C) | 2438           | 1548 (63.5)            | 109.9 ± 39.9       | 45.2 ± 15.1          | 1041 (42.7)          |
| Triglycerides (Trigs)    | 2690           | 1681 (62.5)            | 119.9 ± 87.6       | 44.8 ± 15.1          | 1139 (42.3)          |
| Non-HDL cholesterol (Non-HDL-C) | 2468           | 1576 (63.9)            | 133.9 ± 44.8       | 45.5 ± 15.0          | 1066 (42.3)          |

The data are shown as number (%) and mean ± s.d.
size in AAs as in Caucasians; however, rs645040 in MSL2L1 was associated with larger changes in TG in AAs than in Caucasians—the minor allele was associated with $-8.56 \text{ mg dl}^{-1}$ TG change in AAs, but only $-2.2 \text{ mg dl}^{-1}$ change in Caucasians1 (Table 2).

### Fine-mapping of GLGC loci significantly associated with lipid traits
Second, we tested all genetic variants present on the MetaboChip within loci identified by the GLGC as associated with lipids. Locus fine-mapping yielded additional associations (Table 3), which were not previously reported as leading SNPs from GLGC analyses in Caucasians.

1. **Associations with TC.** One additional PCSK9 variant, rs28362286, associated with TC (Table 3, $P=1.50 \times 10^{-10}$). A well-characterized APOE variant, rs7412, was also associated with TC with genome-wide significance ($P=3.09 \times 10^{-22}$).

2. **Associations with HDL-C.** Six additional variants were associated with HDL-C in AAs—one in LPL and five in CETP. Two SNPs in CETP achieved genome-wide significance: rs7498992 ($P=1.51 \times 10^{-10}$) was previously reported in several Caucasian GWAS studies,32,33 and rs34065661 ($P=1.53 \times 10^{-10}$), a missense variant, was the variant most significantly associated HDL-C in our analyses (Table 3). We queried the lead CETP SNP (rs34065661) and the linked SNPs in Genotype-Tissue Expression (GTEx) database. Rs711752, which associated with rs34065661 ($D^2 = 1.0, R^2 = 0.192$, Supplementary Table 2), was found to be an eQTL variant and carriers of the variant had significantly lower CETP expression ($P=1.6 \times 10^{-5}$) compared to non-carriers (Figure 1a).

3. **Associations with LDL-C.** Three variants were associated with LDL-C. R28362266 ($P=2.08 \times 10^{-5}$) and rs28362286 ($P=1.99 \times 10^{-13}$) in the PCSK9 locus, and rs7412 ($P=2.48 \times 10^{-44}$) in the APOE locus, which accounted for $-22.08 \text{ mg dl}^{-1}$ LDL-C change, comparable to $-22.52 \text{ mg dl}^{-1}$ from NHANES III.34 Using the GTEx database, rs7412 is predicted to significantly alter APOE expression ($P=7.3 \times 10^{-6}$) (Figure 1b).

4. **No additional associations were observed for TG.**

### Associations between variants in the entire MetaboChip and GLGC lipid traits
Third, by analyzing the entire MetaboChip, we identified additional variants significantly associated with lipids (Supplementary Table 3). For each identified gene, we defined a gene region as the gene transcript ± 50 kb, and generated regional association plots (Figure 2). We further identified 8 SNPs in the CETP region associated with HDL-C and seven SNPs in the APOE region associated with LDL-C (Supplementary Table 3). The eight CETP variants are in strong linkage disequilibrium (Supplementary Figure 1). Most associations were attributed to their linkage with the lead SNPs in the regions (Supplementary Table 3). After conditioning on the lead SNPs, two variants remained significantly associated with HDL-C (rs4783961 in CETP region) and rs4389957 in LPL/SLC18A2 region, Table 4), and one variant significantly associated with LDL-C (rs611917 in CELSR2/SORT1 region) (Table 4).

No additional associated variant was identified for TC and TG.

### Genetic predictors for non-HDL-C
We tested the association between all MetaboChip variants and non-HDL-C levels, and identified two loci that contained 5 SNPs
significantly associated with non-HDL-C levels (Table 5). All five SNPs were also associated with either TC or LDL in previous analyses (Table 5).

**Genetic predictors for remnant cholesterol**

We further tested the association with remnant cholesterol, and observed no association in AAs (data not shown).

**Heritability of lipid traits explained by MetaboChip**

By estimating the percentage of heritability explained by available genotypes in MetaboChip, we found that additive genetic components explained 22.52 ± 8.86% of TC (P = 0.0048), 19.22 ± 9.59% of HDL-C (P = 0.023) and 28.51 ± 9.74% of LDL-C (P = 0.0013), 34.35 ± 10.9% of non-HDL-C (P = 0.00063), but only 8.25 ± 8.51% of TG (P = 0.159).

**DISCUSSION**

In this study, we sought to replicate variants associated with lipid traits identified by the GLGC and to fine-map significantly associated loci in ~2800 AAs. There are two major findings of the study: (1) relatively few lipid-associated variants identified by the GLGC replicated in AAs; and (2) we identified additional variants associated with TC (APOE), HDL-C (LPL and CETP) and LDL-C (ApoE), and two loci significantly associated with non-HDL-C (APOE/APOC1/TOMM40 and PCSK9).

The GLGC previously identified 157 loci associated with one or more lipid traits from a cohort of predominantly Caucasians; only a few variants replicated in AAs. Fine-mapping and analyses using the entire MetaboChip identified a few additional genetic associations with GLGC lipid traits in AAs and also 5 SNPs associated with non-HDL-C.

In addition to the complexity of the genetic architecture in individuals of African ancestry, several reasons could explain why so
few GLGC variants replicated in AAs. It is likely that because the 157 loci account for a small portion of overall inter-individual variability of lipids in Caucasians, the contribution of those variants could be even harder to detect if their frequency is lower in AAs than in Caucasians. Nevertheless, we were able to confirm the role of APOE, CETP, PCSK9 and LPL in regulating lipid levels in AA population:

Rs7412 in APOE is one of the most promising predictors of LDL-C and TC levels in AAs. APOE is a ligand for LDL receptor, and therefore is involved in the removal of LDL from the circulation. The rs7412 SNP changes the amino acid at position 158 in APOE from Arg to Cys and in the GTEx database, rs7412 carriers have lower APOE expression compared to non-carriers. Previous reports from NHANES III suggested that rs7412 was associated with 22.52 mg dl⁻¹ lower LDL-C concentrations and 20.68 mg dl⁻¹ lower TC concentrations per minor allele. In AAs, rs7412 was associated with 22.08 mg dl⁻¹ lower LDL-C, similar to findings in other populations.

Human cholesteryl ester transfer (CETP) protein plays a crucial role in lipid metabolism by mediating the transfer of cholesteryl esters from HDL to apolipoprotein (apo) B rich lipoproteins in exchange for TG. High CETP activity contributes to an unfavorable plasma
lipoprotein profile by lowering HDL-C and increasing LDL-C. In the current analyses, variants in the CETP region were identified with genome-wide significance. Further efforts to sequence CETP in AAs may provide additional insight into the gene function and its genetic structure.

PCSK9 is an enzyme that regulates LDL-C levels by promoting the degradation of LDL receptors that are responsible for the removal of LDL particles from the circulation into the liver. Rs28362286 is a degradation of LDL receptors that are responsible for the removal of LDL from the circulation into the liver. This variant results in a truncated PCSK9 protein associated with altered PCSK9 function, lower LDL-C concentrations, and reduced CAD. The variant effects on a truncated PCSK9 protein with reduced activity that is therefore less efficient in degrading LDL receptors; consequently, more LDL is removed from the circulation. In the current cohort, the variant was significantly associated with TC, LDL-C and non-HDL-C concentrations. It is possible that additional ancestry-specific PCSK9 variants contribute further to inter-individual variation in lipid concentrations but high-density genotyping or sequencing will be needed to fully elucidate the function of ancestry-specific variants.

In addition to LDL-C, HDL-C, TC and TG, we further sought to identify the genetic determinants of non-HDL-C. A previous study identified a splice region variant in LDLR associated with lower non-HDL-C in a Caucasian population. We did observe that variants in PCSK9 and APOE associated with other lipid traits were also significantly associated with non-HDL-C. This is not unexpected since non-HDL-C contains LDL-C, TC and remnant cholesterol. The extent to which genetic associations with non-HDL-C are driven by association with LDL-C or TC, or by an unidentified relationship with remnant cholesterol, is unclear.

Furthermore, no genetic predictor was identified for remnant cholesterol. A large-scale Mendelian randomization study suggested that remnant cholesterol is a causal risk factor for ischemic heart disease independent of HDL levels. However, no genetic factor has been reported to affect remnant cholesterol levels exclusively and in clinical practice, remnant cholesterol varies more than other lipid traits over years; therefore, identification of its genetic predictors will be challenging.

Although we replicated some genetic associations for TC, HDL-C and LDL-C, no significant association was observed for TG, even after scanning the entire MetaboChip. Compared to other lipid traits, TG is affected more by environmental factors such as diet and medications. The GLGC previously estimated that known genetic predictors explained ~15% of overall variance for TC, LDL-C and HDL-C, and ~11% for TG. Compared to the other lipid traits, less TG heritability was explained by genotypes on the MetaboChip. However, given that the MetaboChip was designed to capture most loci associated with metabolic traits in Caucasians, it may not be the best tool to estimate overall heritability of lipid traits in AAs.

We acknowledge several limitations: (1) using the MetaboChip, we only fine-mapped regions known to be GWAS-significant regions for metabolic traits. Because most of the early GWAS that informed the construction of the MetaboChip were conducted in Caucasians, African-exclusive loci will be underrepresented. Pilot studies from the African Genome Variation Project have identified a substantial proportion of unshared (11–23%) and novel (16–24%) variants in populations of African ancestry. (2) Power to detect associations with rare variants was limited. With a cohort of ~2800 AA individuals, there was adequate power to detect associations for variants with >1% frequency and an effect size larger than 5 mg dl⁻¹. Nevertheless, power was limited for rare or private variants. Sequencing the whole genome or whole exome in an adequate number of individuals of African ancestry would assist in understanding the genetic architecture of lipid metabolism in African populations.

### Table 4 Additional variants associated with lipid traits by scanning entire MetaboChip

| CHR | SNP    | MAF   | Minor Allele | BETA  | STAT  | P-value | R², to lead SNPs | D', to lead SNPs |
|-----|--------|-------|--------------|-------|-------|---------|----------------|----------------|
| 1   | rs28362286 | 0.009733 | A            | −0.367 | −8.029 | 1.5E-15 | PCSK9 TC         |                |
| 19  | rs7548982  | 0.1274  | A            | −0.08772 | −6.588 | 5.5E-11 | APOE LDL         |                |
| 19  | rs61679753 | 0.11    | A            | −0.1424 | −9.878 | 1.4E-22 | APOE TC          |                |
| 19  | rs7412   | 0.1061  | T            | −0.1719 | −11.95 | 5.6E-32 | APOE TC,LDL      |                |
| 19  | rs75627662 | 0.1643  | T            | −0.09318 | −7.656 | 2.8E-14 | APOE TC,LDL      |                |

### Table 5 Genetic variants associated with non-HDL cholesterol

| CHR | SNP    | MAF   | Minor Allele | BETA  | STAT  | P-value | Gene   | Also Associated with |
|-----|--------|-------|--------------|-------|-------|---------|--------|---------------------|
| 1   | rs28362286 | 0.009733 | A            | −0.367 | −8.029 | 1.5E-15 | PCSK9 TC         |                |
| 19  | rs7548982  | 0.1274  | A            | −0.08772 | −6.588 | 5.5E-11 | APOE LDL         |                |
| 19  | rs61679753 | 0.11    | A            | −0.1424 | −9.878 | 1.4E-22 | APOE TC          |                |
| 19  | rs7412   | 0.1061  | T            | −0.1719 | −11.95 | 5.6E-32 | APOE TC,LDL      |                |
| 19  | rs75627662 | 0.1643  | T            | −0.09318 | −7.656 | 2.8E-14 | APOE TC,LDL      |                |

Abbreviations: BETA, regression coefficient; MAF, minor allele frequency; STAT, coefficient t-statistic.

a|b|c|d|e|f|g|h|i|j|k|l|m|n|o|p|q|r|s|t|u|v|w|x|y|z|

- a All lipid traits were natural log transformed.
- b The linkage to lead SNPs in the region (rs28362286 for PCSK9, rs7412 for APOE, rs34065661 for CETP and rs4389957 for LPL, bold).
- c The associations have been adjusted for age, gender and six PCs.
- d All lipid traits were natural log transformed.
In conclusion, we identified both known and novel genetic variants, which significantly associated lipids traits in AAs. The observations will contribute to understand genetic architecture of lipid, especially in minorities.

CONFLICT OF INTEREST
The authors declare no conflict of interest.

ACKNOWLEDGEMENTS
This study was supported by NIH grant (GM120523, GM109145, HL133786 and ST23GM080178-09), American Heart Association (16TFT0310005 and 16SDG27490014) and Vanderbilt Faculty Research Scholar Fund. The dataset(s) used for the analyses described were obtained from Vanderbilt University Medical Center’s resources, BioVU and the Synthetic Derivative, which are supported by institutional funding and by the Vanderbilt National Center for Advancing Translational Science grant 2UL1 TR000445-06 from NCATS/NIH. Existing genotypes in BioVU were funded by NIH grants RC2GM092618 from NIGMS/OD and U10HG004603 from NHGRI/NIGMS.

1 Teslovich, T. M., Musunuru, K., Smith, A. V., Edmondson, A. C., Stylianou, I. M., Koscielniak, M. et al. Biological, clinical and population relevance of 95 loci for blood lipids. Nature 466, 707–713 (2010).
2 Global Lipids Genetics Consortium. Discovery and refinement of loci associated with lipid levels. Nat. Genet. 45, 1274–1283 (2013).
3 de Miranda Chagas, S. V., Kanaan, S., Chang Hung, H., Cagy, M., Abreu, R. E., da Silva, L. A. et al. Environmental factors, familial aggregation and heritability of total cholesterol, low density lipoprotein-cholesterol and high density lipoprotein-cholesterol in a Brazilian population assisted by the Family Doctor Program. Public Health 125, 329–337 (2011).
4 Filia, G., Chen, W.-M., Scuteri, A., Omri, M., Albi, G., Del, M. et al. Heritability of cardiovascular and personality traits in 6148 Sardinians. PLoS Genet. 2, e132 (2006).
5 Dumitrescu, L., Carty, C. L., Taylor, K., Schumacher, F. R., Hindorff, L. A., Ambite, J. L. et al. Genetic determinants of lipid traits in diverse populations from the population architecture using genomics and epidemiology (PAGE) study. PLoS Genet. 7, e1002138 (2011).
6 Cohen, J., Pertsemlidis, A., Kotsowis, I. K., Graham, R., Garcia, C. K. & Hobbs, H. H. Sequence variations in ANGPTL4 and risk of coronary artery disease. N. Engl. J. Med. 347, 1429–1431 (2007).
7 N. Engl. J. Med. 347, 1429–1431 (2007).
8 Cui, Y., Blumenthal, R. S., Flaws, J. A., Whiteman, M. K., Langenberg, P., Bachorik, P. S. et al. Clinical utility of different lipid measures for prediction of coronary heart disease in men and women. JAMA 298, 776–785 (2007).
9 17 Rana, J. S., Boekholt, S., Mattle, G., Butschart, I. J. P. & Shah, P. K. The role of non-high cholesterol in risk stratification for coronary artery disease. Curr. Atheroscler. Rep. 14, 130–134 (2011).
10 Redondo-Solà, J. E., Geris, E. G., Calle, M. A., Aikens, J., Sorensen, A., Thisted, R. A. et al. A single phase variant in LDLR lowers non-high-density lipoprotein cholesterol and protects against coronary artery disease. PLoS Genet. 11, e1005379 (2015).
11 Nio, P., Pirtorin, J., Thorlund, K., Eggers, H., Agustsdottir, A. B., Norddal, G. L. et al. Variant ASGR1 associated with a reduced risk of coronary artery disease. N. Engl. J. Med. 374, 2131–2141 (2016).
12 Nordestgaard, B. G., Benn, M., Schnoor, P. & Tpyejaer-Hansen, A. Nonfasting triglycerides and risk of myocardial infarction, ischaemic heart disease, and death in men and women. JAMA 298, 2939–2947 (2007).
13 Leme, L. G., Song, J. Y., de Andrade, G., de Andrade, G., de Andrade, G. et al. Inactivating variants in ANGPTL4 and risk of coronary artery disease. J. Lipid Res. 57, 1509 (2015).
14 Cui, Y., Blumenthal, R. S., Flaws, J. A., Whiteman, M. K., Langenberg, P., Bachorik, P. S. et al. Clinical utility of different lipid measures for prediction of coronary heart disease in men and women. JAMA 298, 776–785 (2007).
15 Varbo, A., Benn, M., Thybjerg-Hansen, A., Jorgensen, A. B., Frikke-Schmidt, R. & Nordestgaard, B. G. Remnant cholesterol as a causal risk factor for ischaemic heart disease. J. Am. Coll. Cardiol. 61, 427–436 (2013).
16 Crawford, D. C., Goodlow, R., Brown-Gentry, K., Wilk, S., Roberson, J., Gillani, B. N. et al. Characterization of the MetaboChip in diverse populations from the International HapMap Project in the Epidemiologic Architecture for Genes Linked to Environment (EAGLE) project. Pac. Symp. Biocomput. Pac. Symp. Biocomput. 813–824 (2013). https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3747470/.
17 Voight, B. F., Kang, H. M., Ding, J., Palmer, C. D., Skagard, C., Chines, P. S. et al. The MetaboChip, a custom genotyping array for genetic studies of metabolic, cardiovascular, and anthropometric traits. PLoS Genet. 8, e1002793 (2012).
18 Dumitrescu, L., Ritchie, M. D., Brown-Gentry, K., Wilk, S., Roberson, J., Gillani, B. N. et al. Assessing the accuracy of observer-reported ancestry in a biorepository linked to electronic medical records. Genet. Med. 12, 648–650 (2010).
19 Tuomin, R., Welch, R. P., Sanna, S., Teslovich, T. M., Chines, P. S., Giegi, T. P. et al. LocusZoom: regional visualization of genome-wide association scan results. Bioinformatics 26, 2336–2337 (2010).
20 Yang, J., Lee, S. H., Goddard, M. E. & Visscher, P. M. GCTA: a tool for genome-wide complex trait analysis. Am. J. Hum. Genet. 88, 76–82 (2011).
21 Chamber, A. J., Paré, G., Gora, M., Sojef, J., Cilero, G., Clarke, R. et al. Forty-three loci associated with plasma lipiddensity, concentration, and cholesterol content in genome-wide association studies. PLoS Genet. 8, e1000730 (2009).
22 Zemunik, T., Boban, M., Lauc, G., Jankovic, S., Rot, J., Vataku, Z. et al. Genome-wide association study of biochemical traits in Korcula Island, Croatia. Croat. Med. J. 50, 23–33 (2009).
23 Chang, M., Yesupriya, A., Red, R. M., Mueller, P. W. & Dowling, N. F. Genetic variants associated with fasting blood lipids in the US population: Third National Health and Nutrition Examination Survey. BMC Med. Genet. 11, 62 (2010).
24 Tannock, L. & Bhat, A. In Risk Assessment and Guidelines for the Management of High triglycerides (eds De Groot L. J. et al.) Endocet (Internet), MDText.com, Inc., South Dartmouth (MA, USA), (2000) Available at http://www.ncbi.nlm.nih.gov/books/NBK326749/.
25 Gudasani, D., Cartensen, T., Tekola-Yehe, A., Fagani, L., Tachmazidou, I., Hatzikotoulas, K. et al. The African Genome Variation Project shapes medical genetics in Africa. Nature 517, 327–332 (2015).
26 Wu, Y., Marcelle, A. F., Li, J., Jotte-Chonka, D. C., Ferani, A. B., Kuzawa, C. W. et al. Genetic association with lipids in Filipinos: waist circumference modifies an APOAI effect on triglyceride levels. J. Lipid Res. 54, 3198–3205 (2013).
27 Rasmussen-Torvik, L. J., Kakko, J., Wilke, R. A., Thompson, W. K., Ritchie, M. D. & Koesik, M. N. High density GWAS for LDL cholesterol in African americans using electronic medical records reveals a strong protective variant in APOE. Clin. Transl. Sci. 5, 394–399 (2012).
28 Chen, D. I., Bambazon, F. M., Lovisa, F. B., Nyberg, F. & Riedler, P. M. Genetic determinants of statin-induced low-density lipoprotein cholesterol reduction: the Justification for the Use of Statins in Prevention: an Intervention Trial Evaluating Rosuvastatin (JUPITER) trial. Circ. Cardiovasc. Genet. 5, 257–264 (2012).
29 Kettouren, J., Tukia, T., Saino, A.-P., Ortega-Alano, A., Tikkkanen, E., Liehmanen, L.-P. et al. Genome-wide association study identifies multiple loci influencing human serum metabolite levels. Nat. Genet. 44, 269–276 (2012).

Supplementary Information accompanies the paper on Journal of Human Genetics website (http://www.nature.com/jhg)