**Supplemental Methods**

**Metabolome**

Metabolite profiling of fasting serum samples that had been stored at −80°C since collection at baseline in 1987-1989 among selected African and European Americans in 2010 and 2014 respectively. Detection and quantification of 140 lipid-related traits were completed by Metabolon, Inc. (Durham, North Carolina) by using an untargeted, gas chromatography–mass spectrometry and liquid chromatography–mass spectrometry­–based (GC/LC-MS) metabolomic quantification protocol (Evans *et al.* 2009; Ohta *et al.* 2009). Serum samples were extracted and prepared using Metabolon’s standard solvent extraction method. To remove protein, dissociate small molecules bound to protein or trapped in the precipitated protein matrix, and to recover chemically diverse metabolites, proteins were precipitated with methanol under vigorous shaking for 2 min (Glen Mills, Inc. NJ, USA), followed by centrifugation (Lawton *et al.* 2008; Shin *et al.* 2014). The extract supernatant was split into equal aliquots for analysis on the GC/MS and LC/MS platforms. For LC/MS analysis extract aliquots were reconstituted in 10% methanol and 0.1% formic acid(Evans *et al.* 2009). For GC/MS analysis, the samples were dried under vacuum desiccation for a minimum of 18 h before being derivatized under dried nitrogen using bistrimethylsilyl-trifluoroacetamide (Shin *et al.* 2014). Derivatized samples were separated on a 5% phenyldimethyl silicone column with helium as carrier gas and a temperature ramp from 60° to 340°C within a 17-min period (Shin *et al.* 2014). All samples were analyzed on a Thermo-Finnigan Trace DSQ MS operated at unit mass resolving power (Shin *et al.* 2014). Instrument variability was determined by calculating the median relative standard deviation (SD) for the internal standards that were added to each sample prior to injection into the mass spectrometers. Overall process variability is determined by calculating the median relative SD for all endogenous metabolites (i.e. non-instrument standards) present in 100% of the technical replicate samples. Since the measurements span multiple days, a data normalization step was performed to correct variation arising from instrument inter-day tuning differences: raw area counts for a compound were divided by the median value, setting the medians equal to 1.0 for each day's run (Shin *et al.* 2014). Quantitative values were derived from integrated raw detector counts of the mass spectrometers. This preserved all of the variation between samples, yet allowed compounds of widely different raw peak areas to be compared on a similar scale (Lawton *et al.* 2008). The lipid-related traits were identified by comparison to an in-house generated authentic standard Metabolon's reference library that includes retention time, molecular weight, preferred adducts, in-source fragments, and associated fragmentation spectra of the intact parent ion (Evans *et al.* 2009). Out of the 140 lipid-related metabolites, 124 had missing values ≤ 25%, and 16 metabolites were excluded due to high levels of missingness. We designed a set of 97 samples to measure their lipid-related metabolites at both 2010 and 2014. We calculated the Pearson correlation coefficients between the 97 pairs for these 124 lipid-related compounds, and mild batch effect was observed. For analysis, 22 lipids with poor Pearson correlation coefficient (<0.3) were excluded, therefore, 102 lipids were examined in this study. Among 102 lipid-related metabolites, Pearson correlation coefficientranged from 0.3147 to 0.9948, with mean of 0.7051 and median of 0.7675.

**Whole Genome Sequencing, Variant Calling, Quality Control and Annotation**

WGS data was generated by the Baylor College of Medicine Human Genome Sequencing Center (BCM-HGSC). DNA samples were constructed into Illumina paired-end libraries according to the manufacturer’s protocol (Illumina Multiplexing\_SamplePrep\_Guide\_1005361\_D) and sequenced on the Hiseq 2000 (Illumina, Inc.; San Diego, CA) in a pooled format to generate a minimum of 18 unique aligned Gbp per sample. A detailed description of methods for WGS of ARIC study participants is available in Morrison et al. (Morrison *et al.* 2013). Individuals of African and European ancestry were sequenced at 7.4-fold average depth on Illumina HiSeq instruments and variant calling was completed using goSNAP, (https://sourceforge.net/p/gosnap/git/ci/master/tree/) which employed GATK, SNPTools and GotCloud as callers, each in joint calling mode, and took an ensemble consensus approach to generate the high quality variant call set. The per-sample genotyping and reference panel independent imputation and phasing were done using SNPTools. Across 5,297 samples from the Cohorts for Health and Aging Research in Genomic Epidemiology (CHARGE) Consortium, 72.8 million single nucleotide variants (SNV) were called with a transition-to-transversion (Ti/Tv) ratio of 2.13. A total of 59.7% of the Variants are novel compared to dbSNP v142 as most of the novel variants are rare due to large sample size. Compared to a subset of CHARGE samples with WES (n=4,612), the sensitivity and specificity of the WGS call set is 63.6% and 99.9%, respectively, and compared to an overlapping set of single nucleotide polymorphism (SNP) array data (n=3,533), the false discovery rate (FDR) is 1.6%. Variant-level quality assurance was achieved by excluding variants with site level inbreeding coefficient <-0.9. Variants not meeting Hardy-Weinberg equilibrium exact test expectations in ancestry-specific groups (p-value <1x10-14) were also excluded. Sample-level quality control and quality assurance checks included principal-components analysis (PCA) to identify possible population-substructure and sample abnormalities. The set of variants for PCA was restricted to variants with MAF >5% and linkage disequilibrium between variants of r2<0.30. A total of 40 ARIC AA individuals identified as outliers by PCA were removed from further analyses. Higher-order principal components (PCs) showed minor levels of population structure. After sample-level quality control, a total of 1679 AA and 1458 EA from the ARIC study were available for the genotype-phenotype analyses reported here. The 3’ and 5’ untranslated regions (UTR) of a gene were identified using ANNOVAR (Zhang *et al.* 2010) and the RefSeq gene model (O'Leary *et al.* 2016). The promoter of a gene was defined based on the overlap between the permissive set of CAGE peaks reported by the FANTOM5 project (Consortium *et al.* 2014) and the 5 kb upstream region determined by ANNOVAR annotation based on the RefSeq gene model. The enhancers and the target genes of the enhancers were defined based on the permissive set of enhancers and enhancer-promoter pairs reported by FANTOM5 (Andersson *et al.* 2014). In the case of an undesignated enhancer-gene pair, the enhancer was assigned to the nearest gene.

**Whole Exome Sequencing, Variant Calling, Quality Control**

The DNA samples were constructed into Illumina paired-end pre-capture libraries according to the manufacturer’s protocol. The complete protocol and oligonucleotide sequences are accessible from the Baylor College of Medicine Human Genome Sequencing Center (BCM-HGSC) website (<https://www.hgsc.bcm.edu/content/protocols-sequencing-library-construction>). Two, four or six pre-capture libraries were pooled together and then hybridized to Nimblegen exome capture array (HGSC VCRome 2.1 design (Bainbridge *et al.* 2011) (42Mb, NimbleGen) and sequenced in paired-end mode in a single lane on the Illumina HiSeq 2000 platform. Illumina sequence analysis was performed using the HGSC Mercury analysis pipeline (<https://www.hgsc.bcm.edu/content/mercury>). Pooled samples were de-multiplexed using the Consensus assessment of sequence and variation (CASAVA) software. Reads were then mapped to the Genome Reference Consortium Human Build 37 (GRCh37) human reference sequence using Burrows-Wheeler Aligner (BWA) (Li and Durbin 2009) producing Binary Alignment/Map (BAM) files. Aligned reads were then recalibrated using the Genome Analysis Toolkit (GATK) (DePristo *et al.* 2011) along with BAM sorting, duplicate read marking, and realignment near insertions or deletions (indels). The Atlas2 (Challis *et al.* 2012) suite was used to call variants and produce high-quality variant call files (VCF) (Danecek *et al.* 2011). Each single nucleotide variant (SNV) call was filtered based on the following criteria to produce a high-quality variant list: low SNV posterior probability (<0.95), low variant read count (<3), variant read ratio <0.25 or >0.75, strand-bias of more than 99% variant reads in a single strand direction, or total coverage less than 10-fold. All variant calls filtered by these criteria, and reference calls with less than 10-fold coverage, were set to missing. The variant call filters were the same for indels except a total coverage less than 30-fold was used. Variant-level quality control steps excluded variants outside the exon capture regions (VChrome 2.1), multi-allelic sites, missing rate >20%, and mean depth of coverage >500-fold. Variants not meeting Hardy-Weinberg equilibrium expectations in ancestry-specific groups (*P*<5x10-6) were also excluded. Sample-level quality control metrics were calculated by cohort and ancestry group. A sample was excluded for missingness >20%, or if compared to the other samples it fell less than 6 standard deviations (SD) for mean depth, more than 6 SD for singleton count, or outside of 6 SD for heterozygote to homozygote ratio or a transition-to-transversion (Ti/Tv) ratio.

The study samples for the WGS and WES analysis were 3,137 and 3,180 individuals, respectively, with 2,893 participants found in both samples.

**References**

Andersson, R., C. Gebhard, I. Miguel-Escalada, I. Hoof, J. Bornholdt *et al.*, 2014 An atlas of active enhancers across human cell types and tissues. Nature 507**:** 455-461.

Bainbridge, M. N., M. Wang, Y. Wu, I. Newsham, D. M. Muzny *et al.*, 2011 Targeted enrichment beyond the consensus coding DNA sequence exome reveals exons with higher variant densities. Genome Biol 12**:** R68.

Challis, D., J. Yu, U. S. Evani, A. R. Jackson, S. Paithankar *et al.*, 2012 An integrative variant analysis suite for whole exome next-generation sequencing data. BMC Bioinformatics 13**:** 8.

Consortium, F., R. Pmi, CLST, A. R. Forrest, H. Kawaji *et al.*, 2014 A promoter-level mammalian expression atlas. Nature 507**:** 462-470.

Danecek, P., A. Auton, G. Abecasis, C. A. Albers, E. Banks *et al.*, 2011 The variant call format and VCFtools. Bioinformatics 27**:** 2156-2158.

DePristo, M. A., E. Banks, R. Poplin, K. V. Garimella, J. R. Maguire *et al.*, 2011 A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet 43**:** 491-498.

Evans, A. M., C. D. DeHaven, T. Barrett, M. Mitchell and E. Milgram, 2009 Integrated, Nontargeted Ultrahigh Performance Liquid Chromatography/Electrospray Ionization Tandem Mass Spectrometry Platform for the Identification and Relative Quantification of the Small-Molecule Complement of Biological Systems. Analytical Chemistry 81**:** 6656-6667.

<https://sourceforge.net/p/gosnap/git/ci/master/tree/>.

Lawton, K. A., A. Berger, M. Mitchell, K. E. Milgram, A. M. Evans *et al.*, 2008 Analysis of the adult human plasma metabolome. Pharmacogenomics 9**:** 383-397.

Li, H., and R. Durbin, 2009 Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25**:** 1754-1760.

Morrison, A. C., A. Voorman, A. D. Johnson, X. Liu, J. Yu *et al.*, 2013 Whole-genome sequence-based analysis of high-density lipoprotein cholesterol. Nat Genet 45**:** 899-901.

O'Leary, N. A., M. W. Wright, J. R. Brister, S. Ciufo, D. Haddad *et al.*, 2016 Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. Nucleic Acids Res 44**:** D733-745.

Ohta, T., N. Masutomi, N. Tsutsui, T. Sakairi, M. Mitchell *et al.*, 2009 Untargeted metabolomic profiling as an evaluative tool of fenofibrate-induced toxicology in Fischer 344 male rats. Toxicol Pathol 37**:** 521-535.

Shin, S. Y., E. B. Fauman, A. K. Petersen, J. Krumsiek, R. Santos *et al.*, 2014 An atlas of genetic influences on human blood metabolites. Nat Genet 46**:** 543-550.

Zhang, S., X. Liu, J. Necheles, H. J. Tsai, G. Wang *et al.*, 2010 Genetic and environmental influences on serum lipid tracking: a population-based, longitudinal Chinese twin study. Pediatr Res 68**:** 316-322.