GWAS in Africans identifies novel lipids loci and demonstrates heterogenous association within Africa

Amy R. Bentley¹, PhD; Guanjie Chen¹, MD; Ayo P. Doumatey¹, PhD; Daniel Shriner¹, PhD; Karlijn Meeks¹, PhD; Mateus H. Gouveia¹, PhD; Kenneth Ekoru¹, PhD; Jie Zhou¹, MS; the Africa America Diabetes Mellitus Investigators; Adebawale Adeyemo¹, MD; Charles N. Rotimi¹, PhD

¹Center for Research on Genomics and Global Health, National Human Genome Research Institute, NIH, Bethesda, MD, 20892, USA

Corresponding Authors: Amy R. Bentley (amy.bentley@nih.gov) and Charles N. Rotimi (rotimic@mail.nih.gov)

Center for Research on Genomics and Global Health

National Human Genome Research Institute, NIH

Building 12, Room 4047

12 South Drive, MSC 5635

Bethesda, MD 20892-5635

Phone: (301) 451-2302; Fax: (301) 451-5426

Word Count:

Journal Subject Terms: Genetics, Cardiovascular Disease, Race and Ethnicity

Abstract

Background: Serum lipids are biomarkers of cardiometabolic disease risk, and understanding the genomic factors contributing to their distribution has been of considerable interest. Large genome-wide association studies (GWAS) have identified over 150 lipids loci; however, GWAS of Africans (AF) are rare. Given the genomic diversity among those of African ancestry, it is expected that a GWAS in Africans could identify novel lipids loci. While GWAS have been conducted in African Americans (AA), such studies are not proxies for studies in continental Africans due to the drastically different environmental context. Therefore, we conducted a GWAS of 4,317 Africans enrolled in the Africa America Diabetes Mellitus study.

Methods and Results: We used linear mixed models of the inverse normal transformations of covariate-adjusted residuals of high-density lipoprotein cholesterol (HDL), low-density lipoprotein cholesterol (LDL), total cholesterol (CHOL), triglycerides (TG), and TG/HDL, with adjustment for three principal components and the random effect of relatedness. Replication of loci associated at $p < 5 \times 10^{-8}$ was attempted in 9,542 AA. Meta-analysis of AF and AA was also conducted. We also conducted analyses that excluded the relatively small number of East Africans. We evaluated known lipids loci in Africans using both exact replication and “local” replication, which accounts for interethnic differences in linkage disequilibrium.

In our main analysis, we identified 23 novel associations in Africans. Of the 14 of these that were able to be tested in AA, two associations replicated (*GPXMB*-TG and *ENPP1*-TG). Two additional novel loci were discovered upon meta-analysis with AA (rs138282551-TG and *TLL2*-CHOL). Analyses considering only those with predominantly West African ancestry (Nigeria, Ghana, and AA) yielded new insights: *ORC5*-LDL and chr20:60973327-CHOL.

Conclusions: While functional work will be useful to confirm and understand the biological mechanisms underlying these associations, this study demonstrates the utility of conducting large-scale genomic

analyses in Africans for discovering novel loci. The functional significance of some of these loci in relation to lipids remains to be elucidated, yet some have known connections to lipids pathways. For instance, rs147706369 (intronic, *TLL2*) alters a regulatory motif for sterol regulatory element-binding proteins (SREBPs), which are a family of transcription factors that control the expression of a range of enzymes involved in cholesterol, fatty acid, and triglyceride synthesis.

Key Words: serum lipids, genome-wide association study, Africans

Introduction

Serum lipids are an important biomarker of cardiometabolic disease risk. As such, understanding the genomic factors contributing to their distribution has been of considerable interest. Large meta-analyses have identified over 150 loci associated with the distribution of serum lipids¹⁻⁸; GWAS in Africa are rare⁹. There are several reasons that this lack is concerning. First, as those of African ancestry have the greatest degree of genetic diversity in the world¹⁰, unless these populations are sufficiently represented in genomic research, a great deal of that diversity is not being evaluated for an association with serum lipids, potentially limiting the molecular insights that can be gained in our understanding of these traits. Second, although genome-wide studies of African Americans (AA) have been conducted, the drastically different environments of African Americans and continental Africans may differentially impact genetic associations. For instance, gene-environment interactions with environmental factors that differ between AF and AA have been found to affect serum lipid distribution, which may affect ability to detect a genetic association across populations¹¹⁻¹⁸. As representation of those of African ancestry in genomic research has been relatively limited, it is common to analyze all those of African ancestry together to increase sample size; however, collapsing groups with such genomic and environmental diversity may obscure findings that are unique to populations within this broad category.

Here, we conducted a genome-wide association study of serum lipids in 4,317 Africans from the Africa America Diabetes Mellitus (AADM) study, a case-control study of type 2 diabetes (T2D) with sites in Nigeria, Ghana, and Kenya. Replication of GWAS-significant loci was sought in 9,542 AA. We report 23 novel loci that were genome-wide significant in Africans. Of the 14 of these that were available in the dataset of AA, two replicated. Novel loci were also identified when conducting analyses in West Africans separately, highlighting the importance of considering heterogeneity between African populations when conducting genomic analyses.

Methods

Participants

Participants included in this study were drawn from the Africa America Diabetes Mellitus (AADM) study, which has been previously described¹⁹⁻²¹. Briefly, AADM is a genetic epidemiology study of T2D, enrolling participants from university medical centers in Nigeria (Enugu, Lagos, and Ibadan), Ghana (Accra and Kumasi), and Kenya (Eldoret). Participants underwent a clinical examination that included a medical history, clinical anthropometry, blood pressure measurements and blood sampling.

Serum measurements were made on fasting samples. Serum lipids (triglycerides [TG], total cholesterol [CHOL], low-density lipoprotein cholesterol [LDLC], and high-density lipoprotein cholesterol [HDLC]) were determined enzymatically either with the COBAS Integra 400 Plus or Modular-E analyzers (Roche Diagnostics, Indianapolis, IN). Methods were standardized to in-house and other appropriate reference methods (e.g., CDC reference methods for HDLC, isotope dilution mass spectrometry for CHOL and TG). Missing LDLC measurements were calculated using the Friedewald equation²². Ethical approval was obtained by the National Institutes of Health and from the ethical committees in each study site. All participants gave written informed consent before participation in the study.

Genotyping and Imputation

Genotyping was conducted either the Affymetrix Axiom® PANAFR SNP array or the Illumina Consortium Multi-Ethnic Global Array (MEGA). Quality control was conducted separately for each of the resulting datasets. After technical quality control, sample-level genotype call rate was at least 0.95 for all subjects. Each SNP dataset was filtered for missingness (< 0.05), Hardy-Weinberg equilibrium (HWE, $p < 10^{-6}$) and minor allele frequency (MAF > 0.01). SNPs that passed quality control filters were used as the basis for

imputation. Imputation of all samples was done using the African Genome Resources Haplotype Reference Panel using the Sanger Imputation Service (<https://imputation.sanger.ac.uk/>). Variants were included in the analysis if they had a MAF ≥ 0.01 and an info score ≥ 0.3 . All models were run using the dosage coding of variants.

Statistical Analysis

Analyses were conducted using linear mixed models of the inverse normal transformations of lipids traits using EPACTS (version 3.2.6; <http://genome.sph.umich.edu/wiki/EPACTS>). Models were adjusted for age, age², and sex. Additionally, we included the first three PCs of the genotypes, based on prior work^{21,23} (see figure S4 of ²⁴). Models were run with and without adjustment for body mass index (BMI), as an effect on lipids can occur through and independent of body adiposity. As this study is nested within a case-control study of T2D, models were also run with and without adjustment for T2D to evaluate a potential effect of disease status. All models also included a genetic relationship matrix to account for the random effect of relatedness, as related individuals were included in AADM. The genome-wide statistical significance level α for the main analysis was set at 5×10^{-8} .

Replication in African Americans

We sought replication for our genome-wide statistically significant associations in a combined dataset of African Americans (AA; n=9,542). Included AA samples were drawn from five datasets: the Howard University Family Study (HUFs)²⁵ along with data available through dbGaP for the Atherosclerosis Risk in Communities study (ARIC phs000280.v2.p1, phs000090.v2.p1), the Cleveland Family Study (CFS phs000284.v1.p1), the Jackson Heart Study (JHS phs000286.v4.p1, phs000499.v2.p1), and the Multi-Ethnic Study of Atherosclerosis study (MESA phs000209.v13.p1, phs000420.v6.p3). When participants had more than one study visit, the visit with the most complete lipid measurements was selected, and if there were multiple visits with complete measurements, the most recent was selected (CFS). All studies

obtained ethical approval from the relevant institutions and written informed consent from each participant prior to participation. Genome-wide genotyping data for all studies were imputed using the methods described above. Meta-analysis of AA study results was conducted with the inverse-variance-weighted fixed-effects method as implemented in METAL (<https://genome.sph.umich.edu/wiki/METAL>)²⁶, and replication was assessed based on results of meta-analyses. Heterogeneity between East and West Africans was evaluated by with the p-value for heterogeneity using the “analyze heterogeneity” option in METAL.

Analyses in West Africans

As our study included participants from both West Africa (Nigeria [n=2,207] and Ghana [n=1,376]) and East Africa (Kenya [n=734]), we investigated whether the inclusion of East Africans (EA) may have obscured findings that would have been detectable in WA alone. EA were not considered separately because of their relatively small sample size. Models in West Africans (WA) were run as described above for main analyses. Given the genetic similarity between AA and West Africans (WA), results from the WA analyses were meta-analyzed with AA data to take advantage of increased sample sizes for discovery.

Replication of Known Lipids Loci

To evaluate the degree to which known lipids loci are also observed among Africans, we sought replication for 1,853 previously identified SNP-trait associations from large genome-wide studies of serum lipids¹⁻⁸. First, we conducted “exact” replication: evaluating the index variant reported in previous publications, with replication defined by consistent direction and $p < 0.05$. Additionally, to account for differences in linkage disequilibrium (LD) between the initial reports (reflecting predominantly European ancestry) and our African samples, we conducted “local” replication. Local replication is an LD-based replication technique, investigating all variants with $r^2 > 0.3$ with the index variant among CEU. A locus was included in the local replication analysis if there was no exact

replication of that locus in our data and if there were additional variants with $r^2 > 0.3$ [CEU] with the index variant that passed filtering criteria and were available in our data. Associations were considered replicated locally with consistent direction (based on direction of association of minor allele for index and local variant) and $p < 0.05$ / LD-based effective number of tests in the region (calculated as in ²⁷).

Results

The study population was heavier and had a worse lipid profile than would be expected of a general population sampled in these regions, consistent with the facts that this study was nested within a case-control study of T2D and approximately half of participants were T2D cases (**Table 1**). Additionally, all recruitment was conducted in urban areas, and urban location is associated with worse cardiometabolic outcomes in West Africa ²⁸.

Our main GWAS analysis in AADM identified 23 novel GWAS-significant lipids loci (29 variants) in AADM (**Table 2; Supplementary Figures 1 [Manhattan plots] and 2 [QQ plots]**), with novelty defined by distance from known lipids loci (> 1 MB) and statistical significance set at 5×10^{-8} . Notable among these findings were associations of variants with higher frequency among Africans compared to other worldwide populations, perhaps allowing for their detection in this analysis. For instance, rs138202830 (*GBE1*), rs188701119 (*TINAG*), rs140987192 (*PBX3*), and rs75360819 (*CDH2*) are monomorphic in non-AFR 1000 Genomes populations, and the minor alleles for rs111590558 (*HTR2A*), rs7281821 (*APP*), and rs144086909 (*LARGE*) are at their highest frequencies in the AFR super-population (**Table 2**).

Given the potential impact of both BMI and T2D on associations with serum lipids, particularly in this dataset nested within a case-control study of T2D, we tested models with and without adjustment for these two covariates. In terms of statistical significance, adjustment for BMI and T2D had relatively

minor effects. For 22 of the 23 novel loci, associations that were statistically significant in one model had $p \leq 10^{-7}$ for all other models (for the remaining locus, all models were $p \leq 10^{-6}$; **Table 2**).

To determine the degree to which these associations were also found among AA, we evaluated these loci in 9,542 AA (**Supplemental Table 1**). Of 29 variants in 23 loci investigated, 16 variants in 14 loci passed frequency and quality control criteria and were able to be evaluated (**Supplemental Table 2**). Two of these loci were replicated ($p < 0.05$ with matching direction of association). The minor (G) allele of rs706014 (5' of *GPNMB*, Glycoprotein Nonmetastatic Melanoma Protein B) was associated with lower TG in both Africans (β -0.75 TG, $p = 2.5 \times 10^{-9}$) and in AA (β -0.084 TG, $p = 0.0093$). The minor (T) allele of rs9375831, intronic to *ENPP1* (ectonucleotide pyrophosphatase/phosphodiesterase 1), was also associated with lower TG in Africans (β -0.55 TG, $p = 4.1 \times 10^{-8}$) and in AA (β -0.076 TG, $p = 0.037$). As variants in this gene have been previously associated with insulin resistance and related traits²⁹⁻³⁴, we also investigated this relationship in our data. An association of the minor allele of rs9375831 with HOMA-IR in AF participants without T2D did not reach statistical significance ($p = 0.2$), but was directionally consistent: the minor allele was associated with decreased TG and decreased HOMA-IR. We also evaluated the ratio of TG to HDL, a biomarker of insulin resistance: the minor allele was associated with decreased TG/HDL ($p = 0.0002$), though this association may be largely driven by the association of this variant with TG.

We conducted meta-analyses of Africans and AA, which yielded two additional GWAS-significant loci (**Table 3**). The minor (T) allele of intergenic variant rs138282551 was associated with higher TG in both Africans (β 0.29 TG, $p = 4.5 \times 10^{-7}$) and in AA (β 0.14 TG, $p = 0.0035$). This variant is at much higher frequency among those of African ancestry compared to other populations (among 1000 Genomes super-populations: AFR 0.03, AMR 0, EAS 0, EUR 0.002, and SAS 0). The minor (T) allele for an African

ancestry-specific intronic variant in *TLL2*, rs147706369, was associated with higher total cholesterol, with similar magnitude (β 0.34 CHOL [AFR], β 0.28 CHOL [AA]) and strength ($p=1.5 \times 10^{-4}$ [AFR] and $p=2.6 \times 10^{-5}$ [AA]) of association for both groups.

Our main analyses of Africans included participants from both West Africa (Nigeria [n=2,207] and Ghana [n=1,376]) and East Africa (Kenya [n=734]). Given the genetic and environmental differences among the sampled populations, it is possible that a signal present in participants from only one of these geographic regions could be masked when the two were included together. Additionally, mean values for TG and closely related values (TG/HDLC and CHOL) differed dramatically in East compared to West Africans (TG 167.5 vs. 98.3 mg/dl, respectively; **Supplemental Table 3**). Based on these observations, we conducted a separate analysis of WA; a separate analysis of EA is not presented due to the relatively small sample size. There were no novel statistically significant hits when WA were considered on their own, but known signals that were not apparent in the full GWAS were observed in the analysis of WA only, including the known chromosome 11 locus *BUD13* and, to a lesser extent, *APOE/APOC1* on chromosome 19 (**Figure 1**). As AA have predominantly West African ancestry (79.9%³⁵), meta-analysis of WA and AA was also conducted, and two additional hits were identified (**Table 4**). The minor (C) allele of rs7797481 (*ORC5*) was associated with higher LDLC in both WA and AA. The association was quite different among the EA (β -0.38, $p=0.053$; p for heterogeneity between WA and EA=0.0028), and their inclusion in the meta-analysis reduced the statistical significance of the association (**Figure 2a**). In contrast, for the association between the minor (T) allele of 20:60973327 (chromosome:position [build 37]) and higher TG in the meta-analysis of WA and AA, the effect size was similar among EA but with a large standard error (β 0.13, SE=0.11), making the association above our threshold for statistical significance when they were included in the meta-analysis (**Figure 2b**; p for heterogeneity between WA and EA=0.50).

We sought replication for 1,853 previously identified SNP-trait associations from large genome-wide studies of serum lipids¹⁻⁸. Of 1,534 SNP-trait associations available to be tested in our data, 947 were in a consistent direction (62%), and 144 of those replicated (consistent direction and $p < 0.05$). This represents an exact replication rate of 9.4% (**Supplemental Table 4**). Of the SNP-trait associations that did not replicate in the exact analysis, 702 had variants that were suitable for evaluation using “local” replication (an LD-based replication technique, see Methods). An additional 27 SNP-trait associations in 24 loci were replicated using these criteria (**Supplemental Table 5**). In total, we replicated 83 out of 303 loci (27%).

Discussion

In this GWAS of serum lipids in 4,317 continental Africans, we identified 23 novel lipids loci. Of these, 14 loci were available to be tested in 9,542 African Americans for replication and two of these associations replicated (*GPNMB*-TG and *ENPP1*-TG). Upon meta-analysis of the Africans and African Americans, two additional associations were found (rs138282551-TG and *TLL2*-CHOL). Given genetic and phenotypic diversity between WA and EA, we also conducted analyses excluding EA. Two additional novel associations were identified with meta-analysis of WA and AA: *ORC5*-LDLC and chr20:60973327-CHOL.

Our focus on African populations may have allowed for novel TG findings based on both the understudied environmental and genomic context of these individuals. The vast majority of our findings were for triglycerides. In a recent study assessing transferability of known lipids loci from European ancestry populations to Ugandans, approximately 70% of HDLC and LDLC loci were transferable to Ugandans, whereas only 10.5% of TG loci were transferable. The authors suggested that this difference may have resulted from differences in the environmental context³⁶. While environmental context is a likely contributor to variation in TG levels, differences in genetic context may also be important.

Descriptive data from a recent serum lipids meta-analysis of nearly 100 cohorts under the aegis of the CHARGE consortium offers useful comparisons in terms of ancestry-level differences in TG¹⁶. Mean TG levels in AA are relatively low compared to those of other ancestries (~ 100 mg/dl for all included AA studies, except ARIC [115 mg/dl]) and similar to WA (mean 98.3 mg/dl), despite dramatically different environmental backgrounds. In contrast, the mean values for EA were markedly higher: mean 168 mg/dl. This distribution suggests that genomic factors underlie some of the variation across populations in TG levels.

As the numbers of studies and samples with African ancestry are generally low relative to European ancestry, studies including individuals of African ancestry are generally considered as a single group. However, this grouping may not be appropriate given the genetic diversity among those who could be categorized as having African ancestry and the diverse environmental backgrounds in which those of African ancestry reside. Given that our study population included individuals from both West and East Africa, who are distinct in terms of genetics and diet, we conducted analyses separately for WA. The exclusion of EA from the analyses, leaving the two groups with predominantly West African ancestry, allowed for the discovery of novel loci, presumably because of genetic similarities between the WA and AA, in contrast to the EA. The AA are still not an ideal replication sample for the WA, however, given both their admixture (with an average of 20% European ancestry) and the environmental background, which is quite distinct between those of African ancestry living in West Africa and in the United States. Thus, the differences in observed associations between these groups are likely to represent a combination of factors, including genetic differences, environmental context, and false positives, in unknown proportions.

One emerging potential explanation for differences by ancestry in associations with lipids is through an effect of diet on the gut microbiome³⁷. Gut microbiome features have been shown to underlie some of the interindividual variation in lipids-related protein levels³⁸. Transplant of microbiota from humans with

dyslipidemia to germ-free mice led to a recapitulation of the lipid phenotype³⁹. Diet differs for each of the groups included in this analysis⁴⁰⁻⁴², and this could contribute to differences in the associations observed. Further analyses including individuals of varying degrees of genetic and environmental differences, particularly those including microbiome data, will help to advance our understanding of these relationships in the context of serum lipids.

Higher BMI and T2D are both associated with dyslipidemia, and Mendelian randomization studies suggest that BMI and T2D are causal for dyslipidemia^{43,44}. As such, we sought to establish whether associations were independent of an effect on either of these traits by conducting all analyses with and without adjustment for BMI and T2D. Given that this study is nested within a case-control study for T2D, and, thus, participants had higher proportions of T2D and elevated BMI compared to what would be expected in a general population study, it was of particular importance to account for the potential impact of these factors on our analyses. In our experience, however, adjusting for either of these factors did not alter the findings considerably. For all of our main findings, statistical significance of the associations was quite similar across models, although some of the minor fluctuations meant that some associations in some models met our criteria for statistical significance, while others just missed. For all of our 23 lead associations, all other models yielded associations with $p \leq 10^{-6}$, and for 22 of them, all other models yielded association with $p \leq 10^{-7}$. Based on these findings, while the presence or absence of BMI or T2D in the model may have made some signals easier to detect, they did not wipe out any associations, as would have been expected if an association with lipids was occurring through an effect on BMI or T2D.

The proportion of known lipids loci that replicated among Africans was 27%. Attempted replication of the previously reported index SNPs yielded a replication rate of only 9.8%. This was increased to 27% by using a linkage disequilibrium-based “local” replication method, which accounted for the generally lower linkage disequilibrium among African ancestry individuals by evaluating all of the variants in LD

with the index variant in the previously reported population (with correction for the number of variants tested). These proportions are similar to those reported for replication of known type 2 diabetes loci in Africans, in which 15% replicated exactly and an additional 15% replicated using a local replication method²¹, giving us confidence that these findings are in a reasonable range. While this relatively low proportion reflects in part the smaller sample size present for the Africans compared to the very large meta-analyses in which many of these loci were identified, there are other factors that may have contributed. For instance, there is evidence of differential associations across ancestry⁴⁵. Additionally, the environmental background of the Africans compared to those more well-represented populations could play a role. A recent analysis of the transferability of lipids loci to Ugandans reported that among major lipids loci ($p < 10^{-100}$ in the largest European ancestry meta-analysis), 71% of those that were not transferable to Ugandans showed evidence of pleiotropy with BMI among European ancestry populations, while none of those that were transferable were also associated with BMI (these mostly represented well-established lipids pathways). The authors theorize that these discrepancies may stem from gene-environment interactions, with effects on lipids not shared across different environmental backgrounds. For these major known lipids loci, evidence for and against transferability to our West and East African study populations matched very closely the results in the study in Ugandans (90% concordance)³⁶.

In our analysis, the A allele for rs9375831 (intron, *ENPP1*) was associated with lower TG in Africans, and this association was replicated in AA. This variant is in promoter and enhancer histone marks for liver tissue (Haploreg⁴⁶: <https://pubs.broadinstitute.org/mammals/haploreg/haploreg.php>). This gene is known for a role in the development of insulin resistance, of which hypertriglyceridemia is a hallmark. ENPP1, also known as PC-1, directly interacts with the α subunit of the insulin receptor, inhibiting insulin signaling⁴⁷. Transgenic mouse models of overexpression of *ENPP1* in muscle and liver⁴⁸ and adipose tissue⁴⁹ develop insulin resistance, including hypertriglyceridemia⁴⁹. It has been hypothesized that higher

ENPP1 expression leads to reduced triglyceride storage capacity, increased fat deposition in the liver, and subsequent systemic abnormalities associated with insulin resistance⁴⁹. Increased expression of *ENPP1* in adipose tissue of non-obese, non-diabetic individuals is associated with a reduction in insulin sensitivity⁵⁰. Variants in this gene (not in LD with rs9375831) have been associated with insulin resistance and related traits in multiple analyses²⁹⁻³⁴, with evidence for elevated triglycerides reported for some^{29,34}, but not all studies³⁰⁻³³.

We found that the minor (C) allele for rs706014 (335 bp 5' of *GPPMB*) was associated with decreased serum triglycerides. According to Gene-Tissue Expression project (GTEx) data, this variant has 44 expression quantitative trait loci (eQTL) hits in a variety of tissues. In adipose tissue, this variant is associated with circulating levels of *GPNMB*, *NUPL2*, and *KLHL7-AS1*⁴⁶. *GPNMB* is involved in pro-inflammatory responses and tissue repair. Animals fed high fat diets experienced large increases in *Gpnmb* expression in ascending aorta of rabbits (15-fold)⁵¹ and visceral adipose tissue of mice (30-fold change)⁵². A lysosomal storage disease involving intracellular accumulation of unesterified cholesterol, Niemann-Pick disease, is also associated with elevated levels of *Gpnmb* in tissue and plasma in both model organisms and in humans, with levels increasing with age and disease progression⁵¹. The minor alleles of two variants in moderate LD with rs706014 (rs199351 and rs199347, $R^2 \sim 0.3$ among EUR), which are also eQTLs for *GPNMB* (Haploreg), have been associated with decreased risk of Parkinson's Disease (PD)⁵³⁻⁵⁵. Higher lipid levels, including specifically triglycerides⁵⁶, have been associated with decreased PD risk^{57,58}. Emerging evidence of a complex interplay between lipids and α -synuclein, a protein that is strongly associated with PD, in disease risk has even led some researchers to suggest that PD should be considered a lipidopathy⁵⁹.

We report that the minor allele (T) for rs138282551 (intergenic) is associated with higher TG among Africans and AA. This variant alters motifs for CHD2, E2F, and TCF4⁶⁰. Studies of E2F1 knockout mice

demonstrated that E2F1 is involved in global transcriptional regulation of *de novo* lipid synthesis. In mice, E2F1 expression and activity in the liver increased in response to feeding and insulin levels, and E2F1 has been characterized as a major regulator of lipid metabolism⁶¹. Deletion of *E2f1* in cellular and mouse models leads to decreased *Pcsk9* expression and increased *LDLR* expression, and abnormal cholesterol accumulation in the liver⁶².

In our meta-analysis of AFR and AA, the minor allele of rs147706369 (intronic, *TLL2*) was associated with increased total cholesterol. This variant alters a regulatory motif for sterol regulatory element-binding proteins (SREBPs)⁶⁰, which are indirectly required for cholesterol biosynthesis. SREBPs are a family of transcription factors that control the expression of a range of enzymes involved in cholesterol, fatty acid, and triglyceride synthesis, and have been called the “Master Regulators of Lipid Homeostasis”⁶³.

The minor allele of rs7797481 (*ORC5*, intron) was associated with higher LDLC in WA and AA. This variant is an eQTL for *ORC5L* in whole blood⁶⁴. The origin recognition complex (ORC) proteins play a central role in the initiation of DNA replication. It is unclear how this could be related to the current findings. This variant also alters regulatory motifs for FOXA and GR⁶⁰. Members of the FOXA family have been shown to alter lipid metabolism through regulation of multiple lipids pathways. Foxa1 inhibits the expression of many of the genes involved in triglyceride synthesis and accumulation and in VLDL synthesis, including *APOB* (which encodes the major structural component of LDLC), in human hepatocytes⁶⁵. Foxa1 also inhibits triglyceride synthesis and accumulation in human hepatocytes, and *FOXA1* is downregulated in non-alcoholic fatty liver disease⁶⁶.

Our conclusions should be evaluated in the context of our study’s strengths and weaknesses. One strength of this study is the large sample size for a study of continental Africans, who remain understudied in genomic analyses for serum lipids. Notably, all these individuals were recruited as part of one study, with identical questionnaires and assay methodology, to limit the potential for observed

differences in East vs. West Africans being created by study methods. Differences among African ancestry individuals are explored by considering West and East Africans in stratified analyses rather than treating “African ancestry” as a monolithic category. The dataset of AA, all imputed together using the same reference panel, brought an additional large number of African ancestry samples to the analysis. We used both an exact and a local replication method for assessing transferability of previously identified loci to Africans thereby ensuring that we had accounted for the LD-based differences in the discovery and replication datasets. It would have been ideal to have replication cohorts that were a better match for the environmental and genomic background of the West Africans in our study for replication. With the replication cohorts that we used, we were unable to determine whether association differences represented false positives, differences in environment, or differences in genomic context. We look forward to future studies in which these types of analyses are possible.

In summary, we conducted a genome-wide association study of serum lipids in Africans in Nigeria, Ghana, and Kenya, with meta-analysis with and replication in AA. Among the overall study findings, several of the novel loci are important in TG or cholesterol synthesis, storage or turnover. While functional work will be useful to confirm and to understand the biological mechanisms underlying these associations, this study demonstrates the power of conducting large-scale genomic analyses in Africans.

Table 1. Descriptive Statistics for Included Participants

	<i>West Africans</i>	<i>East Africans</i>	<i>All</i>
N	3583	734	4317
Age (years)	50.6 (13.2)	55.4 (10.2)	51.4 (12.9)
BMI (kg/m ²)	26.6 (5.7)	26.4 (5.4)	26.6 (5.6)
Type 2 Diabetes*	50.7%	50.7%	50.7%
HDLC (mg/dl)	45.3 (16.7)	47.4 (16.8)	45.6 (16.7)
LDLC (mg/dl)	131.1 (46.6)	130.0 (46.3)	130.0 (46.3)
TG (mg/dl)	98.3 (48.3)	167.5 (91.9)	110.0 (63.5)
CHOL (mg/dl)	196.8 (52.9)	206.2 (58.0)	198.4 (53.9)
TG/HDLC	2.63 (2.13)	3.96 (2.77)	2.85 (2.3)

*Mean (SD) reported, except where indicated. *Proportion shown.*

BMI: body Mass Index; TG: triglycerides; HDLC: high density lipoprotein cholesterol; LDLC: Low density lipoprotein cholesterol; CHOL: total cholesterol

Table 2. Genome-Wide Lipids Associations $p < 5E-08$ among Africans (n=4,317)

Variant	Nearest Gene ¹	Chr:Pos (Bld 37)	Effect/ Non- Effect Allele	EA	1000 Genomes AFR/AMR/EAS/EUR/SAS	Beta	StdErr	P-value	Trait	Main	+BMI	+T2D	+T2D, BMI
rs77612115	<i>CDC73</i>	1:194058083	C/A	0.011	-	0.851	0.155	4.1E-08	TG	*			
rs62122280	<i>DDX1</i> (intronic)	2:15768981	G/C	0.019	0.01/0.29/0.01/0.39/0.17	-0.476	0.087	4.1E-08	TG				
rs138202830	<i>GBE1</i>	3:81361244	T/C	0.011	0.001/0/0/0/0	-0.672	0.115	4.6E-09	TG	*			
rs183141928	<i>ARAP2</i>	4:35816686	A/T	0.012	0/0/0/0.01/0	-0.930	0.165	1.9E-08	TG			*	
5:86213255	<i>RP11-72L22.1</i> (32 kb)	5:86213255	C/CT	0.012	-	0.808	0.144	2.2E-08	TG				*
rs62383172	<i>MACROH2A1</i> (78 kb)	5:134591437	C/T	0.011	0.01/0.05/0/0.07/0.01	-0.685	0.125	4.1E-08	TG/HDLC	*			
rs141894016	<i>TINAG</i>	6:54308947	T/A	0.012	0.009/0/0/0/0	-0.570	0.102	2.1E-08	TG/HDLC				*
rs188701119	<i>TINAG</i>	6:54323870	T/C	0.012	0.01/0/0/0/0	-0.533	0.097	3.9E-08	TG				*
rs9375831	<i>ENPP1</i> (intronic)	6:132164687	A/T	0.015	0.02/0.27/0.18/0.28/0.37	-0.547	0.100	4.1E-08	TG	*			
rs706014	<i>GPNMB</i> (335 bp)	7:23275250	C/G	0.010	0.01/0.18/0.23/0.36/0.28	-0.753	0.126	2.5E-09	TG			*	
rs144563873	<i>COL1A2</i>	7:93935161	A/G	0.012	0.01/0.04/0.01/0.01/0.04	-0.656	0.117	1.9E-08	TG	*			
rs75741534	<i>VIRMA</i> (13 kb)	8:95579637	C/T	0.020	0.01/0.14/0.004/0.14/0.20	0.510	0.088	7.3E-09	TG			*	
rs140987192	<i>PBX3</i> (11 kb)	9:128498293	T/C	0.013	0.01/0/0/0/0	0.587	0.103	1.2E-08	TG				*
rs140568748	<i>LOC112268059</i> <i>C10orf35</i> (8.5 kb)	10:9349704	CA/C	0.042	0.04/0.12/0.10/0.19/0.13	0.325	0.058	1.7E-08	TG	*			
rs72801121		10:71401897	G/A	0.011	0.01/0.11/0.09/0.16/0.19	-0.690	0.123	2.0E-08	TG/HDLC				*
11:102724515	<i>MMP3</i> (10 kb)	11:102724515	A/T	0.013	-	0.686	0.122	2.2E-08	TG	*			
rs150742617	<i>TMEM117</i> (intronic)	12:44256788	CA/C	0.014	0.01/0.13/0.23/0.15/0.17	0.601	0.107	1.8E-08	TG		*		
rs75579422	<i>ERICH6B</i> (intronic)	13:46170293	T/C	0.014	0.02/0/0/0/0	-0.547	0.099	3.7E-08	TG				*
rs111590558	<i>HTR2A</i>	13:47737090	T/C	0.033	0.02/0/0.01/0/0	0.346	0.062	2.8E-08	LDL	*			
17:70312798	<i>KCNJ2</i>	17:70312798	A/T	0.048	-	0.319	0.057	2.9E-08	TG/HDLC				*
rs75360819	<i>CDH2</i>	18:25928153	A/G	0.013	0.006/0/0/0/0	-0.593	0.104	1.1E-08	TG				*
rs56946445	<i>DCC</i> (intronic)	18:50836315	T/C	0.016	-	-0.611	0.108	1.5E-08	TG	*			
rs7281821	<i>APP</i> (intronic)	21:27457717	C/T	0.199	0.17/0.11/0/0.05/0.05	0.165	0.029	1.1E-08	TG/HDLC		*		
rs144086909	<i>LARGE</i> (intronic)	22:33925322	A/T	0.011	0.008/0/0/0.003/0	-0.639	0.115	2.7E-08	TG/HDLC			*	

Notes: Red indicates $p < 5E-8$; orange $5E-8 < p \leq 1E-8$; gold $1E-8 < p \leq 1E-7$; yellow $1E-7 < p \leq 1E-6$; models indicated with an asterisk were the most statistically significant results and are shown in the table. ¹Nearest gene to variant is listed, with distance show in parenthesis where < 50 kb.

Table 3. Novel Signals in Meta-Analysis of Africans and African Americans

rsID	Chr:Pos ¹	Eff/Non-Eff Alleles	1000 Genomes MAF AFR/AMR/EAS/EUR/SAS	Trait ²	Analysis	N	EAF	Beta	SE	P-Value
rs138282551 <i>Intergenic</i>	1:81649749	T/C	0.03/0/0/0.002/0	TG	Africans	4292	0.04	0.29	0.058	4.5E-07
					Afr. Americans	8537	0.03	0.14	0.049	3.5E-03
					Meta-analysis	12829	0.035	0.20	0.037	4.6E-08
rs147706369 <i>TLL2</i> <i>Intronic</i>	10:98213172	T/C	0.016/0/0/0/0	CHOL	Africans	4300	0.016	0.34	0.090	1.5E-04
					Afr. Americans	9217	0.015	0.28	0.067	2.6E-05
					Meta-analysis	13517	0.016	0.30	0.054	1.7E-08

¹Bld. 37; ²All traits inverse normal transformed.

Table 4. Novel Signals in Meta-Analysis of West Africans and African Americans

rsID	Chr:Pos ¹	Eff/Non-Eff Alleles	1000 Genomes MAF AFR/AMR/EAS/EUR/SAS	Trait ²	Analysis	N	EAF	Beta	SE	P-Value
rs7797481	7:103770056	C/G	0.02/0.13/0/0.20/0.07	LDLC	West Africans	3566	0.028	0.25	0.075	8.4E-04
<i>ORC5</i>					Afr. Americans	8501	0.057	0.16	0.035	2.5E-06
intronic					Meta-analysis	12067	0.052	0.18	0.031	1.4E-08
-	20:60973327	T/C	--	CHOL	West Africans	3565	0.110	0.21	0.049	2.2E-05
					Afr. Americans	8901	0.105	0.13	0.035	2.2E-04
					Meta-analysis	12466	0.107	0.16	0.029	4.4E-08

¹Bld. 37; ²All traits inverse normal transformed.

Acknowledgements

The authors would like to gratefully acknowledge the AADM collaborators, including Olufemi Fasanmade, Thomas Johnson, Johnnie Oli, Godfrey Okafor, Benjamin A. Eghan Jr., Kofi Agyenim-Boateng, Jokotade Adeleye, Williams Balogun, Clement Adebamowo, Albert Amoah, Joseph Acheampong, and Duncan Ngare (deceased). The views expressed in this manuscript are those of the authors and do not necessarily represent the views of the NIH.

Funding Sources

This project was largely supported by the Intramural Research Program of the National Human Genome Research Institute of the National Institutes of Health (NIH) through the Center for Research on Genomics and Global Health (CRGGH). The CRGGH is also supported by the National Institute of Diabetes and Digestive and Kidney Diseases and the Office of the Director at the NIH (Z01HG200362). Support for participant recruitment and initial genetic studies of the AADM study was provided by NIH grant No. 3T37TW00041-03S2 from the Office of Research on Minority Health.

Disclosures

None

Figure Legends

Figure 1: Manhattan plot for TG in all Africans (A.) and in West Africans only (B.). The grayed points represent variants within 1 MB of known lipids loci.

Figure 2: Novel Associations in Meta-Analysis of West Africans and African Americans. Forest plots

showing the associations of rs7797481 (*ORC5*) and LDLC (A.) and 20:60973327 and TG (B.) across included African Ancestry cohorts.

References

1. Liu, D.J. *et al.* Exome-wide association study of plasma lipids in >300,000 individuals. *Nat Genet* **49**, 1758-1766 (2017).
2. Hoffmann, T.J. *et al.* A large electronic-health-record-based genome-wide study of serum lipids. *Nat Genet* **50**, 401-413 (2018).
3. Klarin, D. *et al.* Genetics of blood lipids among ~300,000 multi-ethnic participants of the Million Veteran Program. *Nat Genet* **50**, 1514-1523 (2018).
4. Do, R. *et al.* Common variants associated with plasma triglycerides and risk for coronary artery disease. *Nat Genet* **45**, 1345-52 (2013).
5. Teslovich, T.M. *et al.* Biological, clinical and population relevance of 95 loci for blood lipids. *Nature* **466**, 707-13 (2010).
6. Willer, C.J. *et al.* Discovery and refinement of loci associated with lipid levels. *Nat Genet* **45**, 1274-1283 (2013).
7. Kim, T., Park, A.Y., Baek, Y. & Cha, S. Genome-Wide Association Study Reveals Four Loci for Lipid Ratios in the Korean Population and the Constitutional Subgroup. *PLoS One* **12**, e0168137 (2017).
8. Peloso, G.M. *et al.* Association of low-frequency and rare coding-sequence variants with blood lipids and coronary heart disease in 56,000 whites and blacks. *Am J Hum Genet* **94**, 223-32 (2014).
9. Bentley, A.R., Callier, S. & Rotimi, C.N. Diversity and inclusion in genomic research: why the uneven progress? *J Community Genet* **8**, 255-266 (2017).
10. Auton, A. *et al.* A global reference for human genetic variation. *Nature* **526**, 68-74 (2015).
11. Garske, K.M. *et al.* Reverse gene-environment interaction approach to identify variants influencing body-mass index in humans. *Nat Metab* **1**, 630-642 (2019).
12. Kim, M., Yoo, H.J., Lee, H.J. & Lee, J.H. Longitudinal interaction between APOA5 -1131T>C and overweight in the acceleration of age-related increase in arterial stiffness through the regulation of circulating triglycerides. *Hypertens Res* **42**, 241-248 (2019).
13. Farook, V.S. *et al.* Genetics of serum carotenoid concentrations and their correlation with obesity-related traits in Mexican American children. *Am J Clin Nutr* **106**, 52-58 (2017).
14. Wu, Y. *et al.* Short-term exposure to air pollution and its interaction effects with two ABO SNPs on blood lipid levels in northern China: A family-based study. *Chemosphere* **249**, 126120 (2020).
15. Khodarahmi, M., Kahroba, H., Jafarabadi, M.A., Mesgari-Abbasi, M. & Farhangi, M.A. Dietary quality indices modifies the effects of melanocortin-4 receptor (MC4R) rs17782313 polymorphism on cardio-metabolic risk factors and hypothalamic hormones in obese adults. *BMC Cardiovasc Disord* **20**, 57 (2020).
16. Bentley, A.R. *et al.* Multi-ancestry genome-wide gene-smoking interaction study of 387,272 individuals identifies new loci associated with serum lipids. *Nat Genet* **51**, 636-648 (2019).
17. de Vries, P.S. *et al.* Multiancestry Genome-Wide Association Study of Lipid Levels Incorporating Gene-Alcohol Interactions. *American Journal of Epidemiology* **188**, 1033-1054 (2019).

18. Kilpeläinen, T.O. *et al.* Multi-ancestry study of blood lipid levels identifies four loci interacting with physical activity. *Nature Communications* **10**, 376 (2019).
19. Rotimi, C.N. *et al.* A genome-wide search for type 2 diabetes susceptibility genes in West Africans: the Africa America Diabetes Mellitus (AADM) Study. *Diabetes* **53**, 838-41 (2004).
20. Rotimi, C.N. *et al.* In search of susceptibility genes for type 2 diabetes in West Africa: the design and results of the first phase of the AADM study. *Ann Epidemiol* **11**, 51-8 (2001).
21. Adeyemo, A.A. *et al.* ZRANB3 is an African-specific type 2 diabetes locus associated with beta-cell mass and insulin response. **10**, 1-12 (2019).
22. Friedewald, W.T., Levy, R.I. & Fredrickson, D.S.J.C.c. Estimation of the concentration of low-density lipoprotein cholesterol in plasma, without use of the preparative ultracentrifuge. **18**, 499-502 (1972).
23. Adeyemo, A.A. *et al.* Evaluation of Genome Wide Association Study Associated Type 2 Diabetes Susceptibility Loci in Sub Saharan Africans. *Front Genet* **6**, 335 (2015).
24. Liu, Z. *et al.* Admixture mapping identifies genetic regions associated with blood pressure phenotypes in African Americans. *PloS one* **15**, e0232048-e0232048 (2020).
25. Adeyemo, A. *et al.* A Genome-Wide Association Study of Hypertension and Blood Pressure in African Americans. *PLOS Genetics* **5**, e1000564 (2009).
26. Willer, C.J., Li, Y. & Abecasis, G.R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics (Oxford, England)* **26**, 2190-2191 (2010).
27. Ramos, E. *et al.* Replication of genome-wide association studies (GWAS) loci for fasting plasma glucose in African-Americans. *Diabetologia* **54**, 783-788 (2011).
28. Agyemang, C. *et al.* Obesity and type 2 diabetes in sub-Saharan Africans – Is the burden in today's Africa similar to African migrants in Europe? The RODAM study. *BMC Medicine* **14**, 166 (2016).
29. Tasic, I. *et al.* The association of PC-1 (ENPP1) K121Q polymorphism with metabolic syndrome in patients with coronary heart disease. *Clin Chim Acta* **377**, 237-42 (2007).
30. Frittitta, L. *et al.* The Q121 PC-1 Variant and Obesity Have Additive and Independent Effects in Causing Insulin Resistance. *The Journal of Clinical Endocrinology & Metabolism* **86**, 5888-5891 (2001).
31. Baratta, R. *et al.* Evidence for genetic epistasis in human insulin resistance: the combined effect of PC-1 (K121Q) and PPAR γ 2 (P12A) polymorphisms. *Journal of Molecular Medicine* **81**, 718-723 (2003).
32. Abate, N. *et al.* Genetic Polymorphism PC-1 K121Q and Ethnic Susceptibility to Insulin Resistance. *The Journal of Clinical Endocrinology & Metabolism* **88**, 5927-5934 (2003).
33. Pizzuti, A. *et al.* A polymorphism (K121Q) of the human glycoprotein PC-1 gene coding region is strongly associated with insulin resistance. **48**, 1881-1884 (1999).
34. Tanyolac, S. *et al.* Genetic variants of the ENPP1/PC-1 gene are associated with hypertriglyceridemia in male subjects. *Metabolic syndrome and related disorders* **7**, 543-548 (2009).
35. Chen, G. *et al.* Development of admixture mapping panels for African Americans from commercial high-density SNP arrays. *BMC Genomics* **11**, 417 (2010).
36. Kuchenbaecker, K. *et al.* The transferability of lipid loci across African, Asian and European cohorts. *Nat Commun* **10**, 4330 (2019).
37. Villette, R. *et al.* Unraveling Host-Gut Microbiota Dialogue and Its Impact on Cholesterol Levels. **11**(2020).
38. Zhernakova, D.V. *et al.* Individual variations in cardiovascular-disease-related protein levels are driven by genetics and gut microbiome. *Nature Genetics* **50**, 1524-1532 (2018).

39. Le Roy, T. *et al.* The intestinal microbiota regulates host cholesterol homeostasis. *BMC Biology* **17**, 94 (2019).
40. Hansen, A.W. *et al.* Dietary patterns, food and macronutrient intakes among adults in three ethnic groups in rural Kenya. *Public Health Nutrition* **14**, 1671-1679 (2011).
41. Galbete, C. *et al.* Food consumption, nutrient intake, and dietary patterns in Ghanaian migrants in Europe and their compatriots in Ghana. *Food Nutr Res* **61**, 1341809 (2017).
42. Rodriguez, L.A. *et al.* Differences in Diet Quality among Multiple US Racial/Ethnic Groups from the Mediators of Atherosclerosis in South Asians Living in America (MASALA) Study and the Multi-Ethnic Study of Atherosclerosis (MESA). *The Journal of Nutrition* **150**, 1509-1515 (2020).
43. Viitasalo, A. *et al.* Abdominal adiposity and cardiometabolic risk factors in children and adolescents: a Mendelian randomization analysis. *Am J Clin Nutr* **110**, 1079-1087 (2019).
44. Wang, N. *et al.* Type 2 Diabetes and Adiposity Induce Different Lipid Profile Disorders: A Mendelian Randomization Analysis. *J Clin Endocrinol Metab* **103**, 2016-2025 (2018).
45. Wojcik, G.L. *et al.* Genetic analyses of diverse populations improves discovery for complex traits. *Nature* **570**, 514-518 (2019).
46. Ward, L.D. & Kellis, M. HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. *Nucleic Acids Research* **40**, D930-D934 (2011).
47. Maddux, B.A. & Goldfine, I.D.J.D. Membrane glycoprotein PC-1 inhibition of insulin receptor function occurs via direct interaction with the receptor alpha-subunit. **49**, 13-19 (2000).
48. Maddux, B.A. *et al.* Overexpression of the insulin receptor inhibitor PC-1/ENPP1 induces insulin resistance and hyperglycemia. **290**, E746-E749 (2006).
49. Pan, W. *et al.* Metabolic consequences of ENPP1 overexpression in adipose tissue. *American journal of physiology. Endocrinology and metabolism* **301**, E901-E911 (2011).
50. Frittitta, L. *et al.* Increased adipose tissue PC-1 protein content, but not tumour necrosis factor- α gene expression, is associated with a reduction of both whole body insulin sensitivity and insulin receptor tyrosine-kinase activity. **40**, 282-289 (1997).
51. Xu, J., Jüllig, M., Middleditch, M.J. & Cooper, G.J.S. Modelling atherosclerosis by proteomics: Molecular changes in the ascending aortas of cholesterol-fed rabbits. *Atherosclerosis* **242**, 268-276 (2015).
52. Choi, M.-S. *et al.* High-fat diet decreases energy expenditure and expression of genes controlling lipid metabolism, mitochondrial function and skeletal system development in the adipose tissue, along with increased expression of extracellular matrix remodelling- and inflammation-related genes. *British Journal of Nutrition* **113**, 867-877 (2015).
53. Nalls, M.A. *et al.* Identification of novel risk loci, causal insights, and heritable risk for Parkinson's disease: a meta-analysis of genome-wide association studies. *Lancet Neurol* **18**, 1091-1102 (2019).
54. Nalls, M.A. *et al.* Large-scale meta-analysis of genome-wide association data identifies six new risk loci for Parkinson's disease. *Nature Genetics* **46**, 989-993 (2014).
55. Murthy, M.N. *et al.* Increased brain expression of GPNMB is associated with genome wide significant risk for Parkinson's disease on chromosome 7p15.3. *neurogenetics* **18**, 121-133 (2017).
56. Vikdahl, M., Bäckman, L., Johansson, I., Forsgren, L. & Håglin, L. Cardiovascular risk factors and the risk of Parkinson's disease. *European Journal of Clinical Nutrition* **69**, 729-733 (2015).
57. Ng, A.S.L. & Tan, E.-K. Linking statins and lipids in Parkinson's disease. **32**, 807-809 (2017).
58. Rozani, V. *et al.* Higher serum cholesterol and decreased Parkinson's disease risk: A statin-free cohort study. *Mov Disord* **33**, 1298-1305 (2018).

59. Fanning, S., Selkoe, D. & Dettmer, U. Parkinson's disease: proteinopathy or lipidopathy? *NPJ Parkinson's disease* **6**, 3-3 (2020).
60. Kheradpour, P. & Kellis, M. Systematic discovery and characterization of regulatory motifs in ENCODE TF binding experiments. *Nucleic Acids Research* **42**, 2976-2987 (2013).
61. Denechaud, P.-D. *et al.* E2F1 mediates sustained lipogenesis and contributes to hepatic steatosis. *The Journal of Clinical Investigation* **126**, 137-150 (2016).
62. Lai, Q. *et al.* E2F1 inhibits circulating cholesterol clearance by regulating Pcsk9 expression in the liver. *JCI Insight* **2**(2017).
63. Eberlé, D., Hegarty, B., Bossard, P., Ferré, P. & Foulle, F. SREBP transcription factors: master regulators of lipid homeostasis. *Biochimie* **86**, 839-48 (2004).
64. Westra, H.-J. *et al.* Systematic identification of trans eQTLs as putative drivers of known disease associations. *Nature Genetics* **45**, 1238-1243 (2013).
65. Moya, M. *et al.* Foxa1 Reduces Lipid Accumulation in Human Hepatocytes and Is Down-Regulated in Nonalcoholic Fatty Liver. *PLOS ONE* **7**, e30014 (2012).
66. Sahini, N. & Borlak, J. Genomics of human fatty liver disease reveal mechanistically linked lipid droplet-associated gene regulations in bland steatosis and nonalcoholic steatohepatitis. *Translational Research* **177**, 41-69 (2016).

Figure 1. Manhattan Plots for TG in West and East Africans combined (A.) and in West Africans only (B.)

A.

B.

Figure 2. Novel Associations in Meta-Analysis of West Africans and African Americans