



Published in final edited form as:

*Circ Cardiovasc Genet.* 2010 August ; 3(4): 358–364. doi:10.1161/CIRCGENETICS.109.914267.

## Fine-Mapping in African Americans of Eight Recently Discovered Genetic Loci for Plasma Lipids: The Jackson Heart Study

Mary E. Keebler, MD<sup>\*</sup>, Rahul C. Deo, MD, PhD<sup>\*</sup>, Aarti Surti, BS, David Konieczkowski, BS, Candace Guiducci, BS, Noel Burt, BS, Sarah G. Buxbaum, PhD, Daniel F. Sarpong, PhD, Michael W. Steffes, MD, James G. Wilson, MD, Herman A. Taylor, MD, MPH, and Sekar Kathiresan, MD

Massachusetts General Hospital (M.E.K., R.C.D, S.K.), Boston, MA; Broad Institute (M.E.K., A.S., D.K., C.G., N.B., S.K.), Cambridge, MA, Jackson State University (S.G.B., D.F.S, H.A.T.), Jackson, MS, University of Minnesota (M.W.S.), Minneapolis, MN, GV (Sonny) Montgomery VA Medical Center (J.G.W.) Jackson, MS, University of Mississippi Medical Center (J.G.W., H.A.T.), Jackson, MS and Tougaloo College (H.A.T.), Tougaloo, MS

### Abstract

**Background**—Genome-wide association studies in cohorts of European descent have identified novel genomic regions as associated with lipids, but their relevance in African Americans remains unclear.

**Methods and Results**—We genotyped 8 index SNPs and 488 tagging SNPs across 8 novel lipid loci in the Jackson Heart Study, a community-based cohort of 4605 African Americans. For each trait, we calculated residuals adjusted for age, sex, and global ancestry and performed multivariable linear regression to detect genotype-phenotype association with adjustment for local ancestry. To explore admixture effects, we conducted stratified analyses in individuals with a high probability of 2 African ancestral alleles or at least 1 European allele at each locus. We confirmed 2 index SNPs as associated with lipid traits in African Americans, with suggestive association for 3 more. However, the effect sizes for 4 of the 5 associated SNPs were larger in the European local ancestry subgroup compared to the African local ancestry subgroup, suggesting that the replication is driven by European ancestry segments. Through fine-mapping, we discovered 3 new SNPs with significant associations, two with consistent effect on triglyceride levels across ancestral groups: rs636523 near *DOCK7/ANGPTL3* and rs780093 in *GCKR*. African LD patterns did not assist in narrowing association signals.

**Conclusions**—We confirm that 5 genetic regions associated with lipid traits in European-derived populations are relevant in African Americans. To further evaluate these loci, fine-mapping in larger African American cohorts and/or resequencing will be required.

---

**Corresponding Author:** Sekar Kathiresan, MD Cardiovascular Research Center and Center for Human Genetic Research Massachusetts General Hospital 185 Cambridge Street, CZPN 5.252 Boston, MA 02115 Ph: 617-643-6186 Fax: 617-507-7766 [skathiresan@partners.org](mailto:skathiresan@partners.org).

<sup>\*</sup>ME Keebler and RC Deo contributed equally to the work.

**Conflict of Interest Disclosures:** None

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

## Keywords

lipids; genetics; epidemiology; risk factors

---

## Introduction

Two principal goals of genome wide association (GWA) studies have been the identification of predictive genetic markers of human disease<sup>1</sup> and the unbiased discovery of novel mechanisms of disease pathogenesis. To date, the majority of GWA studies have been conducted in populations of European ancestry. Although this population homogeneity has facilitated statistical analyses, it remains unclear whether future predictive algorithms or therapies based upon recently discovered variants will be relevant in other racial groups, such as African Americans.

African Americans represent an admixed population, carrying segments of DNA on their chromosomes that are derived from both European American (~20% of chromosomal regions) and West African (~80%) ancestral populations.<sup>2</sup> These ancestral populations were subjected to differing demographic history, resulting in different allele frequencies and linkage disequilibrium patterns in their genomes.<sup>2</sup> Thus, it is possible that specific genetic variants associated with disease in European populations may not be associated in African American individuals. In addition to a potential clinical utility for predicting and treating disease in worldwide populations, it has, in fact, been proposed that carrying out genetic association analysis in different ethnic populations with different patterns of linkage disequilibrium will be useful to narrow genetic loci and potentially identify causal variants.<sup>3</sup>

Using a GWA approach, we and others identified novel loci as associated with LDL-cholesterol (LDL), HDL-cholesterol (HDL), and/or triglycerides (TG) in populations of European descent.<sup>4, 5</sup> We recently confirmed the association of 1 newly discovered locus (rs646776 in 1p13 near *CELSR2/PSRC1/SORT1*) with LDL in the Third National Health and Nutritional Examination Survey (NHANES III), a multiethnic cohort consisting of non-Hispanic black, Mexican American, and non-Hispanic white individuals.<sup>6</sup> However, there was variable replication among the three ethnicities for the remaining loci, which may have been due to limited power to detect associations, differences in linkage disequilibrium patterns, or fundamental differences in the genetic architecture of the quantitative traits. Furthermore, statistical analysis of traits in admixed populations is more complex due to population stratification, which can lead to both false negative and positive associations. We have recently implemented a method that utilizes genotypes from a dense panel of markers differentiated in frequency between European and African populations to allow accurate measurement of ancestry across the genome and correct for issues of stratification.<sup>2</sup>

In the current study, we directly addressed the following issues governing the relevance of specific genetic variants and, more broadly, genetic loci discovered in European-derived populations to African Americans. First, we sought to validate eight recently reported “index” polymorphisms for LDL, HDL, and TG in a large community-based cohort of African American individuals. This analysis was done with stringent control for local ancestry to account for population stratification. Furthermore, we densely tagged the eight genomic regions in linkage disequilibrium with the index polymorphisms to determine whether these loci were relevant to variation in plasma lipids in African Americans. Finally, we used local-ancestry based analysis to explore whether fine-mapping in an African American cohort would serve the additional goal of narrowing the genetic loci of interest and potentially identifying causal variants.<sup>2</sup>

## Methods

### Participants

The Jackson Heart Study (JHS) is a community-based, observational study developed as an extension of the Jackson, Mississippi cohort of the Atherosclerosis Risk in Communities (ARIC) study. The overall cohort consists of 5,301 individuals and was designed to identify risk factors for common cardiovascular diseases in African Americans. Of the 4,605 participants whose consent allowed their inclusion in the current study, 1,448 are part of a family study.

### Lipid Measurements

Fasting LDL, HDL, and TG were measured in each participant using previously described methods.<sup>2</sup>

### Tag SNP Selection

Because the index SNPs were discovered via GWA in cohorts of European ancestry, we first identified the interval in which common variants were in LD with the index SNPs in the HapMap population of northern and western European ancestry (CEU) with a  $r^2$  threshold of  $\geq 0.25$ . Given the power of the original studies, this interval is likely to contain the causal SNP(s) of interest. Using Tagger<sup>7</sup>, we selected tag SNPs along this specified region in the Yoruban West African HapMap population (YRI) with an  $r^2$  threshold of 0.80 and minor allele frequency (MAF)  $\geq 1\%$ . We then forced these SNPs into Tagger to tag the same region in CEU with an  $r^2$  0.80 and MAF  $\geq 1\%$ .<sup>2</sup> In total, 488 SNPs were genotyped (1q42 near *GALNT2* = 13 SNPs, 19p13 near *CILP2/PBX4* = 145 SNPs, 7q11 near *TBL2/MLXIPL* = 33 SNPs, 8q24 near *TRIB1* = 53 SNPs, 1p13 near *CELSR2/PSRC1/SORT1* = 57 SNPs, 1p31 near *ANGPTL3* = 30 SNPs, 12q24 near *KCTD10* = 80 SNPs, 2p23 near *GCKR* = 77 SNPs).

### Genotyping

Genotyping was performed on the Sequenom platform, which utilizes matrix-assisted laser-desorption ionization time-of-flight mass spectroscopy as described previously.<sup>1</sup> SNPs with a genotype call rate  $< 90\%$  ( $n=90$ ) and individuals with genotyping success rates  $< 75\%$  ( $n=90$ ) were excluded from analysis. Hardy-Weinberg equilibrium was not used to judge genotyping quality given the established deviation of SNP frequencies from equilibrium in admixed populations.<sup>8</sup>

### Ancestry estimation

ANCESTRYMAP software<sup>9</sup> was used for estimation of global and local ancestry in JHS as previously described.<sup>2</sup> Briefly, the Hidden Markov Model underlying this program uses genotypes from a panel of densely spaced markers differentiated in frequency between African and European populations for ancestry inference. Two overlapping panels of ancestry informative markers (AIMs) were used for genotyping the 4,464 JHS individuals: 976 were genotyped on an older “Phase 2” Panel of 1,536 markers, and 3,488 were genotyped on an updated “Phase 3” Panel of 1,536 markers. These various Phases represent improvements in the admixture mapping panels in terms of information content, genome coverage, and SNP genotyping efficiency.

### Statistical Analysis

For subjects on lipid lowering medication ( $n=560$ ), an underlying untreated LDL value was estimated using imputation.<sup>10</sup> HDL and TG were log transformed. Residuals for each lipid trait were created using SAS 9.0 (Cary, NC) and were adjusted for age, gender, and global ancestry.

All association analyses were performed using MERLIN to account for family structure among related individuals in JHS and for imputation of missing genotypes based on inheritance patterns.<sup>11</sup> The multivariable adjusted residual was used as the phenotype to test phenotype-genotype association. We utilized two strategies to account for confounding by local ancestry. First, linear regression models were further adjusted for a continuous local ancestry variable derived from output of the ANCESTRYMAP program<sup>9</sup>, which takes as input AIM genotypes and uses a Hidden Markov Model to estimate local ancestry for individuals across the genome. In addition, we performed stratified genotype-phenotype association analyses in individuals with >95% probability of 2 African ancestral chromosomes (JHS\_afr) and in individuals with >95% probability of at least 1 European ancestral allele (JHS\_eur) at each locus of interest,<sup>2</sup> with a different JHS\_afr subpopulation (of ~1600–1800 individuals) and JHS\_eur subpopulation (of ~ 600 individuals) at each locus (e.g. *SORT1* locus). Within each of these subgroups, we evaluate the association of SNP genotype with the three lipid phenotype traits, computing an effect size and p-value.

Because there were 13 comparisons being considered for GWAS SNP replication, we applied a Bonferroni cutoff of 0.05/13 or 0.004 as our threshold of significance of these associations. SNPs were considered “suggestive” if p-values were in the range  $0.05 > p > 0.004$ . For declaration of statistical significance in the fine-mapping analysis, the Benjamini-Hochberg procedure (*p-adjust* in R 2.8.1) was performed to provide conservative estimates of tail-based false discovery rates (FDR). An FDR threshold of 0.15 was chosen for LDL, HDL, and TG.

Power calculations for lipid traits were performed by simulation, using a linear regression model of lipid plasma level (in standardized units) against genotype with minor allele frequencies ranging from 0.05–0.50 and testing a range of effect sizes ranging from 0.02–0.35 standardized units. This was performed for 2 purposes: replication of index SNPs in the entire JHS population, and discovery of new SNPs in the JHS\_afr subpopulation. For replication, the number of individuals was set at 3,300 (corresponding to unrelated members in JHS) with a p-value of 0.05/13=0.004 as threshold. For SNP discovery in JHS\_afr, which involved testing a large number of SNPs with low prior odds of association, the number of individuals was set at 1,700 and the critical p-value for declaring significance set at 0.05/300=0.00017. 500 simulations were performed at each combination of minor allele frequency and effect size (percentage of residual variance explained), and, for each minor allele frequency, the largest effect size at which the significance threshold was exceeded in >80% of the runs was reported.

## Results

Participant characteristics for individuals in JHS are displayed in Table 1. Because local ancestry estimates vary among loci, characteristics of local ancestry subgroups at the chr1p13 locus are shown as a representative sample. After excluding participants with TG  $\geq$  400mg/dL and/or low genotyping rate (<90%), genotype-phenotype data from 4,474 African American individuals were analyzed.

### Associations of Index Loci with Lipid Traits in African Americans

The genetic associations of blood lipids traits in JHS with eight index SNPs identified through GWA are shown in Table 2. Given that several index SNPs were associated with more than one lipid trait in the original studies, we evaluated a total of 13 SNP-lipid trait associations, defining significance as an unadjusted p-value of less than 0.004, with effect in the same direction as the original study. Associations with p-values  $> 0.004$  and  $< 0.05$  with consistent direction of change as the original study were considered suggestive.

At this threshold, 2 of the 8 index SNPs genotyped in the JHS were significant: 1 associated with TG (rs1260326) and 1 with LDL (rs646776). Three SNPs were suggestive (rs12130333, rs17145738, and rs10774708). rs1260326 in *GCKR* was strongly associated with plasma TG levels, decreasing TG by ~ 7.3% per copy of the major allele across the entire cohort ( $p=7.4E-05$ ). Two other loci were more weakly associated with TG: rs12130333 near *DOCK7/ANGPTL3*, and rs17145738 near *TBL2* on chromosome 7. The genotype at each of these SNPs resulted in a ~4% change in TG concentrations ( $p=0.04$  for both). Of the 3 SNPs associated with LDL in GWA studies, only rs646776 was significant in JHS ( $p=1.2E-04$ ), with each copy of the major allele increasing LDL ~3.3 mg/dL across the entire cohort. Finally, rs10774708 in *KCTD10* was modestly associated with HDL, increasing HDL by approximately 1% per copy of the major allele ( $p=0.04$ ).

### Local Ancestry Based Analysis Demonstrates Stronger Effect Sizes on the European Local Ancestral Background

In order to determine whether local ancestry at each index locus influences effect sizes, we performed stratified analyses comparing individuals with a >95% probability of African versus European ancestry in the eight index SNP regions. Representative demographic characteristics for the 1p13 subgroups are shown in Table 1. For 4 of the 5 index SNPs significantly/suggestively associated with lipid traits in African Americans (rs1260326, rs12130333, and rs17145738 with TG and rs646776 with LDL), European ancestry appeared to drive the associations (Table 2). For these SNPs the effect sizes on the European ancestral background exceeded those seen on the African ancestral background, ranging from 1.6 to 3 fold differences in magnitude of effect. Interestingly, rs4846914 near *GALNT2* was not associated with HDL in the overall analysis despite evidence for association on the African chromosomal segment ( $p_{\text{all}}=0.09$ ,  $p_{\text{afr}}=0.02$ ). However, this finding may be spurious and warrants replication in additional samples.

### Fine-Mapping Genomic Regions in African Americans

To discover potentially more informative variants in African Americans, we undertook dense fine mapping for each of the eight genetic loci within JHS. Controlling the false-discovery rate at <15% gave 1 significant association for LDL, 0 significant associations for HDL, and 2 significant associations for TG as detailed in Table 3 (see Methods).

For two of the six genomic regions densely genotyped and tested for association with TG, we identified SNPs with stronger associations with TG in African Americans than their corresponding index GWAS SNPs. Depicted in Figures 1 and 2 are SNPs in these two regions, 2p23 containing *GCKR* and 1p31 containing *DOCK7/ANGPTL3*, that were significant at an FDR of 0.15. rs780093, an intronic SNP in *GCKR*, was most strongly associated with TG concentrations, with each copy of the major allele resulting in a ~8% reduction in TG ( $p=4.6E-06$ ) and with similar effects on the African and European local ancestral backgrounds. rs780093 is in moderate/strong LD with rs1260326, the index SNP at this locus, in both YRI and CEU HapMap populations ( $r^2=0.65$  and 0.93, respectively). In fact rs780093 has subsequently been found to strongly associated with TG levels in a large meta-analysis of >19,000 individuals of European descent, with  $p=3.1E-29$ .<sup>12</sup> Although rs1260326 is a missense coding variant in the *GCKR* gene, our data suggest that rs780093 may either be the causal variant, or be more tightly linked to the causal variant at this locus by showing greater strength of association ( $p=4.6E-06$  vs.  $7.4E-05$ ). Although we cannot rule out statistical fluctuation, it is interesting that this nominally stronger association is predominantly driven by a more significant association on the African local ancestry background ( $p=0.004$  vs.  $p=0.12$ ) and is supported by a consistent effect size of 8% per allele in both African and European ancestry subgroups. As described above, such a similarity in effect size is lacking for rs1260326.



Within the 1p31 region near *DOCK7* and *ANGPTL3*, rs636523 had a markedly stronger association with TG than the index SNP rs12130333 ( $p=4.6E-04$  vs.  $p=0.04$ ). Each copy of the risk allele affected TG concentrations by ~4%. rs636523 is in moderate LD with the index SNP rs12130333 in the CEU Hapmap population ( $r^2=0.45$ ) but in weak LD in YRI ( $r^2=0.04$ ). The subsequent meta-analysis described above has in fact found a stronger association of rs636523 with TG ( $p=2.91E-07$ ), than the index SNP rs12130333 ( $p=0.0004$ ), in keeping with our results. Moreover, rs636523 demonstrates comparable effect sizes on both ancestral backgrounds rendering it a more plausible candidate as a causal variant. Although our work cannot shed light on the gene most likely regulated by this SNP, the *ANGPTL3* gene is known to harbor rare variants that regulate TG metabolism<sup>13</sup>; therefore, it is likely the causal gene at this locus.

rs629301 at the 1p13 locus was nominally more strongly associated with LDL compared with the index SNP rs646776 ( $p=5.4E-05$  vs.  $p=1.2E-04$ ; Figure 3). rs629301 (in blue) is in strong LD with rs646776 (in red) in both the CEU and Yoruban (YRI) HapMap populations ( $r^2=0.92$  in YRI and  $r^2=1.0$  in CEU); hence, the effect sizes per copy of the major allele are similar to those seen with rs646776 (~ 3.5 mg/dL), with similar disparities in effect sizes between the African and European local ancestry subgroups. We cannot exclude statistical fluctuation as driving the subtle differences in p-values. rs629301 was also found to be subsequently strongly associated with LDL in the large meta-analysis of European Ancestry individuals described above ( $p=2.147E-41$ )

Fine mapping analysis provided no further support for the importance of the other chromosomal regions in determining lipid profiles in African Americans.

### Using African Linkage Disequilibrium Patterns to Narrow GWA Signals

Applying false discovery rates of 0.15, we found no variants for LDL, HDL or TG that met our significance threshold.

### Discussion

In the Jackson Heart Study, we confirmed that five SNPs discovered in European-derived populations were associated with lipid traits in African Americans at a significant or suggestive level. Furthermore, we identified three SNPs with stronger associations in African Americans than the index discovery SNPs at their respective loci. Taken together, our findings suggested that the same genomic regions that predict lipid traits in European-derived populations are likely relevant in African Americans, but that the specific variants within each locus may differ among ethnic groups. Finally, we found that fine-mapping genomic regions on an African ancestral background has limited utility to narrow genomic regions of association *in an admixed cohort of this size*.

Three index GWA SNPs did not meet our significance threshold in JHS. This failure to replicate may have occurred for a variety of reasons. Our analysis has modestly reduced power (Supplemental Table) due to sample size relative to the original reports, which were meta-analyses of ~8000 individuals. If the effect sizes in Kathiresan et al<sup>4</sup>, which were typically less than 0.12 standard units for all SNPs, are applicable to JHS, we had 80% power to detect associations with three variants at  $p<0.004$ : rs1260326 (2p23, *GCKR*) with TG; rs646776 (1p13, *CELSR2/PSRC1/SORT1*); and rs17321515 (8q24, *TRIB1*). In addition to a smaller sample size, the reduced power arises in part from the fact that minor allele frequencies at 4 of the 8 loci were lower in African Americans as compared to the discovery cohorts of European ancestry, further decreasing power to confirm association (rs4846914: MAF=0.12 vs 0.40; rs17145738, 0.08 vs 0.13; rs17321515, 0.44 vs 0.49; and rs12130333, 0.10 vs 0.22). Interestingly, although we replicated associations at 2p23 and 1p13, we saw

no association at 8q24, despite adequate power. Although the reason for this result is unclear, a possible explanation could be differing patterns in linkage disequilibrium in African Americans, which would weaken the observed association with a non-causal SNP. Additionally, unidentified gene-x-gene or gene-x-environment interactions may cause fundamental differences in effect sizes between European and African Americans.<sup>2</sup>

We confirmed that local ancestry data at each locus provides additional insight into genetic associations in African Americans. For 4 of the 5 associated index GWA SNPs, the effect sizes on the European ancestry background exceeded those on the African background. Although the effect size differences for any single SNP can only be considered suggestive, this finding supports a pattern of systematically weaker effect sizes on the African ancestral background. Because members of the JHS are from the same community, it is unlikely that unmeasured differential environmental exposures contributed to the apparent systematic difference in effect sizes. A more likely explanation is that the index SNPs are markers for the major causal SNP at each locus and that smaller effect sizes in the African local ancestry subgroup result from the weaker SNP-SNP correlations seen in the ancestral West African population.

Fine-mapping yielded two variants with consistent effects on plasma TG levels irrespective of local ancestry - rs780093 in *GCKR* and rs636523 near *DOCK7/ANGPTL3* - making them credible candidates as the causal variants at their respective loci. Neither SNP is a nonsynonymous coding variant: rs780093 is intronic within the *GCKR* gene, and rs636523 is near *DOCK7* and *ANGPTL3*. It currently remains unclear how these noncoding variants might influence phenotypic variation. Ultimately, experimental validation in a faithful disease model will be required to further elucidate their mechanistic action. While further replication of these SNPs in additional African American samples is needed, our findings suggest that these SNPs may have predictive value in African Americans at these two loci. In contrast, while rs629301 was nominally more strongly associated with plasma LDL cholesterol than its corresponding index SNP rs646776, there remained a disparity in the effect on the African background compared to the European background. A suitable genetic marker at this locus for LDL in African Americans thus remains to be identified.

Fine-mapping genomic regions is one proposed next step to further characterize loci discovered through GWA.<sup>14</sup> It has been suggested that the shorter linkage disequilibrium (LD) patterns in historically larger populations such as West Africans may allow narrowing the interval in which causal variants may lie.<sup>3</sup> However, fine-mapping does not necessarily provide an independent pattern of LD because of the large amount of European chromosomal segments arising from admixture. Therefore, we performed a stratified association analysis based upon local ancestry, limiting ourselves to individuals with only a high probability of 2 alleles of African ancestry at the locus (JHS\_AFR<sub>2</sub>). In principle, analysis in this subpopulation will provide an entirely independent pattern of LD as there is no contribution from European ancestral segments. For each of the eight loci, this resulted in a subgroup of ~2500 individuals. Using this approach, we did not find any significant associations for LDL, HDL, or TG.

While theoretically attractive, our work highlights several challenges for tag-based fine-mapping in African Americans. First, power to detect novel associations, even in previously defined loci, is limited. Power calculations for SNP discovery within the African Ancestry subpopulations (Supplemental Table) show that we are substantially underpowered to detect the typical effect sizes for SNPs seen in European populations. Genome-wide association studies have shown that common variants impart modest effects on their associated traits; thus, identifying these variants often requires meta-analyses of large cohort consortia. Unfortunately, there are fewer available cohorts of African ancestry, limiting the number of

individuals available for genetic analysis. Furthermore, analyses in admixed populations such as African Americans are complicated by the mixture of African and European segments across the genome. By excluding individuals with European ancestry at a given locus from analysis in order to narrow the association signal, sample size is further diminished (by approximately 45% in our study). In addition, it is plausible that the genetic loci discovered in European-derived populations in fact harbor multiple functional loci along a single haplotype and that in the absence of similar LD structure, the association signal may be much weaker. Finally, tag-based strategies typically utilize highly correlated common variants from the HapMap to capture genetic variation across a genomic region. The second version of HapMap contains only ~30% of the common SNPs that are present in the genome,<sup>14</sup> so coverage is currently incomplete. The 1000 Genomes Project, an effort currently underway to resequence the entire genome of >1000 individuals, should significantly improve upon the currently available datasets of common variants. Next generation genomic resequencing technology holds great promise by providing the ability to resequence genomic regions, identifying both common and rare variants in a large number of individuals at a relatively modest financial cost.

Our study has several strengths and limitations. First, the Jackson Heart Study is a well-phenotyped, community based study and is one of the largest available cohorts of African American individuals. Next, we accounted for the genetic architecture of both African and European-derived populations when we selected our tag SNPs. In our analyses, we accounted for population stratification, which can confound genotype-phenotype associations, by adjusting for both global and local ancestry. In addition, we performed stratified analyses by local ancestry and examined whether genetic association signals could be narrowed on the African ancestry background. While many have postulated that this may be a useful approach, few have tested this hypothesis. A key limitation to our study is statistical power. As mentioned earlier, most common variants discovered to date impart small effects, mandating large sample sizes to detect genetic associations. Unfortunately, genetic research in African Americans is currently limited by the relatively few cohorts available as compared to groups of European descent. In addition, tag SNP selection is limited to known common variants.

In conclusion, we investigated eight genomic regions recently reported as associated with lipids in populations of European descent to determine their relevance in African Americans. Of the eight regions, we confirmed that five of the eight index SNPs were significantly/suggestively associated with plasma levels of LDL, HDL, and/or TG levels in African Americans. Moreover, at three of the five confirmed loci, we found SNPs with stronger associations with lipids in African Americans compared to the index SNPs discovered in European-derived cohorts, demonstrating that although associated regions may be relevant to many ethnic groups, specific markers of association can be unique to individual groups. Finally, we demonstrated the limitations of using a cross-ethnic fine-mapping approach to narrow association signals and suggest resequencing these loci in multi-ethnic groups as an alternate approach.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

We gratefully acknowledge the contributions of the participants and staff of the Jackson Heart Study.

**Funding:** Research support for JHS studies was provided by R01-HL-084107 (JGW) from the National Heart, Lung, and Blood Institute and contracts N01-HC-95170, N01-HC-95171, and N01-HC-95172 from the National



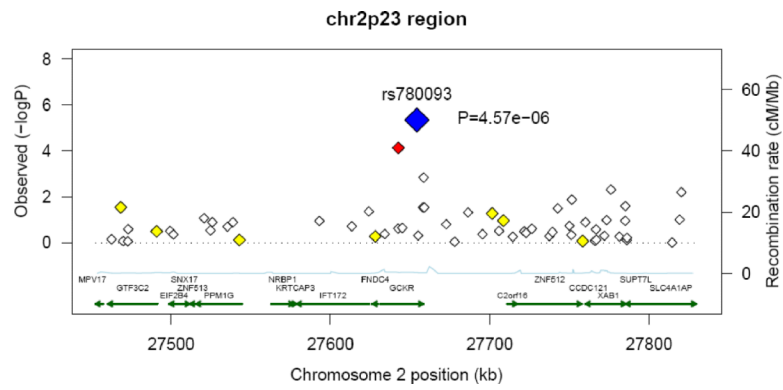
Heart, Lung, and Blood Institute and the National Center on Minority Health and Health Disparities. M.K. was supported by the Charles A. King Trust, Bank of America, Co-Trustee. S.K. is supported by a Doris Duke Charitable Foundation Clinical Scientist Development Award, a charitable gift from the Fannie E. Rippe Foundation, the Donovan Family Foundation, a career development award from the United States National Institutes of Health (NIH), and institutional support from the Department of Medicine and Cardiovascular Research Center at Massachusetts General Hospital.

## References

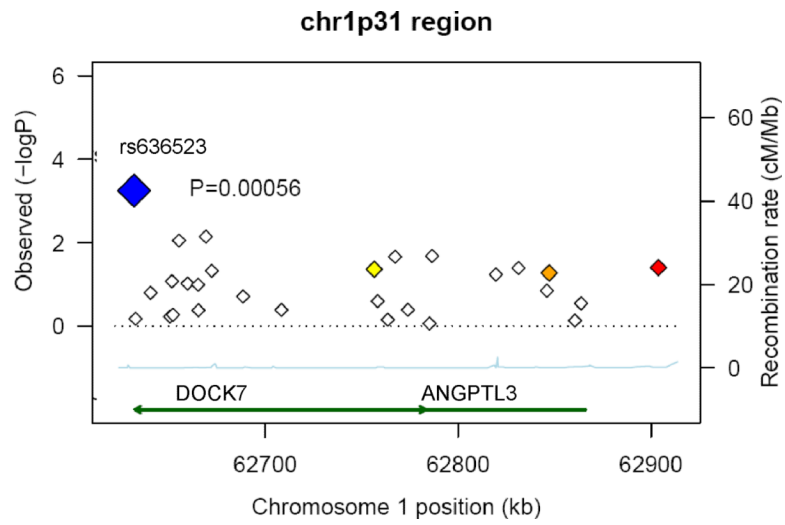
1. Kathiresan S, Melander O, Anevski D, Guiducci C, Burt NP, Roos C, Hirschhorn JN, Berglund G, Hedblad B, Groop L, Altshuler DM, Newton-Cheh C, Orho-Melander M. Polymorphisms associated with cholesterol and risk of cardiovascular events. *N Engl J Med*. 2008; 358:1240–1249. [PubMed: 18354102]
2. Deo RC, Reich D, Tandon A, Akyzbekova E, Patterson N, Waliszewska A, Kathiresan S, Sarpong D, Taylor HA Jr, Wilson JG. Genetic differences between the determinants of lipid profile phenotypes in African and European Americans: the Jackson Heart Study. *PLoS Genet*. 2009; 5:e1000342. [PubMed: 19148283]
3. McCarthy MI, Hirschhorn JN. Genome-wide association studies: potential next steps on a genetic journey. *Hum Mol Genet*. 2008; 17:R156–165. [PubMed: 18852205]
4. Kathiresan S, Melander O, Guiducci C, Surti A, Burt NP, Rieder MJ, Cooper GM, Roos C, Voight BF, Havulinna AS, Wahlstrand B, Hedner T, Corella D, Tai ES, Ordovas JM, Berglund G, Vartiainen E, Jousilahti P, Hedblad B, Taskinen MR, Newton-Cheh C, Salomaa V, Peltonen L, Groop L, Altshuler DM, Orho-Melander M. Six new loci associated with blood low-density lipoprotein cholesterol, high-density lipoprotein cholesterol or triglycerides in humans. *Nat Genet*. 2008; 40:189–197. [PubMed: 18193044]
5. Willer CJ, Sanna S, Jackson AU, Scuteri A, Bonnycastle LL, Clarke R, Heath SC, Timpson NJ, Najjar SS, Stringham HM, Strait J, Duren WL, Maschio A, Busonero F, Mulas A, Albai G, Swift AJ, Morken MA, Narisu N, Bennett D, Parish S, Shen H, Galan P, Meneton P, Hercberg S, Zelenika D, Chen WM, Li Y, Scott LJ, Scheet PA, Sundvall J, Watanabe RM, Nagaraja R, Ebrahim S, Lawlor DA, Ben-Shlomo Y, Davey-Smith G, Shuldiner AR, Collins R, Bergman RN, Uda M, Tuomilehto J, Cao A, Collins FS, Lakatta E, Lathrop GM, Boehnke M, Schlessinger D, Mohlke KL, Abecasis GR. Newly identified loci that influence lipid concentrations and risk of coronary artery disease. *Nat Genet*. 2008; 40:161–169. [PubMed: 18193043]
6. Keebler ME, Sanders CL, Surti A, Guiducci C, Burt NP, Kathiresan S. Association of blood lipids with common DNA sequence variants at 19 genetic loci in the multiethnic United States National Health and Nutrition Examination Survey III. *Circ Cardiovasc Genet*. 2009; 2:238–243. [PubMed: 20031591]
7. de Bakker PI, Yelensky R, Pe'er I, Gabriel SB, Daly MJ, Altshuler D. Efficiency and power in genetic association studies. *Nat Genet*. 2005; 37:1217–1223. [PubMed: 16244653]
8. Deng HW, Chen WM, Recker RR. Population admixture: detection by Hardy-Weinberg test and its quantitative effects on linkage-disequilibrium methods for localizing genes underlying complex traits. *Genetics*. 2001; 157:885–897. [PubMed: 11157005]
9. Patterson N, Hattangadi N, Lane B, Lohmueller KE, Hafler DA, Oksenberg JR, Hauser SL, Smith MW, O'Brien SJ, Altshuler D, Daly MJ, Reich D. Methods for high-density admixture mapping of disease genes. *Am J Hum Genet*. 2004; 74:979–1000. [PubMed: 15088269]
10. Levy D, DeStefano AL, Larson MG, O'Donnell CJ, Lifton RP, Gavras H, Cupples LA, Myers RH. Evidence for a gene influencing blood pressure on chromosome 17. Genome scan linkage results for longitudinal blood pressure phenotypes in subjects from the framingham heart study. *Hypertension*. 2000; 36:477–483. [PubMed: 11040222]
11. Abecasis GR, Cherny SS, Cookson WO, Cardon LR. Merlin--rapid analysis of dense genetic maps using sparse gene flow trees. *Nat Genet*. 2002; 30:97–101. [PubMed: 11731797]
12. Kathiresan S, Willer CJ, Peloso GM, Demissie S, Musunuru K, Schadt EE, Kaplan L, Bennett D, Li Y, Tanaka T, Voight BF, Bonnycastle LL, Jackson AU, Crawford G, Surti A, Guiducci C, Burt NP, Parish S, Clarke R, Zelenika D, Kubalanza KA, Morken MA, Scott LJ, Stringham HM, Galan P, Swift AJ, Kuusisto J, Bergman RN, Sundvall J, Laakso M, Ferrucci L, Scheet P, Sanna S, Uda M, Yang Q, Lunetta KL, Dupuis J, de Bakker PI, O'Donnell CJ, Chambers JC, Kooner JS,

Hercberg S, Meneton P, Lakatta EG, Scuteri A, Schlessinger D, Tuomilehto J, Collins FS, Groop L, Altshuler D, Collins R, Lathrop GM, Melander O, Salomaa V, Peltonen L, Orho-Melander M, Ordovas JM, Boehnke M, Abecasis GR, Mohlke KL, Cupples LA. Common variants at 30 loci contribute to polygenic dyslipidemia. *Nat Genet.* 2009; 41:56–65. [PubMed: 19060906]

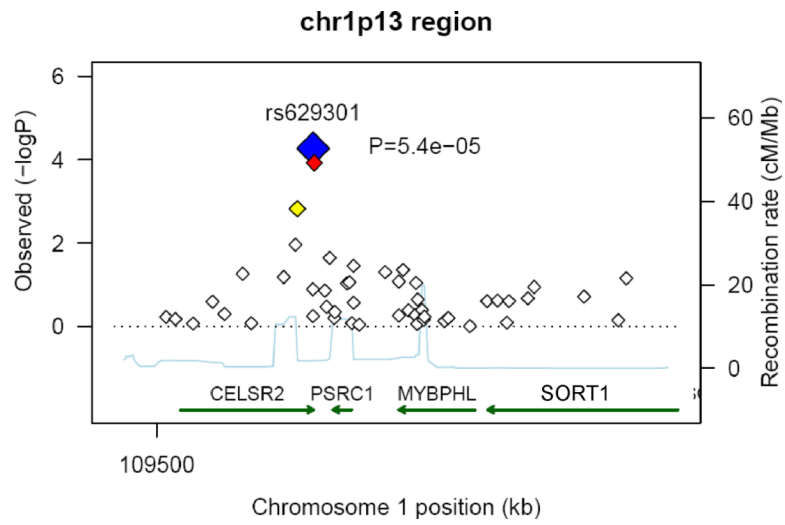
13. Romeo S, Yin W, Kozlitina J, Pennacchio LA, Boerwinkle E, Hobbs HH, Cohen JC. Rare loss-of-function mutations in ANGPTL family members contribute to plasma triglyceride levels in humans. *J Clin Invest.* 2009; 119:70–79. [PubMed: 19075393]
14. Ioannidis JP, Thomas G, Daly MJ. Validating, augmenting and refining genome-wide association signals. *Nat Rev Genet.* 2009; 10:318–329. [PubMed: 19373277]



**Figure 1.** Associations of triglycerides with fine-mapping genotypes at the 2p23 locus.



**Figure 2.** Associations of triglycerides with fine-mapping genotypes at the 1p31 locus.



**Figure 3.**  
Associations of LDL with fine-mapping genotypes at the 1p13 locus.



**Table 1**

Sample characteristics for members of the cohort analyzed at the chr1p13 locus.

Characteristic	JHS_all* N=4,474	JHS_afr <sup>†</sup> N=2,471	JHS_eur <sup>‡</sup> N=743
Mean age, yrs	55 (13)	55(13)	54 (12)
Male, %	37.2	36.3	37.6
On lipid lowering medications, %	13.7	13.6	14.9
Body mass index, kg/m <sup>2</sup>	32 (7)	32 (7)	32 (7)
Low-density lipoprotein cholesterol, mg/dL <sup>§</sup>	130 (36)	130 (36)	131 (35)
High-density lipoprotein cholesterol, mg/dL	51 (14)	52 (14)	51 (15)
Triglycerides, mg/dL	110 (92)	109 (98)	113 (120)
Systolic blood pressure, mmHg	127 (18)	128 (18)	127 (19)
Diabetes mellitus, %	19.2	19.6	17.0
Overall African Ancestry, %	83 (9)	85 (7)	76 (11)

\* JHS\_all = JHS cohort.

<sup>†</sup> JHS\_afr = 2 African ancestral alleles at chr1p13.

<sup>‡</sup> JHS\_eur =  $\geq 1$  European ancestral allele at chr1p13. Values are represented as the mean (standard deviation) or as percentages, where applicable.

<sup>§</sup> Includes imputed LDL values for subjects on lipid lowering medications.

Table 2

Associations of GWA index polymorphisms with blood lipids.

SNP	Locus	Alleles maj/min	Modeled allele all (f_all) <sup>*†</sup>	Modeled allele afr (f_afr) <sup>*‡</sup>	Modeled allele eur (f_eur) <sup>*§</sup>	Associated Lipid Trait(s) through GWA	p_all <sup>†</sup>	p_afr <sup>‡</sup>	p_eur	Effect all ±SE <sup>¶</sup>	Effect afr <sup>‡</sup> ±SE <sup>¶</sup>	Effect eur <sup>§</sup> ±SE <sup>¶</sup>
rs1260326	Chr 2p23, In <i>GCKR</i> , Nonsynonymous	C/T	C (0.86)	C (0.93)	C (0.70)	TG	<b>7.4E-05</b>	0.12	0.0099	-7.3±1.6	-4.4±2.9	-7.0±2.8
rs646776	1p13 near <i>CELSR2/PSRC1/SORT1</i>	T/C	T (0.63)	T (0.61)	T (0.69)	LDL	<b>1.2E-04</b>	0.04	0.018	3.3±0.85	2.3±1.1	5.0±2.1
rs10774708	12q24 in <i>KCTD10</i> , intronic	G/A	G (0.72)	G (0.77)	G (0.59)	HDL	<b>0.04</b>	0.08	0.46	1.3 ±0.60	1.6±0.90	1.0±1.4
rs12130333	Chr 1p31, near <i>DOCK7/ANGPTL3</i>	C/T	C (0.90)	C (0.92)	C (0.85)	TG	<b>0.04</b>	0.36	0.048	4.1±1.9	2.6±2.8	7.7±3.9
rs17145738	Chr 7q11, near <i>TBL2</i>	C/T	C (0.92)	C (0.92)	C (0.90)	HDL	0.62	0.78	0.95	0.50±1.0	0.40±1.4	0.10±2.0
rs4846914	Chr 1q42 in <i>GALNT2</i> , intronic	G/A	G (0.88)	G (0.96)	G (0.65)	TG	<b>0.04</b>	0.43	0.23	4.4±2.0	2.2±2.8	4.7±4.0
rs17321515	Chr 8q24 near <i>TRIB1</i>	G/A	G (0.56)	G (0.58)	G (0.51)	LDL	0.09	0.02	0.97	-1.5±0.90	-4.5±2.0	0.10±1.6
rs16996148	Chr 19p13 near <i>CILP2/PBX4</i>	G/T	G (0.89)	G (0.83)	G (0.87)	TG	0.16	0.37	0.72	2.5±1.5	3.6±4.0	-1.2±3.2
						HDL	0.20	0.72	0.16	1.0±0.82	0.37±1.1	2.6±1.8
						TG	0.89	0.87	0.57	0.10±0.60	0.10±0.80	-0.70±1.3
						LDL	0.18	0.87	0.40	-1.5±1.2	0.20±1.5	-2.2±2.6
						LDL	0.29	0.19	0.26	1.2±1.1	1.8±1.4	3.0±2.7
						TG	0.97	0.27	0.074	0.10±1.6	-2.2±1.9	7.0±3.9

\* C = major allele frequency.

† f\_all = entire JHS cohort.

‡ f\_afr = 2 African local ancestral alleles.

§ f\_eur = ≥ 1 European local ancestral allele

¶ For LDL associations, the effect size is in mg/dL. Because HDL and TG were analyzed with log transformation, the effect sizes for these traits represent percentage change in trait value.

**Table 3**

Significant associations of fine-mapping tag SNPs and lipid traits

SNP	Locus	Position	Alleles maj/min	Modeled allele all (f_all) <sup>†</sup>	Modeled allele afr (f_afr) <sup>‡</sup>	Modeled allele eur (f_eur) <sup>§</sup>	Associated Lipid Trait through GWA	p_all <sup>†</sup>	p_afr <sup>‡</sup>	Effect all ±SE <sup>  </sup>	Effect afr <sup>‡</sup> ±SE <sup>  </sup>	Effect eur <sup>§</sup> ±SE <sup>  </sup>
rs780093	Chr 2p23, in <i>GCKR</i> , intronic	27654254	G/A	G (0.85)	G (0.92)	G (0.70)	TG	4.6E-06	0.004	-8.2±1.6	-8.7±2.8	-8.6±2.8
rs636523	Chr 1p31, near <i>DOCK7/ANGPTL3</i>	62632029	T/C <sup>#</sup> /C/T	T (0.53)	C <sup>#</sup> (0.50)	T (0.62)	TG	5.6E-04	0.002	4.1±1.1	-4.7±1.5*	4.3±2.6
rs629301	1p13 in <i>CELSR2</i> , untranslated	109530348	A/C	A (0.63)	A (0.61)	A (0.69)	LDL	5.4E-05	0.02	3.5±0.86	2.6±1.2	6.0±2.1

\* f = major allele frequency.

<sup>†</sup> \_all = entire JHS cohort.

<sup>‡</sup> \_afr = 2 African local ancestral alleles.

<sup>§</sup> \_eur = ≥ 1 European local ancestral allele

<sup>||</sup> For LDL associations, the effect size is in mg/dL. Because HDL and TG were analyzed with log transformation, the effect sizes for these traits represent percentage change in trait value.

<sup>#</sup> Due to genotype frequencies in the JHS\_afr subgroup, the C allele was modeled as the major allele for rs636523