Association of Exome Sequences With Cardiovascular Traits Among Blacks in the Jackson Heart Study

Gina M. Peloso, PhD; Leslie A. Lange, PhD; Tibor V. Varga, PhD; Deborah A. Nickerson, PhD; Joshua D. Smith, PhD; Michael E. Griswold, PhD; Solomon Musani, PhD; Linda M. Polfus, PhD; Hao Mei, PhD; Stacey Gabriel, PhD; Rakale Collins Quarells, PhD; David Altshuler, MD, PhD; Eric Boerwinkle, PhD; Mark J. Daly, PhD; Benjamin Neale, PhD; Adolfo Correa, MD, PhD; Alex P. Reiner, MD; James G. Wilson, MD; Sekar Kathiresan, MD

Background—The correlation of null alleles with human phenotypes can provide insight into gene function in humans. In individuals of African ancestry, we set out to identify null and damaging missense variants, and test these variants for association with a range of cardiovascular phenotypes.

Methods and Results—We performed whole-exome sequencing in 3223 black individuals from the Jackson Heart Study and found a total of 729,666 variant sites with minor allele frequency <5%, including 17,263 null variants and 49,929 missense variants predicted to be damaging by in silico algorithms. We tested null and damaging missense variants within each gene for association with 36 cardiovascular traits. We found 3 associations that met our prespecified level of significance (α=1.1×10−7). Null and damaging missense variants in PCSK9 were associated with 36 mg/dL lower-density lipoprotein cholesterol (P=3×10−21). Three individuals in their 50s with complete PCSK9 deficiency (each compound heterozygote for PCSK9 p.Y142X and p.C679X) were identified, with one having a coronary artery calcification score in the 83rd percentile despite a low-density lipoprotein cholesterol of 32 mg/dL. A damaging missense variant in HBQ1 (p.G52A) was associated with a 2 pg/cell lower mean corpuscular hemoglobin (P=9×10−13) and rare damaging missense variants in VPS13A with higher red blood cell distribution width (P=9.9×10−4).

Conclusions—A limited number of null/damaging alleles with a large effect on cardiovascular traits were detectable in ≈3000 black individuals. (Circ Cardiovasc Genet. 2016;9:368-374. DOI: 10.1161/CIRCGENETICS.116.001410.)

Key Words: alleles ■ exome ■ lipids ■ missense ■ mutation

A compelling therapeutic target for lowering low-density lipoprotein cholesterol (LDL-C) emerged from human genetic studies—PCSK9 (the proprotein convertase subtilisin/kexin type 9 gene).1 Null alleles (also termed loss-of-function protein-coding sequence variants) in PCSK9 were identified in blacks2 and shown to associate with lower plasma LDL-C levels3-4 and reduced risk for coronary heart disease (≤88% reduction).5,6 On the basis of this human genetic evidence and corroborating functional studies, several pharmaceutical companies have established drug development programs targeting PCSK9,7 and 2 inhibitors have been approved for reducing LDL-C in individuals with heterozygous familial hypercholesterolemia and in individuals with clinical atherosclerotic cardiovascular disease.5,9 On the basis of the PCSK9 example, it has been suggested that low-frequency or rare mutations of large effect may be paradigmatic for therapeutic target discovery.10

Clinical Perspective on p 374

To address whether additional such examples can be readily identified, we sequenced the exomes of 3223 individuals from the JHS (Jackson Heart Study), a prospective cohort of blacks living in Jackson, Mississippi, and catalogued null and damaging
null and damaging missense mutations across 18,465 genes. Subsequently, we performed an association study of these variants with a range of quantitative and qualitative cardiovascular traits.

Methods

Study Participants
The JHS is a community-based, longitudinal, cohort study located in the Jackson, Mississippi metropolitan area designed to investigate the determinants of cardiovascular disease in blacks. JHS recruited 5301 blacks, aged between 35 and 84 years, between September 2000 and March 2008. The Institutional Review Board of the University of Mississippi Medical Center approved the study protocol, and all participants provided written informed consent.

Exome Sequencing
Exome sequencing was performed at 3 sequencing centers (the Broad Institute [n=2317], University of Washington [n=481], and Baylor College of Medicine [n=475]) across 5 projects (The US National Heart, Lung, and Blood Institute’s Exome Sequencing Project, Myocardial Infarction Genetics Consortium Exome Sequencing Project, CHARGE-S, Type 2 Diabetes Genetic Exploration by Next-generation sequencing in multi-Ethnic Samples, and Minority Health Genomics and Translational Research Bio-Repository Database; Table I in the Data Supplement). The sequencing reads (ie, fastq files) from exomes were aligned to the human genome reference (hg19) using Burrows-Wheeler Transform on a per-lane basis and bam files were obtained from the 3 sequencing centers. The Genome Analysis Toolkit v3.1 HaplotypeCaller algorithm was used for joint variant discovery and genotyping on both exomes and flanking 50 bp of intronic sequence (http://www.broadinstitute.org/gatk/guide/article?id=3893). Single-sample gVCFs were created using the Genome Analysis Toolkit HaplotypeCaller with the options -emitRefConfidence GVCF, –variant_index_type LINEAR, and –variant_index_parameter 128000. Then, batches of =200 gVCFs were merged into a single gVCF using the CombineGVCF command in Genome Analysis Toolkit. Finally, GenotypeGVCFs was run on the combined gVCFs to create the raw SNP and indel VCFs. Because a majority of individuals were sequenced at the Broad Institute, we limited analysis to the sequence intervals captured by the Broad’s exome-sequencing platform.

Variant Quality Control
Genome Analysis Toolkit Variant Quality Score Recalibration (VQSR) was used with the recommended resources to filter variants. The SNP VQSR model was trained using HapMap3.3 and 1KG Omni 2.5 SNP sites and a 99.5% sensitivity threshold was applied to filter variants, whereas the INDEL VQSR model was trained using the Mills 1000G gold standard and Axiom Exome Plus sites for insertions/deletions and a 99.0% sensitivity threshold was applied to filter INDEL sites. Variants were filtered to VQSR PASS and quality depth ≥2. (Table II in the Data Supplement). Individual genotypes were set to missing if depth <5.

Sample Quality Control
We performed quality control on the jointly called samples. Individuals were checked for total number of variants, observed number of singletons and doubletons, Ti/Tv ratio, Het/Hom ratio, missingness, contamination with VerifyBamID, and nonreference concordance with available genotype data from the Illumina HumanExome BeadChip v1.0. Individuals that were outliers (>±3*interquartile range) on at least one metric were excluded (Table I and Figure I in the Data Supplement). Population structure was assessed using the multidimensional scaling algorithm in the PLINK software and 10 principal components of ancestry were obtained (Figure II in the Data Supplement).

Annotation
All variant sites were annotated with the Variant Effect Predictor algorithm (VEP; http://useast.ensembl.org/info/docs/tools/vep) and

Figure. Cardiovascular traits tested for association with null and damaging missense variants. Thirty-six cardiovascular traits were tested for association with null and damaging missense variants in 3223 black individuals from the Jackson Heart Study. BMI indicates body mass index; CAC, coronary artery calcification; CHD, coronary heart disease; DBP, diastolic blood pressure; HDL-C, high-density lipoprotein cholesterol; HOMA, homeostatic model assessment; hsCRP, high-sensitivity C-reactive protein; LDL, low-density lipoprotein cholesterol; MCH, mean corpuscular hemoglobin; MCHC, mean corpuscular hemoglobin concentration; MCV, mean corpuscular volume; RDW, red blood cell distribution width; and sCort, serum Cortisol.
Circ Cardiovasc Genet August 2016

Analysis was limited to variants predicted to be null (nonsense, splice, frame-shift) plus missense variation predicted to be damaging in at least 5 of the following 7 variation prediction tools:

- LRT
- Mutation Taster
- PolyPhen2 (HumDiv)
- PolyPhen2 (HumVar)
- SIFT
- MutationAssessor
- FATHMM

Phenotypes

We analyzed 36 cardiovascular traits (Figure) available in the Jackson Heart Study Vanguard Center data package. For participants who were taking antihypertensive medication, we added 10 mm Hg to observed systolic blood pressure values and 5 mm Hg to diastolic blood pressure values. For participants on lipid-lowering medication, we replaced their total cholesterol values by total cholesterol divided by 0.8. No adjustment was made on high-density lipoprotein cholesterol or triglycerides. Only fasting lipid measures were used, and LDL-C was calculated using the Friedewald equation for those with triglycerides <400 mg/dL, using the lipid-adjusted total cholesterol for those on treatment.

Individuals with diabetes mellitus were excluded in analyses of fasting plasma glucose, fasting insulin, homeostatic model assessment-IR, homeostatic model assessment-B, and glycated hemoglobin. Only fasting lipid measures were used, and LDL-C was calculated using the Friedewald equation for those with triglycerides <400 mg/dL, using the lipid-adjusted total cholesterol for those on treatment.

Nonnormality of the following raw traits was resolved by a natural log transform before analysis: triglycerides, leptin, high-sensitivity C-reactive protein, endothelin, renin, aldosterone, and adiponectin.

Table 1. Descriptive Statistics of Jackson Heart Study Participants With Exome Sequences

| Trait                        | N       | Statistic            |
|------------------------------|---------|----------------------|
| Demographic                  |         |                      |
| Female                       | 3223    | 1211 (37.6%)         |
| Age, y                       | 3223    | 55.59±12.82          |
| Current smoking status       | 3195    | 428 (13.4%)          |
| Anthropometrics              |         |                      |
| Body mass index, kg/m²       | 3216    | 31.99±7.37           |
| Weight, kg                   | 3218    | 91.37±21.71          |
| Height, cm                   | 3218    | 169.06±9.25          |
| Waist circumference, cm      | 3216    | 101.36±16.26         |
| Neck circumference, cm       | 3219    | 38.72±3.76           |
| Hypertension                 |         |                      |
| Hypertension, yes            | 3223    | 2012 (62.4%)         |
| Systolic blood pressure, mmHg | 3217    | 132.11±19.87         |
| Diastolic blood pressure, mmHg | 3217   | 81.50±10.80          |
| Antihypertensive treatment   | 2619    | 1655 (63.2%)         |
| Lipids                       |         |                      |
| LDL-C†, mg/dL                | 2950    | 131.8±39.29          |
| HDL-C, mg/dL                 | 2980    | 51.58±14.76          |
| Triglycerides, mg/dL         | 2979    | 107.61±82.77         |
| Total cholesterol†, mg/dL    | 2395    | 206.08±43.44         |
| Lipid-lowering treatment     | 2619    | 367 (14%)            |
| Coronary heart disease       |         |                      |
| Coronary heart disease status | 3223   | 251 (7.8%)           |
| CAC score, Agatston units    | 1795    | 176.93±550.8         |
| CAC>0                        | 1795    | 882 (49.1%)          |
| CAC>100                      | 1795    | 439 (24.5%)          |
| Diabetes mellitus            |         |                      |
| Diabetic status              | 3220    | 745 (23.1%)          |
| Fasting insulin, plasma IU/mL‡ | 2388   | 15.88±9.22           |
| HOMA-B†, mmol/L              | 2357    | 215.75±107.75        |
| HOMA-IR†, mmol/L             | 2386    | 3.59±2.29            |
| Fasting plasma glucose level, mg/dL‡ | 2390 | 90.53±8.97          |
| Hemoglobin Hba1c, %‡          | 2429    | 5.51±0.47            |
| Biomarkers                   |         |                      |
| Leptin, serum ng/mL          | 3198    | 28.39±23.98          |
| High-sensitivity C-reactive protein, serum mg/dL | 3214 | 0.53±1                |
| Endothelin-1, serum pg/mL    | 3214    | 1.34±0.6             |
| Aldosterone, serum ng/dL     | 3213    | 5.81±4.92            |
| Renin activity RIA, plasma ng/mL/h | 1509 | 1.72±6.45           |
| Cortisol levels, serum µg/dL | 3213    | 9.87±4.13            |
| Adiponectin, plasma ng/mL    | 3166    | 5345.18±4236.78      |

Table 1. Continued

| Trait                        | N       | Statistic            |
|------------------------------|---------|----------------------|
| ECG                          |         |                      |
| QT interval, ms              | 3008    | 413.34±30.74         |
| QRS interval, ms             | 2802    | 92.08±9.95           |
| Blood                        |         |                      |
| Hematocrit level, %          | 3110    | 39.27±4.2            |
| Hemoglobin, g/dL             | 3109    | 13.04±1.48           |
| Mean corpuscular hemoglobin, pg | 2781 | 28.88±2.51           |
| Mean corpuscular hemoglobin concentration, % | 2781 | 33.16±0.91 |
| Mean corpuscular volume, fl  | 2781    | 86.97±6.41           |
| Red blood cell distribution width, % | 2780 | 13.70±1.38          |
| Red cell count, m/cmm        | 2781    | 4.53±0.51            |

Statistic provided as mean±SD for continuous variables and n (%) for categorical variables. CAC indicates coronary artery calcification; Hba1c, glycated hemoglobin; HDL-C, high-density lipoprotein cholesterol; HOMO, homeostatic model assessment; LDL-C, low-density lipoprotein cholesterol; and RIA, radioimmunoassay.

*Values were adjusted for individuals on blood pressure-lowering medication.
†Values were adjusted for individuals on lipid-lowering medication.
‡Values for nondiabetic individuals.

dbhSF14 (https://sites.google.com/site/jpopgen/dbNSFP). Analysis was limited to variants predicted to be null (nonsense, splice, frame-shift) plus missense variation predicted to be damaging in at least 5 of the following 7 variation prediction tools: LRT, Mutation Taster, PolyPhen2 (HumDiv), PolyPhen2 (HumVar), SIFT, MutationAssessor, and FATHMM.

Phenotypes

We analyzed 36 cardiovascular traits (Figure) available in the Jackson Heart Study Vanguard Center data package (https://www.jacksonheartsudy.org/jhsinfo/ForResearchers/VanguardCenters/tabid/171/Default.aspx). For participants who were taking antihypertensive medication, we added 10 mm Hg to observed systolic blood pressure values and 5 mm Hg to diastolic blood pressure values. We adjusted the total cholesterol values for individuals on lipid-lowering medication by replacing their total cholesterol values by total cholesterol divided by 0.8. No adjustment was made on high-density lipoprotein cholesterol or triglycerides. Only fasting lipid measures were used, and LDL-C was calculated using the Friedewald equation for those with triglycerides <400 mg/dL, using the lipid-adjusted total cholesterol for those on treatment.

Individuals with diabetes mellitus were excluded in analyses of fasting plasma glucose, fasting insulin, homeostatic model assessment-IR, homeostatic model assessment-B, and glycated hemoglobin. Individuals with QRS ≥120, atrial fibrillation, or coronary heart disease were excluded for analysis of QT interval. Individuals with end-stage renal disease defined as estimated glomerular filtration rate <15 or reporting being on dialysis, hemoglobinopathy defined as being homozygous for rs334, or myelotoxic drug use were excluded from the blood cell trait analyses.

Nonnormality of the following raw traits was resolved by a natural log transform before analysis: triglycerides, leptin, high-sensitivity C-reactive protein, endothelin, renin, aldosterone, and adiponectin.
Association Analysis
We performed gene-based analyses of 36 cardiovascular phenotypes. We limited analysis to null mutations plus missense variants predicted to be damaging by at least 5 of 7 in silico prediction algorithms (LRT, Mutation Taster, PolyPhen2 (HumDiv), PolyPhen2 (HumVar), SIFT, MutationAssessor, and FATHMM). We aggregated variants with minor allele frequency (MAF) ≤5% within each gene using 4 sets of variants: (1) null mutations only, (2) null mutations plus missense variants predicted to be damaging by 7 of 7 in silico prediction algorithms, (3) null mutations plus missense variants predicted to be damaging in at least 6 of 7 in silico prediction algorithms, and (4) null mutations plus missense variants predicted to be damaging in at least 5 of 7 in silico prediction algorithms. All associations were performed using the EPACTS (Efficient and Parallelizable Association Container Toolbox; http://genome.sph.umich.edu/wiki/EPACTS) software. EPACTS is a software pipeline to perform statistical tests of association using sequence data. It implements the EMMA24 (Efficient Mixed Model Association eXpedited) model, a mixed model association approach that captures pedigree, cryptic relatedness, and population structure by using a covariance matrix estimated from genome-wide data. To apply the EMMAX model, we used the epacts-group command with the emmaxCMC test option to perform collapsing burden gene-based tests. The single command with the q.emmax test option in EPACTS was used to obtain the single variant results for each variant going into the gene-based test. We used an additive genetic model. A kinship matrix of all individuals was created with EPACTS and used in analyses. All analyses were adjusted for age, sex, and 4 principal components of ancestry. Analyses for QT interval and QRS additionally included adjustments for height and BMI.

We excluded results with ≤10 minor alleles contributing to the gene-based test to ensure robust association statistics. We set our significance threshold to 1.1×10−7 (0.05/36 traits in silico prediction algorithms; Table II in the Data Supplement). Of the 18,465 genes sequenced, 14,058 have a null or damaging missense variant with MAF ≤5%. On average, we observe 5 null or damaging missense variants per gene and an average of 7 null or damaging missense alleles per gene. Each individual carries, on average, a total of 15 null or damaging missense variants with MAF ≤5%.

We found 3 gene-based associations that met our prespecified significance threshold of 1.1×10−7 (Table 2; Tables III and IV in the Data Supplement). The most significant association was between LDL-C and PCSK9. Participants who carried null or damaging missense mutations in PCSK9 had 36 mg/dL lower LDL-C compared with noncarriers (P=2.9×10−21). Of note, we identified 3 individuals with complete PCSK9 deficiency (each compound heterozygote for PCSK9 p.Y142X and p.C679X; Table 3). These individuals had a lower median LDL-C (64.2 mg/dL) compared with individuals who carry only one null mutation (85.7 mg/dL; n=77; P=0.044; Figure III in the Data Supplement). The 3 PCSK9 null compound heterozygotes did not differ from heterozygotes in any other cardiometabolic trait tested except QT interval (Table V in the Data Supplement). Compound heterozygotes had a lower QT interval (mean=369; range=362–380) compared with individuals who carried only one null PCSK9 variant (mean=413; P=0.006 using a Wilcoxon rank-sum test). Individuals carrying one null PCSK9 variant had similar QT intervals compared with noncarriers (mean=413), suggesting a recessive effect. Two individuals carrying both PCSK9 p.Y142X and p.679X had a CAC greater than the 80th percentile for their age and sex. A 52-year-old man had a CAC of 24.9, which is in the 83rd percentile for age and sex, despite an LDL-C of 32 mg/dL (Table 3).

The second most significant gene association was between mean corpuscular hemoglobin and hemoglobin subunit beta 1 (HBB1). Individuals carrying a damaging missense variant (p.G52A) in HBB1 had lower mean corpuscular hemoglobin compared with noncarriers (P=8.4×10−13). One additional association passed our significance threshold. Rare damaging missense variants in Vacular Protein Sorting-Associated Protein 13A (VPS13A) were associated with an increase in red blood cell distribution width (P=7.1×10−4). Of the 9 variants that contributed to the association between VPS13A and red blood cell distribution width, 6 were singletons, 1 a doubleton, 1 with 4 carriers (p.S2673L), and 1 with 22 minor allele carriers (p.K2672N) (Table IV in the Data Supplement).

Table 2. Genes With Rare Variant Association Signals Meeting a Significance Threshold <1.1×10−7 in the Jackson Heart Study

| Outcome                  | Gene  | Chrm | Best Test* | No. of Variant Sites | MAC | % Carriers | β±SE  | P Value |
|--------------------------|-------|------|------------|---------------------|-----|------------|-------|---------|
| LDL, mg/dL               | PCSK9 | 1    | Null+≥6/7 damaging missense | 7      | 119 | 3.9%       | −35.8±3.8 | 2.9×10−21 |
| Mean corpuscular hemoglobin, pg/cell | HBB1 | 16   | Null+≥5/7 damaging missense | 1      | 88  | 3.1%       | −2.0±0.3  | 8.9×10−13 |
| Red blood cell distribution width, % | VPS13A | 9    | Null+≥5/7 damaging missense | 9      | 34  | 1.2%       | 1.3±0.2  | 7.1×10−4  |

*Best test indicates the group of variants that provided the most significant results for the gene. Null variants are defined as nonsense, splice-site, and frameshift variants. Damaging missense variants were classified according to the following 7 in silico prediction algorithms: LRT, Mutation Taster, PolyPhen2 (HumDiv), PolyPhen2 (HumVar), SIFT, MutationAssessor, and FATHMM. Chrm indicates chromosome; MAC (minor allele count), number of minor alleles across the variant sites; no. of variant sites, number of sites going into the gene-based test; and % carriers, percent of individuals carrying a null or damaging missense variant tested.
Supplement). VPS13A showed evidence for association with other hematologic phenotypes, including lower hemoglobin levels ($P=7.0 \times 10^{-4}$; Table VI in the Data Supplement).

Li et al. recently reported 10 gene-based associations aggregating null variants with a $P<4.4 \times 10^{-6}$. Individuals of African Ancestry contributed to 7 of these associations. We attempted to replicate these 7 associations in our data (Table VII in the Data Supplement). We replicated the association of total cholesterol with $PCSK9$ ($\beta=-39$ mg/dL; $P=6.6 \times 10^{-12}$) and of triglycerides with apolipoprotein C-III ($P=1.0 \times 10^{-5}$).2,28–30 We found suggestive evidence for the association of fasting glucose with thioredoxin domain containing 5 ($TXNDC5$), consistent with the report by Li et al; carriers of null alleles in $TXNDC5$ had higher fasting glucose compared with noncarriers ($P=0.07$).

For 3223 individuals and a significance level of $1.1 \times 10^{-7}$, we had 99% statistical power to detect a 1-SD unit effect with a 1% cumulative MAF, and 64% statistical power to detect a 1-SD unit effect with a 0.5% cumulative MAF. Analysis of Mendelian lipid genes as a positive control shows several genes where a burden of null/damaging mutations alters the expected plasma lipid fraction in the appropriate direction (eg, $LDLR$ and higher LDL-C [$P=4.7 \times 10^{-5}$], $CETP$ and higher high-density lipoprotein cholesterol [$P=0.0001$; Table VIII in the Data Supplement]). However, even an analysis of positive controls is limited by the number of carriers, with the majority of the Mendelian lipid genes having <10 observed null alleles.

### Discussion

We set out to discover null or damaging missense variants that lead to a large effect on any of a range of cardiovascular traits. In a study of 3223 blacks, we found 3 associations that met our prespecific significance threshold.
We report 1 new observation that of VPS13A associated with an increase in red blood cell distribution width. Red blood cell distribution width is a measure of the range of variation in red blood cells and higher values can indicate certain disorders such as anemia. Mutations in VPS13A have been reported to cause chorea-atachycytosis, an autosomal-recessive neurodegenerative disorder that causes red blood cells to appear spiky. Ten VPS13A variants are reported in ClinVar with chorea-atachycytosis listed as the condition. We did not find any of the reported ClinVar variants in our data nor any carriers of rare damaging recessive variants in VPS13A. Here, in a sample of individuals unselected for disease state, we report a milder phenotype resulting from heterozygous mutations in VPS13A. Similar to VPS13A, Mendelian lipid genes having a large effect on plasma lipid levels have been shown to harbor common variants with smaller effects on phenotype.

We found 3 individuals who are compound heterozygous for null mutations in PCSK9. Previously, only 2 individuals with PCSK9 deficiency have been reported. Both of the previously reported individuals were young (21 and 31 years old) and had low circulating LDL-C (14–16 mg/dL). The 3 individuals we have identified here are older (50–52 years old) and have higher circulating LDL-C (32–72 mg/dL). One of the 3 individuals had a CAC score in the 83rd percentile with an increase in red blood cell distribution width. Red blood cell distribution width is a measure of the range of variation in red blood cells and higher values can indicate certain disorders such as anemia. Mutations in VPS13A have been reported to cause chorea-atachycytosis, an autosomal-recessive neurodegenerative disorder that causes red blood cells to appear spiky. Ten VPS13A variants are reported in ClinVar with chorea-atachycytosis listed as the condition. We did not find any of the reported ClinVar variants in our data nor any carriers of rare damaging recessive variants in VPS13A. Here, in a sample of individuals unselected for disease state, we report a milder phenotype resulting from heterozygous mutations in VPS13A. Similar to VPS13A, Mendelian lipid genes having a large effect on plasma lipid levels have been shown to harbor common variants with smaller effects on phenotype.

Some limitations deserve mention. The association between VPS13A and red blood cell distribution width needs to be confirmed in an independent study. Furthermore, sequencing will be required for replication; none of the variants driving the novel gene-based association were available on the widely-used exome genotyping array. The few results passing our prespecified significance level could be explained by statistical power given our sample size and the limited number of observed null alleles per gene. We also note that we have used a stringent significance threshold given the multiple testing burden inherent in our study design. In conclusion, a limited number of null/damaging alleles with a large effect on cardiovascular traits were detectable from the exome sequences of 3000 black individuals.

Sources of Funding
G.M.P. is supported by the National Heart, Lung, and Blood Institute of the National Institutes of Health under Award Number K01HL125751. S.K. is supported by a Research Scholar award from the Massachusetts General Hospital (MGH), the Howard Goodman Fellowship from MGH, the Donovan Family Foundation, R01HL107816, and a grant from Fondation Leducq. The Jackson Heart Study is supported by contracts HHSN268201300046C, HHSN268201300047C, HHSN268201300048C, HHSN268201300049C, and HHSN268201300050C from the National Heart, Lung, and Blood Institute and the National Institute on Minority Health and Health Disparities. The MH-GRID Network (Investigators: Rakale C. Quarells, Gary H. Gibbons, Donna K. Arnett, Robert L. Davis, Suzanne M. Leal, Deborah A. Nickerson, James Perkins, Charles N. Rotimi, Joel H. Saltz, Herman A. Taylor, and James G. Wilson) was supported, in part, by a grant from the National Institute on Minority Health and Health Disparities (grant #1RC4MD005964).

Disclosures
None.

References
1. Abifadel M, Varret M, Rabès JP, Allard D, Ouqueruam K, Devillers M, et al. Mutations in PCSK9 cause autosomal dominant hypercholesterolemia. Nat Genet. 2003;34:154–156. doi: 10.1038/ng1161.
2. Cohen J, Pertsemlidis A, Kottowsky JK, Graham R, Garcia CK, Hobbs HH. Low LDL cholesterol in individuals of African descent resulting from frequent nonsense mutations in PCSK9. Nat Genet. 2005;37:161–165. doi: 10.1038/ng1509.
3. Shioji K, Mannami T, Kubo Y, Inamoto N, Takagi S, Goto Y, et al. Genetic variants in PCSK9 affect the cholesterol level in Japanese. Hum Genet. 2004;99:10–14. doi: 10.1007/s00439-003-0114-3.
4. Chen SN, Bailly chapter CM, Goto AM Jr, Tan Y, Willerson JT, Marian AJ. A common PCSK9 haplotype, encompassing the E670G coding single nucleotide polymorphism, is a novel genetic marker for plasma low-density lipoprotein cholesterol levels and severity of coronary ath- erosclerosis. J Am Coll Cardiol. 2005;45:1611–1619. doi: 10.1016/j.jacc.2005.01.051.
5. Cohen JC, Boerwinkle E, Mosley TH Jr, Hobbs HH. Sequence variations in PCSK9, low LDL, and protection against coronary heart disease. N Engl J Med. 2006;354:1264–1272. doi: 10.1056/NEJMoa054013.
6. Kathiresan S; Myocardial Infarction Genetics Consortium. A PCSK9 missense variant associated with a reduced risk of early-onset myocar-dial infarction. N Engl J Med. 2008;358:2299–2300. doi: 10.1056/NEJMoa0707445.
7. Stein EA, Mellis S, Yancopoulos GD, Stahl N, Logan D, Smith WB, et al. Effect of a monoclonal antibody to PCSK9 on LDL cholesterol. N Engl J Med. 2012;366:1108–1118. doi: 10.1056/NEJMoa1105803.
8. Robinson JG, Kastelein JJ. PCSK9 inhibitors and cardiovascular events. N Engl J Med. 2015;373:774. doi: 10.1056/NEJMoa1508222.
9. Sabatine MS, Wasserman SM, Stein EA. PCSK9 inhibitors and cardiovas-cular events. N Engl J Med. 2015;373:774–775.
10. Cohen JC, Hobbs HH. Genetics. Simple genetics for a complex disease. Science. 2013;340:689–690. doi: 10.1126/science.1239101.
11. Fuqua SR, Wyatt SB, Andrew ME, Sarpong DF, Henderson FR, Cumming-ham MF, et al. Recruiting African-American research participation in the Jackson Heart Study: methods, response rates, and sample description. Ethn Dis. 2015;15(4 suppl 6):S6–S18.
12. Jun G, Flickringer M, Hetrick NK, Roman JM, Doheny KF, Abecasis GR, et al. Detecting and estimating contamination of human DNA samples in sequen-cing and array-based genotype data. Am J Hum Genet. 2012;91:839–848. doi: 10.1016/j.ajhg.2012.09.004.
13. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007;81:559–575. doi: 10.1086/519795.
14. Liu X, Jian X, Boerwinkle E. Dbsnp v2.0: a database of human non-synonymous snvs and their functional predictions and annotations. Hum Mutat. 2013;34:E2393–E2402.
15. Purcell SM, Moran JL, Froomer M, Ruderfer D, Solovieff N, Roussos P, et al. A polygenic burden of rare disruptive mutations in schizophrenia. Nature. 2014;506:185–190. doi: 10.1038/nature12975.
16. Chun S, Fay JC. Identification of deleterious mutations within three human genomes. Genome Res. 2009;19:1553–1561. doi: 10.1101/gr.092619.109.
17. Schwartz JM, Cooper DN, Schuelke M, Seelow D. MutationTaster2: muta-tion prediction for the deep-sequencing age. Nat Methods. 2014;11:361–362. doi: 10.1038/nmeth.2990.
18. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, et al. A method and server for predicting damaging missense mutations. Nat Methods. 2010;7:248–249. doi: 10.1038/nmeth0410-248.
19. Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nat Protoc. 2012;7:839–846. doi: 10.1038/nprot.2012.408.
20. Reva B, Antipin Y, Sandler C. Predicting the functional impact of protein mutations: application to cancer genomics. Nucleic Acids Res. 2011;39:e18. doi: 10.1093/nar/gkr407.
21. Shihab HA, Gough J, Cooper DN, Stenson PD, Barker GL, Edwards KJ, et al. Predicting the functional, molecular, and phenotypic consequences of amino acid substitutions using hidden Markov models. Hum Mutat. 2013;34:57–65. doi: 10.1002/humu.22225.
22. Cui JS, Hopper JL, Harrap SB. Antiinflammatory treatments obscure familial contributions to blood pressure variation. Hypertension. 2003;41:207–210.
23. Peloso GM, Auer PL, Bis JC, Vooraman A, Morrison AC, Stützel NO, et al; NHLBI GO Exome Sequencing Project. 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.* 2014;94:223–232. doi: 10.1016/j.ajhg.2014.01.009.
24. Kang HM, Sul JH, Service SK, Zaitlen NA, Kong SY, Freimer NB, et al. Variance component model to account for sample structure in genome-wide association studies. *Nat Genet.* 2010;42:348–354. doi: 10.1038/ng.548.
25. McClelland RL, Chung H, Detrano R, Post W, Kronmal RA. Distribution of coronary artery calcium by race, gender, and age: results from the Multi-Ethnic Study of Atherosclerosis (MESA). *Circulation.* 2006;113:30–37. doi: 10.1161/CIRCULATIONAHA.105.580696.
26. Auer PL, Johnsen JM, Johnson AD, Logsdon BA, Lange LA, Nalls MA, et al. Imputation of exome sequence variants into population-based samples and blood-cell-trait-associated loci in African Americans: NHLBI GO Exome Sequencing Project. *Am J Hum Genet.* 2012;91:794–808. doi: 10.1016/j.ajhg.2012.08.031.
27. Li AH, Morrison AC, Kovar C, Cupples LA, Polfus LM, et al. Analysis of loss-of-function variants and 20 risk factor phenotypes in 8,554 individuals identifies loci influencing chronic disease. *Nat Genet.* 2015;47:640–642. doi: 10.1038/ng.3270.
28. Pollin TI, Damcott CM, Shen H, Ott SH, Shelton J, Horenstein RB, et al. A null mutation in human APOC3 confers a favorable plasma lipid profile and apparent cardioprotection. *Science.* 2008;322:1702–1705. doi: 10.1126/science.1161524.
29. TG and HDL Working Group of the Exome Sequencing Project, National Heart, Lung, and Blood Institute, Crosby J, Peloso GM, Auer PL, Crosslin DR, Sützlief NO, et al. Loss-of-function mutations in apoc3, triglycerides, and coronary disease. *N Engl J Med.* 2014;371:22–31.
30. Jørgensen AB, Frikke-Schmidt R, Nordestgaard BG, Tybjærg-Hansen A. Loss-of-function mutations in APOC3 and risk of ischemic vascular disease. *N Engl J Med.* 2014;371:32–41. doi: 10.1056/NEJMoa1308027.
31. Dobson-Stone C, Danek A, Rampoldi L, Hardie RJ, Chalmers RM, Wood NW, et al. Mutational spectrum of the CHAC gene in patients with chorea-acanthocytosis. *Eur J Hum Genet.* 2002;10:773–781. doi: 10.1038/sj.ejhg.5200866.
32. Johansen CT, Wang J, Lanktree MB, Cao H, McIntyre AD, Ban MR, et al. Excess of rare variants in genes identified by genome-wide association study of hypertriglyceridemia. *Nat Genet.* 2010;42:684–687. doi: 10.1038/ng.628.
33. Teslovich TM, Musunuru K, Smith AV, Edmondson AC, Stylianou IM, Koseki M, et al. Biological, clinical and population relevance of 95 loci for blood lipids. *Nature.* 2010;466:707–713. doi: 10.1038/nature09270.
34. Gloyn AL, McCarthy MI. Variation across the allele frequency spectrum. *Nat Genet.* 2010;42:648–650. doi: 10.1038/ng0810-648.
35. Hooper AJ, Marais AD, Tanyanyiwa DM, Burnett JR. The C679X mutation in PCSK9 is present and lowers blood cholesterol in a Southern African population. *Atherosclerosis.* 2007;193:445–448. doi: 10.1016/j.atherosclerosis.2006.08.039.
36. Zhao Z, Tuakli-Wosornu Y, Lagace TA, Kinch L, Grishin NV, Horton JD, et al. Molecular characterization of loss-of-function mutations in PCSK9 and identification of a compound heterozygote. *Am J Hum Genet.* 2006;79:514–523. doi: 10.1086/507488.

**CLINICAL PERSPECTIVE**

The correlation of null alleles with human phenotypes can provide insight into gene function in humans. Here, we performed whole-exome sequencing in 3223 black individuals living in Jackson, Mississippi, to identify null and damaging missense variants and test these variants for association with 36 cardiovascular traits. We replicated the association of null and damaging missense variants in PCSK9 with low-density lipoprotein cholesterol and found 3 individuals in their 50s each compound heterozygous for PCSK9. Of note, one of these 3 individuals had a coronary artery calcification score in the 83rd percentile despite a low-density lipoprotein cholesterol of 32 mg/dL. We also found that individuals with rare damaging missense variants in VPS13A had higher red blood cell distribution width compared with noncarriers. Mutations in VPS13A have been previously reported to cause chorea-acanthocytosis, an autosomal-recessive neurodegenerative disorder that causes red blood cells to appear spiky. Only a limited number of null/damaging alleles with a large effect on cardiovascular traits were detectable in ≈3000 black individuals.