Common genetic variation near the connexin-43 gene is associated with resting heart rate in African Americans: A genome-wide association study of 13,372 participants

R. Deo, MD, MTR, M.A. Nalls, PhD, C.L. Avery, PhD, J.G. Smith, MD, D.S. Evans, PhD, MPH, M.F. Keller, BS, A.M. Butler, MS, S.G. Buxbaum, PhD, G. Li, MS, P. Miguel Quibrera, MD, E.N. Smith, PhD, T. Tanaka, PhD, E.L. Akylbekova, MS, A. Alonso, MD, PhD, E.R. Fox, MD, M.F. Keller, BS, A.M. Butler, MS, S.G. Buxbaum, PhD, E.N. Smith, PhD, T. Tanaka, PhD, E.L. Akylbekova, MS, A. Alonso, MD, PhD, E.R. Fox, MD, M.F. Keller, BS, A.M. Butler, MS, S.G. Buxbaum, PhD, E.N. Smith, PhD, T. Tanaka, PhD, E.L. Akylbekova, MS, A. Alonso, MD, PhD, E.R. Fox, MD, M.F. Keller, BS, A.M. Butler, MS, S.G. Buxbaum, PhD, E.N. Smith, PhD, T. Tanaka, PhD, E.L. Akylbekova, MS, A. Alonso, MD, PhD.
BACKGROUND Genome-wide association studies have identified several genetic loci associated with variation in resting heart rate in European and Asian populations. No study has evaluated genetic variants associated with heart rate in African Americans.

OBJECTIVE To identify novel genetic variants associated with resting heart rate in African Americans.

METHODS Ten cohort studies participating in the Candidate-gene Association Resource and Continental Origins and Genetic Epidemiology Network consortia performed genome-wide genotyping of single nucleotide polymorphisms (SNPs) and imputed 2,954,965 SNPs using HapMap YRI and CEU panels in 13,372 participants of African ancestry. Each study measured the RR interval (ms) from 10-second resting 12-lead electrocardiograms and estimated RR-SNP associations using covariate-adjusted linear regression. Random-effects meta-analysis was used to combine cohort-specific measures of association and identify genome-wide significant loci ($P \leq 2.5 \times 10^{-8}$).

RESULTS Fourteen SNPs on chromosome 6q22 exceeded the genome-wide significance threshold. The most significant association was for rs9320841 (+13 ms per minor allele; $P = 4.98 \times 10^{-15}$). This SNP was approximately 350 kb downstream of GJA1, a locus previously identified as harboring SNPs associated with heart rate in Europeans. Adjustment for rs9320841 also attenuated the association between the remaining 13 SNPs in this region and heart rate. In addition, SNPs in MYH6, which have been identified in European genome-wide association study, were associated with similar changes in the resting heart rate as this population of African Americans.

CONCLUSIONS An intergenic region downstream of GJA1 (the gene encoding connexin 43, the major protein of the human myocardial gap junction) and an intragenic region within MYH6 are associated with variation in resting heart rate in African Americans as well as in populations of European and Asian origin.

KEYWORDS African Americans; Heart rate; Single nucleotide polymorphisms; Meta-analysis.
Abstract

The study was approved by the institutional review boards at each participating center. Written informed consent was obtained from all participants.

Methods

Study populations

The CARE\(^1\)\(^,\)\(^9\) and COGENT\(^20\) consortia included 13,372 self-reported African Americans meeting inclusion criteria. The participants originated in 10 cohort studies: the Atherosclerosis Risk in Communities study (ARIC; n = 2391); Baltimore Longitudinal Study of Aging (BLSA; n = 155); Bogalusa Heart Study (BHS; n = 148); Cardiovascular Health Study (CHS; n = 674); Cleveland Family Study (CFS; n = 267); the Health, Aging, and Body Composition Study (Health ABC; n = 1054); the Healthy Aging in Neighborhoods of Diversity across the Life Span Study (HANDLS; n = 945); Jackson Heart Study (JHS; n = 1962); Multi-Ethnic Study of Atherosclerosis (MESA; n = 1627); and Women’s Health Initiative clinical trials (WHI; n = 4149). Additional information is provided in the Supplemental Methods, including cohort-specific genotype and imputation quality control methods (see Online Supplements 1 and 2). Participants with missing covariates, poor-quality ECGs, pacemakers or implantable cardioverter-defibrillators, paroxysmal or persistent atrial fibrillation, heart failure, myocardial infarction, second- or third-degree atrioventricular block, and extremes of heart rate (>100 or <50 beats/min) were excluded. Participants on medications altering nodal or atrioventricular conduction (beta-blockers, nondihydropyridine calcium channel blockers, digoxin, type I or III antiarhythmics) were also excluded.

ECC recordings

A standard 10-second, resting ECG was obtained and recorded digitally on all participants from the 10 cohorts included in this analysis. Standard 12-lead positions were recorded at baseline in all cohort studies using a Marquette MAC PC, MAC6, or MAC1200 ECG machine system (GE Healthcare, Milwaukee, WI). The RR interval (ms) was measured electronically as the unit-corrected inverse of heart rate (beat/min). All ECGs were processed automatically using GE Marquette 12-SL version 2001 running under GE Magellan Research Work Station or MC Means. The ECG software is Food and Drug Association approved. Heart rate was calculated from the median RR interval during the 10-second recording. Since ECG recordings were simultaneous in all 12 leads, the rate was not affected by the lead from which the RR interval was recorded. The automated nature of calculating heart rate from the median RR interval ensures the highest repeatability with no inter- or intraobserver variability. Poor-quality ECGs were excluded by software algorithms. As an added quality control measure, all ECGs were visually checked.

After a filtering process that results in signal conditioning and averaging, the program generates a median complex. All QRs of the same shape are aligned in time and the interval measurements depend on the proper identification of fiduciary points, which are determined from an analysis of all 12 leads simultaneously. The intervals are then measured according to published standards.\(^21\)

Genotyping and quality control

Genome-wide SNP genotyping was performed within each cohort using genotyping arrays from Affymetrix or Illumina (Online Supplement 2). Studies underwent similar quality control procedures (specific details in the Online Supplemental Materials). DNA samples with an array-wide genotyping success rate <95% were excluded. Autosomal heterozygosity rates were estimated to identify and exclude samples with poor DNA quality or contamination. Duplicated or contaminated samples were identified from identity by descent estimates and excluded. In addition, SNPs with a genotyping success rate <90% per SNP within each cohort, SNPs that map to multiple locations, SNPs where missingness could be predicted from surrounding haplotypes, and SNPs...
associated with chemistry plates were excluded. African ancestry was confirmed through either principal components or multi-dimensional scaling analyses. Population-based (ie, non-family-based) studies used identity-by-descent (IBD) estimates to exclude cryptically related individuals. Subsequent identical SNP filters after imputation and GWAS analyses were applied to summary statistics at the meta-analysis level.

**Imputation and quality control**

SNP imputation was performed in each cohort to facilitate the combination of results from different genotyping platforms and to increase genotype coverage. Genotyped SNPs passing quality control metrics described above and reference haplotypes from HapMap Phase 2 (release 22 on NCBI build 36) were used to impute approximately 2.5 million SNPs using MACH v1.16 or BEAGLE. Untyped SNPs were imputed using a 1:1 ratio of CEU/YRI HapMap reference haplotypes based on consistency across all CARe-COGENT studies. Imputed SNPs were excluded if imputation quality was below 0.30 as reported by MACH or BEAGLE.

**Statistical analysis**

GWAS analysis was performed in either PLINK (ARIC, BHS, CHS, JHS, WHI), R (HANdls, Health ABC, MESA), ProbABEL (WHI), or MERLIN (BLSA) using linear regression with an additive genetic model based on allelic dosages accounting for imputation uncertainty. The family-based CFS study was analyzed using linear mixed-effects models as implemented in the GWAF package for R.24 Pedigrees for CFS were confirmed using identity by state or IBD estimates from PREST-Plus (http://www.ustat.utoronto.ca/Sun/Software/PresT/). Previously published analyses indicated that the inclusion of related individuals from the JHS family-based subcohort had little effect on P-value inflation.20 As a result, these related individuals were included in the present analysis. Eigenvectors were used to adjust for global ancestry in population substructures. Principal components were used to adjust for global ancestry in population stratification.

Cohort-specific genome-wide association was examined on a SNP-by-SNP basis using simple linear models regressing RR on allele dosage, age, sex, body mass index, global measures of African ancestry, and, when relevant, study site. Cohort-specific SNP association estimates were combined using fixed- and random-effects meta-analysis, the latter to examine potential effects of among-cohort heterogeneity on the combined estimates and the extent to which it can support qualitative inference to other African American populations. Given evidence of greater genetic and geographical diversity across African American cohorts compared to Europeans and initial evidence of heterogeneity across studies, random-effects estimates, which have wider 95% confidence intervals than do fixed-effects estimates, were reported in the current meta-analysis. Genomic control methods were applied when study-specific and combined distributions of test statistics suggested early departure from the null (λ > 1). Genomic inflation factors were evaluated in each cohort before the random-effects meta-analysis and in the combined results.25 We calculated X^2 estimates of homogeneity (Cochran’s Q) using METAL and I^2 estimates with R. Prior to conducting meta-analyses, SNP results with a minor allele frequency <0.01 or imputation quality scores <0.3 were excluded. In addition, SNPs not seen in >2 studies were excluded from the meta-analyses.

To confirm that the random-effects model was not overly conservative, standard fixed-effects meta-analyses were conducted on SNP association estimates for each cohort using METAL (and incorporating genomic control at the meta-analysis level). For the meta-analysis, we prespecified a genome-wide significance threshold of 2.5 × 10^{-8} as suggested for populations of African ancestry,26 accounting for approximately 2 million independent common variant tests. Other polymorphisms that were detected at the same locus as the initial SNP were subsequently analyzed in conditional regression models to assess statistical independence. Finally, SNPs that have been identified in prior GWAS but not in the discovery phase of our analysis were evaluated using a less stringent threshold. Specifically, we evaluated 13 genome-wide significant SNPs described by prior RR GWAS in individuals of European and Asian ancestry using a significance level of 3.85 × 10^{-3} (Bonferroni corrected P value calculated as 0.05/13).

**Results**

This GWAS of the RR interval included 13,372 adults of African descent from 10 cohort studies. Each study contributed a widely varying number of participants (range 148–4149). The ARIC, JHS, and WHI studies accounted for the majority of participants in this analysis: 8502 (64%) of 13,372. On average, the study population was middle-aged (mean 56.5 years; range 35–73 years).

| Table 1 Description of contributing African American cohort studies |
|---|---|---|---|---|
| **Cohort study** | **n** | **Age (y)** | **Sex (M): (%)** | **BMI (kg/m^2)** | **HR (beat/min)** | **RR interval (ms)** | **λ** |
| ARIC | 2391 | 53.3 (5.8) | 39 | 29.4 (6.1) | 67 (10) | 896 | 1.023 |
| BLSA | 155 | 64.4 (11.4) | 37 | 28.3 (5.2) | 63 (8) | 952 | 1.050 |
| BHS | 148 | 35.7 (4.8) | 33 | 31.7 (8.9) | 68 (11) | 882 | 1.004 |
| CHS | 674 | 72.8 (5.6) | 35 | 28.4 (5.5) | 67 (11) | 896 | 1.005 |
| CFS | 267 | 42.7 (14.9) | 43 | 34.4 (9.3) | 69 (9) | 875 | 1.070 |
| Health ABC | 1054 | 73.4 (2.9) | 45 | 28.1 (5.3) | 66 (8) | 909 | 0.996 |
| HANDLS | 945 | 48.5 (9.0) | 44 | 29.9 (8.1) | 67 (11) | 896 | 1.007 |
| JHS | 1962 | 49.3 (11.8) | 37 | 32.4 (5.8) | 66 (10) | 909 | 1.007 |
| MESA | 1627 | 61.5 (10.1) | 46 | 30.2 (5.9) | 65 (9) | 923 | 1.003 |
| WHI | 4149 | 61.7 (6.9) | 0 | 31.6 (6.2) | 66 (8) | 909 | 1.017 |
| All studies* | 13372 | 56.5 | 29 | 30.8 | 66.3 | 906 | 1.029 |

Mean (standard deviation) is tabulated for age, body mass index (BMI), and heart rate (HR).

ARIC = Atherosclerosis Risk in Communities study; BHS = Bogalusa Heart Study; BLSA = Baltimore Longitudinal Study on Aging; CFS = Cleveland Family Study; CHS = Cardiovascular Health Study; HANDLS = Healthy Aging in Neighborhoods of Diversity across the Life Span Study; Health ABC = Health and Body Composition; JHS = Jackson Heart Study; MESA = Multi-Ethnic Study of Atherosclerosis; WHI = Women’s Health Initiative clinical trials.

*Sum (n), % (male), and weighted mean (age; BMI; HR; RR interval; λ) across studies.
and overweight (mean body mass index 30.8 kg/m²) and 71% were women.

Genomic inflation was minimal in most studies and modest in the family-based CFS ($\lambda$ 1.070) and JHS ($\lambda$ 1.071) (Table 1). Specifically, the lambda estimates from the random-effects meta-analysis did not suggest inflation of the test statistic (0.868), and the secondary fixed-effects modeling did not show a significant departure from null expectations ($\lambda$ 1.017) (Figures 1A and 1B).

A total of 2,954,965 SNPs were incorporated into this meta-analysis after data quality control. Fourteen SNPs at a largely intergenic region on chromosome 6q22 (Figure 2) reached genome-wide significance. The most significant association at this locus was for rs9320841 (+13 ms per minor allele, standard error 1.7 ms, random effects $P = 4.98 \times 10^{-15}$). This SNP is located in a noncoding region, 350 kb downstream from $GJA1$ and 64 kb upstream from $HMGB3P18$. The magnitude and direction of the association were similar across most cohorts ($P_{\text{heterogeneity}} = .45$) as shown in Figure 3. None of the other 13 SNPs in this region were independent variants associated with resting heart rate. The results for the regional association plot at the $GJA1$ locus are depicted in Figure 4. This plot covers 1000 kb of the genomic region associated with the $GJA1$ locus and demonstrates strong linkage disequilibrium (LD) with other SNPs in this gene cluster that were associated with variations in heart rate. Adjustment for rs9320841, however, eliminated the significance of these additional SNPs.

We also evaluated a series of SNPs from the chromosome 6q22 locus that were identified in prior European and Asian GWAS. Both rs9398652 and rs12110693 in the 6q22 locus were associated with the RR interval, which were similar to
estimates reported in prior studies of Asian14 and European13 populations; however, only rs9398652 reached genome-wide significance in the current meta-analysis. The rs9398652 SNP was approximately 30 kb downstream and in high linkage disequilibrium with the leading SNP from the present study (rs9320841; CEU $r^2 = 1.00$; YRI $r^2 = .81$). In addition, rs12110693 was in strong linkage disequilibrium with rs9320841 (rs9320841; CEU $r^2 = 1.00$; YRI $r^2 = .76$) (Table 2). The final reported SNP from the 6q22 locus, rs11154022, did not reach genome-wide significance, was the greatest distance from rs9320841 (approximately 365 kb upstream), and not in LD with it (CEU $r^2 = .01$; YRI $r^2 = .01$).

Other variants that were identified from prior European and Asian GWAS were also tested (Table 2). The 2 SNPs that have previously been identified at the MYH6 locus (rs452036 and rs365990) were associated with resting heart rate in African Americans using the replication threshold ($3.85 \times 10^{-3}$). These variants are associated with a similar increase in the sinus cycle length across Europeans, Asians, and African Americans. We were unable to confirm associations for several previously published loci at replication thresholds: CD46 on chromosome 1, SLC35F1 on chromosome 6, SLC12A9 and UfSp1 on chromosome 7, FADS1 on chromosome 11, an intergenic region on chromosome 12, GPR133 on chromosome 12, and MYH7 on chromosome 14 (Table 2). These findings were consistent across the different cohorts analyzed through the CARe-COGENT consortium (Figure 5). Further evaluation of these loci (the 1 Mb regions, 500 kb upstream and downstream of the SNPs in Table 2) did not identify any other genome-wide significant RR-SNP associations despite having adequate power (power $> 0.8$).

**Discussion**

In a large GWAS of African Americans, we generalized a previously reported association between a variant on chromosome 6q22.3 and resting heart rate to a population of African descent. The present findings suggest that rs9320841, which is located in an intergenic region 350 kb downstream from GJA1, is the leading SNP at this locus associated with heart rate. In addition,
rs9320841 is in high LD with other intergenic SNPs from this region previously associated with heart rate in GWAS in populations of European and Asian ancestry.13,14

Multiple studies including the current report have demonstrated intergenic SNPs in proximity to rs9320841 that are associated with variation in heart rate among individuals of Asian, European, and African ancestry. The closest putative transcript to rs9320841 on chromosome 6q22.31 is HMGB3P18, which has no known function. However, GJA1, which is approximately 350 kb upstream of this SNP, encodes connexin 43, the main cardiac gap junction channel that is found throughout the heart and is responsible for intercellular conductance in the atria and ventricles.27 Connexin 43 is expressed abundantly in the atria and permits the node to conduct impulses to the surrounding muscle.28 Experimental models have demonstrated that the deletion of various gap junction subunits results in a sick sinus syndrome phenotype with bradycardia, sinus dysrhythmia, and sinus node exit block.29,30 As a result, these intergenic variants in the 6q22 locus, which are in close proximity to GJA1 and have been identified across different populations, may reduce sinus automaticity.

Although rs9320841 and previously identified 6q22.31 loci are 300–500 kb away from and in low LD with SNPs in GJA1, recent studies suggest that variations in intergenic regions may regulate transcription factor binding and chromatin modification.31 Functional and translational studies focused on this intergenic region on chromosome 6q22 will be required to understand its potential effect on GJA1.

In the portion of our study that restricted the analysis to previously identified variants, we observed an association between 2 SNPs located within the MYH6 gene and resting heart rate. MYH6 encodes one of the myosin heavy chain subunits in the cardiac sarcomere and is a major component of the cardiac contractile system. In addition, MYH6 encodes a cardiac-specific microRNA, miR-208a, which is a key regulatory molecule that is necessary for normal cardiac conduction.32 Specifically, miR-208a regulates expression of connexin 40, a gap junction protein that is implicated in sinus automaticity and cardiac arrhythmias.29–33 As a result, changes in the MYH6 genetic architecture could alter microRNA production, gap junction formation, and sinus node function. Prior GWAS in European populations have identified common variants in this gene to be associated with resting heart rate13,15 and rare variants, located 0.3–4.4 kb from these SNPs, to be associated with sick sinus syndrome.34 Although SNPs at the MYH6 locus were not identified in the discovery phase of our analysis at genome-wide significance thresholds ($P < 2.5 \times 10^{-8}$), the similar magnitude and direction of the point estimates in our analysis suggest that the MYH6 gene affects sinus node automaticity in diverse populations.

While we were unable to replicate associations for other previously published loci at a threshold level of $3.85 \times 10^{-3}$ (0.05/13), the similar magnitude and direction of the point estimates suggest consistency across ancestries. Specifically, SLC12A9 and UfSp1 on chromosome 7 and the MYH7 region on chromosome 14 had effects on heart rate similar to those described by prior studies. Compared to individuals of European ancestry, however, African Americans have greater genetic diversity,18 which may lower the frequency of a particular allele and subsequently reduce the statistical likelihood of detecting an effect on the RR interval. In addition, linkage disequilibrium is commonly lower in African Americans35 and subsequently reduces the likelihood that a common SNP is in linkage disequilibrium with a causal variant.
Furthermore, these analyses were conducted in a population that was predominantly woman, middle-aged, and overweight. This demographic profile differs from that of prior studies and may have influenced the results.

A common limitation of meta-analyses is among-study phenotype heterogeneity; however, the current study followed similar electrocardiographic and clinical protocols when measuring heart rate and its correlates. In addition, the statistical assessment of heterogeneity did not suggest large variation in SNP effects across studies. Moreover, the random-effects meta-analysis of these effects was weighted for both their within- and among-study variation. Another limitation of GWAS is potential for population stratification, including confounding by ancestry. However, we attempted to minimize bias from population structure by excluding participants of non-African ancestry, adjusting for principal components in study-specific regression models, and applying genomic control methods.

Conclusions

In summary, the genome-wide significance of an association linking resting heart rate and the GJA1 locus previously described in European and Asian populations has now been generalized to African Americans. In addition, this analysis has replicated associations initially discovered in Europeans between common variants within the MYH6 gene and a reduction in heart rate to an African American population. Generalizability across global populations and biological plausibility of the heart rate-GJA1 and heart rate-MYH6 associations highlight the potential importance of these loci in the intrinsic (nodal and myocardial) determination of resting heart rate.

Appendix A

Supplementary data

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.hrthm.2012.11.014.

References

Supplement 1: Description of Cohort Studies

**ARIC**: The Atherosclerosis Risk in Communities study is a prospective population-based study of atherosclerosis and cardiovascular disease that included 15,792 men and women between 45 and 64 years from four US communities. Of these participants, 4,314 individuals were self-reported black Americans from two of the four communities (Jackson, MS and Forsyth County, NC). Electrocardiographic recordings used in the present study were performed at baseline examinations between 1987-1989. Specifically, participants were at rest and underwent, standard supine 12-lead ECG with each tracing consisting of 10 of each of the 12 leads simultaneously. After exclusions, data on 2,391 African American individuals with genotypes, information on all covariates and informed consent remained for analyses.

**BLSA**: The Baltimore Longitudinal Study of Aging is a population-based study aimed to evaluate contributors of healthy aging in the older population residing predominantly in the Baltimore-Washington DC area. Healthy volunteers aged 18 and older were enrolled in the study starting in 1958. Participants are community-residing volunteers who tend to be well-educated, with above-average income and access to medical care. These subjects visit the Gerontology Research Center at regular intervals for two days of medical, physiological, and psychological testing. A resting 12-lead ECG was performed on all participants. There are approximately 1,100 active participants enrolled in the study. After exclusions, 155 African Americans with genotype and phenotype data were included in the analysis. The BLSA has continuing approval from the Institutional Review Board of the Johns Hopkins Bayview Medical Center, the Gerontology Research Center, and Medstar Research Institute. Informed consent was obtained from all participants.
**BHS:** The Bogalusa Heart Study is a biracial, community-based investigation of cardiovascular disease. Between 1973 and 2008, nine cross-sectional surveys of children aged 4-17 years and 10 cross-sectional surveys of adults aged 18-48 years, who had been examined as children, were conducted for cardiovascular disease risk factor examinations in Bogalusa, Louisiana. A standard resting, 12 lead electrocardiogram was performed on all participants. There are 1,202 Caucasian and African American participants who have been examined on multiple occasions from childhood to adulthood with DNA available for genotyping. After exclusions, 148 African Americans with genotype and phenotype data were included in the analysis. Study protocols were approved by the Institutional Review Board of the Tulane University Medical Center. Informed Consent was obtained from all participants.

**CHS:** The Cardiovascular Health Study is a community-based study of cardiovascular disease risk in ambulatory older adults. The study started recruiting participants in 1989 from Medicare eligibility lists in Forsyth County, North Carolina; Sacramento County, California; Washington County, Maryland; and Allegheny County, Pennsylvania. Initially, 244 African American adults aged 65 years or older were sampled; an additional 687 black participants were recruited between 1992 and 1993. Physical and laboratory evaluations were performed at baseline to identify the presence and severity of cardiovascular risk factors including hypertension, hypercholesterolemia, and glucose intolerance. Electrocardiographic recording was also performed at rest during the baseline clinical visit. In 2010, genotyping was performed for 844 African American participants who consented to genetic testing and had DNA available. After exclusions, 674 African Americans with genotype and phenotype data were included in the analysis. The institutional review board at each of the study sites approved the study protocols. In addition, written consent was obtained from all participants.
**CFS**: The Cleveland Family Study (CFS) is a family-based, longitudinal study designed to study the risk factors for sleep apnea. Participants include first-degree or selected second-degree relatives of a proband with either laboratory diagnosed obstructive sleep apnea or neighborhood control of an affected proband. Resting ECG recordings used for the present study were performed at the final exam cycle conducted in a Clinical Research Unit between 2001 and 2006. Families were selected for genotyping on the basis of genetic informativity, including multigenerational data or individuals from the extremes of the distribution of apnea phenotype. The 632 African Americans with available DNA were genotyped as part of CARe. After exclusions, 267 African Americans with genotype and phenotype data were included in the analysis. The institutional review board approved the study, and written informed consent was obtained from all participants.

**Health ABC**: The Health, Aging, and Body Composition Study is a prospective study designed to investigate the effect of age-related changes in body composition and health among well-functioning elderly persons between the ages of 70 and 79 years. Between 1997 and 1998, 3,075 adults were recruited using mass mailing with telephone follow-up. African American participants were recruited from two clinical centers in Pittsburgh, Pennsylvania and Memphis, Tennessee. The participants were relatively functional and had no reported difficulty in walking a quarter of a mile, climbing 10 steps without resting, or performing mobility-related activities of daily living. Resting, twelve-lead ECGs were collected at the baseline examination and sent to the St. Louis University Core ECG Laboratory (St. Louis, Missouri) for analysis. In addition, a baseline examination with blood specimen collection was performed on all participants. After exclusions, 1,054 African Americans with genotype and phenotype data were included in the analysis. The institutional review boards at both clinical centers approved the study, and written informed consent was obtained from all participants.
HANDLS: The Healthy Aging in Neighborhoods of Diversity across the Life Span Study is a prospective longitudinal study that evaluates the effect of race and socioeconomic status on the development of age-related health disparities in overall longevity, cardiovascular disease, and cognitive decline among socioeconomically diverse African Americans and whites in Baltimore. A total of 2,200 African American participants between 30 and 64 years of age were recruited as a fixed cohort of participants by household screenings from an area probability sample of Baltimore based on the 2000 Census. Data including resting, twelve-lead electrocardiograms were collected from household interviews and an in-depth examination in a mobile medical research vehicle. Genotyping was performed for 1,024 participants who self-report as African Americans. After exclusions, 945 African Americans with genotype and phenotype data were included in the analysis. The study protocol was approved by the human subjects review boards at both MedStar Research Institute and the University of Delaware. All participants provided written informed consent.

JHS: The Jackson Heart Study (JHS) is a prospective, community-based study that evaluates causal factors for the high prevalence of cardiovascular disease in African Americans. Between 2000 and 2004, 5,301 self-reported African Americans were recruited from the Jackson, Mississippi region. This study includes a subsample of unrelated participants (35-84 years) and a nested family-based subcohort. Resting ECG recordings used in the present study and blood collection for DNA extraction were performed at baseline examinations between 2000-2004. Genotype data were available in 3,030 individuals including 885 who were also included in the ARIC study. After exclusions, 1,962 individuals with ECG data and not included in ARIC were included in the present analysis. The institutional review board approved the study protocol, and written informed consent was obtained from all participants.
MESA: The Multi-Ethnic Study of Atherosclerosis (MESA) is a population-based cohort designed to investigate subclinical cardiovascular disease and the risk factors associated with progression to clinical disease. Between 2000 and 2002, the study enrolled 6,814 individuals (28% African Americans) aged 45-84 from six US field centers. Those with a history of cardiovascular disease were excluded from participation. Resting electrocardiographic recordings used in the present study and blood sampling for DNA extraction were obtained at baseline visits. After exclusions, 1627 African Americans with genotype and phenotype data remained in the study. The institutional review boards at each study site approved the study, and written informed consent was obtained from all participants.

WHI: The Women’s Health Initiative is a large, prospective study to evaluate the risk factors for common diseases including cardiovascular disease, cancer, and osteoporotic fractures. The WHI comprises both randomized clinical trials and an observational study. This study is limited to WHI clinical trial participants with standard twelve-lead ECGs obtained at rest. ECGs were not available for the observational study participants. The clinical trials evaluated estrogen with or without progestin treatment, calcium/vitamin D supplementation, and dietary modification on the risk of breast and colorectal cancer, cardiovascular disease and bone fractures in post-menopausal women. Between 1993-1998, the trials enrolled 68,132 postmenopausal women aged 50-79 years. From the WHI emerged the WHI SNP Health Association Resource (SHARe) GWAS project, which includes 12,157 (8,515 self-identified African American and 3,642 self-identified Hispanic) women who consented to genetic research. After exclusions, 4,149 African Americans with genotype and phenotype data were included in the analysis. The Institutional Review Boards at all participating institutions approved study protocols and consent forms.
### Supplement 2

Supplement 2. Genotyping characteristics for ten studies of 13,372 African American participants.

| Characteristic              | ARIC     | BLSA     | BHS      | CFS      | CHS       | Health ABC | HANDLS    | JHS      | MESA     | WHI      |
|----------------------------|----------|----------|----------|----------|-----------|------------|-----------|----------|----------|----------|----------|
| Genotyping array           | Affymetrix 6.0 | Illumina Human610, HumanCVD BeadChip | Illumina HumanOmni1-Quad_v1 BeadChip system | Affymetrix 6.0 | Illumina 1M | Illumina 1M | Affymetrix 6.0 | Affymetrix 6.0 | Affymetrix 6.0 |
| Genotype calling software  | Birdseed v1.33 | BeadStudio | BeadStudio | Birdseed v1.33 | BeadStudio | BeadStudio | Birdseed v1.33 | Birdseed v1.33 | Birdseed v2 |
| Sample call rate exclusion | <95%     | ≤98.5%   | <99%     | <95%     | <97%     | <97%       | ≤95%      | <95%     | <95%     | <95%     |
| SNP call rate exclusion    | <90%     | ≤99.5%   | <90%     | <90%     | <97%     | <97%       | ≤95%      | <90%     | <90%     | <95%     |
| SNP MAF exclusion*         | <1%      | <1%      | NA       | <1%      | None      | <1%        | ≤1%       | <1%      | <1%      | <1%      |
| SNP HWE P-value exclusion* | NA       | ≤10^{-4} | NA       | NA       | <10^{-5} | <10^{-6}   | ≤10^{-7}  | NA       | NA       | <10^{-6} |
| Imputation software        | MACH v1.16 | MACH 1.0 | MACH v1.0.16 | MACH v1.16 | MACH v1.16 | MACH v1.16 | MACH v1.16 | MACH v1.16 | MACH v1.16 |
| NCBI imputation build      | 36       | 36       | 36       | 36       | 36        | 36         | 36        | 36       | 36       |
| Genotyped autosomal SNPs passing QC | 796,384 | 501,704  | 608,756  | 867,495  | 963,248  | 1,007,948  | 907,763  | 868,969  | 881,666  | 829,370  |

* NA indicates no filter was applied. †Filter applied pre-imputation. Genotyped SNP results used to replace imputed SNP results were not filtered by HWE P-value.
Supplemental References


