



# LUND UNIVERSITY

## Genome-Wide Association Studies of the PR Interval in African Americans.

Smith, Gustav; Magnani, Jared W; Palmer, Cameron; Meng, Yan A; Soliman, Elsayed Z; Musani, Solomon K; Kerr, Kathleen F; Schnabel, Renate B; Lubitz, Steven A; Sotoodehnia, Nona; Redline, Susan; Pfeufer, Arne; Müller, Martina; Evans, Daniel S; Nalls, Michael A; Liu, Yongmei; Newman, Anne B; Zonderman, Alan B; Evans, Michele K; Deo, Rajat; Ellinor, Patrick T; Paltoo, Dina N; Newton-Cheh, Christopher; Benjamin, Emelia J; Mehra, Reena; Alonso, Alvaro; Heckbert, Susan R; Fox, Ervin R

Published in:  
PLoS Genetics

DOI:  
[10.1371/journal.pgen.1001304](https://doi.org/10.1371/journal.pgen.1001304)

2011

[Link to publication](#)

### Citation for published version (APA):

Smith, G., Magnani, J. W., Palmer, C., Meng, Y. A., Soliman, E. Z., Musani, S. K., Kerr, K. F., Schnabel, R. B., Lubitz, S. A., Sotoodehnia, N., Redline, S., Pfeufer, A., Müller, M., Evans, D. S., Nalls, M. A., Liu, Y., Newman, A. B., Zonderman, A. B., Evans, M. K., ... Fox, E. R. (2011). Genome-Wide Association Studies of the PR Interval in African Americans. *PLoS Genetics*, 7(2), Article e1001304. <https://doi.org/10.1371/journal.pgen.1001304>

Total number of authors:  
28

### General rights

Unless other specific re-use rights are stated the following general rights apply:

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: <https://creativecommons.org/licenses/>

### Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117  
221 00 Lund  
+46 46-222 00 00

# Genome-Wide Association Studies of the PR Interval in African Americans

J. Gustav Smith<sup>1,2\*</sup>, Jared W. Magnani<sup>3,4</sup>, Cameron Palmer<sup>2</sup>, Yan A. Meng<sup>2</sup>, Elsayed Z. Soliman<sup>5</sup>, Solomon K. Musani<sup>6</sup>, Kathleen F. Kerr<sup>7</sup>, Renate B. Schnabel<sup>8</sup>, Steven A. Lubitz<sup>9,10</sup>, Nona Sotoodehnia<sup>11</sup>, Susan Redline<sup>12</sup>, Arne Pfeufer<sup>13,14</sup>, Martina Müller<sup>15,16</sup>, Daniel S. Evans<sup>17</sup>, Michael A. Nalls<sup>18</sup>, Yongmei Liu<sup>19</sup>, Anne B. Newman<sup>20</sup>, Alan B. Zonderman<sup>21</sup>, Michele K. Evans<sup>21</sup>, Rajat Deo<sup>22</sup>, Patrick T. Ellinor<sup>10,23</sup>, Dina N. Paltoo<sup>24</sup>, Christopher Newton-Cheh<sup>2,10,23</sup>, Emelia J. Benjamin<sup>3,4,25,26,27</sup>, Reena Mehra<sup>28</sup>, Alvaro Alonso<sup>29</sup>, Susan R. Heckbert<sup>30</sup>, Ervin R. Fox<sup>6</sup>, Candidate-gene Association Resource (CARE) Consortium<sup>†</sup>

**1** Department of Cardiology, Faculty of Medicine, Lund University, Lund, Sweden, **2** Program for Medical and Population Genetics, Broad Institute of Harvard and Massachusetts Institute of Technology, Cambridge, Massachusetts, United States of America, **3** Section of Cardiovascular Medicine, Boston University School of Medicine, Boston, Massachusetts, United States of America, **4** National Heart, Lung and Blood Institute's Framingham Heart Study, Framingham, Massachusetts, United States of America, **5** Epidemiological Cardiology Center (EPICARE), Department of Epidemiology and Prevention, Wake Forest University School of Medicine, Winston-Salem, North Carolina, United States of America, **6** Department of Medicine, Division of Cardiovascular Disease, University of Mississippi Medical Center, Jackson Heart Study, Jackson, Mississippi, United States of America, **7** Department of Biostatistics, School of Public Health, University of Washington, Seattle, Washington, United States of America, **8** Medical Clinic 2 (Cardiology), Johannes-Gutenberg University, Mainz, Germany, **9** Division of Preventive Medicine, Brigham and Women's Hospital, Boston, Massachusetts, United States of America, **10** Cardiovascular Research Center, Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts, United States of America, **11** Division of Cardiology, Department of Medicine, University of Washington, Seattle, Washington, United States of America, **12** Pediatrics, Case Western Reserve University, Cleveland, Ohio, United States of America, **13** Institute of Human Genetics, Klinikum Rechts der Isar der Technischen Universität München, München, Germany, **14** Institute of Human Genetics, Helmholtz Centrum München, Neuherberg, Germany, **15** Institute of Medical Informatics, Biometry, and Epidemiology and Department of Medicine I, University Hospital Grosshadern, Ludwig-Maximilians-Universität, Munich, Germany, **16** Institute of Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany, **17** Diabetes Center, University of California San Francisco, San Francisco, California, United States of America, **18** Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health, Bethesda, Maryland, United States of America, **19** Department of Epidemiology and Prevention, Division of Public Health Sciences, Wake Forest University, Winston-Salem, North Carolina, United States of America, **20** Department of Epidemiology, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, Pennsylvania, United States of America, **21** National Institute on Aging, National Institute on Aging/NIH–Intramural Research Program, NIH Biomedical Research Center, Baltimore, Maryland, United States of America, **22** Division of Cardiovascular Medicine, University of Pennsylvania, Philadelphia, Pennsylvania, United States of America, **23** Center for Human Genetic Research, Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts, United States of America, **24** Division of Cardiovascular Sciences, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, Maryland, United States of America, **25** Department of Epidemiology, School of Public Health, Boston University, Massachusetts, United States of America, **26** Division of Preventive Medicine, Boston University School of Medicine, Boston, Massachusetts, United States of America, **27** Whitaker Cardiovascular Institute, Boston University School of Medicine, Boston, Massachusetts, United States of America, **28** Department of Medicine, Case School of Medicine, Cleveland, Ohio, United States of America, **29** Division of Epidemiology and Community Health, School of Public Health, University of Minnesota, Minneapolis, Minnesota, United States of America, **30** Department of Epidemiology, University of Washington, Seattle, Washington, United States of America

## Abstract

The PR interval on the electrocardiogram reflects atrial and atrioventricular nodal conduction time. The PR interval is heritable, provides important information about arrhythmia risk, and has been suggested to differ among human races. Genome-wide association (GWA) studies have identified common genetic determinants of the PR interval in individuals of European and Asian ancestry, but there is a general paucity of GWA studies in individuals of African ancestry. We performed GWA studies in African American individuals from four cohorts ( $n=6,247$ ) to identify genetic variants associated with PR interval duration. Genotyping was performed using the Affymetrix 6.0 microarray. Imputation was performed for 2.8 million single nucleotide polymorphisms (SNPs) using combined YRI and CEU HapMap phase II panels. We observed a strong signal (rs3922844) within the gene encoding the cardiac sodium channel (*SCN5A*) with genome-wide significant association ( $p < 2.5 \times 10^{-8}$ ) in two of the four cohorts and in the meta-analysis. The signal explained 2% of PR interval variability in African Americans ( $\beta = 5.1$  msec per minor allele, 95% CI = 4.1–6.1,  $p = 3 \times 10^{-23}$ ). This SNP was also associated with PR interval ( $\beta = 2.4$  msec per minor allele, 95% CI = 1.8–3.0,  $p = 3 \times 10^{-16}$ ) in individuals of European ancestry ( $n = 14,042$ ), but with a smaller effect size ( $p$  for heterogeneity  $< 0.001$ ) and variability explained (0.5%). Further meta-analysis of the four cohorts identified genome-wide significant associations with SNPs in *SCN10A* (rs6798015), *MEIS1* (rs10865355), and *TBX5* (rs7312625) that were highly correlated with SNPs identified in European and Asian GWA studies. African ancestry was associated with increased PR duration (13.3 msec,  $p = 0.009$ ) in one but not the other three cohorts. Our findings demonstrate the relevance of common variants to African Americans at four loci previously associated with PR interval in European and Asian samples and identify an association signal at one of these loci that is more strongly associated with PR interval in African Americans than in Europeans.

**Citation:** Smith JG, Magnani JW, Palmer C, Meng YA, Soliman EZ, et al. (2011) Genome-Wide Association Studies of the PR Interval in African Americans. *PLoS Genet* 7(2): e1001304. doi:10.1371/journal.pgen.1001304

**Editor:** Peter M. Visscher, Queensland Institute of Medical Research, Australia

**Received:** July 14, 2010; **Accepted:** January 11, 2011; **Published:** February 10, 2011

This is an open-access article distributed under the terms of the Creative Commons Public Domain declaration which stipulates that, once placed in the public domain, this work may be freely reproduced, distributed, transmitted, modified, built upon, or otherwise used by anyone for any lawful purpose.

**Funding:** The following four parent studies have contributed parent study data, ancillary study data, and DNA samples through the Massachusetts Institute of Technology - Broad Institute (N01-HC-65226) to create this genotype/phenotype database for wide dissemination to the biomedical research community. Atherosclerosis Risk in Communities (ARIC): The ARIC study is carried out as a collaborative study supported by National Heart, Lung, and Blood Institute contracts to the University of North Carolina at Chapel Hill (N01-HC-55015), Baylor Medical College (N01-HC-55016), University of Mississippi Medical Center (N01-HC-55021), University of Minnesota (N01-HC-55019), Johns Hopkins University (N01-HC-55020), University of Texas, Houston (N01-HC-55022), University of North Carolina, Forsyth County (N01-HC-55018); Cleveland Family Study (CFS): Case Western Reserve University (NIH HL 46380, M01RR00080); Jackson Heart Study (JHS): Jackson State University (N01-HC-95170), University of Mississippi (N01-HC-95171), Tougaloo College (N01-HC-95172); Multi-Ethnic Study of Atherosclerosis (MESA): University of Washington (N01-HC-95159), Regents of the University of California (N01-HC-95160), Columbia University (N01-HC-95161), Johns Hopkins University (N01-HC-95162, N01-HC-95168), University of Minnesota (N01-HC-95163), Northwestern University (N01-HC-95164), Wake Forest University (N01-HC-95165), University of Vermont (N01-HC-95166), New England Medical Center (N01-HC-95167), Harbor-UCLA Research and Education Institute (N01-HC-95169), Cedars-Sinai Medical Center (R01-HL-071205), University of Virginia (subcontract to R01-HL-071205). The Health, Aging, and Body Composition Study (Health ABC) was supported by NIA contracts N01AG62101, N01AG62103, and N01AG62106. The genome-wide association study was funded by NIA grant 1R01AG032098-01A1 to Wake Forest University Health Sciences and genotyping services were provided by the Center for Inherited Disease Research (CIDR). CIDR is fully funded through a federal contract from the National Institutes of Health to The Johns Hopkins University, contract number HHSN268200782096C. This research was supported in part by the Intramural Research Program of the NIH, National Institute on Aging. The Healthy Aging in Neighborhoods of Diversity across the Life Span Study (HANDLS) study was in part supported by the intramural research program of the National Institute on Aging and the National Center for Minority Health and Health Disparities, National Institutes of Health. This research was supported by the Intramural Research Program of the NIH, National Institute on Aging and the National Center on Minority Health and Health Disparities (contract # Z01-AG000513 and human subjects protocol # 2009-149). Data analyses for the HANDLS study utilized the high-performance computational capabilities of the Biowulf Linux cluster at the National Institutes of Health, Bethesda, Md. (<http://biowulf.nih.gov>). JGS was supported by the Swedish Heart Lung Foundation. JWM was supported by American Heart Association grant 09FTF2190028. KFK was supported by NHLBI contract 268200900055C-0-0-1. SAL was supported by NIH T32HL007575. AP was supported by the German National Genome Research Network grants NGFN 01GR0803 and the German Federal Ministry of Research BMBF 01EZ0874. RD was supported by Grant Number KL2RR024132 from the National Center for Research Resources. PTE was supported by 1R01HL092577, 1R21DA027021. CN-C was supported by NIH 1K23HL080025, 1R01HL098283, Doris Duke Charitable Foundation Clinical Scientist Development Award, and Burroughs Wellcome Fund Career Award for Medical Scientists. EJB was supported by NIH awards 1R01HL092577, 1R01HL092577, 1R01HL102214, 1R01 AG028321. RM was supported by NIH K23 HL079114, American Heart Association National Scientist Development Award 0530188N, Central Society of Clinical Research Award, and NCI 1U54CA116867. AA was supported by NIH awards 1RC1HL099452 and 1RC1HL101056 and American Heart Association grant 09SDG2280087. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

\* E-mail: [gustav.smith@med.lu.se](mailto:gustav.smith@med.lu.se)

† These authors contributed equally to this work.

‡ Information on the Candidate-gene Association Resource (CARE) Consortium is provided in the Acknowledgments.

## Introduction

The electrocardiogram is an important clinical tool that provides a graphical representation of the electrical activity of the heart as captured by skin surface electrodes. The electrocardiographic PR interval represents conduction through the atria and AV node, and impaired conduction through these tissues is a central pathogenic mechanism in many cardiac arrhythmias. Prolongation of the PR interval is a risk factor for long-term atrial fibrillation, heart block, and all-cause mortality [1]. A substantial proportion of the variability of the PR interval is explained by genetic factors with heritability estimates ranging between 30 and 50% in populations of European and Asian ancestry [2–7]. Identification of the specific alleles underlying the heritability of the PR interval might provide novel insights into molecular electrophysiology, lead to novel drug targets for arrhythmia treatment, and facilitate genetic prediction of arrhythmia risk and response to antiarrhythmics. Genome-wide association (GWA) studies have identified nine loci associated with PR interval duration in populations of European [6,8] or South Asian ancestry [9]. Four of these associations have been replicated [6,9]. However, there is a general paucity of GWA studies performed in populations of African ancestry [10] and none have been performed for PR interval duration which has been suggested to be longer in individuals of African than European ancestry [11]. Samples of African ancestry typically have shorter-range linkage disequilibrium (LD) than samples of European or Asian ancestry, potentially reducing genomic coverage in GWA studies using fixed arrays. However, samples of African ancestry offer greater genetic diversity; European and Asian populations have experienced a loss of allelic

diversity since the out-of-Africa founding events [12,13]. Thus, unique alleles influencing complex traits may be present in samples of African ancestry and shorter-range LD can facilitate narrowing of association signals. We studied the association of genetic ancestry with PR interval and performed a GWA study in 6,247 African Americans from four cohorts as part of the Candidate-gene Association Resource (CARE) consortium including the Atherosclerosis Risk in Communities (ARIC) study, the Cleveland Family Study (CFS), the Jackson Heart Study (JHS) and the Multi-Ethnic Study of Atherosclerosis (MESA). Replication of genome-wide significant associations was attempted by direct genotyping in 2,022 African Americans from the Health, Aging, and Body Composition Study (Health ABC) and the Healthy Aging in Neighborhoods of Diversity across the Life Span Study (HANDLS).

## Results

### Phenotype modeling

Baseline characteristics of the study populations are shown in Table 1. After adjustment for covariates, residual standard deviations for PR interval ranged from 22.1 to 26.9 msec in the four cohorts. The proportion of variation in phenotype explained by covariates ( $r^2$ ) was modest, ranging from 0.04 to 0.15, as shown in Table 1.

### Genetic ancestry and PR interval

The distribution of European ancestry relative to African ancestry was similar across cohorts; ARIC (median 0.15, IQR 0.11–0.22), CFS (median 0.18, IQR 0.13–0.26), JHS (median

## Author Summary

We performed genome-wide association studies in African American participants from four population-based cohorts to identify genetic variation that correlates with variation in PR interval duration, an electrocardiographic measure of conduction through the atria and atrioventricular node. We observed a strong signal within the gene encoding the cardiac sodium channel, *SCN5A*, with genome-wide significant association ( $p < 2.5 \times 10^{-8}$ ) in two cohorts and in a meta-analysis of four cohorts with African Americans. We replicated this association in two additional cohorts of African Americans and in Europeans ( $p = 3 \times 10^{-16}$ ). The signal explains 2% of PR duration variability in African Americans and 0.5% in Europeans. In further meta-analysis, we observed genome-wide significant associations for single nucleotide polymorphisms in *SCN10A*, *MEIS1*, *TBX5*, corresponding to signals observed in people of European and Asian descent. We found an association of genetic ancestry and PR interval in one but not the other three cohorts. Our findings provide the first demonstration of the relevance of these loci to individuals of African ancestry and identify an association signal from *SCN5A* that is more strongly associated with PR interval in African Americans.

0.16, IQR 0.12–0.21) and MESA (median 0.19, IQR 0.12–0.30). African ancestry was associated with longer PR interval than European ancestry in ARIC (13.3 msec longer for complete African compared to complete European ancestry,  $p = 0.009$ ) but not in CFS (–13.2 msec,  $p = 0.44$ ), JHS (1.0 msec,  $p = 0.89$ ) or MESA (7.2 msec,  $p = 0.11$ ).

## Meta-analysis of genome-wide association studies

Genomic inflation was minimal in ARIC ( $\lambda_{GC} = 1.02$ ) and MESA ( $\lambda_{GC} = 1.01$ ) and modest in the family-based CFS

( $\lambda_{GC} = 1.10$ ) and JHS ( $\lambda_{GC} = 1.08$ ) studies, likely due to relatedness of some participants. Quantile-quantile plots of p-values were consistent with null distributions at all but the lowest p-values (Figure S1). Genomic control was applied to test statistics in all cohorts. Genomic inflation in the meta-analysis was modest ( $\lambda_{GC} = 1.04$ ); genomic control was applied again (Figure 1). The distribution of association results across the genome is shown in Figure 2. Twenty-four single nucleotide polymorphisms (SNPs) at four genomic loci reached genome-wide significance as shown in Table S1. We clustered these SNPs by pairwise correlation in the HapMap phase II YRI panel and observed five weakly correlated linkage disequilibrium bins (all pairwise  $r^2$  were  $< 0.3$ ). Results for the SNP with the lowest p-value in each cluster are shown in Table 2 and the correlation of these SNPs to SNPs reported in previous GWA studies of European and Asian ancestry are shown in Table 3.

The top association result was for rs3922844 at *SCN5A* (5.1 msec per minor allele, 95% CI = 4.1–6.1,  $p = 3 \times 10^{-23}$ ), explaining 2% of variation in PR interval. The SNP (rs3922844) reached genome-wide significance ( $p < 2.5 \times 10^{-8}$ ) in the larger cohorts, ARIC and JHS, considered individually. The SNP had a large effect estimate (5.1–6.1 msec or 0.19–0.24 standard deviations per C-allele) consistently across ARIC, JHS and MESA as shown in Table 2. The effect estimate was smaller in CFS, but CFS had a smaller sample of African Americans and hence wider confidence intervals than was observed in other larger samples. We observed no evidence of heterogeneity across the four African American cohorts ( $p$  for heterogeneity = 0.41). The SNP was well imputed in all cohorts as shown in Table S2. It was not correlated in HapMap CEU or YRI panels with the SNPs at this locus that have been identified in previous GWA studies of European or Asian ancestry ( $r^2 < 0.10$ , Table 3).

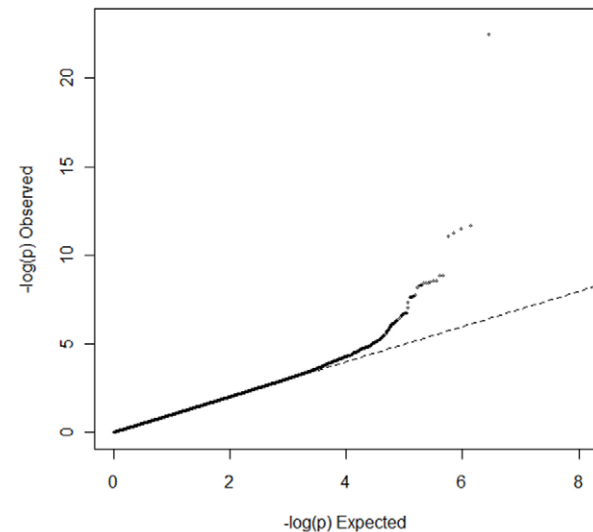
We observed an additional association signal in *SCN5A* (rs6599222,  $p = 2 \times 10^{-8}$ ), for which the strongest SNP was in LD with the SNP identified in a previous European GWA study (rs11708996,  $r^2 = 0.61$ ) and with an effect size (+3.4 msec per minor allele, 95% CI = 2.2–4.6) that was similar to the European

**Table 1.** Description of African American study samples.

|                                      | ARIC             | CFS              | JHS              | MESA             |
|--------------------------------------|------------------|------------------|------------------|------------------|
| Age (years)                          | 53.2<br>(8.8)    | 44.3<br>(15.2)   | 49.3<br>(11.7)   | 62.1<br>(10.1)   |
| Sex (% male)                         | 38.1%            | 42.2%            | 38.7%            | 45.5%            |
| Body mass index (kg/m <sup>2</sup> ) | 29.5<br>(6.1)    | 34.5<br>(9.2)    | 32.3<br>(7.8)    | 30.2<br>(5.9)    |
| Height (cm)                          | 168.1<br>(9.0)   | 168.9<br>(9.9)   | 169.7<br>(9.4)   | 168.4<br>(9.6)   |
| Systolic blood pressure (mm Hg)      | 128.1<br>(20.7)  | 126.1<br>(14.4)  | 124.6<br>(17.8)  | 131.6<br>(21.6)  |
| RR duration (msec)                   | 923.0<br>(149.8) | 902.6<br>(131.4) | 948.8<br>(148.2) | 975.4<br>(154.9) |
| PR duration (msec)                   | 172.4<br>(27.5)  | 169.0<br>(25.7)  | 170.6<br>(26.0)  | 170.8<br>(26.0)  |
| PR sd (msec)                         | 26.9             | 22.1             | 25.4             | 25.1             |
| PR coefficient of determination      | 0.04             | 0.15             | 0.05             | 0.07             |
| Sample size                          | 2391             | 267              | 1962             | 1627             |

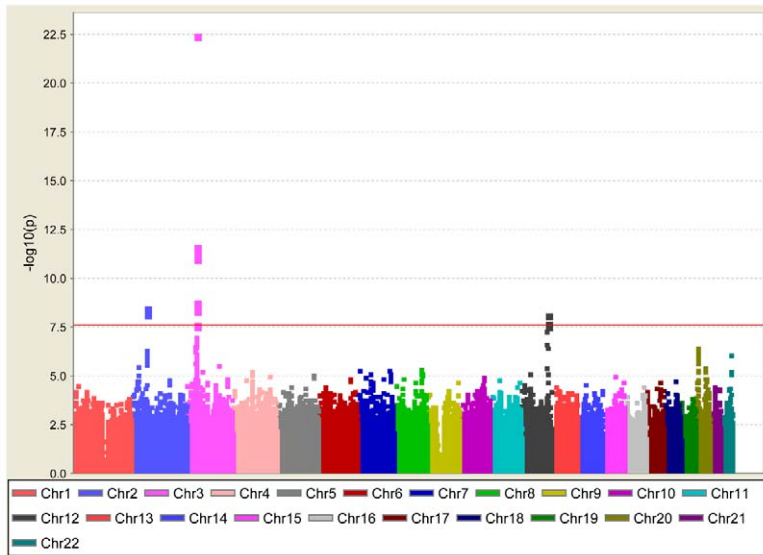
Mean and standard deviations are presented for continuous measures and percentages for categorical variables. PR sd refers to standard deviation of PR interval after covariate adjustment.  $\lambda_{GC}$  refers to the genomic inflation factor. PR coefficient of determination refers to the proportion of variability explained by covariates in the model (age, sex, RR interval, body mass index, height, systolic blood pressure and study site where relevant).

doi:10.1371/journal.pgen.1001304.t001



**Figure 1.** Quantile-quantile plot for meta-analysis of genome-wide association studies of PR interval duration in African Americans. Plotted on the x-axis are expected p-values under the null hypothesis and on the y-axis the observed p-values after genomic control has been applied.

doi:10.1371/journal.pgen.1001304.g001



**Figure 2. PR interval association results for 2.8 million SNPs across 22 autosomes.** Each dot represents one SNP. Plotted on the x-axis is physical position by chromosome and on the y-axis  $-\log_{10}(p\text{-value})$ . doi:10.1371/journal.pgen.1001304.g002

study [8]. However, when analyzed conditional on rs3922844, the SNP rs6599222 was only borderline significantly associated with PR interval as shown in Table S3. The SNPs in *SCN5A* previously reported to be associated with PR interval in a cohort of Micronesians [5] were not associated with PR interval in African Americans ( $p = 0.57$ ). The association results, correlations and locations of all studied SNPs in *SCN5A* are plotted against recombination rates in Figure 3.

We further observed significant associations for a cluster of ten highly correlated SNPs in *SCN10A* that included the missense SNP (rs6795970) reported previously in GWA studies of individuals of European and Asian ancestry [6,8,9]. The missense SNP did not have the lowest p-value in the cluster but had the highest effect estimate, the lowest minor allele frequency (Table 2) and a somewhat lower imputation quality than other SNPs. Significant associations also were seen with SNPs in *TBX5* and *MEIS1*. These

correlated with the strongest reported SNPs from the GWA study performed in individuals of European ancestry at the loci shown in Table 3. The SNPs with the lowest p-values in each cluster were all well imputed as shown in Table S2. We observed no evidence for heterogeneity of effect across cohorts for any SNP reported in Table 2 ( $p > 0.05$ ).

As shown in Table 4, additional adjustment for local ancestry estimates did not impact any of the genome-wide significant associations.

Of the other SNPs reaching genome-wide significance in previous studies [6,8], rs3807989 in *CAVI* showed some evidence of association in the same direction (2.1 msec per A allele,  $p = 6 \times 10^{-5}$ ) as did rs7692808 (−1.7 msec per A allele,  $p = 0.004$ ) and rs7660702 (−1.2 msec per C allele,  $p = 0.02$ ) in *ARHGAP24* but not rs251253 near *NKX2-5* ( $p = 0.98$ ), rs4944092 in *WNT11* ( $p = 0.93$ ) or rs11047543 in *SOX5* ( $p = 0.20$ ).

**Table 2. SNPs with the lowest p-values in each cluster of SNPs reaching genome-wide significance ( $p < 2.5 \times 10^{-8}$ ).**

| SNP        | Gene          | Position          | Allele | CAF  | ARIC         |                    | CFS          |      | JHS          |                     | MESA         |                    | N <sub>eff</sub> | Meta-analysis |                     | Q          | I <sup>2</sup> |
|------------|---------------|-------------------|--------|------|--------------|--------------------|--------------|------|--------------|---------------------|--------------|--------------------|------------------|---------------|---------------------|------------|----------------|
|            |               |                   |        |      | Beta (SE)    | P                  | Beta (SE)    | P    | Beta (SE)    | P                   | Beta (SE)    | P                  |                  | Beta (SE)     | P                   |            |                |
| rs3922844  | <i>SCN5A</i>  | Chr3:38,599,257   | C/T    | 0.41 | 5.08 (0.85)  | $2 \times 10^{-9}$ | 0.91 (2.58)  | 0.72 | 6.14 (0.90)  | $7 \times 10^{-12}$ | 4.57 (0.94)  | $1 \times 10^{-6}$ | 5726             | 5.11 (0.51)   | $3 \times 10^{-23}$ | $p = 0.41$ | 0.25           |
| rs6798015  | <i>SCN10A</i> | Chr3:38,773,840   | C/T    | 0.14 | 4.04 (1.18)  | $6 \times 10^{-4}$ | 6.31 (3.23)  | 0.05 | 5.43 (1.25)  | $1 \times 10^{-5}$  | 5.41 (1.28)  | $2 \times 10^{-5}$ | 6080             | 4.98 (0.70)   | $2 \times 10^{-12}$ | $p = 0.81$ | 0              |
| rs10865355 | <i>MEIS1</i>  | Chr2:66,618,501   | A/G    | 0.43 | 2.55 (0.79)  | 0.001              | 3.64 (2.32)  | 0.11 | 3.40 (0.87)  | $9 \times 10^{-5}$  | 2.67 (0.89)  | 0.003              | 6247             | 2.89 (0.48)   | $3 \times 10^{-9}$  | $p = 0.93$ | 0.46           |
| rs7312625  | <i>TBX5</i>   | Chr12:113,284,357 | A/G    | 0.71 | −2.27 (0.92) | 0.01               | −3.71 (2.60) | 0.15 | −4.11 (0.97) | $2 \times 10^{-5}$  | −3.34 (1.02) | 0.001              | 5718             | −3.23 (0.55)  | $7 \times 10^{-9}$  | $p = 0.67$ | 0              |
| rs6599222  | <i>SCN5A</i>  | Chr3:38,623,066   | C/T    | 0.23 | 3.24 (0.99)  | 0.001              | −0.67 (3.02) | 0.83 | 4.38 (1.06)  | $4 \times 10^{-5}$  | 3.07 (1.13)  | 0.007              | 5616             | 3.40 (0.61)   | $2 \times 10^{-8}$  | $p = 0.99$ | 0              |

Positions refer to NCBI build 36. Alleles refer to coded/noncoded. CAF refers to the coded allele frequency in the pooled samples. N<sub>eff</sub> refers to the effective sample size in the meta-analysis, defined as  $N \times R^2$ , and reflects the loss of power with poor imputation compared to the total sample size shown in the first row. Beta coefficients per coded allele and p-values are given per cohort and for the meta-analysis. Genomic control has been applied in all samples. For heterogeneity, Q refers to the p-value from Cochran's Q test and I<sup>2</sup> refers to the I<sup>2</sup> statistic. All five SNPs are intronic. doi:10.1371/journal.pgen.1001304.t002

**Table 3.** Correlation of the SNPs with the lowest p-value in each cluster of SNPs reaching genome-wide significance in CARE to SNPs reaching genome-wide significance in previous European and Asian studies.

| CARE SNP   | Locus         | SNP in previous studies             | Reference                               | CEU r2              | YRI r2              | ASW r2              | LWK r2              | MKK r2       |
|------------|---------------|-------------------------------------|---|---------------------|---------------------|---------------------|---------------------|--------------|
| rs3922844  | <i>SCN5A</i>  | rs11708996<br>rs7638909             | Pfeufer 2009<br>Smith 2009              | 0.09<br>0.04        | Mono<br>0.004       | N/A<br>0.03         | N/A<br>0.006        | N/A<br>0.01  |
| rs6798015  | <i>SCN10A</i> | rs6795970<br>rs6795970<br>rs6800541 | Chambers 2009 Holm<br>2009 Pfeufer 2009 | 0.72<br>aaa<br>0.75 | 0.19<br>aaa<br>0.19 | 0.40<br>aaa<br>0.40 | 0.36<br>aaa<br>0.36 | 0.48<br>0.51 |
| rs10865355 | <i>MEIS1</i>  | rs11897119                          | Pfeufer 2009                            | 1.00                | 0.59                | 0.78                | 0.71                | 0.86         |
| rs7312625  | <i>TBX5</i>   | rs1896312<br>rs3825214              | Pfeufer 2009<br>Holm 2009               | -<br>0.70           | -<br>0.42           | -<br>0.59           | -<br>0.44           | -<br>0.20    |
| rs6599222  | <i>SCN5A</i>  | rs11708996<br>rs7638909             | Pfeufer 2009<br>Smith 2009              | 0.61<br>0.07        | Mono<br>0.003       | N/A<br><0.001       | N/A<br>0.04         | N/A<br>0.008 |

Correlations were examined in the following samples from HapMap phase II (release 22) or HapMap phase III (release 2): CEU – Utah residents with northern and western European ancestry from the CEPH collection, YRI – Yoruba in Ibadan, Nigeria, ASW – African ancestry in southwestern USA, LWK – Luhya in Webuye, Kenya, MKK – Maasai in Kinyawa, Kenya. The strongest SNP in *SCN5A* in European studies (rs11708996) is monomorphic in YRI and was not genotyped in HapMap phase III. The SNP downstream of *TBX5* described in Pfeufer et al is >500 kb from the other two SNPs, that are located within the gene. The correlation of the two previously described SNPs in *SCN10A* is 0.97 in CEU. The correlation of our two SNPs in *SCN5A* is 0.001 in CEU, 0.27 in YRI, 0.19 in ASW, 0.03 in LWK and 0.05 in MKK. doi:10.1371/journal.pgen.1001304.t003

### Replication in African Americans

The strongest SNPs at loci reaching genome-wide significance were genotyped in Health ABC and HANDLS. As shown in Table 4, we observed evidence of replication for the top SNPs in *SCN5A*, *SCN10A* and *MEIS1* and borderline significant replication for *TBX5*.

### Association study in individuals of European ancestry

We examined the result for rs3922844 in a large, European GWA study (CHARGE-PR) [8], where it was imputed in four cohorts (n = 14,042). We observed a significant association with the PR interval (beta 2.40 msec per major allele, 95% CI = 1.83–2.97, p = 3 × 10<sup>-16</sup>). As in HapMap reference panels, allele frequencies were reversed in Europeans compared to African Americans, with a minor allele frequency of 0.30 for the T-allele. The effect size was significantly different in CHARGE-PR as compared to CARE (p for heterogeneity <0.001, I<sup>2</sup> = 0.99).

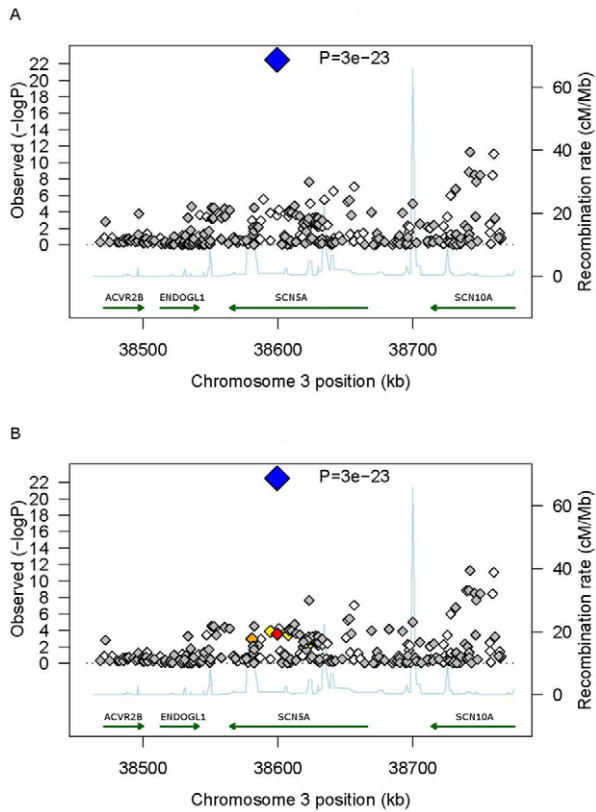
### Discussion

We performed a meta-analysis of GWA studies of the PR interval duration in individuals of African ancestry from four cohorts. We used a stringent significance threshold of 2.5 × 10<sup>-8</sup>, as previous genetic analyses have suggested twice as many independent SNPs in samples of African ancestry compared to European ancestry [14].

In two cohorts and in the meta-analysis we observed a genome-wide significant association with the SNP rs3922844 in *SCN5A* that explains approximately 2% of variability in PR interval duration. The SNP was not correlated with rs11708996, another SNP in *SCN5A* associated with PR interval duration in prior GWA studies of subjects of European ancestry [8], and had a larger effect estimate than rs11708996. We, therefore, examined the result for rs3922844 in a previously reported GWA study of PR interval in subjects of European ancestry [8] and observed genome-wide significant association but with a smaller effect estimate as compared to our study (5.1 msec, 95% CI = 4.1–6.1 versus 2.4 msec, 95% CI = 1.8–3.0, p for heterogeneity <0.001). Thus, we have evaluated genetic variants associated with PR interval duration in subjects of African American ancestry and identified an association signal from *SCN5A* with a larger effect size in individuals of African American compared to European ancestry.

We consider direct prolongation of the PR interval by rs3922844 unlikely given the differing effect sizes in individuals of African American and European ancestry. No SNP catalogue, including HapMap, used as reference panel for imputation in our study, includes all common sequence variants. We consider it more likely that rs3922844 is correlated with one or several causal variants and that the varying effect sizes result from differences in LD patterns across populations. For example, two variants with opposing effects could exist at the locus, with overlapping signals in Europeans but with isolation of individual signals in African Americans due to breakdown of linkage disequilibrium. The SNP rs3922844 was not directly genotyped in our samples but was imputed with high imputation quality as measured by the average ratio of observed variance to that expected by the Hardy-Weinberg equilibrium. We note that no SNPs included on our genotyping platform had an r<sup>2</sup> > 0.2 with rs3922844 in the African (YRI) HapMap panel. A small cluster of genotyped SNPs were well correlated with rs3922844 in the European (CEU) HapMap panel (r<sup>2</sup> = 0.6–0.8) as shown in Figure 3 but none of these reached a p-value less than 10<sup>-4</sup>. These observations suggest that the association was detected due to correlation of rs3922844 with multiple, weakly correlated, genotyped SNPs. Alternatively, a strong gene-environment interaction for rs3922844 could result in large differences in effect estimates across populations subject to differential environmental exposures or other genetic effects. Large-scale resequencing or imputation from large-scale resequencing efforts such as the 1000 Genomes Project, and functional experiments will ultimately be necessary to identify the causal variants.

Although allelic heterogeneity between families is well established for monogenic diseases, little is known about common variants across different populations. Recently identified findings include different association signals at the locus containing *KCNQJ* with type 2 diabetes in European and Asian cohorts [10], at *MYH9* with focal segmental glomerulosclerosis [15] and a *MYBPC3* variant that causes cardiomyopathy and is only polymorphic in Asian populations [16]. Allelic heterogeneity is conceptually important and highlights the value of genetic association studies spanning diverse racial and ethnic populations. Studies across multiple populations may identify novel disease loci and genetic factors underlying differences in risk by race. Our results do not



**Figure 3. Regional association plots for PR interval duration at the *SCN5A* locus.** The plot covers the genomic region from 100 kb upstream of *SCN5A* to 100 kb downstream of *SCN5A*. White diamonds represent genotyped SNPs and grey diamonds represent imputed SNPs. The large blue diamond represents the SNP with the lowest p-value (rs3922844). Diamond color represents pairwise correlation between directly genotyped SNPs with the strongest SNP; red indicates strong correlation ( $r^2 > 0.8$ ), orange indicates moderate correlation ( $0.5 > r^2 > 0.2$ ), and yellow indicates weak correlation ( $0.5 > r^2 > 0.2$ ). Recombination rate is plotted in the background and known genes are shown in the bottom of the plot. Positions refer to NCBI build 36. SNP correlations and estimates of recombination rates were obtained from HapMap phase II. Panel A uses LD patterns from the YRI sample and Panel B from the CEU sample for results in individuals of African American ancestry.

doi:10.1371/journal.pgen.1001304.g003

provide support for allelic heterogeneity across populations for PR interval, although limited power may have failed to detect such effects.

Genetic admixture studies have estimated that, on average, modern African American genomes have approximately 80–90% African and 10–20% European ancestry [17], which is well in line with our global ancestry analyses. It is important to note that there is great genetic diversity across the African continent. Our study included African Americans, who predominantly have an ancestral origin in Central and West Africa. It is possible that GWA studies in other samples of African ancestry might identify additional association signals.

The clinical impact of our findings merit further exploration. For example, a missense variant in *SCN5A*, S1103Y (rs7626962), which is polymorphic in African Americans but very rare in Europeans, has been associated with a markedly increased risk of ventricular arrhythmias [18]. This variant was not correlated with rs3922844 in the YRI panel of HapMap phase III ( $r^2 = 0.04$ ). Furthermore, African Americans have been shown to have a lower incidence of atrial fibrillation than Europeans. Longer PR interval duration (difference in means 11.8 msec) has also been reported in African Americans compared to Europeans in the ARIC study [11]. We were able to confirm this observation in ARIC using genetic ancestry, with a similar magnitude (13.3 msec longer with complete African compared to complete European ancestry,  $p = 0.009$ ). However, we did not observe differences by genetic ancestry in CFS, JHS or MESA. This discrepancy could be due to differences in trans-African ancestry or other factors. Our results further indicate that common variants are unlikely to explain a large part of any ancestry-related variation in atrial and atrioventricular nodal conduction. The strongest variant observed here explained 2% of PR interval variability. It seems more likely that any ancestry-related differences in atrial conduction are due to rare variants, but this awaits further testing. Additional studies in populations with atrial fibrillation or sudden cardiac death will be necessary to determine the impact of the SNPs identified in our study on clinical outcomes.

In our meta-analysis, we also observed genome-wide significant association signals of SNPs in *SCN10A*, *TBX5* and *MEIS1*, which are correlated with SNPs reported in previous GWA studies in individuals of European ancestry. The lowest p-value in the *SCN10A* cluster of SNPs was not observed in the previously described missense variant (rs6795970) [6,9] but in an intronic SNP (rs6798015). The effect estimate was highest for the missense SNP but it also had the lowest minor allele frequency in the cluster, a lower allele frequency than in European (CEU) and

**Table 4. Replication and local ancestry analyses of SNPs with the lowest p-value in each cluster of SNPs reaching genome-wide significance.**

| CARE SNP   | Locus         | Local ancestry                         |                          |   |  |                         | Replication            |
|------------|---------------|--|--------------------------|---|--|-------------------------|------------------------|
|            |               | ARIC                                   | CFS                      | JHS                                     | MESA                                   | Meta-analysis           |                        |
| rs3922844  | <i>SCN5A</i>  | 4.80 (0.87),<br>$p = 3 \times 10^{-8}$ | 0.30 (2.52), $p = 0.90$  | 6.30 (0.89),<br>$p = 1 \times 10^{-12}$ | 3.67 (0.98),<br>$p = 2 \times 10^{-4}$ | $p = 7 \times 10^{-21}$ | $p = 7 \times 10^{-7}$ |
| rs6798015  | <i>SCN10A</i> | 3.42 (1.24),<br>$p = 0.006$            | 5.93 (3.29), $p = 0.07$  | 5.72 (1.28),<br>$p = 8 \times 10^{-6}$  | 4.08 (1.34),<br>$p = 0.002$            | $p = 7 \times 10^{-10}$ | $p = 0.0002$           |
| rs10865355 | <i>MEIS1</i>  | 2.57 (0.79),<br>$p = 0.001$            | 3.61 (2.22), $p = 0.10$  | 3.45 (0.83),<br>$p = 4 \times 10^{-5}$  | 2.64 (0.88),<br>$p = 0.003$            | $p = 5 \times 10^{-10}$ | $p = 0.007$            |
| rs7312625  | <i>TBX5</i>   | -2.27 (0.91),<br>$p = 0.01$            | -3.63 (2.48), $p = 0.14$ | -4.13 (0.93),<br>$p = 1 \times 10^{-5}$ | -3.34 (1.01),<br>$p = 0.001$           | $p = 1 \times 10^{-9}$  | $p = 0.05$             |

Results refer to inverse variance weighted meta-analyses of CARE cohorts adjusted for local ancestry estimates and for replication by direct genotyping in Health ABC and HANDLS cohorts. For local ancestry analyses, beta (standard error) and p-values are shown.

doi:10.1371/journal.pgen.1001304.t004

Asian (JPT and CHB) HapMap panels. No evidence that the missense variant explains the association with PR interval has been published yet. Although it is possible that another SNP mediates the association, we find it more likely that the weaker p-value for the intronic SNP in our meta-analysis reflects differences in imputation quality or chance difference in the strength of the observed association. To ultimately identify the causal variants in each population, large-scale resequencing and functional experiments will be necessary.

Four genome-wide association studies for PR interval duration have been published previously. The first study identified an association with SNPs in *SCN5A* in a Micronesian founder population, but no SNP reached genome-wide significance [5]. A large study in individuals of European ancestry identified nine loci of genome-wide significance [8], including *ARHGAP24*, *CAVI/CAV2*, *MEIS1*, *NKX2-5*, *SCN5A*, *SCN10A*, *SOX5*, *TBX5/TBX3* and *WNT11*. A third study in Icelanders identified and replicated significant associations with four of these loci [6]; *ARHGAP24*, *CAVI*, *SCN10A* and *TBX5/TBX3*. Finally, a study in individuals of Asian ancestry identified and replicated association with SNPs in *SCN10A* [9]. Here, we have reported genome-wide significant associations in African Americans with SNPs in strong LD with the previously reported SNPs at four of these loci; *MEIS1*, *SCN5A*, *SCN10A* and *TBX5/TBX3*. These findings demonstrate the relevance of these common alleles to individuals of African ancestry.

Our study has a number of limitations. All GWA studies suffer risk of false positive associations due to multiple testing, which is higher in samples of African ancestry where two million independent SNPs have been described [14]. We therefore applied a stringent significance threshold ( $2.5 \times 10^{-8}$ ) to account for the approximately two million common variant tests estimated to exist across the genome. Large sample sizes are required to achieve such stringent significance levels. Our significant findings are thus unlikely to be false positives, especially as association signals at the same loci have been observed in European studies. However, even in the meta-analysis our study was underpowered to detect both common variants of small effect and rare variants. Measurement error in PR duration and heterogeneity between cohorts may further contribute to loss of power. Another potential problem encountered in genetic association studies is population stratification or confounding by ancestry. We sought to address this issue by filtering out individuals with cryptic relatedness or unlikely to be of African ancestry, adjusting for principal components of genome-wide data that describe within-population genetic variation, using family-based methods where appropriate and applying genomic control.

In conclusion, our results suggest that genetic loci underlying complex traits may be in large part the same across ancestral populations. However, association signals may show different effects across ethnic and racial populations and multiple association signals may exist at these loci. We anticipate that large-scale resequencing and functional experiments will ultimately identify the causal variants specific to each population.

## Materials and Methods

### Ethics statement

The study was approved by the Institutional Review Board at all participating institutions. Only individuals who provided informed consent to genetic testing were included.

### Study samples

Individuals of self-reported African American ancestry from four cohort studies were genotyped as part of the Candidate-gene

Association Resource (CARE) [19] (Lettre G et al, submitted). Detailed descriptions of these cohorts have been published previously [20–23]. Additional details are shown in Table S4. Details on replication samples (Health ABC, HANDLS) are shown in Table S5.

**ARIC.** The Atherosclerosis Risk in Communities (ARIC) 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, 4,314 of whom were self-reported black Americans from two of the four communities (Jackson, MS and Forsyth County, NC) [20]. Electrocardiographic recordings used in the present study were performed at baseline examinations between 1987–1989. After exclusions, data on 2,391 black individuals with genotypes, information on all covariates and informed consent remained for analyses.

**CFS.** The Cleveland Family Study (CFS) is a family-based, longitudinal study designed to study the risk factors for sleep apnea [21]. The 632 African Americans with available DNA were genotyped as part of CARE. Electrocardiographic recordings used for the present study were performed at the final exam cycle conducted in a Clinical Research Unit between 2001–2006. After exclusions, 267 individuals remained for analyses.

**JHS.** The Jackson Heart Study (JHS) is a prospective, community-based study of the causes of the high prevalence of cardiovascular disease in African Americans that includes 5,301 self-reported African Americans recruited between 2000–2004 [22]. The cohort comprises residents of a tri-county area near Jackson, MS (Hinds County, Rankin County and Madison County). The study includes a subsample of unrelated participants (35–84 years) and a nested family-based subcohort ( $\geq 21$  years). Electrocardiographic 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.

**MESA.** The Multi-Ethnic Study of Atherosclerosis (MESA) is a population-based study of the characteristics of subclinical cardiovascular disease that included 6,814 individuals (28% African Americans) free from known cardiovascular disease between 45–84 years recruited from 6 field centers in the US [23]. Electrocardiographic recordings used in the present study and blood sampling for DNA extraction were performed at baseline visits between 2000–2002. African Americans constituted 28% of the total sample, of whom 1,627 with genotype and phenotype data remained after exclusions.

### ECG recordings

12-lead electrocardiograms with standard lead placements were recorded during ten seconds in all cohorts using Marquette MAC PC, MAC6 or MAC1200 machines (GE Healthcare, Milwaukee, WI, USA). PR interval duration was measured electronically using either the Marquette 12SL algorithm or the MC MEANS algorithm (Table S4).

### Genotyping and quality control

All genotyping was performed at the Broad Institute of Harvard and MIT using the Affymetrix Genome-Wide Human SNP Array 6.0, which interrogates 906,600 SNPs, according to the manufacturers recommendations. Sample processing and quality control has been described previously [19] (Lettre G et al, submitted). Briefly, 1  $\mu$ g of genomic DNA was equally interleaved on 96-well



master plates. Quantity of double stranded DNA was ascertained using PicoGreen (Molecular Probes, Oregon, USA). Restriction enzyme digestion and PCR amplification using universal primers was performed to generate fragments of between 200–1,100 base pairs in length which were further fragmented to 25–50 bp and labeled with biotinylated nucleotides. Labeled fragments were hybridized to the microarray, washed and detected. Genotypes were called using Birdseed v1.33 and quality control steps were performed in PLINK, EIGENSTRAT and PREST-Plus [24–28]. To confirm sample identity, genotype concordance between 24 SNPs genotyped in all samples both on the microarrays and using Sequenom iPLEX were monitored. DNA samples with 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 (IBD) estimates and excluded. Samples that were present in both JHS and ARIC were excluded based on IBD concordance. We also excluded samples unlikely to be African American (ARIC = 8), based on principal component analyses incorporating carefully curated reference populations of 1,178 European Americans and 756 Nigerians in EIGENSTRAT as previously described (Lettre G et al, submitted).

We excluded SNPs with genotyping success rate <90%, minor allele frequency <1%, SNPs that map to multiple locations, SNPs where missingness could be predicted from surrounding haplotypes and SNPs associated with chemistry plates (i. e. differential genotyping by plate). SNPs with unusually high numbers of Mendelian errors were excluded from family-based samples. No exclusions were performed based on Hardy-Weinberg equilibrium given admixture of the sample. In total, about 800,000 SNPs passed quality control in all cohorts.

### Imputation and quality control

Imputation of additional ungenotyped SNPs was performed using hidden Markov models as implemented in MACH v1.16 [29] for a total of 2,801,419 million SNPs. Phased reference haplotypes from combined HapMap phase II CEU and YRI panels were used with equal proportions. Imputation was performed in two steps: in the first step, related individuals (as identified by proportion of genomic variation shared IBS/IBD ( $\hat{\pi}$ ) > 0.05) were filtered and error and recombination rates were estimated per sample in MACH using 30 iterations. In the second step, estimated rates were used to estimate allele dosage for all samples in one iteration. The average ratio of observed variance over that expected under Hardy-Weinberg equilibrium per SNP was also estimated as quality metric ( $R^2$ ). All samples were genotyped for 50,000 SNPs using the ITMAT-Broad-CARE (IBC) array [30] and the concordance rate ( $1 - 1/2 \times [\text{imputed dosage} - \text{chip dosage}]$ ) of imputed genotypes with IBC genotypes was high (~95%) in all samples. We excluded poorly imputed SNPs (ratio of observed over expected variance < 0.3), SNPs with minor allele frequency < 1% and non-autosomal SNPs. Sample outliers from missingness clustering analyses in PLINK were also removed.

### Phenotype modeling

We excluded from all analyses individuals with missing covariates, <18 years, pacemaker implant, Wolf-Parkinson-White pattern, extreme PR values (<80 ms and >320 ms), prevalent atrial fibrillation, heart failure, myocardial infarction, or 2<sup>nd</sup> or 3<sup>rd</sup> degree AV block. PR interval duration was regressed linearly on age, sex, RR interval, BMI, height, systolic blood pressure and study site where relevant using SAS 9.2 (SAS Institute, Cary, North Carolina). Residuals were used as phenotype for association analyses.

### Genome-wide association analyses and meta-analysis

For all cohorts except CFS, GWA analysis was performed in PLINK, using a linear regression model with an additive genetic model. The family-based CFS study was analyzed using linear mixed-effects models as implemented in the GWA package for R [31]. Although JHS contains a family-based subcohort, little effect on inflation of using family-based methods were observed in preliminary analyses for a set of traits. Pedigrees for CFS were confirmed using identity by state (IBS) or IBD estimates from PREST-Plus. In all association analyses, we adjusted for population stratification by including the first ten principal components in addition to genome-wide SNP data. Cohort-specific GWA results were combined using fixed effects meta-analysis with inverse variance weights. Genomic inflation factors were evaluated in each cohort prior to meta-analysis and in the combined results [32]. We prespecified a genome-wide significance threshold of  $2.5 \times 10^{-8}$  as suggested for populations of African ancestry by Pe'er et al [14], assuming 2 million independent common variant tests in genomes of individuals of African American ancestry with  $p < 0.05$  genome-wide. Heterogeneity across samples was assessed by Cochran's  $\chi^2$  test of heterogeneity with 4 degrees of freedom. To identify independent signals at loci we clustered all genome-wide significant hits by pairwise correlations  $\geq 0.5$  in the YRI HapMap panel. All analyses were performed using PLINK or R version 2.10.1 (R Foundation for Statistical Computing, Vienna, Austria). We examined the results for any SNPs of genome-wide significance that were uncorrelated with signals previously reported in other ancestral populations in a large published GWA study of PR interval in Europeans ancestry [8].

### Genetic ancestry analyses

Individual estimates of global ancestry (proportion of European ancestors relative to African ancestors) and local ancestry were calculated using ANCESTRYMAP [33] and HAPMIX [34], as previously described (Lettre G et al, submitted). To study the association of genetic ancestry and PR interval, global ancestry estimates were regressed linearly on PR residuals. To assess the impact of local ancestry on SNP associations, all associations with genome-wide significance were adjusted for local ancestry.

### Supporting Information

**Figure S1** Quantile-quantile plots for genome-wide association studies of PR interval duration in African Americans.

Found at: doi:10.1371/journal.pgen.1001304.s001 (0.34 MB TIF)

**Table S1** All SNPs reaching genome-wide significance ( $p < 2.5 \times 10^{-8}$ ) in a meta-analysis of four genome-wide association studies of PR interval duration in African Americans. Positions refer to NCBI build 36. Genomic control has been applied to standard errors and p-values. SNP clusters as defined by  $r^2$  greater than or equal to 0.5 are indicated. Neff refers to the effective sample size, defined as  $N \times R^2$ , and reflects the loss of power with poor imputation compared to the total sample sizes shown in Table 1.

Found at: doi:10.1371/journal.pgen.1001304.s002 (0.07 MB DOC)

**Table S2** Imputation quality and effective sample size for the SNPs with the lowest p-values in each linkage equilibrium bin of genome-wide significant SNPs.  $R^2$  refers to the average ratio of the observed variance over that expected under Hardy-Weinberg equilibrium and indicates quality of imputation as poorly imputed SNPs show less variability than expected based on allele frequency.

Neff refers to the effective sample size, defined as  $N \times R^2$ , and reflects the loss of power with poor imputation compared to the total sample sizes shown in Table 1.

Found at: doi:10.1371/journal.pgen.1001304.s003 (0.03 MB DOC)

**Table S3** Conditional association analysis of the secondary signal in *SCN5A*. Shown are association results for the SNP rs6599222 in the *SCN5A* gene adjusted for the top SNP in the gene (rs3922844).

Found at: doi:10.1371/journal.pgen.1001304.s004 (0.03 MB DOC)

**Table S4** Details on cohorts in the genome-wide association analysis. \* African American participants in ARIC were all recruited at centers in Forsyth County, MS and Jackson, MS. Individuals who participated in both JHS and ARIC were excluded from JHS based on identity-by-descent estimates from genome-wide data.

Found at: doi:10.1371/journal.pgen.1001304.s005 (0.03 MB DOC)

**Table S5** Details on replication cohorts. Mean and standard deviations are presented for continuous variables and percentages for categorical variables.

## References

- Cheng S, Keyes MJ, Larson MG, McCabe EL, Newton-Cheh C, et al. (2009) Long-term outcomes in individuals with prolonged PR interval or first-degree atrioventricular block. *JAMA* 301: 2571–2577.
- Havlik RJ, Garrison RJ, Fabsitz R, Feinleib M (1980) Variability of heart rate, P-R, QRS and Q-T durations in twins. *J Electrocardiol* 13: 45–48.
- Newton-Cheh C, Guo CY, Wang TJ, O'Donnell C J, Levy D, et al. (2007) Genome-wide association study of electrocardiographic and heart rate variability traits: the Framingham Heart Study. *BMC Med Genet* 8 Suppl 1: S7.
- Pilia G, Chen WM, Scuteri A, Orru M, Albai G, et al. (2006) Heritability of cardiovascular and personality traits in 6,148 Sardinians. *PLoS Genet* 2: e132. doi:10.1371/journal.pgen.0020132.
- Smith JG, Lowe JK, Kovvali S, Maller JB, Salit J, et al. (2009) Genome-wide association study of electrocardiographic conduction measures in an isolated founder population: Kosrae. *Heart Rhythm* 6: 634–641.
- Holm H, Gudbjartsson DF, Arnar DO, Thorleifsson G, Thorgeirsson G, et al. (2010) Several common variants modulate heart rate, PR interval and QRS duration. *Nat Genet* 42: 117–122.
- Hanson B, Tuna N, Bouchard T, Heston L, Eckert E, et al. (1989) Genetic factors in the electrocardiogram and heart rate of twins reared apart and together. *Am J Cardiol* 63: 606–9.
- Pfeuffer A, van Noord C, Marciante KD, Arking DE, Larson MG, et al. (2010) Genome-wide association study of PR interval. *Nat Genet* 42: 153–159.
- Chambers JC, Zhao J, Terracciano CM, Bezzina CR, Zhang W, et al. (2010) Genetic variation in *SCN10A* influences cardiac conduction. *Nat Genet* 42: 149–152.
- Rosenberg NA, Huang L, Jewett EM, Szpiech ZA, Jankovic I, et al. (2010) Genome-wide association studies in diverse populations. *Nat Rev Genet* 11: 356–366.
- Soliman EZ, Prineas RJ, Case LD, Zhang ZM, Goff DC, Jr. (2009) Ethnic distribution of ECG predictors of atrial fibrillation and its impact on understanding the ethnic distribution of ischemic stroke in the Atherosclerosis Risk in Communities (ARIC) study. *Stroke* 40: 1204–11.
- The International HapMap Investigators (2005) A haplotype map of the human genome. *Nature* 437: 1299–1320.
- Tishkoff SA, Williams SM (2002) Genetic analysis of African populations: human evolution and complex disease. *Nat Rev Genet* 3: 611–621.
- Pe'er I, Yelensky R, Altshuler D, Daly MJ (2008) Estimation of the multiple testing burden for genome-wide association studies of nearly all common variants. *Genet Epidemiol* 32: 381–385.
- Kopp JB, Smith MW, Nelson GW, Johnson RC, Freedman BI, et al. (2008) MYH9 is a major-effect risk gene for focal segmental glomerulosclerosis. *Nat Genet* 40: 1175–1184.
- Dhandapani PS, Sadayappan S, Xue Y, Powell GT, Rani DS, et al. (2009) A common MYBPC3 (cardiac myosin binding protein C) variant associated with cardiomyopathies in South Asia. *Nat Genet* 41: 187–191.
- Tishkoff SA, Reed Floyd A, Friedlaender FR, Ehret C, Ranciaro A (2009) The genetic structure of Africans and African Americans. *Science* 324: 1035–44.
- Splawski I, Timothy KW, Tateyama M, Clancy CE, Malhotra A, et al. (2002) Variant of *SCN5A* sodium channel implicated in risk of cardiac arrhythmia. *Science* 297: 1333–1336.
- Musunuru K, Lettre G, Young T, Farlow DN, Pirruccello JP, et al. (2010) Candidate Gene Association Resource (CARE): Design, Methods, and Proof of Concept. *Circ Cardiovasc Genet* 3: 267–75.
- The ARIC Investigators (1989) The Atherosclerosis Risk in Communities (ARIC) Study: design and objectives. *Am J Epidemiol* 129: 687–702.
- Redline S, Tishler PV, Tosteson TD, Williamson J, Kump K, et al. (1995) The familial aggregation of obstructive sleep apnea. *Am J Respir Crit Care Med* 151: 682–7.
- Taylor HA, Jr., Wilson JG, Jones DW, Sarpong DF, Srinivasan A, et al. (2005) Toward resolution of cardiovascular health disparities in African Americans: design and methods of the Jackson Heart Study. *Ethn Dis* 15: S6–4–17.
- Bild DE, Bluemke DA, Burke GL, Detrano R, Diez Roux AV, et al. (2002) Multi-ethnic study of atherosclerosis: objectives and design. *Am J Epidemiol* 156: 871–881.
- Korn JM, Kuruvilla FG, McCarroll SA, Wysoker A, Nemesh J, et al. (2008) Integrated genotype calling and association analysis of SNPs, common copy number polymorphisms and rare CNVs. *Nat Genet* 40: 1253–1260.
- McCarroll SA, Kuruvilla FG, Korn JM, Cawley S, Nemesh J, et al. (2008) Integrated detection and population-genetic analysis of SNPs and copy number variation. *Nat Genet* 40: 1166–1174.
- Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, et al. (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 81: 559–575.
- Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, et al. (2006) Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* 38: 904–909.
- McPeck MS, Sun L (2000) Statistical tests for detection of misspecified relationships by use of genome-screen data. *Am J Hum Genet* 66: 1076–1094.
- Li Y, Willer C, Sanna S, Abecasis G (2009) Genotype imputation. *Annu Rev Genomics Hum Genet* 10: 387–406.
- Keating BJ, Tischfield S, Murray SS, Bhargale T, Price TS, et al. (2008) Concept, design and implementation of a cardiovascular gene-centric 50 k SNP array for large-scale genomic association studies. *PLoS ONE* 3: e3583. doi:10.1371/journal.pone.0003583.
- Chen MH, Yang Q (2010) GWAf: an R package for genome-wide association analyses with family data. *Bioinformatics* 26: 580–581.
- Devlin B, Roeder K (1999) Genomic control for association studies. *Biometrics* 55: 997–1004.
- Patterson N, Hattangadi N, Lane B, Lohmueller KE, Hafler DA, et al. (2004) Methods for high-density admixture mapping of disease genes. *Am J Hum Genet* 74: 979–1000.
- Price AL, Tandon A, Patterson N, Barnes KC, Rafaels N, et al. (2009) Sensitive detection of chromosomal segments of distinct ancestry in admixed populations. *PLoS Genet* 5: e1000519. doi:10.1371/journal.pgen.1000519.

Found at: doi:10.1371/journal.pgen.1001304.s006 (0.04 MB DOC)

## Acknowledgments

The authors wish to acknowledge the support of the National Heart, Lung, and Blood Institute and the contributions of the involved research institutions, study investigators, field staff, and study participants of ARIC, CFS, JHS, and MESA in creating the Candidate-gene Association Resource for biomedical research (CARE; <http://public.nhlbi.nih.gov/GeneticsGenomics/home/care.aspx>).

The authors also wish to thank the staff and participants of Health ABC and HANDLS for their important contributions and the CHARGE consortium for access to results from a large European GWA study of PR interval.

## Author Contributions

Conceived and designed the experiments: JGS JWM CNC EJB RM AA SRH EF. Performed the experiments: JGS CP YAM. Analyzed the data: JGS CP YAM CNC. Contributed reagents/materials/analysis tools: JGS CP YAM AP MM DSE MAN YL ABN ABZ MKE CNC. Wrote the paper: JGS JWM. Interpreted the results and revised the paper: EZS SKM KK RBS SAL NS SR AP MM DSE MAN RD PE DNP CNC RM EJB AA SRH EF.