

# Characterization of Genome-Wide Association-Identified Variants for Atrial Fibrillation in African Americans

Jessica T. Delaney<sup>1</sup>, Janina M. Jeff<sup>2,3</sup>, Nancy J. Brown<sup>1,4</sup>, Mias Pretorius<sup>5</sup>, Henry E. Okafor<sup>6</sup>, Dawood Darbar<sup>1</sup>, Dan M. Roden<sup>1,4</sup>, Dana C. Crawford<sup>2,3\*</sup>

**1** Department of Medicine, Vanderbilt University, Nashville, Tennessee, United States of America, **2** Department of Molecular Physiology and Biophysics, Vanderbilt University, Nashville, Tennessee, United States of America, **3** Center for Human Genetics Research, Vanderbilt University, Nashville, Tennessee, United States of America, **4** Department of Pharmacology, Vanderbilt University, Nashville, Tennessee, United States of America, **5** Department of Anesthesiology, Vanderbilt University, Nashville, Tennessee, United States of America, **6** Department of Medicine, Meharry Medical College, Nashville, Tennessee, United States of America

## Abstract

**Background:** Despite a greater burden of risk factors, atrial fibrillation (AF) is less common among African Americans than European-descent populations. Genome-wide association studies (GWAS) for AF in European-descent populations have identified three predominant genomic regions associated with increased risk (1q21, 4q25, and 16q22). The contribution of these loci to AF risk in African American is unknown.

**Methodology/Principal Findings:** We studied 73 African Americans with AF from the Vanderbilt-Meharry AF registry and 71 African American controls, with no history of AF including after cardiac surgery. Tests of association were performed for 148 SNPs across the three regions associated with AF, and 22 SNPs were significantly associated with AF ( $P < 0.05$ ). The SNPs with the strongest associations in African Americans were both different from the index SNPs identified in European-descent populations and independent from the index European-descent population SNPs ( $r^2 < 0.40$  in HapMap CEU): 1q21 rs4845396 (odds ratio [OR] 0.30, 95% confidence interval [CI] 0.13–0.67,  $P = 0.003$ ), 4q25 rs4631108 (OR 3.43, 95% CI 1.59–7.42,  $P = 0.002$ ), and 16q22 rs16971547 (OR 8.1, 95% CI 1.46–45.4,  $P = 0.016$ ). Estimates of European ancestry were similar among cases (23.6%) and controls (23.8%). Accordingly, the probability of having two copies of the European derived chromosomes at each region did not differ between cases and controls.

**Conclusions/Significance:** Variable European admixture at known AF loci does not explain decreased AF susceptibility in African Americans. These data support the role of 1q21, 4q25, and 16q22 variants in AF risk for African Americans, although the index SNPs differ from those identified in European-descent populations.

**Citation:** Delaney JT, Jeff JM, Brown NJ, Pretorius M, Okafor HE, et al. (2012) Characterization of Genome-Wide Association-Identified Variants for Atrial Fibrillation in African Americans. *PLoS ONE* 7(2): e32338. doi:10.1371/journal.pone.0032338

**Editor:** Amanda Ewart Toland, Ohio State University Medical Center, United States of America

**Received:** September 23, 2011; **Accepted:** January 25, 2012; **Published:** February 23, 2012

**Copyright:** © 2012 Delaney et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** This work was funded in part by the National Institutes of Health Pharmacogenetics Research Network (U01 HL65962 and U01 HL92217), NIH training grant (GM007569), NIH grant R01 HL092217, and an American Heart Association Established Investigator Award. 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: dana.crawford@chgr.mc.vanderbilt.edu

## Introduction

Atrial fibrillation (AF) is the most frequent arrhythmia in clinical practice, with over 2.3 million Americans diagnosed with the disease [1]. The incidence of AF varies by race, occurring more frequently in European Americans than in African Americans [2,3]. Paradoxically, risk factors for AF including hypertension (HTN), obesity, diabetes mellitus (DM), and congestive heart failure (CHF) are more prevalent in African Americans [3–7]. As a result, traditional risk factors used to assess AF risk among African Americans perform poorly in this population.

Over the last few years, several genetic variants have been associated with AF risk in genome-wide association studies (GWAS) performed in populations of European-descent [8]. To date, three genomic regions/candidate genes have been identified: 4q25 near *PITX2*, in *ZFHX3* on 16q22, and in *KCNJ3* on 1q21 [9–12]. For the 4q25 locus, rs2200733 has been consistently the SNP with the highest risk in multiple GWAS cohorts and in a fine

mapping study of the region [13]. The 4q25 locus also includes other SNPs that predict risk and are not in linkage disequilibrium with rs2200733 [13]. For both *ZFHX3* and *KCNJ3*, the reported index SNP has varied among studies likely reflecting the strong linkage disequilibrium observed for these genomic regions in European-descent populations [9–14]. As no GWAS has yet been performed for AF in populations of African-descent, it remains to be determined whether there is an association with these regions and if there is, whether the index European variants from these three genomic regions generalize to African Americans or if there are additional variants in the same genomic regions more strongly associated with AF in populations of non-European-descent.

African-descent populations such as African Americans are often used for fine-mapping experiments based on the different allelic architecture and lower levels of linkage disequilibrium of these populations compared with European Americans [15]. African Americans, however, are also a genetically heterogeneous population, with varying European and African contributions. It is

possible that the proportion of European ancestry (EA) can modify the risk of AF [16–18]. Indeed, genetic ancestry has been associated with disease risk for many diseases including obesity, hypertension (HTN), diabetes mellitus (DM), kidney disease, asthma, rheumatoid arthritis, breast cancer, and prostate cancer [19–29]. Recently, the proportion of European ancestry in African Americans was associated with the development of AF in individuals enrolled in the Cardiovascular Health Study (CHS) and Atherosclerosis Risk in Communities (ARIC) Study [30].

We report here a targeted genetic association study of AF in African Americans ascertained at Vanderbilt University in Nashville, TN. We first sought to generalize in African Americans three regions associated with AF risk in European-descent populations. We then calculated estimates of genetic ancestry and tested for associations with AF in African Americans. We demonstrate that for the three regions discovered in European-descent populations do generalize with respect to direction of effect, although the best index genetic variant associated with AF in African Americans is different. Our data also suggest that in contrast to previous reports based on genome-wide data, AF is not associated with admixture in this dataset for these three genomic regions.

## Methods

### Ethics Statement

The study protocol was approved by the Vanderbilt University Institutional Review Board. A written informed consent was obtained from all patients prior to enrollment.

### Study Population

Cases were selected from patients consecutively enrolled in the Vanderbilt-Meharry AF registry who self-reported African American race. Details of the Vanderbilt-Meharry AF registry have been previously published [31]. In brief, from 2002 to present, the registry has been prospectively enrolling patients with newly diagnosed AF or a documented history of AF, excluding those who developed atrial fibrillation in the post-operative period after cardiac surgery. Clinical data and blood samples were obtained at the time of enrollment.

Controls were selected from the Vanderbilt Cardiac Surgery registry, a registry of consecutive patients undergoing cardiac surgery [32]. Controls were selected among those who did not develop AF in the post-operative period from cardiac surgery and self-reported African American ancestry. Patients with a personal or family history of AF were excluded. The study protocol was approved by the Vanderbilt University Institutional Review Board, and participants were enrolled after informed consent was obtained.

### SNP selection

We selected SNPs based on three index associations observed in European-descent populations for AF. The three index SNPs include rs2200733, an intergenic SNP located at 4q25 near *PITX2*; rs13376333, an intronic SNP located at 1q21 in *KCNK3*; and rs2106261 an intronic SNP located at 16q22 in *ZFX* [9–12]. Regions of interest from the index SNP were identified in African American samples from the southwest (ASW) in the International HapMap Project using the Genome Variation Server (<http://gvs.gs.washington.edu/GVS/>). We implemented LDselect to identify tagSNPs in the three regions of interest for genotyping [33]. In brief, all SNPs in the region were filtered by minor allele frequency (MAF 5%) in the ASW sample. Common SNPs were then binned by linkage disequilibrium ( $r^2$ ) for each of the three regions based on the ASW data. A tagSNP was selected from each bin with the criteria that the tagSNP was in strong LD with the other SNPs in

the bin ( $r^2$  threshold of 0.80) but independent (i.e. not in strong LD) with other tagSNPs across bins. If multiple tagSNPs were available per bin, priority was given to the tagSNP identified in prior European descent AF GWAS studies followed by those with the highest PhastCons Conservation Score. A total of 51 tagSNPs around rs13376333, 50 tagSNPs around rs2200733, and 47 tagSNPs around rs2106261 were included for genotyping. Due to varying linkage disequilibrium, the genomic distance covered by selected tagSNPs varied: 400 kb for rs2200733, 200 kb for rs13376333, and 160 kb for rs2106261.

### SNP Genotyping

Genotyping was performed by the Vanderbilt DNA Resources Core with either Sequenom iPLEX Gold® assay on the MassARRAY® Platform or Applied Biosystem's TaqMan. All SNP call rates were >95%, with an average call rate of 99%. Three samples were excluded from the analysis due to low genotyping efficiency (<95%). Tests of Hardy-Weinberg equilibrium (HWE) were calculated separately for cases and controls, and no SNP deviated from HWE ( $p < 0.0003$ ).

### Statistical Methods

Differences between cases and controls were identified using a chi-square test or a Student's t-test for variables listed in Table 1. Using logistic regression implemented in PLINK [34], tests of association were performed both unadjusted and adjusted for age, body mass index (BMI), prevalent congestive heart failure (CHF), prevalent diabetes mellitus (DM), prevalent HTN, and prevalent coronary artery disease (CAD) assuming an additive genetic model for all SNPs that passed quality control described above (148 SNPs). All adjusted tests of association are shown in Table S1. Comparison of the allele frequencies for the ten most significantly associated SNPs of cases and controls to HapMap Phase II CEU and West African (YRI) samples was performed with chi-square test. The fixation index  $F_{ST}$ , a measure of population differentiation, was calculated using the Weir and Cockerham algorithm [35] implemented in the Platform for the Analysis, Translation, and Organization of large-scale data (PLATO) [36]. Correction for multiple comparisons by permutation testing and calculation of the False Discovery Rate (FDR) was performed in (PLINK) [34].

**Table 1.** Study Population Characteristics.

Trait	Case (n=73)	Control (n=71)	p-value
Age, y	50.2±1.7	53.7±1.5	0.06
Gender, %female	42.4	56.3	0.09
Body Mass Index, kg/m <sup>2</sup>	35.1±1.1	28.8±0.8	<0.001
Left Atrial Size	44.4±1.5	42.4±1.3	0.58
Left Ventricular Ejection Fraction, %	46.4±2.1	46.3±1.7	0.19
Hypertension, %	84.9	84.5	0.94
Diabetes, %	26	35.2	0.23
Congestive heart failure, %	40.9	25.4	0.05
Stroke, %	19.1	17.9	0.89
Coronary artery disease, %	26.8	71.4	<0.001
Current smoker, %	19.4	31.4	0.11
% European Ancestry	23.6	23.8	0.64

Mean (± standard errors) or proportions are given for each variable among cases and controls. A chi-square test or a Student's t-test was performed to test for differences between cases and controls, where appropriate.  
doi:10.1371/journal.pone.0032338.t001

Using the risk estimate of AF reported in Europeans (OR = 1.2), percent European ancestry within the three loci was calculated and averaged over all independent SNPs ( $r^2 < 0.8$ ) using ANCESTRYMAP [37]. In addition, local estimates of European ancestry were calculated using HAPMIX [38]. This method determines the likelihood that a haplotype from an admixed individual is derived from one reference population using phased haplotype information from the International HapMap Project. For our African American samples, we used phased haplotype data from CEU and YRI HapMap Phase II. We calculated the probability for each individual having 0, 1, or 2 copies of a European allele at each locus. For each individual, we averaged the number of copies of European alleles over all SNPs for each chromosome separately. Logistic regressions with AF case-control status as the dependent variable and estimates of European ancestry as the independent variable were performed. Logistic regression with adjustment for ancestry in addition to the covariates discussed above was also performed for all SNPs.

Using Quanto we calculated our power to detect GWAS-identified associations in our African American study population [39,40]. We assumed the reported effect size (OR) in Europeans to calculate the power to detect this effect given our sample size and minor allele frequencies in African Americans.

## Results

### Study Population

Of the 1,624 patients enrolled in the Vanderbilt-Meharry AF registry, 73 were self-identified as African American. From the

Vanderbilt Cardiac Surgery registry, 71 controls were identified for comparison. Cases and controls did not differ with respect to mean age or proportion of females (Table 1). Cases had a higher mean body mass index (BMI; 35.1 kg/m<sup>2</sup>) and higher prevalence of CHF (40.9%) compared with controls (28.8 kg/m<sup>2</sup> and 25.4%, respectively,  $p < 0.001$ ; Table 1). Controls had an increased prevalence of CAD compared with cases (71.4% versus 26.8%,  $p < 0.001$ ), which is consistent with their selection from a post-operative cardiac surgery registry.

### Association results

A total of 22/148 SNPs were associated with AF in adjusted models at  $P < 0.05$  (Table 2). Further adjustment for European ancestry did not appreciably alter the results (data not shown). The greatest proportion of significant associations among tested SNPs was observed in the region of 4q25 (15/50; 30%), followed by 1q21 (6/51; 11.8%), and 16q22 (1/47; 2.1%). The most significant SNP in each region in this African American dataset differed from the index SNPs from previously published European-descent populations: rs4631108 (4q25), rs16971547 (1q21), and rs4845396 (16q22) (Figures S1,S2,S3). Similar to the index SNPs identified in European Americans, the African American index SNPs were either intergenic (rs4631108) or intronic (rs16971547 and rs4845396).

Of the three European-descent index SNPs, both rs2200733 and rs1337633 were significantly associated with AF among African Americans, while rs2106261 was not. Regardless of significance, the direction of effect for all three index SNPs in African Americans was consistent with that previously reported for

**Table 2.** SNPs associated with atrial fibrillation in African Americans.

SNP	CHR	Chromosomal Location	Coded Allele	OR	95% Confidence Interval	P-value
rs4631108	4	111773867	A	3.430	1.587–7.417	0.002
rs4845396	1	154828409	A	0.298	0.134–0.663	0.003
rs2200733	4	111929618	T	3.283	1.495–7.207	0.003
rs1906602	4	111713323	C	5.926	1.795–19.570	0.004
rs4845397	1	154832304	C	0.321	0.147–0.699	0.004
rs2634071	4	111669220	A	2.810	1.353–5.836	0.006
rs4605724	4	111685081	A	4.723	1.535–14.530	0.007
rs2723334	4	111688752	G	0.376	0.183–0.774	0.008
rs6843082	4	111718067	G	2.629	1.284–5.383	0.008
rs12647316	4	111649251	T	3.760	1.399–10.100	0.009
rs10516563	4	111677722	G	3.283	1.332–8.097	0.010
rs11264275	1	154825270	G	0.385	0.181–0.820	0.013
rs4277843	4	111777749	G	2.280	1.173–4.434	0.015
rs6426987	1	154815257	C	0.454	0.240–0.861	0.016
rs16971547	16	73075708	C	8.162	1.467–45.420	0.016
rs6838973	4	111765495	T	0.365	0.157–0.852	0.019
rs13376333	1	154814353	T	2.292	1.124–4.672	0.022
rs4285153	4	111778733	A	0.506	0.265–0.966	0.039
rs11098092	4	111798201	A	2.143	1.037–4.432	0.040
rs3866823	4	111782436	T	2.281	1.033–5.040	0.041
rs11930528	4	111660194	T	2.103	1.025–4.315	0.043
rs1984285	1	154796895	G	0.287	0.083–0.988	0.048

SNP, location (NCBI.36), coded allele, odds ratios, 95% confidence intervals, and p-value are shown for significant associations ( $p < 0.05$ ) after adjusting for age, BMI, history of CAD, history of CHF, history of diabetes, and history of hypertension.  
doi:10.1371/journal.pone.0032338.t002

European-descent populations (Figure 1). In addition to the index SNPs, three additional SNPs also previously reported to be strongly associated with AF within the selected region were evaluated: rs10033464, rs6843082, and rs7193343. Only rs6843082 was significantly associated in this African American sample. Again, regardless of significance, all three of these SNPs maintained a consistent direction of effect compared with European-descent populations (Figure 1).

### Genetic Ancestry

Given that African Americans have a greater risk for AF but the incidence is lower compared with European Americans, we tested the hypothesis that genomic regions associated with AF in African Americans would have a higher percentage of European ancestry compared with genome-wide averages [2,3]. This was tested using two strategies. First, we calculated  $F_{ST}$  between CEU and case samples, and we compared minor allele frequencies in cases and controls to HapMap CEU and YRI samples for the ten most significant SNPs (Table S2). Case allele frequencies differed from CEU allele frequencies for all ten SNPs (100%), while only 40% (4/10) differed in comparison to YRI. For the controls, 50% (5/10) of SNPs had different allele frequencies when compared to CEU, while 60% (6/10) differed when compared to YRI.  $F_{ST}$  estimates did not reveal major differences between cases and CEU reference samples (Table S2).

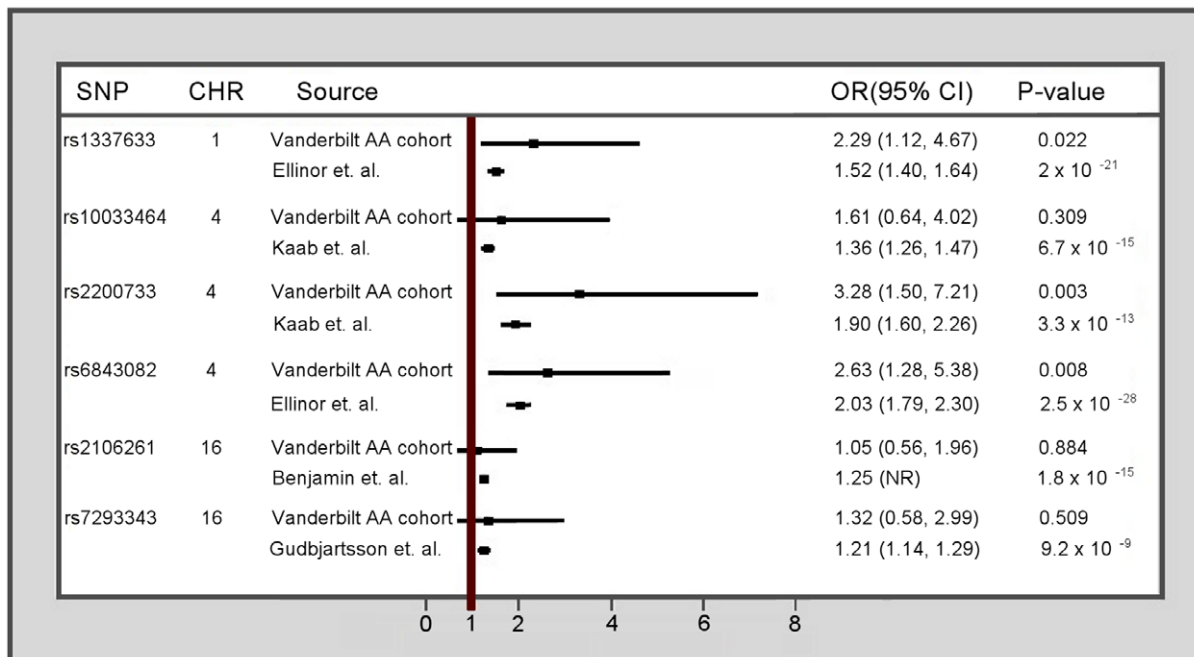
Second, we estimated admixture in cases and controls using ANCESTRYMAP and HAPMIX, respectively, and tested these estimates for an association with AF. Estimates of admixture from ANCESTRYMAP did not differ between cases and controls (23.6% and 23.8%, respectively; Table 1), and thus did not predict case-control status in this dataset. Local estimates of European ancestry for all three regions were calculated using HAPMIX. For chromosome 1, on average, only 2% of cases had two copies of the European derived chromosome in this region while no controls

had two copies. On chromosome 4, 50% of cases had two copies of the European chromosomes at this region compared to 53% of controls. On chromosome 16, ~2% of cases and controls had two copies of the European derived chromosome. Based on two estimates of admixture, cases and controls did not differ in proportion of European ancestry at any of the loci tested.

### Discussion

We demonstrate here that the European index SNPs from three genomics regions associated with AF are not the index SNPs associated with AF in African Americans. Using two estimates of admixture, we also show that European admixture at any of the genomic regions tested is not associated with increased AF risk in this African American cohort. Our results suggest that AF remains a complex genetic disease as the discordant findings of lower disease prevalence in African Americans, despite the greater incidence of AF clinical risk factors, cannot be explained by European admixture within the 1q21, 4q25, and 16q22 regions.

Of the six SNPs previously reported to be strongly associated with AF in GWA studies of European-descent populations, only three were found to be associated in African Americans, and their effect was reduced in comparison to other SNPs within the regions of interest. SNP rs2200733 in the 4q25 region has consistently replicated in European American populations as the index SNP, whereas in our African American cohort, rs4631108 was found to have the strongest association. A recent fine mapping study of the 4q25 region in European-descent populations identified two additional tagSNPs with a strong association with AF [13]. Of these two SNPs, one (rs3853445) was included in our SNP selection. In adjusted analyses, this SNP was not associated with AF in African Americans ( $P=0.164$ ; Table S1). Similar to the 4q25 region, the index SNP observed in African Americans differed from that reported in European-descent populations for both 1q21 and 16q22 genomic regions.



**Figure 1. A comparison of previously identified significant SNPs from literature to African Americans from Vanderbilt AF cohort.** The case/control sample size for the previously published studies are: Ellinor et. al.  $n=14,179$  (11); Kaab et. al.  $n=36,196$  (9); Benjamin et. al.  $n=46,736$  (10); and Gudbjartsson et. al.  $n=39,014$  (50). AA = Subjects with self-reported African American race. doi:10.1371/journal.pone.0032338.g001

The SNPs included in this study are located in intronic and intergenic regions, both of which are often thought of as regions without obvious function. However, variants in these regions may modify regulatory elements, ultimately resulting in disturbed protein expression. The nearest gene target for the intergenic SNPs found in the 4q25 region is *PITX2*, a gene believed to be involved in cardiac development, specifically sinus node development and susceptibility to atrial arrhythmias [41,42]. The 1q21 loci is intronic to *KCNK3*, a gene encoding a voltage-independent calcium-activated potassium channel, which has been found in a rabbit model to modify the action potential duration of atrial tissue resulting in increased AF susceptibility, yet, the contribution of this channel to repolarization in humans is unknown [43]. The SNPs located at 16q22 are intronic to *ZFX3*, a transcription factor with unknown cardiac involvement. It has been hypothesized that *ZFX3* may be a regulatory factor for the JAK/STAT signaling cascade and this cascade may be involved in atrial fibrillation susceptibility; however, further study is needed to clarify this relationship [44,45].

Fine mapping in African Americans has been proposed and employed to further refine signals from GWAS in European-descent populations with the premise that the lower levels of LD in the former population will be more likely to reveal the true causal SNP in a region of high LD observed in the latter population [46]. Also, the differences in allelic architecture known for these ancestral populations may reveal additional risk SNPs monomorphic or rare in one population but common in another. In this study, we attempted to capitalize on the lower levels of LD in African Americans compared with European Americans by densely genotyping independent SNPs for these three AF-associated regions. Previous studies of breast cancer and lipid traits have already capitalized on the differences in LD patterns, and, like this study, have identified new index SNPs in African Americans [47,48]. Future studies incorporating new variation data from the 1000 Genomes Project [49] promise to identify the full spectrum of “risk” variants for AF and other traits in these cross-population fine-mapping experiments.

Unlike a previous report in the literature [30], European ancestry was not associated with increased AF risk in our African American cohort. Marcus et. al. genotyped 1,747 ancestry informative markers across the genome from the Illumina custom ITMAT-Broad-CARE array, while our study focused on 148 SNPs in three genomic regions previously identified in GWAS. The conflicting results may be due to the different study designs, such as the choice of markers (AIMs versus region specific). Also, the study by Marcus et. al. was larger in sample size compared with the present study.

The primary limitation of this study is the small sample size, which is a reflection of the observed reduced prevalence of AF among African Americans compared with European Americans. Assuming the European genetic effect size and the African American MAF, we were powered (>80%) to replicate two of the three index SNPs from prior GWA studies (rs13376333 and rs2200733). For the 145 non-index SNPs, we were underpowered to detect an effect size of <2.0 to 3.3 (dependent on MAF) at a liberal p-value threshold of 0.05. The power calculations presented here should be interpreted with caution given that the European genetic effect size was assumed to be equal to the African American effect size for the same variant, the latter of which there are no available data.

The limited sample size of the present study is also represented in the confidence intervals, which are wide. *ZFX3* rs16971547, for example, was associated with AF with an odds ratio of 8.162 but with an upper bound 95% confidence interval of 45.420

( $p = 0.016$ ). This variant is monomorphic in HapMap CEU but common in YRI (6%) and ASW (9%). Because of the limited case count and the low minor allele frequency (5%), the association identified between rs16971547 and AF in African Americans would benefit from further study.

Overall, the p-values presented here and in supplementary material were not corrected for multiple testing. Correction for multiple testing is complicated given that a fraction of the SNPs represented replication (that is, the index SNPs) while the others represented novel tests of association. Also, several methods for correction are available, each with limitations. For example, the Bonferroni correction is the most widely implemented method but it is often criticized for being overly conservative for datasets with even low levels of linkage disequilibrium between SNPs. Other methods of correction include the false discovery rate (FDR) and permutation testing [50]. None of the tests of association reported here survive correction for multiple testing when either Bonferroni or FDR are applied (data not shown). However, all SNPs except rs3866823 with an uncorrected p-value < 0.05 have an empirical p-value < 0.05 after permutation testing. Replication of the associations reported here in an independent dataset is warranted given the small sample sizes and variability observed when implementing various corrections for multiple testing.

In conclusion, despite the overall limited power our study strongly suggests that the 1q21, 4q25, and 16q22 regions remain important in assessing AF risk among both European- and African-descent populations. These data also suggest the underlying biology identified by GWAS in European populations also applies to African American subjects.

## Supporting Information

**Figure S1 Locus Zoom plot for 1q21 region on chromosome 1.** Tests of association were performed for each SNP adjusted for age, body mass index, coronary artery disease, congestive heart failure, diabetes mellitus, and hypertension and are represented as circles or a diamond in the figure. SNPs are plotted based on chromosomal location (x-axis) and significance level (y-axis). Recombination rates are given on the opposing y-axis. The index association (rs484596) is denoted by the diamond. For this region, both the recombination rates, represented by the right axis, and linkage disequilibrium (based on HapMap phase II YRI), represented by the dot color in the SNP positions, are low. (TIF)

**Figure S2 Locus Zoom plot for 4q25 region on chromosome 4.** Tests of association were performed for each SNP adjusted for age, body mass index, coronary artery disease, congestive heart failure, diabetes mellitus, and hypertension and are represented as circles or a diamond in the figure. SNPs are plotted based on chromosomal location (x-axis) and significance level (y-axis). Recombination rates are given on the opposing y-axis. The index association (rs4631108) is denoted by the diamond. For this region, both the recombination rates, represented by the right axis, and linkage disequilibrium (based on HapMap phase II YRI), represented by the dot color in the SNP positions, are low. (TIF)

**Figure S3 Locus Zoom plot for 16q22 region on chromosome 16.** Tests of association were performed for each SNP adjusted for age, body mass index, coronary artery disease, congestive heart failure, diabetes mellitus, and hypertension and are represented as circles or a diamond in the figure. SNPs are plotted based on chromosomal location (x-axis) and significance



level (y-axis). Recombination rates are given on the opposing y-axis. The index association (rs16971547) is denoted by the diamond. For this region, both the recombination. (TIF)

**Table S1 Tests of association for atrial fibrillation in African Americans.** All tests of association are shown here regardless of significance, adjusted for age, body mass index, coronary artery disease, congestive heart failure, diabetes mellitus, and hypertension. (DOC)

**Table S2 Comparison of allele frequencies between cases and controls and reference samples.** Coded allele frequencies for the ten most significant SNPs are shown for cases, controls, HapMap CEU, and HapMap YRI samples. Allelic

frequency comparisons were performed with chi-square test and  $F$ -statistic ( $F_{ST}$ ). (DOCX)

## Acknowledgments

The Vanderbilt University Center for Human Genetics Research, Computational Genomics Core provided computational and/or analytical support for this work. Genotyping was performed by the Vanderbilt DNA Resources Core.

## Author Contributions

Conceived and designed the experiments: JTD JMJ NJB MP HEO DD DMR DCC. Performed the experiments: JTD JMJ DD DCC. Analyzed the data: JTD JMJ DD DMR DCC. Contributed reagents/materials/analysis tools: JMJ DCC. Wrote the paper: JTD JMJ NJB MP HEO DD DMR DCC.

## References

- Go AS, Hylek EM, Phillips KA, Chang Y, Henault LE, et al. (2001) Prevalence of Diagnosed Atrial Fibrillation in Adults: National Implications for Rhythm Management and Stroke Prevention: the AnTicoagulation and Risk Factors In Atrial Fibrillation (ATRIA) Study. *JAMA* 285: 2370–2375.
- Alonso A, Agarwal SK, Soliman EZ, Ambrose M, Chamberlain AM, et al. (2009) Incidence of atrial fibrillation in whites and African-Americans: the Atherosclerosis Risk in Communities (ARIC) study. *Am Heart J* 158: 111–117.
- Ruo B, Capra AM, Jensvold NG, Go AS (2004) Racial variation in the prevalence of atrial fibrillation among patients with heart failure: the Epidemiology, Practice, Outcomes, and Costs of Heart Failure (EPOCH) study. *J Am Coll Cardiol* 43: 429–435.
- Benjamin EJ, Chen P-S, Bild DE, Mascette AM, Albert CM, et al. (2009) Prevention of atrial fibrillation: report from a national heart, lung, and blood institute workshop. *Circulation* 119: 606–618.
- Cutler JA, Sorlie PD, Wolz M, Thom T, Fields LE, et al. (2008) Trends in hypertension prevalence, awareness, treatment, and control rates in United States adults between 1988–1994 and 1999–2004. *Hypertension* 52: 818–827.
- LaRosa JC, Brown CD (2005) Cardiovascular risk factors in minorities. *Am J Med* 118: 1314–1322.
- Ogden CL, Carroll MD, Curtin LR, McDowell MA, Tabak CJ, et al. (2006) Prevalence of Overweight and Obesity in the United States, 1999–2004. *JAMA* 295: 1549–1555.
- Milan DJ, Lubitz SA, Käab S, Ellinor PT (2010) Genome-wide association studies in cardiac electrophysiology: recent discoveries and implications for clinical practice. *Heart Rhythm* 7: 1141–1148.
- Käab S, Darbar D, van Noord C, Dupuis J, Pfeufer A, et al. (2009) Large scale replication and meta-analysis of variants on chromosome 4q25 associated with atrial fibrillation. *Eur Heart J* 30: 813–819.
- Benjamin EJ, Rice KM, Arking DE, Pfeufer A, van Noord C, et al. (2009) Variants in ZFXH3 are associated with atrial fibrillation in individuals of European ancestry. *Nat Genet* 41: 879–881.
- Gudbjartsson DF, Holm H, Gretarsdottir S, Thorleifsson G, Walters GB, et al. (2009) A sequence variant in ZFXH3 on 16q22 associates with atrial fibrillation and ischemic stroke. *Nat Genet* 41: 876–878.
- Ellinor PT, Lunetta KL, Glazer NL, Pfeufer A, Alonso A, et al. (2010) Common variants in KCNN3 are associated with lone atrial fibrillation. *Nat Genet* 42: 240–244.
- Lubitz SA, Sinner MF, Lunetta KL, Makino S, Pfeufer A, et al. (2010) Independent Susceptibility Markers for Atrial Fibrillation on Chromosome 4q25. *Circulation* 122: 976–984.
- Gudbjartsson DF, Arnar DO, Helgadóttir A, Gretarsdottir S, Holm H, et al. (2007) Variants conferring risk of atrial fibrillation on chromosome 4q25. *Nature* 448: 353–357.
- 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.
- Rybicki BA, Iyengar SK, Harris T, Liptak R, Elston RC, et al. (2002) The distribution of long range admixture linkage disequilibrium in an African-American population. *Hum Hered* 53: 187–196.
- Reiner AP, Carlson CS, Ziv E, Iribarren C, Jaquish CE, et al. (2007) Genetic ancestry, population sub-structure, and cardiovascular disease-related traits among African-American participants in the CARDIA Study. *Hum Genet* 121: 565–575.
- Parra EJ, Marcini A, Akey J, Martinson J, Batzer MA, et al. (1998) Estimating African American admixture proportions by use of population-specific alleles. *Am J Hum Genet* 63: 1839–1851.
- Cheng C-Y, Kao WHL, Patterson N, Tandon A, Haiman CA, et al. (2009) Admixture mapping of 15,280 African Americans identifies obesity susceptibility loci on chromosomes 5 and X. *PLoS Genet* 5: e1000490.
- Cheng C-Y, Reich D, Coresh J, Boerwinkle E, Patterson N, et al. (2010) Admixture mapping of obesity-related traits in African Americans: the atherosclerosis risk in communities (ARIC) Study. *Obesity* 18: 563–572.
- Behar DM, Rosset S, Tzur S, Selig S, Yudkovsky G, et al. (2010) African ancestry allelic variation at the MYH9 gene contributes to increased susceptibility to non-diabetic end-stage kidney disease in Hispanic Americans. *Hum Mol Genet* 19: 1816–1827.
- Vergara C, Caraballo L, Mercado D, Jimenez S, Rojas W, et al. (2009) African ancestry is associated with risk of asthma and high total serum IgE in a population from the Caribbean Coast of Colombia. *Hum Genet* 125: 565–579.
- Fejerman L, John EM, Huntsman S, Beckman K, Choudhry S, et al. (2008) Genetic ancestry and risk of breast cancer among U.S. Latinas. *Cancer Res* 68: 9723–9728.
- Fejerman L, Romieu I, John EM, Lazcano-Ponce E, Huntsman S, et al. (2010) European ancestry is positively associated with breast cancer risk in Mexican women. *Cancer Epidemiol. Biomarkers Prev* 19: 1074–1082.
- Lai C-Q, Tucker KL, Choudhry S, Parnell LD, Mattei J, et al. (2009) Population admixture associated with disease prevalence in the Boston Puerto Rican health study. *Hum Genet* 125: 199–209.
- Ruiz-Narváez EA, Bare L, Arellano A, Catanese J, Campos H (2010) West African and Amerindian ancestry and risk of myocardial infarction and metabolic syndrome in the Central Valley population of Costa Rica. *Hum Genet* 127: 629–638.
- Hughes LB, Morrison D, Kelley JM, Padilla MA, Vaughan LK, et al. (2008) The HLA-DRB1 shared epitope is associated with susceptibility to rheumatoid arthritis in African Americans through European genetic admixture. *Arthritis Rheum* 58: 349–358.
- Bock CH, Schwartz AG, Ruterbusch JJ, Levin AM, Neslund-Dudas C, et al. (2009) Results from a prostate cancer admixture mapping study in African-American men. *Hum Genet* 126: 637–642.
- Fejerman L, Haiman CA, Reich D, Tandon A, Deo RC, et al. (2009) An admixture scan in 1,484 African American women with breast cancer. *Cancer Epidemiol Biomarkers Prev* 18: 3110–3117.
- Marcus GM, Alonso A, Peralta CA, Lettre G, Vittinghoff E, et al. (2010) European ancestry as a risk factor for atrial fibrillation in African Americans. *Circulation* 122: 2009–2015.
- Darbar D, Motsinger AA, Ritchie MD, Gainer JV, Roden DM (2007) Polymorphism modulates symptomatic response to antiarrhythmic drug therapy in patients with lone atrial fibrillation. *Heart Rhythm* 4: 743–749.
- Motsinger AA, Donahue BS, Brown NJ, Roden DM, Ritchie MD (2006) Risk factor interactions and genetic effects associated with post-operative atrial fibrillation. *Pac Symp Biocomput* 584–595.
- Carlson CS, Eberle MA, Rieder MJ, Yi Q, Kruglyak L, et al. (2004) Selecting a maximally informative set of single-nucleotide polymorphisms for association analyses using linkage disequilibrium. *Am J Hum Genet* 74: 106–120.
- Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, et al. (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 81: 559–575.
- Weir BS, Cockerham CC (1984) Estimating F-Statistics for the Analysis of Population Structure. *Evolution* 38: 1358–1370.
- Grady BJ, Torstenson E, Dudek SM, Giles J, Sexton D, et al. (2010) Finding unique filter sets in plato: a precursor to efficient interaction analysis in gwas data. *Pac Symp Biocomput*. pp 315–326.
- 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.

39. Gauderman WJ (2002) Sample size requirements for association studies of gene-gene interaction. *Am J Epidemiol* 155: 478–484.
40. Gauderman WJ (2002) Sample size requirements for matched case-control studies of gene-environment interaction. *Stat Med* 21: 35–50.
41. Postma AV, Christoffels VM, Bezzina CR (2011) Developmental aspects of cardiac arrhythmogenesis. *Cardiovasc Res* 91: 243–251.
42. Kirchhof P, Kahr PC, Kaese S, Piccini I, Vokshi I, et al. (2011) PITX2c Is Expressed in the Adult Left Atrium, and Reducing Pitx2c Expression Promotes Atrial Fibrillation Inducibility and Complex Changes in Gene Expression/Clinical Perspective. *Circ Cardiovasc Genet* 4: 123–133.
43. Özgen N, Dun W, Sosunov EA, Anyukhovskiy EP, Hirose M, et al. (2007) Early electrical remodeling in rabbit pulmonary vein results from trafficking of intracellular SK2 channels to membrane sites. *Cardiovasc Res* 75: 758–769.
44. Li C, Wang F, Yang Y, Fu F, Xu C, et al. (2010) Significant association of SNP rs2106261 in the ZFX3 gene with atrial fibrillation in a Chinese Han GeneID population. *Hum Genet* 129: 239–246.
45. Tsai C-T, Lin J-L, Lai L-P, Lin C-S, Huang SKS (2008) Membrane translocation of small GTPase Rac1 and activation of STAT1 and STAT3 in pacing-induced sustained atrial fibrillation. *Heart Rhythm* 5: 1285–1293.
46. Udler MS, Meyer KB, Pooley KA, Karlins E, Struwing JP, et al. (2009) FGFR2 variants and breast cancer risk: fine-scale mapping using African American studies and analysis of chromatin conformation. *Hum Mol Genet* 18: 1692–1703.
47. Ruiz-Narváez EA, Rosenberg L, Cozier YC, Cupples LA, Adams-Campbell LL, et al. (2010) Polymorphisms in the TOX3/LOC643714 locus and risk of breast cancer in African-American women. *Cancer Epidemiol Biomarkers Prev* 19: 1320–1327.
48. Keebler ME, Deo RC, Surti A, Konieczkowski D, Guiducci C, et al. (2010) Fine-Mapping in African Americans of 8 Recently Discovered Genetic Loci for Plasma Lipids: The Jackson Heart Study. *Circ Cardiovasc Genet* 3: 358–364.
49. Durbin RM, Abecasis GR, Altshuler DL, Auton A, Brooks LD, et al. (2010) A map of human genome variation from population-scale sequencing. *Nature* 467: 1061–1073.
50. Rice TK, Schork NJ, Rao DC (2008) Methods for handling multiple testing. *Adv Genet* 60: 293–308.