Candidate gene approach to identifying rare genetic variants associated with lone atrial fibrillation

Peter Weeke, MD,[*†] Babar Parvez, MD,[*] Marcia Blair, MS,[*] Laura Short, MS,[*] Christie Ingram, BS,[*] Gayle Kucera, RN,[*] Tanya Stubblefield, RN,[*] Dan M. Roden, MD,[‡] Dawood Darbar, MD[*]

From the [*]Department of Medicine, Vanderbilt University, Nashville, Tennessee, [†]Department of Cardiology, Copenhagen University Hospital, Gentofte, Denmark, and [‡]Departments of Medicine and Pharmacology, Vanderbilt University, Nashville, Tennessee.

BACKGROUND Rare variants in candidate atrial fibrillation (AF) genes have been associated with AF in small kindreds. The extent to which such polymorphisms contribute to AF is unknown.

OBJECTIVE The purpose of this study was to determine the spectrum and prevalence of rare amino acid coding (AAC) variants in candidate AF genes in a large cohort of unrelated lone AF probands.

METHODS We resequenced 45 candidate genes in 303 European American (EA) lone AF probands (186 lone AF probands screened for each gene on average [range 89–303], 63 screened for all) identified in the Vanderbilt AF Registry (2002–2012). Variants detected were screened against 4300 EAs from the Exome Sequencing Project (ESP) to identify very rare (minor allele frequency $\leq$0.04%) AAC variants and these were tested for AF co-segregation in affected family members where possible.

RESULTS Median age at AF onset was 46.0 years [interquartile range 33.0–54.0], and 35.6% had a family history of AF. Overall, 63 very rare AAC variants were identified in 60 of 303 lone AF probands, and 10 of 19 (52.6%) had evidence of co-segregation with AF. Among the 63 lone AF probands who had 45 genes screened, the very rare variant burden was 22%. Compared with the 4300 EA ESP, the proportion of lone AF probands with a very rare AAC variant in *CASQ2* and NKX2-5 was increased 3-5-fold ($P < .05$).

CONCLUSION No very rare AAC variants were identified in ~80% of lone AF probands. Potential reasons for the lack of very rare AAC variants include a complex pattern of inheritance, variants in as yet unidentified AF genes or in noncoding regions, and environmental factors.

KEYWORDS Atrial fibrillation arrhythmia; Candidate genes; Family study; Genetic variation; Genetic epidemiology; Proarrhythmia; Rare variants

ABBREVIATIONS AAC = amino acid coding; **AF** = atrial fibrillation; **BMI** = body mass index; **ECG** = electrocardiogram; **ESP** = Exome Sequencing Project; **GWAS** = genome-wide association study; **IQR** = interquartile range; **LVEDD** = left ventricular end-diastolic diameter; **LVEF** = left ventricular ejection fraction; **LVESD** = left ventricular end-systolic diameter; **MAF** = minor allele frequency; **SNP** = single nucleotide polymorphism

Introduction

Atrial fibrillation (AF) is the most common cardiac arrhythmia in clinical practice and is associated with considerable morbidity and mortality.[1] Whereas traditional risk factors for AF such as advancing age, coronary artery disease, and congestive heart disease are well described, there is accumulating evidence that genetic factors also play a role in the pathogenesis of AF. Epidemiologic studies have shown that offspring with one parent affected with AF have a three- to five-fold increase in the risk of developing AF.[2]

Multiple genes and genetic loci affecting AF susceptibility have been identified using positional cloning and candidate gene approaches.[3–5] More recently, genome-wide association studies (GWAS) have revealed multiple novel AF susceptibility loci, although these findings only explain a small fraction of the interindividual risk for AF and the majority of common single nucleotide polymorphisms (SNPs) identified by GWAS reside in noncoding regions in which functional characterization remains a major challenge.[6–9]

Rare variants in candidate genes encoding cardiac ion channels, gap junction proteins, and signaling molecules have previously been reported in isolated AF cases and small kindreds.[6–11] Rare variants that co-segregate with AF in extended families also demonstrate abnormal electrophysiologic

This work was supported by National Institutes of Health Grants U19 HL65962 and R01 HL092217, and an American Heart Association Established Investigator Award (0940116N). Dr. Weeke was funded by an unrestricted research grant from the Tryg Foundation (J.nr. 7343-09, TrygFonden, Denmark). **Address reprint requests and correspondence:** Dr. Dawood Darbar, Division of Cardiovascular Medicine, Vanderbilt University School of Medicine, 2215B Garland Ave, Room 1285A MRB IV, Nashville, TN 37323-6602. E-mail address: dawood.darbar@vanderbilt.edu.

properties *in vitro*.[12] Although previous studies have identified common loci in unrelated AF individuals, we still do not have a good estimate of the overall prevalence of rare variants in candidate AF genes among lone AF patients.[6–9] Here, we determined the spectrum and prevalence of rare variants in AF candidate genes in a large lone AF cohort and tested for disease co-segregation.

Methods

Vanderbilt AF Registry

Since November 2002, subjects with AF have been prospectively enrolled in the Vanderbilt AF Registry, a clinical and genetic database. Patients were recruited from the Vanderbilt Cardiology and Arrhythmia Clinics, the emergency department, and in-patient services. Individuals older than 18 years with a confirmed diagnosis of AF on an electrocardiogram (ECG), rhythm strip, event recorder, or Holter monitor were included in the Vanderbilt AF Registry. The proband in each family was defined by having lone AF (i.e., early AF onset [$\leq$65 years old]) and no predisposing risk factors for AF including hypertension) and being the first family member encountered in the clinic. Familial AF was defined as the presence of AF in one or more first-degree relatives of the proband.

We defined paroxysmal AF as self-terminating AF lasting $>$30 seconds and persistent AF as AF lasting $>$7 days that was terminated by either pharmacologic intervention or electrical cardioversion. Permanent AF was defined as AF resistant to any type of cardioversion or allowed to continue.

Study cohort

All probands in the Vanderbilt AF Registry of European American (EA) ancestry who presented with lone AF (2002–2012) were identified. Lone AF was defined as AF occurring at age $\leq$65 years with no evidence of structural heart disease, hypertension, or thyroid dysfunction as determined by clinical examination, ECG, echocardiography, and thyroid function tests. For the present study we also gathered echocardiographic information and body mass index (BMI, kg/m^2) at the time of enrollment. Information on left atrial (LA; mm) size measurements was available in 254 lone AF probands, left ventricular ejection fraction (LVEF; %) in 270, left ventricular end-diastolic diameter (LVEDD; mm) in 241, left ventricular end-systolic diameter (LVESD; mm) in 228, and BMI in 291.

Screening of candidate genes

Resequencing of 45 candidate genes was performed in lone AF probands (2002–2012) (Online Supplemental Table 1) and in family members who provided consent. Lone AF probands enrolled in the Vanderbilt AF Registry were screened for genes associated with AF at the time of enrollment. Hence, genes that have only recently been implicated with AF (e.g., *MYH6*) will tend to have a greater proportion of lone AF proband individuals screened compared to genes that were screened earlier in development of the resource (e.g., *KCNQ1*).[8,13]

In brief, coding and flanking regions were amplified by polymerase chain reaction using primers designed to obtain fragments of appropriate size for each gene. Polymerase chain reaction–amplified DNA fragments were analyzed using the Reveal Discovery System (based on temperature gradient capillary electrophoresis) to identify aberrant conformers, which were then directly sequenced.

Rare variants and segregation analysis

Very rare amino acid coding (AAC) variants (i.e., missense, nonsynonymous, splice site, insertions, and deletions) in screened candidate genes among the lone AF probands were analyzed using the filtering steps shown in Figure 1. First, all AAC variants in screened candidate genes were identified. Second, we screened all of the identified AAC variants against dbSNP134 and the 4300 EAs in the National Heart, Lung, Blood Institute (NHLBI) Exome Sequencing Project (ESP) of 6500 individuals, excluding variants with a reported minor allele frequency (MAF) $>$0.04%.[14] Third, because structural variants may be less well reported in publically available databases compared to singletons, we calculated the MAF among the lone AF probands excluding insertions and deletions with an internal MAF$>$2%. Thus, we assumed that if an insertion or deletion is common among the lone AF probands, it is also likely to be common in the general population even though it may be absent from publically available databases (Figure 1). We assessed conservation and predicted functional effects for all variants with SIFT and PhastCons, GERP, Grantham scores, and PolyPhen2 using the SeattleSeq Genomic Variation Server (http://snp.gs.washington.edu/SeattleSeqAnnotation134/).

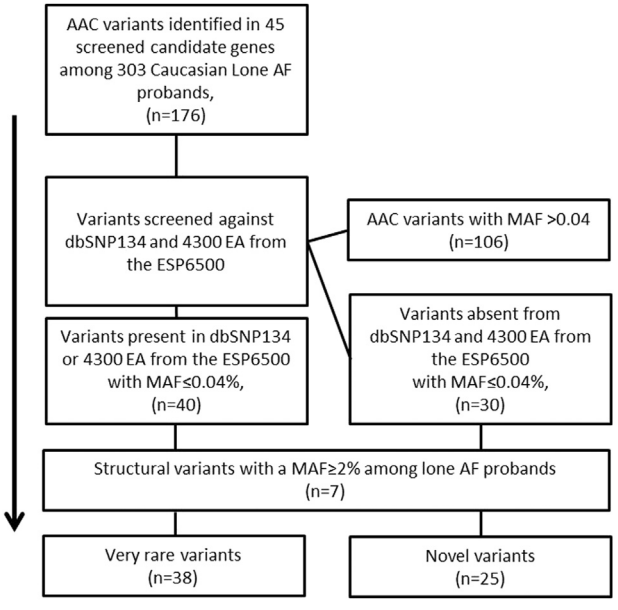

Figure 1 Flow chart identifying rare variants in screened candidate genes. AAC = amino acid coding; AF = atrial fibrillation; ESP6500 = Exome Sequencing Project of 6500 individuals using the reported minor allele frequency (MAF) for 4300 European Americans (EA).