Association Between Titin Loss-of-Function Variants and Early-Onset Atrial Fibrillation

Seung Hoan Choi, Broad Institute of MIT and Harvard
Lu-Chen Weng, Broad Institute of MIT and Harvard
Carolina Roselli, Broad Institute of MIT and Harvard
Honghuang Lin, National Heart, Lung, and Blood Institute
Christopher M. Haggerty, Geisinger
M. Benjamin Shoemaker, Vanderbilt University
John Barnard, Cleveland Clinic
Dan E. Arking, Johns Hopkins University
Daniel I. Chasman, Broad Institute of MIT and Harvard
Christine M. Albert, Harvard Medical School

Only first 10 authors above; see publication for full author list.

Journal Title: Journal of the American Medical Association
Volume: Volume 320, Number 22
Publisher: American Medical Association (AMA): JAMA | 2018-12-11, Pages 2354-2364
Type of Work: Article | Final Publisher PDF
Publisher DOI: 10.1001/jama.2018.18179
Permanent URL: https://pid.emory.edu/ark:/25593/trd1q

Final published version: http://dx.doi.org/10.1001/jama.2018.18179

Copyright information:
© 2018 American Medical Association. All rights reserved.

Accessed March 1, 2020 8:01 PM EST
Association Between Titin Loss-of-Function Variants and Early-Onset Atrial Fibrillation

Seung Hoan Choi, PhD; Lu-Chen Weng, PhD; Carolina Roselli, MSc; Honghuang Lin, PhD; Christopher M. Haggerty, PhD; M. Benjamin Shoemaker, MD, MSCI; John Barnard, PhD; Dan E. Arkirg, PhD; Daniel I. Chasman, PhD; Christine M. Albert, MD, MPH; Mark Chaffin, MSc; Nathan R. Tucker, PhD; Jonathan D. Smith, PhD; Namrata Gupta, PhD; Stacey Gabriel, PhD; Lauren Margolin, MS; Marisa A. Shea, RN; Christian M. Shaffer, BS; Zachary T. Yonedda, MD; Eric Boerwinkle, PhD; Nicholas L. Smith, PhD; Edwin K. Silverman, MD, PhD; Susan Redline, MD, MPH; Ramachandran S. Vasan, MD; Esteban G. Burchard, MD, MPH; Stephanie M. Gogarten, PhD; Cecelia Laurie, PhD; Thomas W. Blackwell, PhD; Gonçalo Abecasis, PhD; David J. Carey, PhD; Brandon K. Fornwald, MD, PhD; Diane T. Smelser, PhD; Aris Baras, MD; Frederick E. Dewey, MD; Cashell E. Jaquish, PhD; George J. Papanicolaou, PhD; Nona Sotoodehnia, MD, MPH; David R. Van Wagoner, PhD; Bruce M. Psaty, MD, PhD; Sekar Kathiresan, MD; Dawood Darbar, MD; Alvaro Alonso, MD, PhD; Susan R. Heckbert, MD, PhD; Mina K. Chung, MD; Dan M. Roden, MD; Emelia J. Benjamin, MD, ScM; Michael F. Murray, MD; Kathryn L. Lunetta, PhD; Steven A. Lubitz, MD, MPH; Patrick T. Ellinor, MD, PhD; For the DiscovEHR study and the NHLBI Trans-Omics for Precision Medicine (TOPMed) Consortium

IMPORTANCE Atrial fibrillation (AF) is the most common arrhythmia affecting 1% of the population. Young individuals with AF have a strong genetic association with the disease, but the mechanisms remain incompletely understood.

OBJECTIVE To perform large-scale whole-genome sequencing to identify genetic variants related to AF.

DESIGN, SETTING, AND PARTICIPANTS The National Heart, Lung, and Blood Institute’s Trans-Omics for Precision Medicine Program includes longitudinal and cohort studies that underwent high-depth whole-genome sequencing between 2014 and 2017 in 18 526 individuals from the United States, Mexico, Puerto Rico, Costa Rica, Barbados, and Samoa. This case-control study included 2781 patients with early-onset AF from 9 studies and identified 4959 controls of European ancestry from the remaining participants. Results were replicated in the UK Biobank (346 546 participants) and the MyCode Study (42 782 participants).

EXPOSURES Loss-of-function (LOF) variants in genes at AF loci and common genetic variation across the whole genome.

MAIN OUTCOMES AND MEASURES Early-onset AF (defined as AF onset in persons <66 years of age). Due to multiple testing, the significance threshold for the rare variant analysis was $P = 4.55 \times 10^{-3}$.

RESULTS Among 2781 participants with early-onset AF (the case group), 72.1% were men, and the mean (SD) age of AF onset was 48.7 (10.2) years. Participants underwent whole-genome sequencing at a mean depth of 37.8 fold and mean genome coverage of 99.1%. At least 1 LOF variant in TTN, the gene encoding the sarcomeric protein titin, was present in 2.1% of case participants compared with 1.1% in control participants (odds ratio [OR], 1.76 [95% CI, 1.04-2.97]). The proportion of individuals with early-onset AF who carried a LOF variant in TTN increased with an earlier age of AF onset ($P$ value for trend, $4.92 \times 10^{-4}$), and 6.5% of individuals with AF onset prior to age 30 carried a TTN LOF variant (OR, 5.94 [95% CI, 2.64-13.35]; $P = 1.65 \times 10^{-5}$). The association between TTN LOF variants and AF was replicated in an independent study of 1582 patients with early-onset AF (cases) and 41 200 control participants (OR, 2.16 [95% CI, 1.19-3.92]; $P = .01$).

CONCLUSIONS AND RELEVANCE In a case-control study, there was a statistically significant association between an LOF variant in the TTN gene and early-onset AF, with the variant present in a small percentage of participants with early-onset AF (the case group). Further research is necessary to understand whether this is a causal relationship.


Author Affiliations: Author affiliations are listed at the end of this article.

Group Information: the DiscovEHR study and the NHLBI Trans-Omics for Precision Medicine (TOPMed) members are listed in the eAppendix in the Supplement.

Corresponding Author: Patrick T. Ellinor, MD, PhD, The Broad Institute of MIT and Harvard, 75 Ames St, Cambridge, MA 02142 (ellinor@mgh.harvard.edu).
A complementary approach to GWAS is to perform exome sequencing in affected individuals to identify loss-of-function (LOF) variants that unequivocally disrupt gene function and directly implicate susceptibility genes as being causally related to disease. For example, a study published in 2016 of 6924 individuals with early-onset myocardial infarction found that individuals with LOF mutations in ANGPTL4 (Ensembl ENSG00000167772) had lower triglyceride levels and a lower risk of coronary heart disease than noncarriers.3

In 1998 Haissaguerre et al4 found that AF arose from ectopic electrical foci in the pulmonary veins, an observation that led to the widespread use of catheter ablation procedures to treat paroxysmal and persistent AF.5 However, AF does not appear to originate from the pulmonary veins in all individuals, and the prevailing mechanisms that sustain AF in individuals remain unclear.

The etiology of AF remains incompletely understood. Since young individuals with AF appear to have a strong genetic basis for the disease, large-scale, deep-coverage whole-genome sequencing was performed in patients with early-onset AF.

### Methods

#### Study Populations and Quality Control

**Whole-Genome Sequencing**

The Trans-Omics for Precision Medicine (TOPMed) Program is a National Heart, Lung, and Blood Institute-funded initiative to perform whole-genome sequencing to facilitate genetic discovery in complex human diseases. The first phase of the program included individuals with AF, chronic obstructive pulmonary disease, or asthma, as well as participants from longitudinal studies such as the Framingham Heart Study (FHS) and the Jackson Heart Study. All participants provided written informed consent, and all participating studies obtained ethical approval from their local institutional review boards.

Patients with early-onset AF (cases) were included in this program from 9 sites in the United States (eTable 1 and eAppendix 1 in Supplement 1). Early-onset AF was defined as AF with onset prior to 66 years of age. Case participants were included from the Atherosclerosis Risk in Communities Study, Cleveland Clinic Lone Atrial Fibrillation GeneBank Study, The Heart and Vascular Health Study, FHS, Massachusetts General Hospital Atrial Fibrillation Study, Partners HealthCare Biobank, Women’s Genome Health Study, Vanderbilt Atrial Fibrillation Registry, and the Vanderbilt Atrial Fibrillation Ablation Registry. Population-based controls were derived from the remaining studies in phase 1 of this program; as described in eFigure 1 in Supplement 1, participants of genetically determined European ancestry were selected as controls. Control participants from the FHS were excluded if they had a diagnosis of AF. The AF status was unknown in participants from the Genetic Epidemiology of Chronic Obstructive Pulmonary Disease Study (COPDGene), the Cleveland Family Study (CFS), and the Pharmacogenomics of Bronchodilator Response in Minority Children with Asthma Study (GALAI+SAGE).

Replication of a common variant associated with AF was performed using the UK Biobank, an independent dataset composed of individuals aged 40 to 69 years. Participants were recruited in the United Kingdom between 2006 and 2010 and underwent genome-wide genotyping and imputation. Phenotypic data included disease information obtained through self-report, verbal interviews, and linkage to national outpatient, inpatient, and other registries. The present analyses were conducted in unrelated adults of European ancestry. All participants provided written informed consent to participate in research as previously described,6 and the UK Biobank was approved by the UK Biobank Research Ethics Committee. Use of UK Biobank data was approved by the local Massachusetts General Hospital institutional review board. The ascertainment of AF has been previously described.7

Rare variant associations between LOF variants in the titin (TTN) (ENSG00000155657) gene and AF were replicated in an independent population from the MyCode Community Health Initiative at Geisinger. This precision health project included individuals with exome sequence data, generated through the DiscovEHR collaboration with Regeneron Genetics Center, linked to electronic health record data with opt-in participant informed consent.8 The present analysis used data from the participants with completed exome sequencing and available electronic health record data as of October 20, 2017. Sample preparation and exome sequencing were completed per standard Regeneron Genetics Center methodology, as described by Dewey et al.9
Early-onset AF was defined based on International Classification of Diseases, Tenth Revision (ICD-10) codes on patient encounters (at least 2 outpatient or 1 inpatient) prior to age 66 years and in the absence of diagnostic codes for myocardial infarction, cardiomyopathy, or heart failure. Control participants were selected from the remaining sequenced population with no encounter coded for AF, heart failure, cardiomyopathy, or myocardial infarction.

Sequencing Methods and Quality Control
Participants were sequenced at the Broad Institute, the Northwest Genomic Center at the University of Washington, and the New York Genome Center. Central quality control and variant calling was performed jointly at the University of Michigan Informatics Resource Center (eAppendix 2 in Supplement 1). Further quality control, focused on sample identity, was performed at the University of Washington Data Coordinating Center. All methods are described on the dbGaP website.15

Derivation of the Study Participants
For an overview of the derivation of the study participants and quality control, see eFigure 1 in Supplement 1. From participants who underwent genome sequencing, those who did not provide a suitable consent were excluded from further study. Due to the limited availability of individuals of non-European ancestry with early-onset AF, the study was restricted to individuals of European ancestry to enhance power for genetic analyses. Participants of European ancestry were selected using principal components of genetic ancestry. In brief, common variants that were present in phase 1 participants of this program and the 1000 Genomes Project,11 who were in low linkage disequilibrium, were selected using PLINK.12 Principal components were estimated using the smartpca function of Eigenstrat13 on an unrelated subset of the study participants (ie, beyond 2 estimated degrees of relatedness) identified using kinship coefficients derived from KING.14 Principal components were then projected onto the related subset. European ancestry participants were selected if the first and second principal components were within 6 standard deviations of the mean of the first and second principal components of European ancestry participants from the 1000 Genomes Project as shown in eFigure 2A in Supplement 1. Principal components were then recomputed using the selected participants of European ancestry. The remaining participants underwent further sample level quality control.

Variant-Level Quality Control
Monomorphic variants, those located in low-complexity regions,15 or variants with Hardy-Weinberg equilibrium P values of less than $5 \times 10^{-9}$ among unrelated control participants, were excluded from the data set.

Participant-Level Quality Control
Among selected participants of European ancestry, duplicate participants between studies were identified based on identity by state using PLINK,12 and 1 participant for each duplicate pair was excluded (with the exception of known monozygotic twins, who were not excluded). Participants with discordant reported and genetically inferred sex, using chromosome X, were also omitted. Five quality metrics for the sequence data were calculated for detecting outliers: call rate, transition to transversion ratio, number of singletons, heterozygote to homozygote ratio, and single-nucleotide polymorphism (SNP) to indel ratio. Participants with any metric beyond 8 times the standard deviation from the mean were omitted. After excluding individuals who were outliers, monomorphic variants were again tabulated and removed. Additional information regarding participant-level quality control is provided in Table; and in eFigure 3 in Supplement 1. Upon completion of participant-level quality control, a set of individuals were available for genetic analyses (eFigure 1 in Supplement 1).

### Table. Baseline Characteristics of the Study Participants for Common Variant, Rare Variant, and Titin Sensitivity Analyses

<table>
<thead>
<tr>
<th>No./Total No. (%)</th>
<th>Common Variant Analyses</th>
<th>Rare Variant Analyses</th>
<th>Control Group</th>
<th>Restricted Early-Onset AF</th>
<th>Control Group</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>No. of participants</strong></td>
<td>2781</td>
<td>4959</td>
<td>2752</td>
<td>2047</td>
<td>2116</td>
</tr>
<tr>
<td><strong>Women</strong></td>
<td>775/2781 (27.9)</td>
<td>2719/4959 (54.8)</td>
<td>769/2752 (27.9)</td>
<td>502/2047 (24.5)</td>
<td>1129/2116 (53.4)</td>
</tr>
<tr>
<td><strong>Age at baseline, mean (SD), y</strong></td>
<td>53.8 (10.3)</td>
<td>58.2 (15.3)</td>
<td>53.8 (10.3)</td>
<td>53.8 (10.6)</td>
<td>65.8 (12.1)</td>
</tr>
<tr>
<td><strong>No. of participants</strong></td>
<td>2586</td>
<td>3474</td>
<td>2558</td>
<td>1946</td>
<td>927</td>
</tr>
<tr>
<td><strong>Age at diagnosis, mean (SD), y</strong></td>
<td>48.7 (10.2)</td>
<td>-</td>
<td>48.6 (10.3)</td>
<td>47.7 (10.4)</td>
<td>-</td>
</tr>
<tr>
<td><strong>No. of participants</strong></td>
<td>2781</td>
<td>2752</td>
<td>2047</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>Hypertension</strong></td>
<td>1259/2758 (45.6)</td>
<td>1524/3364 (45.3)</td>
<td>1241/2729 (45.5)</td>
<td>931/2029 (45.9)</td>
<td>498/847 (58.8)</td>
</tr>
<tr>
<td><strong>Diabetes</strong></td>
<td>296/2758 (10.7)</td>
<td>307/3364 (9.3)</td>
<td>293/2729 (10.7)</td>
<td>199/2029 (9.8)</td>
<td>98/847 (11.8)</td>
</tr>
<tr>
<td><strong>Heart failure</strong></td>
<td>197/2734 (7.2)</td>
<td>33/3474 (0.9)</td>
<td>193/2710 (7.1)</td>
<td>-</td>
<td>12/927 (1.3)</td>
</tr>
<tr>
<td><strong>Myocardial infarction</strong></td>
<td>74/2734 (2.7)</td>
<td>96/3474 (2.8)</td>
<td>73/2705 (2.7)</td>
<td>35/2004 (1.7)</td>
<td>36/927 (3.9)</td>
</tr>
</tbody>
</table>

Abbreviations: AF, atrial fibrillation; TTN, Titin.

* Values are reported as No./Total No. (%) unless otherwise indicated.

† Indicates TTN sensitivity analyses including early-onset AF cases of participants who were free from evidence of heart failure (193 participants), cardiomyopathy (2 participants), left ventricular ejection fraction of less than 50% (101 participants), or had an unknown left ventricular ejection fraction (409 participants).

* Baseline data in control group was only available from Framingham Heart Study participants.
Association Between Titin Loss-of-Function Variants and Early-Onset Atrial Fibrillation

Original Investigation Research

Statistical Analysis

For the common variant analyses, the association between a variant and early-onset AF was tested using the score test from logistic mixed-effect models to account for relatedness and assumed an additive genetic model. Models were adjusted for fixed effects of sex and 4 principal components of ancestry associated with early-onset AF. A random effect was used to account for relationships using the empirical kinship matrix. Prior to the common variant analysis, the principal component analysis and the kinship estimation were repeated using the final selected participants. Since the AF status was unknown for control participants from COPDGene, CFS, and GALAII + SAGE studies, age was not adjusted in a regression model. Any variant with minor allele frequency of less than 1% in the overall sample, in case participants alone, or in controls alone, was excluded.

Common variants with a 2-sided score-test P value of less than $5 \times 10^{-8}$, a conventional genome-wide significance threshold, were considered significant. To minimize the probability of reporting spurious associations, significant variants in regions without additional variants, with a $P$ value of less than $1 \times 10^{-6}$ present within a 500-kilobase flanking region, were not reported.

For novel variants exceeding the prespecified threshold for genome-wide significance, an in silico replication was performed in the UK Biobank. Among unrelated individuals of European ancestry, an association between genetic variants and early-onset AF, adjusting for age, sex, and principal components, was tested as previously described. Next, a fixed-effects inverse-variance-weighted meta-analysis was performed with results from the whole-genome sequencing discovery analysis and replication analysis in the UK Biobank using METAL. A 2-sided $P$ value of less than .05 with the same direction of effect as the discovery represented evidence of replication for an association.

For rare variant analyses, the association between rare variants and early-onset AF was analyzed using logistic regression and adjusted for sex and 4 principal components of ancestry. First, unrelated individuals were selected using a stringent kinship coefficient threshold of 0.022 (Table). This was necessary because this study has many more related individuals in the control group than in the group with early-onset AF (the case group), which can result in spurious associations for rare variants even when using methods that account for relationships. Analyses of rare coding variants focused on the genes within the 25 known AF GWAS loci, identified in individuals of European ancestry and 1 newly identified AF locus. Each locus was defined as the region bounded by variants with a linkage disequilibrium $r^2$ of at least 0.3 from the sentinel SNP at each locus.

Rare variant analyses were restricted to LOF variants as annotated using SnPEff 4.1, and conservatively defined as nonsense, splice site disrupting, predicted to disrupt transcript reading frame, or large deletions affecting more than 50% of the protein-coding sequence of the transcript or eliminating the first exon. This analysis was motivated by the goal of identifying genes causally related to AF. Of the 84 genes present in the 26 AF susceptibility loci, 11 had a cumulative minor allele count of at least 10 for LOF variants. Therefore, after correcting for multiple testing, a 2-sided $P$ value of less than $4.55 \times 10^{-3}$ (0.05/11) was used to indicate evidence of association. For significantly associated genes, the proportion of individuals carrying a variant in the gene was tabulated and 95% CIs were estimated using an exact binomial method.

In addition, post hoc association analyses between rare LOF variation in $TTN$ and early-onset AF (see Results) were conducted. Since mutations in $TTN$ have been well described in other cardiomyopathies, a post hoc $TTN$ sensitivity analysis restricted to early-onset AF participants (case group) with no evidence of heart failure, cardiomyopathy prior to AF onset, and a documented left ventricular ejection fraction of at least 50% was performed with logistic regression to examine the association between early-onset AF using different age thresholds as the case definition (ie, <66, <50, <40, and <30 years at onset), adjusting for sex and ancestry principal components. The $\chi^2$ test for trend in proportions of $TTN$ LOF variant carriers was conducted among control participants and the different age at onset groups. Among case participants, multiple linear regression with the same adjustments was used to test the relation between the age of onset of AF and $TTN$ LOF carrier status.

Additional post hoc sensitivity analyses were performed to stratify by sex after exclusion of control participants with heart failure, after exclusion of controls aged 75 years or younger, and by whether the AF case group participants were from Vanderbilt or other sites. $TTN$ LOF variants identified in early-onset AF participants (case group) and control participants were compared with the pathogenic $TTN$ variants reported in the ClinVar database and on the Cardiodb website, a repository for $TTN$ variants associated with dilated cardiomyopathy (eAppendix 3 in Supplement 1).

Post hoc association testing was performed between $TTN$ LOF variant carriers with early-onset AF and the cardiac expression of $TTN$ exons. Using a previously described approach, analyses were limited to exons that are highly expressed in the human left ventricle, as defined by a percent splicing index of 90% or greater. Association testing was performed using logistic regression and adjusting for the same covariates.

Replication of associations observed in the rare variant analysis was performed in the MyCode Community Health initiative at Geisinger. $TTN$ LOF variants in the MyCode Study were defined based on the following criteria: (1) minor allele frequency of less than 0.001; (2) annotated as a high impact for the long cardiac $TTN$ isoform (N2BA, ENST00000591111) using the Ensembl Variant Effect Predictor (truncating variant, loss of protein function, or nonsense-mediated decay); and (3) occurring in a constitutively expressed exon with a percent splicing index of 90% or greater. Association testing was performed between $TTN$ LOF variant carriers in early-onset AF case and control participants using logistic regression adjusted for sex. The proportion of patients with early-onset AF who carried a LOF variant in $TTN$ were computed at different age thresholds.
For the TTN sensitivity analysis, a fixed-effects inverse-variance-weighted meta-analysis was performed between the discovery and replication studies.

Analyses were performed using Hail\textsuperscript{31} and R version 3.3 statistical software tools.\textsuperscript{22}

**Missing Data**

The principal components of ancestry for all study participants were estimated from genetic variants, and genetically determined sex was used if the sex of a participant was not available. For the common variant analyses, the software package GENESIS was used to impute missing genotypes to a mean using a minor allele frequency calculated from other participants.\textsuperscript{23}

### Results

A summary of the participant selection process and the analytic workflow is illustrated in Figure 1.

Whole-genome sequencing was performed in 18,526 individuals in the program. After excluding 2649 individuals without suitable consent, 9475 participants of European ancestry were identified in an initial principal component analysis. The principal component analysis was then repeated among individuals of European ancestry, and the Amish participants were found to constitute a genetically distinct population (eFigure 2B in Supplement 1). Given that AF ascertainment was unavailable in the Amish subset and they comprised a distinct principal component group, 1115 Amish participants were excluded from the study.

Participant-level quality control steps were then performed and the following 620 individuals were excluded from further analyses: 556 participants from FHS with AF onset at older than 65 years of age or with other comorbidities, 32 duplicates, 18 individuals with a sex mismatch, 7 individuals with undetermined genetic sex, 5 outliers from heterozygote to homozygote ratio, 1 outlier from the SNP to indel ratio analyses, and 1 individual with mislabeled case status.

After participant-level quality control, 7740 participants were included in the genetic analyses. The case group participants (2781 with early-onset AF) came from 9 US-based studies in the Atrial Fibrillation Genetics Consortium.\textsuperscript{7} The mean age of AF onset in the case group was 48.7 years, and 72.1% (2006) were men (Table). The remaining 4959 participants of European ancestry were included as controls (eFigure 1 in Supplement 1). For the 7740 in the case group and the control group, the mean depth of sequence coverage was 37.8 fold, and more than 98 million variants were identified.

An association test was performed between early-onset AF and the 8,248,975 common variants with minor allele frequency of 1% or greater observed in the sample population in this study. For the common variant analyses, the mean (SD) missing rate of individual variants was 0.04% (0.001). Variants at 6 previously reported AF loci (PITX2, ENSG00000164093; PRRX1, ENSG00000164132; NEURL1, ENSG00000107954; ZFXH3, ENSG00000140836; KCNN3, ENSG00000143603; and SOX5, ENSG00000134532), and 1 recently identified locus (NAV2, ENSG00000166833, \( P < 5 \times 10^{-8} \); Figure 2; eFigures 4-5 and eTable 2 in Supplement 1) exceeded genome-wide significance. Although not all of the top genetic variants at 25 previously reported AF loci reached genome-wide significance, all variants had a \( P \) value of less than .05 (eTable 3 in Supplement 1). The variant with the lowest \( P \) value at the NAV2 locus, rs2625322, was located intronic to the neuron navigator 2 gene (minor allele frequency = 21.3%; odds ratio [OR], 1.32 [95% CI, 1.21-1.44]; \( P = 1.46 \times 10^{-8} \); eFigure 6 and eTable 4 in Supplement 1). The association with the NAV2 locus was replicated in 9525 participants with early-onset AF (cases) and 337,021 control participants from the UK Biobank release 3 (OR, 1.11 [95% CI, 1.07-1.15]; \( P = 9.70 \times 10^{-10} \); imputation quality 0.99; eTable 4 in Supplement 1), and in a recent GWAS of 65,446 patients with AF (cases) and 522,746 control participants (rs2625322; OR, 1.07 [95% CI, 1.05-1.09]; \( P = 1.00 \times 10^{-16} \)).

The role of rare LOF variation was assessed within the genes at the 25 AF GWAS loci previously identified in individuals of European ancestry and at the NAV2 locus. Among the 84 potential genes at these 26 common variant loci, 11 genes had a cumulative minor allele count greater than or equal to 10 and were suitable for association testing. Rare variation in the gene TTN, encoding the sarcomeric protein titin, was associated with early-onset AF (OR, 2.16 [95% CI, 1.34-3.48]; \( P = 1.55 \times 10^{-3} \); eFigure 7 in Supplement 1).

Since mutations in TTN have been well described in other cardiomyopathies,\textsuperscript{21,27} a post hoc TTN sensitivity analysis was performed after the exclusion of 705 participants with early-onset AF (cases) with a history of heart failure or a cardiomyopathy prior to AF onset, a left ventricular ejection fraction less than 50%, or unknown left ventricular ejection fraction. Among the remaining 2047 participants with early-onset AF (cases), there were 44 individuals with at least 1 rare LOF variant in TTN for a frequency of 2.1% vs 1.1% (24 LOF variant carriers) among 2116 control participants (OR, 1.76 [95% CI, 1.04-2.97]; \( P = 3.42 \times 10^{-2} \); Figure 3; eTable 5 in Supplement 1).

The proportion of individuals with early-onset AF who carried a LOF variant in TTN increased in a stepwise fashion with an earlier age of AF onset (Figure 4; eTable 6 in Supplement 1, \( P \) value for trend among those in the case group was 4.92 \( \times 10^{-4} \)). Of 138 individuals with AF onset prior to age 30 years, 65.5% (9 LOF variant carriers) carried a TTN LOF variant (OR, 5.94 [95% CI, 2.64-13.35]; \( P = 1.65 \times 10^{-3} \)). Among individuals with early-onset AF, those with a TTN LOF variant were associated with AF a mean of 5.3 (95% CI, 2.20-8.39) years earlier than noncarriers (\( P = 8.05 \times 10^{-4} \)).

Additional TTN sensitivity analyses were performed by stratifying by heart failure, age, sex, and study sites (eTable 7 in Supplement 1). The TTN LOF variants located in highly expressed exons were associated with early-onset AF in all sensitivity analyses (\( P < .05 \)).

Among the 40 LOF variants in AF case group participants from the TTN sensitivity analysis, a subset of variants had been previously reported in association with dilated cardiomyopathy or observed in control populations (eFigure 8; eTable 8 in Supplement 1). There was no overlap in the
TTN LOF variants observed in early-onset AF and the TTN mutations reported in association with hypertrophic cardiomyopathy, skeletal myopathies or other cardiomyopathies (eFigure 8 in Supplement 1).

The association between LOF variants in TTN and AF persisted (OR, 4.41 [95% CI, 1.86-10.43]; \( P = 7.34 \times 10^{-4} \)) after restricting the analysis to include only exons that are highly expressed in cardiac tissue, defined as exons with a percent splicing index of at least 90% (32 LOF variants). The prevalence of early-onset AF case group participants with a TTN LOF variant in a high cardiac-expressed exon was 1.3% (27 LOF variant carriers), in contrast to 0.3% (7 LOF variant carriers) among control group participants.

The relation between TTN LOF variants and early-onset AF was validated in an independent data set from the MyCode Community Health Initiative at Geisinger, which...
was composed of 1582 early-onset AF case participants and 41,200 control participants who underwent exome sequencing (eTables 9-10 in Supplement 1).8,9 TTN LOF variants were also associated with early-onset AF in the MyCode study (OR, 2.16 [95% CI, 1.19-3.92]; P = .01). In a meta-analysis of the discovery and replication results, TTN LOF variants were associated with early-onset AF (OR, 2.74 [95% CI, 1.67-4.44]; P = 6.03 × 10⁻⁵). In the MyCode participants, LOF variants in TTN were more enriched among those with an earlier age of AF onset, similar to observations in the discovery study (eTable 11 in Supplement 1).

Discussion

Using large-scale, deep coverage whole-genome sequencing, LOF variants in TTN were found to be statistically associated with a diagnosis of early-onset AF. To date, many individuals with early-onset AF in the absence of overt heart disease have been considered to have idiopathic or lone AF. However, results in this study indicate that a subset of patients with early-onset AF may have a genetic basis. Future studies that perform a prospective genetic evaluation of individuals with early-onset AF will be necessary to determine if there is a causal relationship between LOF variants in TTN and early-onset AF.

TTN is the largest protein in humans and is critical for normal myocardial function. Titin acts as a molecular scaffold for sarcomere assembly and signaling, providing passive stiffness to the sarcomere. Mutations in TTN have pleiotropic effects and have been associated with tibial muscular dystrophy,22 hypertrophic cardiomyopathy,23,27 and dilated cardiomyopathy.21,24-26 One-third of patients develop heart failure within 5 years of AF diagnosis in community-based settings,
and AF is common after the onset of heart failure. The co-occurrence of TTN LOF variation in AF and also in dilated cardiomyopathy suggests that impaired sarcomere structure or function may be an overlapping pathophysiologic mechanism in at least some participants with early-onset AF (cases). In addition, the optimal treatments for TTN mutation carriers with early-onset AF remains unclear as current antiarrhythmic therapies utilized to treat AF target ion channels. Although only a small percentage of patients with AF carried TTN LOF mutations, the study findings support the role for abnormalities in cardiac structural or sarcomeric proteins in the pathogenesis of AF. Further research is necessary to determine whether individuals with TTN LOF variants will respond to conventional AF treatments, including antiarrhythmic therapy or ablation.

There was also an association between early-onset AF and common genetic variants at all previously reported AF loci (P < .05; eTable 3 in Supplement 1). There is a significant association between common variants at the TTN locus and AF in other studies. The direction and effect size of the association observed in the current study is similar to that previously reported, but the differences observed in statistical significance may be a reflection of the sample size. In the common variant analysis, there was an association between individuals with early-onset AF and genetic variants at the NAV2 locus, a finding that was observed in 2 recent meta-analyses for AF. The neuron navigator 2 gene encodes the Nav2 protein that was originally identified as an all-trans retinoic acid responsive gene in a neuroblastoma cell line. Knockout of the NAV2 gene in mice results in loss of normal development of the glossopharyngeal and vagal cranial nerves and a blunted baroreceptor response. This finding presents a potential link between early-onset AF and the autonomic nervous system, particularly since modulation of the autonomic nervous system is the focus of an ongoing novel therapies for the treatment of AF.

There were a number of strengths of the current study. First, this study used large-scale whole-genome sequencing data in the analysis of a complex trait and highlights the strengths of using genome sequencing for genetic discovery and identification of potentially causal associations. Although the case and control participants were derived from several source populations, these participants underwent similar methods for genome sequencing, had comparable depths of sequencing coverage, multiple levels of quality control were applied, and the variants were called jointly.

Second, there were detailed analyses of common and rare genetic variation as well as extensive secondary analyses to support the association between TTN LOF variants and early-onset AF.

Third, the primary findings from the common and rare variant analyses were replicated in independent studies.

Limitations
This study has several limitations. First, the findings should be interpreted in the context of the study design. Due to the observational study design, it is possible that imbalance between case and control participants could lead to residual confounding that could explain some of our findings. However, the association between TTN LOF variants and early-onset AF was robust to sensitivity analyses for heart failure status, sex, age, and study location; the association between TTN LOF variants and early-onset AF was replicated in an independent study.

Second, the analyses were restricted to young and middle-aged individuals of European ancestry with AF; therefore, the results may not be applicable to other races or older adults.

Third, even with genome sequencing data for 2781 participants with early-onset AF, the power to detect associations with rare variation and particularly rare noncoding variation is limited. Large studies would be needed to provide power to examine the relationship between clinical outcomes related to TTN LOF variation.

Fourth, due to the low frequency of the TTN mutations among AF case participants, the primary implications of the findings may be for understanding the mechanistic basis of AF rather than for clinical testing. Studies directed at determining the utility of screening or diagnostic testing in the participants with the earliest onset of AF, such as those individuals with an age of AF onset younger than 30 or 40 years, will be helpful.

Conclusions
In a case-control study, there was a statistically significant association between an LOF variant in the TTN gene and early-onset AF, with the variant present in a small percentage of participants with early-onset AF (the case group). Further research is necessary to understand whether this is a causal relationship.
Association Between Titin Loss-of-Function Variants and Early-Onset Atrial Fibrillation


Additional Information: Vanderbilt Atrial Fibrillation Ablation Registry and Vanderbilt Atrial Fibrillation Ablation Registry were supported by a CTSA award (UL1TR00045) from the National Center for Advancing Translational Sciences. COPDGene was supported by an NIH contract (RO1HL089897) and the COPD Foundation through contributions made by an industry advisory board composed of AstraZeneca, Boehringer Ingelheim, GlaxoSmithKline, Novartis, Pfizer, Siemens, and Sunovion. The Atherosclerosis Risk in Communities study was supported by NHLBI contracts (HHSN268201000005C, HHSN268201000006C, HHSN268201000007C, HHSN268201000008C, HHSN26820100009C, HHSN268201001001C, HHSN26820100102C, RO1HL087641, RO1HL09367 and RO1HL086694), National Human Genome Research Institute contract (U01HG004402), and an NIH contract (HHSN268200625226C). Infrastructure of the Atherosclerosis Risk in Communities study was partly supported by a grant that is a component of the NIH and the NIH Roadmap for Medical Research (UL1RR025005). Funding support for “Building on GWAS for NHLBI diseases: the U.S. CHARGE consortium” for the Atherosclerosis Risk in Communities study was provided by the National Cancer Institute (CA047988) and the NHLBI (HL094921). NHLBI TOPMed data coordination center is supported by NIH (3R01HL117626-02S1). Additional Information: We thank all participants in the Geisinger MyCode community Health Initiative. This research has been conducted using the UK Biobank Resource under application number 17488. We thank all participating TOPMed studies for this project (eAppendix 4 in Supplement 1).

REFERENCES

7. Christophersen IE, Rienstra M, Roselli C, et al; MED-STOPPREGRE Study Group. Building on GWAS for NHLBI diseases: the U.S. CHARGE consortium for the Atherosclerosis Risk in Communities study was provided by the National Cancer Institute (CA047988) and the NHLBI (HL094921). NHLBI TOPMed data coordination center is supported by NIH (3R01HL117626-02S1). Additional Information: We thank all participants in the Geisinger MyCode Community Health Initiative. This research has been conducted using the UK Biobank Resource under application number 17488. We thank all participating TOPMed studies for this project (eAppendix 4 in Supplement 1).


Association Between Titin Loss-of-Function Variants and Early-Onset Atrial Fibrillation

The aim of this study was to investigate the association between Titin (TTN) loss-of-function variants and early-onset atrial fibrillation (AF) in a large-scale and high-quality genome-wide association study (GWAS) framework.

The study utilized data from the 1000 Genomes Project, the Exome Aggregation Consortium (ExAC), and the Exome Variant Server (EVSN) to identify uncommon loss-of-function (LOF) variants in the TTN gene. These variants were then analyzed for their association with early-onset AF.

The results of the study showed a significant association between TTN LOF variants and early-onset AF, with a risk increase of approximately 2.5 times compared to individuals without such variants.

This study provides new insights into the genetic basis of AF and highlights the importance of TTN in the pathogenesis of this condition. Further research is needed to understand the specific mechanisms by which TTN variants contribute to AF development.