A Common Connexin-40 Gene Promoter Variant Affects Connexin-40 Expression in Human Atria and Is Associated With Atrial Fibrillation

Robert C. Wirka, MD; Shamone Gore, BS; David R. Van Wagoner, PhD; Dan E. Arking, PhD; Steven A. Lubitz, MD; Kathryn L. Lunetta, PhD; Emelia J. Benjamin, MD, ScM; Alvaro Alonso, MD, PhD; Patrick T. Ellinor, MD, PhD; John Barnard, PhD; Mina K. Chung, MD; Jonathan D. Smith, PhD

Background—A common single-nucleotide polymorphism (SNP) in the promoter of the Connexin-40 (Cx40) gene GJA5 was suggested to affect Cx40 promoter activity and the risk of atrial fibrillation (AF), but the role of other common Cx40 polymorphisms is unknown.

Methods and Results—Eight SNPs within the Cx40 gene region were tested for association with Cx40 levels measured in atrial tissue from 61 individuals. The previously described Cx40 promoter SNP (rs35594137, −44G→A) was not associated with Cx40 mRNA levels. However, a common SNP (rs10465885) located in the TATA box of an alternative Cx40 promoter was strongly associated with Cx40 mRNA expression (P<0.0001) and displayed strong and consistent allelic expression imbalance in human atrial tissue. A promoter-luciferase assay in cultured murine cardiomyocytes demonstrated reduced activity of the promoter containing the minor allele of this SNP (P<0.0001). Both rs35594137 and rs10465885 were tested for association with early-onset lone AF (≤60 years of age) in 384 cases and 3010 population control subjects. rs10465885 was associated with the AF phenotype (odds ratio, 1.18; P=0.046). This result was confirmed in a meta-analysis including 2 additional early-onset lone AF case-control cohorts (odds ratio, 1.16, P=0.022). rs35594137 was not associated with the lone AF phenotype in any of the cohorts studied or in a combined analysis.

Conclusions—A previously described Cx40 promoter SNP was not found to influence Cx40 expression or risk of AF. We describe an alternate promoter polymorphism that directly affects levels of Cx40 mRNA in vivo and is associated with early-onset lone AF. (Circ Arrhythm Electrophysiol. 2011;4:87-93.)

Key Words: atrial fibrillation ■ ion channels ■ genetics ■ allelic expression imbalance

Cardiac connexin-40 (Cx40) is found exclusively in the atria and conduction system. Along with connexin-43, it comprises the gap junctions that electrically couple atrial myocytes. As Cx40 is the only connexin found exclusively in the atria, several studies have investigated the effect of genetic alterations in the Cx40 gene gap junction protein, alpha 5 (GJA5) on the development of atrial conduction abnormalities and atrial fibrillation (AF). Complete absence of Cx40 in Cx40−/− mice leads to altered propagation and ectopic rhythms in the atria and arrhythmias on surface ECG. In 15 patients with idiopathic AF, Gollob et al discovered 3 somatic and 1 germ line mutation in the Cx40 coding region, resulting in dominant-negative Cx40 activity.

Clinical Perspective on p 93

The Cx40 gene GJA5 contains 2 alternative first exons (exons 1A and 1B), using separate promoters (promoters A and B), resulting in 2 alternative transcripts (Cx40 transcripts A and B) that share a common second coding exon. In 2003, Groenewegen et al identified a common single nucleotide polymorphism (SNP) in the Cx40 promoter A (rs35594137, −44G→A) that was later associated with decreased Cx40 expression, and a group of researchers suggested that this SNP might affect Cx40 levels and influence the risk of AF. The results of the study by Wirka et al support this hypothesis, as they found that the −44G→A SNP is strongly associated with Cx40 mRNA expression in atrial tissue and is significantly associated with early-onset lone AF. These findings provide new insights into the role of Cx40 in the development of AF and suggest that genetic alterations in the Cx40 gene region may play a role in the risk of AF.
promoter A activity in vitro, atrial vulnerability and AF.\textsuperscript{8} We hypothesized that additional common Cx40 polymorphisms may also affect Cx40 expression in human atrial tissues, and that these polymorphisms may predispose to AF.

We characterized a SNP, rs10465885, in the Cx40 promoter B that alters a putative TATA box element and strongly influences expression of Cx40 transcript B and total Cx40 expression in human atria. We further determined that the promoter B SNP acts in cis, influencing expression of the same-strand transcript. The promoter B SNP was significantly associated with early-onset lone AF in a case-control cohort study, and this finding was validated in a meta-analysis including 2 additional cohorts. In contrast to previous reports, we found that the promoter A SNP was not significantly associated with Cx40 expression in vivo, nor was it associated with the AF phenotype.

Methods

Study Subjects

Atrial appendage tissues were obtained from 61 patients undergoing coronary artery bypass grafting, valve surgery, or the Maze surgical AF ablation procedure. Most patients (85.2\%) from whom tissues were obtained had a history of AF, and 49.2\% of patients had AF at the time of surgery. Supplemental Table 1 (online Data Supplement) provides a detailed description of this cohort.

Three AF case-control cohorts were used to analyze the association of the Cx40 promoter A and B SNPs with the lone AF phenotype. All case subjects had lone AF, defined as AF in the absence of structural heart disease, and all subjects were of European ancestry. In addition, subjects were deemed to have early-onset lone AF if their age at AF diagnosis was $\leq 60$ years.

Subjects from the Cleveland Clinic Lone Atrial Fibrillation GeneBank (CCAF) (n=596, with 384 early-onset cases) were used as the discovery cohort along with 3010 population control subjects. Other methods are in the online Data Supplement.

Results

Haplotype Analysis of the GJA5 Gene

Seven contiguous SNPs from the Illumina Hap550 BeadChip region located within the GJA5 gene and proximal promoter region, plus rs11552588, a component of the previously described Cx40 promoter A haplotype,\textsuperscript{6} were studied. Genotypes at these 8 SNPs in the 61 atrial tissue samples were loaded into Haplovie v.4.1 (Broad Institute) to calculate their LD relationships (Figure 1A). From these data, it was determined that rs12408178 was in perfect LD ($r^2=1$) with rs11552588, the SNP at +71 in exon 1A, which is in perfect LD with the promoter A SNP rs35594137 ($-44G\rightarrow A$).\textsuperscript{7,8} This established rs12408178 as a perfect proxy for the promoter A SNP. In contrast, the promoter A SNP proxies were only moderately linked to the promoter B SNP ($r^2=0.33$).

The Promoter B SNP Alters A Putative TATA Box

The sequence of the proximal promoter of human Cx40 transcript B, along with the homologous sequence in other mammalian species, is shown in Figure 1B. The promoter B SNP resides at $-26$ bp relative to the start site of transcription, is denoted. The overlapping TATA box sequences are underlined, and graphical representations of their respective nucleotide frequency matrices are shown above the sequence.

Figure 1. Cx40 gene region and features of the promoter B SNP. A, The 2 alternative transcripts (A and B) of the Cx40 gene GJA5 are illustrated, both sharing a common second coding exon. The positions of the promoter B SNP (rs10465885), the promoter A SNP proxy rs12408178 and each SNP used as an allelic expression imbalance indicator (AEI) are denoted. Linkage disequilibrium relationships ($r^2$) between SNPs are illustrated. B, Sequence alignment of the promoter region near the promoter B SNP in various mammalian species. The promoter B SNP site (A/G), located at $-26$ relative to the start site of transcription, is denoted. The overlapping TATA box sequences are underlined, and graphical representations of their respective nucleotide frequency matrices are shown above the sequence.
Genotypes at each of the 8 SNPs shown in Figure 1A were measured by separate probes on Illumina Human Ref-8 v2 expression bead chips in human atrial tissue from 61 individuals. A were measured by real-time polymerase chain reaction in 31 atrial tissue samples. The promoter B SNP genotype was strongly associated with levels of Cx40 mRNA expression. Levels of Cx40 transcript B was performed on cDNA from 31 of the 61 atrial tissue samples. The promoter B SNP genotype was not significantly associated with levels of Cx40 mRNA expression. Levels of Cx40 transcript A were measured on Illumina Human Ref-8 v2 expression bead chips in human atrial tissue from 61 individuals.

The promoter B SNP genotype was significantly associated with total Cx40 mRNA levels (P=0.013). B, The promoter B SNP genotype was not significantly associated with levels of Cx40 transcript A (P=0.78). C, Relative levels of Cx40 transcript B were quantified by real-time polymerase chain reaction in 31 atrial tissue samples. The promoter B SNP genotype was significantly associated with Cx40 transcript B levels (P=0.0001).

The 5'-shifted TATA box sequence had a summed nucleotide frequency score of 0.855. The 5'-shifted TATA box sequence score is dependent on the genotype of the promoter B SNP. The “A” allele of this SNP results in a summed nucleotide frequency score of 0.776, whereas the “G” allele results in a poorer score of 0.620.

The Promoter B SNP Genotype Is Associated With Cx40 Transcript Levels In Vivo

cDNAs from 61 human atrial tissue samples were run on Illumina Human Ref-8 v2 expression bead chips, and the signals for total Cx40 and Cx40 transcript A were analyzed. Genotypes at each of the 8 SNPs shown in Figure 1A were tested for an association with Cx40 expression levels. Only the promoter B SNP (rs10465885) genotype was significantly associated with total Cx40 mRNA levels in the 61 atrial tissue samples (Figure 2A, P=0.013, additive genetic model), with the presence of 1 “G” allele at the promoter B SNP genotype associated with a 16% decrease in total Cx40 mRNA expression. As expected, the promoter B SNP was not significantly associated with Cx40 transcript A levels (Figure 2B, P=0.78, 1-way ANOVA). None of the other 7 SNPs within the Cx40 gene region were significantly associated with total Cx40 mRNA levels (1-way ANOVA, data not shown).

Because none of the probes on the microarray directly measured Cx40 transcript B, quantitative polymerase chain reaction of Cx40 transcript B was performed on cDNA from 31 of the 61 atrial tissue samples. The promoter B SNP genotype was strongly associated with Cx40 transcript B expression in a dose-dependent manner (Figure 2C, P<0.0001, 1-way ANOVA), with the presence of 1 “G” allele associated with 4.3-fold lower levels of Cx40 transcript B. In a linear regression analysis of the additive genetic model, the promoter B SNP genotype was associated with 49% of the variation in Cx40 transcript B expression in the 31 atrial tissue samples, with a corresponding quantitative log of the odds ratio (LOD) score of 4.85.

The Promoter A SNP Does Not Affect Cx40 Levels In Vivo

The genotype at the promoter A SNP proxy was not significantly associated with levels of Cx40 transcript A in the microarray assay of the 61 tissue samples (Figure 3, P=0.55, additive genetic model). Similarly, quantitative polymerase chain reaction of Cx40 transcript A in the atrial tissue samples failed to show any association of the promoter A SNP proxy with Cx40 transcript A expression (P=0.66, additive genetic model, Supplemental Figure 1). In addition, the promoter A SNP proxy was not significantly associated with levels of total Cx40 mRNA (P=0.82, additive genetic model, data not shown).

The Promoter B SNP Acts in Cis to Affect Cx40 Transcript B Expression In Vivo

Because the promoter B SNP alters the sequence of one of an overlapping pair of TATA box elements, we hypothesized that the effect of the promoter B SNP on Cx40 transcript B expression occurs in cis through modulation of transcription. Therefore, we tested for allelic expression imbalance, comparing Cx40 transcript B levels produced from each allele in atrial tissue samples from subjects heterozygous for both the promoter B SNP and an indicator SNP within Cx40 transcript B. In all of these atrial tissue samples, the amount of mRNA produced from the minor G allele at the promoter B SNP was strongly and significantly decreased compared with the amount produced from the major A allele, with a mean decrease of 3.3-fold and a range of 2.2- to 6.3-fold (Figure 4A).

Representative sequence tracings from both cDNA and gDNA in the region surrounding the indicator SNP rs1043806 are shown in Figure 4C, illustrating the differing allelic ratios. As a control, 1 sample homozygous at the
promoter B SNP but heterozygous for an indicator SNP was assayed, and was found to have no significant allelic expression imbalance (Figure 4A, first bar marked as (−)).

The Promoter A SNP Is Incompletely Associated With Allelic Expression Imbalance
We measured allelic expression imbalance of Cx40 transcript A in atrial tissue samples from subjects heterozygous for the promoter A SNP and the indicator SNP rs11552588 in exon 1 of transcript A. Most samples exhibited small but significant allelic expression imbalance favoring the −44A allele, with a mean increase in Cx40 transcript A expression of 1.3-fold (Figure 4B). However, significant allelic expression imbalance was not observed in 2 samples, and another 2 samples exhibited significant allelic expression imbalance in the opposite direction. This indicates that the promoter A SNP may be in partial LD with the native Cx40 promoter B SNP but heterozygous at an indicator SNP was included as a negative control (−). B, Thirty of the promoter A SNP heterozygotes demonstrated weak allelic expression imbalance of Cx40 transcript B favoring the “A” allele. One sample homozygous for the promoter B SNP but heterozygous at an indicator SNP was included as a negative control (−).

The Promoter B SNP Affects Promoter Activity in Transfected HL-1 Murine Cardiomyocytes
To exclude the influence of other potential genetic variations in LD with the promoter B SNP on Cx40 transcript B expression, we used a promoter-driven reporter gene assay in cultured cells. HL-1 murine atrial cardiomyocytes were transiently transfected with a promoterless firefly luciferase reporter construct (negative control) or constructs driven by the native Cx40 transcript B promoter containing either the A or G allele at the promoter B SNP site. Each transfection included a plasmid expressing β-galactosidase as a transfection efficiency control. HL-1 cells transfected with the reporter construct containing the G allele at the promoter B SNP site exhibited 2.4-fold lower luciferase activity compared with cells transfected with the construct containing the A allele (Figure 5, P<0.0001, Student t test).

The Promoter B SNP Is Significantly Associated With AF
The promoter A SNP proxy (rs12408178) and the promoter B SNP (rs10465885) were tested for association with the early-onset lone AF phenotype (age of diagnosis <60 years) using 384 CCAF cases and 3010 healthy population control subjects in a logistic regression analysis under a log-additive genetic model. The promoter B SNP was significantly associated with early-onset lone AF status with an odds ratio (OR) of 1.18 (95% confidence interval, 1.00 to 1.39; P=0.046; Figure 6), with the underexpressed “G” allele overrepresented in AF cases. In contrast, the promoter A SNP proxy was not significantly associated with early-onset lone AF (OR=1.11, P=0.26, Figure 6).

Two additional case-control cohorts were used to confirm the findings from the discovery CCAF cohort (Figure 6). The promoter B SNP had an OR of 1.25 for AF in 45 early-onset AF cases and 7300 non-AF control subjects from the ARIC cohort, but because of the small number of cases this was not statistically significant (P=0.31). The promoter A SNP proxy was not significantly associated with early-onset AF in the ARIC cohort (OR=1.06, P=0.81). The promoter B SNP had an OR of 1.09 for lone AF in 334 early-onset AF cases from the MGH cohort and 1067 healthy control subjects from the FHS cohort, which was in the same direction as the other 2 studies, but not statistically significant (P=0.42). The promoter A SNP proxy was not significantly associated with early-onset AF in this group (OR=0.98, P=0.90).
A meta-analysis including all 3 cohorts confirmed the finding of the discovery cohort, revealing a significant association between the promoter B SNP and early-onset lone AF (OR, 1.16; 95% confidence interval, 1.02 to 1.31; P=0.022; Figure 6). A similar meta-analysis failed to find a significant association between the promoter A SNP proxy and early-onset lone AF (OR, 1.07; P=0.37; Figure 6).

The meta-analyses of the early-onset lone AF case-control studies, including all 3 cohorts, had 80% power to detect a relative risk of 1.16 for the minor allele of the promoter B SNP and a relative risk of 1.18 for the minor allele of the promoter A SNP in the additive genetic model. None of the meta-analyses displayed any significant heterogeneity.

Neither the promoter A nor the promoter B SNP was significantly associated with the lone AF phenotype in any of the cohorts when subjects of all ages were included in their respective AF case cohorts (Supplemental Table 3), demonstrating the expected stronger genetic effects in younger cases.

**Discussion**

We used mRNA expression data and analysis of mRNA allelic expression imbalance in human atrial tissue to characterize a novel regulatory polymorphism in the Cx40 gene GJA5. This polymorphism, rs10465885, is located in the promoter region of Cx40 transcript B, where it alters the configuration of 1 of 2 partially overlapping TATA box elements. The strong correlation between the promoter B SNP and Cx40 mRNA levels and its consistent and profound effect on Cx40 transcript B allelic mRNA expression ratios demonstrate that this polymorphism is strongly associated in cis with transcription of Cx40 transcript B in vivo. Its location in a TATA box element and the results of the promoter-luciferase reporter gene assay provide strong evidence that the promoter B SNP directly affects transcription of Cx40 transcript B.

Comparison of the promoter region altered by the promoter B SNP with a nucleotide frequency matrix of the human TATA box sequence revealed 2 possible overlapping TATA box sequences (Figure 1B). The similarity of the 5'-shifted TATA box to the consensus TATA sequence is substantially improved by the G→A transition at the promoter B SNP site. The base affected by the promoter B SNP is not strictly conserved in mammals and, interestingly, the “non-A” allele appears to be present in the sequences of at least 5 other mammalian species. In addition, the chimpanzee genome at this site contains the G allele, suggesting that the A allele could be a gain-of-function mutation unique to humans.

Cx40<sup>−−</sup>/ knockout mice exhibit atrial ectopy and arrhythmias on surface ECG, and rare dominant-negative defects in the Cx40 coding sequence are associated with AF in humans. However, most SNPs associated with complex traits identified by genome-wide association studies to date are not located within gene coding regions or open reading frames and are instead expected to affect gene expression. Therefore, we tested for an association between Cx40 promoter polymorphisms and AF.

We found that the newly characterized promoter B SNP was significantly associated with early-onset lone AF in our discovery cohort. Although the association of the promoter B SNP with AF did not reach statistical significance in the 2 individual replication cohorts, these analyses were based on smaller sample sizes and were not adequately powered to detect the effect size observed in the discovery cohort. However, the effect of the promoter B SNP observed in each of the replication cohorts was consistent with the effect observed in the discovery cohort with increased risk associated with the G allele. Thus, a meta-analysis including all 3 cohorts confirmed the significant association between the promoter B SNP and early-onset lone AF. In addition, the G allele resulting in diminished Cx40 expression was overrepresented in the AF cases, suggesting that lower Cx40 levels may also be associated with AF. This is consistent with previous observations<sup>2–4</sup> that have suggested that impaired Cx40 expression or function predisposes to atrial conduction abnormalities and AF. The observed odds ratio of 1.16 per risk allele in the meta-analysis is modest and is expected under natural selection given the high minor allele frequency (0.47) of this SNP. However, the presence of 2 risk alleles is associated with an odds ratio of 1.35, which underscores the potential importance of this SNP in predicting AF risk.

Because each of the 2 replication cohorts did not achieve statistical significance on their own, how strong is the evidence supporting the association of rs10465885 with early-onset lone AF? First, all 3 studies had odds ratios in the same direction, showing increased risk with the underexpressed G allele. Second, the combined meta-analysis with the greatest power was more highly significant than the discovery cohort, demonstrating that as the sample size
increased the association grew stronger. Third, other genetic association studies with weak effects (e.g., odds ratio of \( \approx 1.2 \)) are often not replicated in the individual validation cohorts, but when combined, the effects are highly significant. For example, the recently identified SNP associated with abdominal aortic aneurysm was assessed in 8 separate follow-up cohorts, and 6 of these failed to find a significant association.\(^{15}\) However, combining these 8 cohorts together yielded a significant association.\(^{15}\) A prior genome-wide association study meta-analysis for lone AF in subjects 65 years of age or younger detected and validated 2 genome-wide significant loci but did not find a genome-wide significant association for the Cx40 gene.\(^{16}\) However, many real associations are present that fail to meet the strict standard of genome-wide significance, which is very stringent because of the correction for hundreds of thousands of SNPs tested. For example, genome-wide significant loci for height, which is a highly heritable trait, account for only 5% of the variance in height in studies with tens of thousands of subjects, but 45% of the variance can explained by considering all SNPs simultaneously in a smaller cohort.\(^{17}\) Thus, many more SNPs add to the heritability of height that cannot be identified with certainty using current statistical methods. We conclude that the evidence on hand supports the association of the Cx40-B promoter SNP rs10465885 with early-onset lone AF, although we cannot definitively rule out that this finding is a false-positive.

Interestingly, the promoter B SNP was not significantly associated with lone AF when subjects of all ages were included in the analyses. The prevalence of conditions that predispose to AF such as hypertension, coronary artery disease, valve disease, and heart failure increases with age. Therefore, although these analyses were performed in subjects with lone AF, the inclusion of older subjects (with AF caused by other “nongenetic” factors) in the AF case cohort may dilute the effect of a potential genetic variant. We confirmed the lack of association of the promoter B SNP with lone AF in subjects of all ages in a recent meta-analysis of 5 case-control cohorts comprising 1335 lone AF cases and 12,844 unaffected subjects, which overlaps the ARIC and MGH/FHS cohorts of the current study.\(^{18}\) In this study, rs10465885 was not significantly associated with AF (OR, 1.064; 95% confidence interval, 0.93 to 1.22; \( P = 0.64 \)). Age-stratified analysis was not performed in this mega-analysis.

A common SNP in the promoter of Cx40 transcript A, first identified cosegregating with an SCN5A mutation in familial atrial standstill,\(^{6}\) has also been associated with atrial vulnerability and AF in a small sample of 30 individuals.\(^{7}\) The association with AF was later replicated in a Taiwanese population of 173 AF patients and 232 control subjects,\(^{7,8}\) yielding an odds ratio of 1.54 for the minor allele of the promoter A SNP in the additive genetic model. Promoter-luciferase assays in cell culture showed varying (\(<20\%\) to 65\%\)) reduction in the activity of the promoter containing the minor allele of the promoter A SNP. A further study suggested that the minor allele of this promoter SNP negatively affected binding of GATA4, a major cardiac-specific transcription factor, to its binding site in the Cx40 transcript A promoter.\(^{8}\) We sought to extend these in vitro observations by studying the effect of this SNP on Cx40 mRNA expression levels in vivo in human atrial tissues. However, we were unable to detect a significant effect of the promoter A SNP on Cx40 transcript A or total Cx40 mRNA expression. Similarly, analysis of allelic expression imbalance in individuals heterozygous at the promoter A SNP revealed minimal and inconsistent allelic expression imbalance of Cx40 transcript A. For a candidate polymorphism to be considered a true cis-acting regulatory polymorphism, allelic expression imbalance must be consistent. Thus, our data argues against a true cis-acting regulatory role for the promoter A SNP.

In contrast to previous reports,\(^{7,8}\) the promoter A SNP was not significantly associated with AF in any of the case-control cohorts, or in the combined meta-analysis. The discordance of our findings with the Taiwanese study could in part be due to race-specific genetic effects or differences in the baseline patient characteristics. However, our meta-analysis had a markedly greater sample size, and was sufficiently powered to detect a relative risk of 1.18 for the minor allele of the promoter A SNP. Our study demonstrated that there was no association between the promoter A SNP and lone AF with an effect size of 1.18 or greater in populations of European ancestry.

In conclusion, we determined that a previously described Cx40 promoter A polymorphism, rs35594137 (\( -44G\rightarrow A \)), was not associated with Cx40 mRNA levels, allelic expression imbalance in human atrial tissue, or the lone AF phenotype. However, we identified a different SNP, rs10465885, as a common Cx40 promoter B variant with a direct cis effect on Cx40 gene expression. Importantly, this promoter B SNP was also significantly associated with early-onset lone AF. Additional studies are needed to more fully characterize the effects of this SNP in atrial tissue and its link to AF pathophysiology. However, this study represents an important step toward understanding the genetic determinants of Cx40 expression, as well as the potential role of variation in Cx40 expression as a determinant of AF risk.

**Acknowledgments**

We thank Dr William Claycomb (LSU Health Science Center) for his gift of HL-1 cells. We also thank the staff and participants of the ARIC study for their important contributions.

**Sources of Funding**

This work was supported by National Institutes of Health (NIH) grant HL090620 to Drs Chung, Barnard, Smith, and Van Wagoner. It was also supported in part by NIH/NCCR, CTSA 1UL-R0024989, Cleveland, OH. RobertWirka was a Howard Hughes Medical Institute Medical Research Training Fellow. The Atherosclerosis Risk in Communities Study is carried out as a collaborative study supported by National Heart, Lung, and Blood Institute contracts N01-HC-55015, N01-HC-55016, N01-HC-55018, N01-HC-55019, N01-HC-55020, N01-HC-55021, N01-HC-55022, R01HL087641, R01HL59367, and R01HL086694; National Human Genome Research Institute contract U01HG004402; and NIH contract HHSN268200625226C. Infrastructure was partly supported by grant No. UL1RR025005, a component of the NIH and NIH Roadmap for Medical Research. Dr Alonso was partially supported by NIH grant 1RC1HL099452 and American Heart Association grant 09SDG2280087. Dr Lubitz was supported by a training grant in the Epidemiology of Cardiovascular Disease from the NIH (T32HL007575). This work was also supported by grants from the NIH: HL092577 to Drs Ellinor and Benjamin, RC1-HL101056 to Dr Benjamin, and DA027021 to Dr Ellinor.
Disclosures
None.

References
A Common Connexin-40 Gene Promoter Variant Affects Connexin-40 Expression in Human Atria and Is Associated With Atrial Fibrillation


_Circ Arrhythm Electrophysiol_. 2011;4:87-93; originally published online November 13, 2010; doi: 10.1161/CIRCEP.110.959726

_Circulation: Arrhythmia and Electrophysiology_ is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2010 American Heart Association, Inc. All rights reserved.
Print ISSN: 1941-3149. Online ISSN: 1941-3084

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://circep.ahajournals.org/content/4/1/87

Data Supplement (unedited) at:
http://circep.ahajournals.org/content/suppl/2010/11/13/CIRCEP.110.959726.DC1

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in _Circulation: Arrhythmia and Electrophysiology_ can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to _Circulation: Arrhythmia and Electrophysiology_ is online at:
http://circep.ahajournals.org//subscriptions/
A common connexin-40 gene promoter variant affects connexin-40 expression in human atria and is associated with atria fibrillation


Methods

Study Subjects

Cleveland Clinic Atrial Tissue Bank. Atrial appendage tissues were obtained from 61 patients undergoing coronary artery bypass grafting, valve surgery or the Maze surgical AF ablation procedure. Supplemental Table 1 describes the clinical characteristics of these subjects.

Cleveland Clinic Lone Atrial Fibrillation GeneBank (CCAF). The discovery lone AF case cohort consisted of individuals from CCAF, which enrolled subjects >18 years of age with a history of AF and no significant coronary artery disease (≤50% coronary artery stenosis by coronary angiography, a normal stress test or no presumed coronary artery disease by clinical criteria; stress testing or coronary angiography was required for inclusion in males >age 50 years or females >age 55 years,) and normal left-ventricular function (left ventricular ejection fraction≥50%). Subjects were excluded from the study if they had significant valvular disease (>2+ mitral regurgitation, >2+ tricuspid regurgitation, >2+ aortic insufficiency, >2+ pulmonic insufficiency or any mitral,
tricuspid, aortic or pulmonic stenosis), coronary artery disease (>50% coronary artery stenosis, prior myocardial infarction, prior percutaneous coronary intervention, or coronary artery bypass surgery), a history of current congenital heart disease (except isolated patent foramen ovale), or a history of or current structural heart disease.

**Illumina Controls.** Healthy population subjects were identified from the Illumina iControl Database. Search criteria for the iControl Database were (1) Caucasian subjects of any age and any gender; (2) genotyping platform= Hap550 or Hap610. These criteria yielded 3180 subjects from studies 64, 65, 66 and 67 ([www.illumina.com](http://www.illumina.com)) carried out in New York, NY and Philadelphia, PA. Two of these studies contained only adults and two studies contained only children; thus, the age distribution of the population controls was bimodal, with one peak at a mean age of 50.2 years and another peak at a mean age of 9.2 years.

**Atherosclerosis Risk in Communities (ARIC) study.** Cases of early-onset AF were identified prospectively from the ARIC cohort. Cases were subjects aged less than 66 years with AF documented by electrocardiography, or AF documented on hospital discharge or death certificate in the absence of clinical evidence of structural heart disease. Referent subjects from the ARIC study had no history of AF at baseline or follow up, and no history of MI, heart failure or valve disease at baseline.

**The Massachusetts General Hospital Atrial Fibrillation Study (MGH)** prospectively enrolled subjects with early-onset AF defined as an age of onset less than 66 years and AF documented by electrocardiography. Subjects with structural heart disease, hyperthyroidism, myocardial infarction (MI) or heart failure were excluded from the MGH cohort.
Framingham Heart Study (FHS). 3-5 Referent subjects for use with the MGH cohort were selected from the Original and Offspring cohorts of the FHS, and included unrelated subjects 18-74 years of age and no history of AF at blood draw or in follow up, and no history of myocardial infarction, heart failure or valve disease at baseline.

Preparation of genomic DNA, total RNA and cDNA from human atrial tissue and blood

Atrial genomic DNA (gDNA) and total RNA were prepared from 61 atrial tissue samples using Qiagen kits. cDNA was produced with the iScript Select cDNA Synthesis Kit (BioRad) using an oligo(dT)20 primer. Buffy coat gDNA was prepared from blood samples of 500 subjects in the LAFGB using the MasterPure DNA Purification Kit for Blood Version II (Epicentre Biotechnologies).

Genotyping of human atrial tissue genomic DNA

Atrial gDNA was genotyped for rs11552588, a perfect proxy for the promoter A SNP6,7 by the TaqMan (Applied Biosystems) method in 61 human atrial tissue samples. Genotype data for the promoter B SNP rs10465885 and 6 additional SNPs, shown in Figure 1a, were obtained from Illumina HumanHap 550 BeadChip data on the same samples.

Genotyping of human blood DNA

Genotypes for the promoter B SNP and another perfect proxy for the promoter A SNP (rs12408178, r²=1) were obtained from genome wide SNP genotyping arrays. Genotyping platforms for each case-control cohort are outlined in Supplemental Table 2.
Because the Affymetrix 5.0 array used for the FHS subjects did not contain the promoter A or B SNPs, imputed values were used in the analyses of the MGH/FHS cohort. Imputation was performed using the program Mach with the CEU Hapmap Phase 2 panel. The imputed SNPs used in the analyses had good imputation quality scores (for example the promoter B SNP had an observed/expected variance of 1.019), which was expected due to the existence of good proxies on the Affymetrix 5.0 array. No SNPs deviated from Hardy Weinberg equilibrium using a p-value cutoff of 0.005.

Measurement of Cx40 mRNA levels

cDNAs from the 61 human atrial tissue samples were analyzed for total Cx40 and transcript A expression on Illumina Human Ref-8 v2 Expression Beads Chips. The expression data was log2 transformed and quantile normalized. Standard quality control and probe filtering were performed. 31 samples with sufficient RNA remaining were subjected to quantitative real-time polymerase chain reaction (qPCR) for measurement of Cx40 transcript A and B levels. PCR primers for specific amplification of Cx40 transcript B were upper primer 5’-GGAGAACACAGACAGGCAGAG-3’ and lower primer 5’-CCAGCACGAGCATAACGGAATA-3’. PCR primers for specific amplification of Cx40 transcript A were upper primer 5’-AAAAAGCGTGGGCAGTTGGAG-3’ and lower primer 5’-CCAGCACGAGCATAACGGAATA-3’. The VIC-labeled TaqMan probe for both Cx40 transcripts was 5’-AGGAAGCTCCAATCGC-3’. A commercially designed FAM-labeled TaqMan assay was used to measure cardiac muscle alpha-actin (ACTC1) cDNA as an internal control. Multiplex PCR reactions were performed in an iCycler IQ real-time PCR instrument (BioRad) using the following thermal protocol: 95°C for 9
minutes, then 40 cycles of: 95°C for 15 seconds, 60°C for 1 minute, and 72°C for 45 seconds (image capture step). Raw data for Cx40 transcript A or B and ACTC1 reactions were extracted as threshold cycle numbers. ACTC1-adjusted Cx40 transcript A or B threshold cycle values were calculated relative to the median threshold cycle value, and were used to calculate relative Cx40 transcript A and Cx40 transcript B expression levels using the $2^{-\Delta\Delta Ct}$ method. The association of the promoter B SNP (rs10465885) genotype with Cx40 transcript B levels was confirmed through a LOD score analysis using the residual sum of squares (rss) method, with the equation LOD=$(n/2) \times \log_{10}(\text{rss}_{0}/\text{rss}_{\text{model}})$, where ‘n’ refers to the number of samples.

**Quantitative sequencing for analysis of allelic expression imbalance**

**Sample selection** - Genotypes at the 8 SNPs within the GJA5 gene region (shown in Fig. 1A) were determined in 61 atrial tissue samples, and unphased genotypes at each SNP from the 61 individuals were assembled into 122 phased haplotypes using PHASE v.2.1.1. For each tissue sample used in the allelic expression imbalance assay, the phased haplotypes were analyzed to determine the allelic relationship between the Cx40 promoter SNPs and the indicator SNPs within the GJA5 coding region.

**Cx40 transcript B** - Cx40 transcript B cDNA-specific PCR amplification primers were 5’-AGACAGGCAGAGGATTACAACACA-3’ and 5’-GGGCCTCCATAGCTGTCATCA-3’. gDNA specific PCR amplification primers were 5’-GGGGCAAGAGCAGAATCCATAT-3’ and 5’-GGGCCTCCATAGCTGTCATCA-3’.

**Cx40 transcript A** - Cx40 transcript A cDNA-specific PCR amplification primers were 5’-GGTGGAAGAGGAACAACTGA-3’ and 5’-CATACGGAATATGAAGAGGACA-
3’. gDNA-specific PCR amplification primers were 5’-
GGTGGAAAGAGGAAACAACGTGA-3’ and 5’-GAAATAGCGGGAGGGTAGAAG-3’.
Both gDNA and cDNA for each transcript were PCR amplified for 40 cycles and PCR
products were purified using the QIAquick PCR Purification Kit according to the
manufacturer’s instructions (Qiagen). The purified products were sequenced using
internal primers (5’-TGGCAGTCAGCAAAGGAAGTAG-3’ for Cx40 transcript B
products, 5’-GAAGCACGCGAGTGTGAAGA-3’ for Cx40 transcript A products) on
an Applied Biosystems 3730xl DNA analyzer. Raw .ABI sequence files were analyzed
using PeakPicker2, a program that calculates allelic expression ratios based on peak
heights at the indicator SNP site in both cDNA and gDNA PCR products. PeakPicker2
was then used to normalize each allelic cDNA expression ratio to the average of the
allelic gDNA ratios, assuming a gDNA allele ratio of 1:1. This resulted in a normalized
allelic mRNA expression ratio for each sample.

Luciferase reporter gene assay
A promoter fragment ranging from -765bp to +94bp relative to the first base pair of exon
1B was PCR amplified from genomic DNA of patients homozygous for either the A or G
allele at the Cx40 transcript B promoter SNP, using the upper primer 5’-
GGAGCTCTTGACCCATCTTTCCCATATAA-3’ and lower primer 5’-
GCTCGAGTTGCTGCTTTGTTGTGATCTCCTC-3’. HindIII and XhoI sites are
underlined in the upper and lower primers, respectively. Using the HindIII and XhoI
restriction enzymes, the A and G allele-containing promoter fragments were then cloned
into a promoter-less pXPII vector directly upstream of the firefly luciferase gene. The
sequence of both the A and G allele-containing vectors was verified by direct sequencing to ensure that the Cx40-B promoter SNP site was the only site of sequence discordance between the two vectors. HL-1 murine atrial cardiomyocytes were cultured at 37°C and 5% CO₂ in Claycomb medium (Sigma) supplemented with 10% fetal bovine serum, 0.1 mM norepinephrine, 2 mM L-glutamine and 100 U/mL each penicillin and streptomycin. HL-1 cells were cultured in antibiotic-free medium in 6-well plates for 24 hours prior to transfection with 5 µL Lipofectamine LTX (Invitrogen). Cells were transiently co-transfected with 1µg of either the A or G allele-containing promoter-luciferase plasmid and 1µg of the β-galactosidase-containing plasmid pCH110. Additional cells were co-transfected with the empty pXPII vector and pCH110 as a background control. The promoter driven luciferase constructs were transfected in sextuplicate (one well of the A allele got contaminated and was excluded from the analysis). 48 hours later, cell lysates were assayed in duplicate for luciferase and β-galactosidase activity with the Dual-Light reporter gene assay system (Tropix), and the mean background-subtracted (wells with no transfection) luciferase/β-galactosidase levels were calculated for each well.

**Statistical Analyses**

We used both one-way ANOVA and linear regression with the additive genotype model to correlate SNP genotypes with Cx40 expression values. The linear regression model included covariates to correct for age, atrial rhythm at the time of surgery, and expression assay batch effects. The promoter A and B SNPs were tested for association with early onset lone AF (including only subjects with age at AF diagnosis ≤60 years) and lone AF (including subjects of all ages) in the case-control cohorts using logistic regression in a
log-additive genetic model. Cohorts from each center were analyzed individually and in a meta-analysis. For the analysis of the CCAF cohort, sex was included as a covariate and the EIGENSTRAT method of population stratification adjustment was used to help account for the different origins of the CCAF case and control sets. The meta-analyses used a fixed-effects model with invariance-variance weights of the genotype regression effects (log odds ratios). Power calculations used a minimum power of 0.80 and an alpha of 0.05. Statistical analyses were carried out using JMP 7, GraphPad PRISM, and R version 2.11 software. Meta-analyses were performed using the ‘metagen’ function of the R package Meta, version 1.6-0. Power calculations were performed using the htPower.cc function of the DGCgenetics software package, version 1.2.
References


Supplemental Table 1. Characteristics of the 61 Atrial Tissue Donors

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Count (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Age (SD)</td>
<td>62.5 (11.2)</td>
</tr>
<tr>
<td>Sex</td>
<td>N (%)</td>
</tr>
<tr>
<td>Male</td>
<td>37 (60.7)</td>
</tr>
<tr>
<td>Female</td>
<td>24 (39.3)</td>
</tr>
<tr>
<td>Race</td>
<td></td>
</tr>
<tr>
<td>European American</td>
<td>56 (91.8)</td>
</tr>
<tr>
<td>African American</td>
<td>3 (4.9)</td>
</tr>
<tr>
<td>Other</td>
<td>2 (3.2)</td>
</tr>
<tr>
<td>AF Burden</td>
<td></td>
</tr>
<tr>
<td>No AF</td>
<td>9 (14.8)</td>
</tr>
<tr>
<td>Paroxysmal</td>
<td>9 (14.8)</td>
</tr>
<tr>
<td>Persistent</td>
<td>8 (13.1)</td>
</tr>
<tr>
<td>Permanent</td>
<td>31 (50.8)</td>
</tr>
<tr>
<td>Unknown</td>
<td>4 (6.5)</td>
</tr>
<tr>
<td>AF Type</td>
<td></td>
</tr>
<tr>
<td>Lone AF</td>
<td>5 (9.6)</td>
</tr>
<tr>
<td>Mixed AF</td>
<td>47 (90.4)</td>
</tr>
<tr>
<td>AF on PreOp ECG</td>
<td>30 (49.2)</td>
</tr>
<tr>
<td>Hypertension</td>
<td>28 (45.9)</td>
</tr>
<tr>
<td>Diabetes</td>
<td>7 (11.5)</td>
</tr>
<tr>
<td>CHF</td>
<td>24 (39.3)</td>
</tr>
<tr>
<td>MI</td>
<td>13 (21.3)</td>
</tr>
<tr>
<td>Surgery Type</td>
<td></td>
</tr>
<tr>
<td>Valve</td>
<td></td>
</tr>
<tr>
<td>Repair/Replace</td>
<td>45 (73.8)</td>
</tr>
<tr>
<td>Maze</td>
<td>51 (83.6)</td>
</tr>
<tr>
<td>CABG</td>
<td>13 (21.3)</td>
</tr>
<tr>
<td>Medications</td>
<td></td>
</tr>
<tr>
<td>Digitalis</td>
<td>23 (37.7)</td>
</tr>
<tr>
<td>Beta Blocker</td>
<td>39 (63.9)</td>
</tr>
<tr>
<td>Ca Channel Blocker</td>
<td>13 (21.3)</td>
</tr>
<tr>
<td>Class 1a AA</td>
<td>1 (1.6)</td>
</tr>
<tr>
<td>Class 1b AA</td>
<td>4 (6.6)</td>
</tr>
<tr>
<td>Class 1c AA</td>
<td>4 (6.6)</td>
</tr>
<tr>
<td>Class 3 AA</td>
<td>7 (11.5)</td>
</tr>
<tr>
<td>Amiodarone</td>
<td>4 (6.6)</td>
</tr>
<tr>
<td>ACE Inhibitor</td>
<td>36 (59.0)</td>
</tr>
<tr>
<td>Statin</td>
<td>22 (36.1)</td>
</tr>
</tbody>
</table>

SD, standard deviation; MI, myocardial infarction; CHF, congestive heart failure; AA, Antiarrhythmic
**Supplemental Table 2. Characteristics of Lone AF Case-Control Study Cohorts**

<table>
<thead>
<tr>
<th>Study</th>
<th>Genotyping Platform</th>
<th>Cleveland Clinic</th>
<th>ARIC</th>
<th>Massachusetts General Hospital &amp; Framingham Heart Study</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Cases</td>
<td>Controls</td>
<td>Cases</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CCAF iControlDB Study</td>
<td>iControlDB Study</td>
<td>ARIC</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Illumina 550/610</td>
<td>Illumina 550/610</td>
<td>Illumina 550/610</td>
</tr>
<tr>
<td><strong>All Ages Lone AF</strong></td>
<td>Number</td>
<td>596</td>
<td>1423</td>
<td>1587</td>
</tr>
<tr>
<td></td>
<td>Age at enrollment, mean ± SD</td>
<td>59.1 ± 10.7</td>
<td>50.2 ± 11.1</td>
<td>9.2 ± 5.5†</td>
</tr>
<tr>
<td></td>
<td>Sex, Male, n (%)</td>
<td>450 (75.5)</td>
<td>314 (22.1)</td>
<td>821 (51.7)</td>
</tr>
<tr>
<td><strong>Early Onset Lone AF</strong></td>
<td>Number</td>
<td>384</td>
<td>1423</td>
<td>1587</td>
</tr>
<tr>
<td></td>
<td>Age at AF diagnosis (cases) or enrollment (controls), mean ± SD</td>
<td>47.3 ± 9.9</td>
<td>50.2 ± 11.1</td>
<td>9.2 ± 5.5†</td>
</tr>
<tr>
<td></td>
<td>Sex, Male, n (%)</td>
<td>306 (79.7)</td>
<td>314 (22.1)</td>
<td>821 (51.7)</td>
</tr>
</tbody>
</table>

*Platform included the 50K Human Gene Focused Panel
† Two studies from the Illumina database contained pediatric patients
Supplemental Table 3. Association of promoter A and B SNPs with lone AF, all ages, additive genetic model

<table>
<thead>
<tr>
<th>Study</th>
<th>AF (n)</th>
<th>No AF (n)</th>
<th>Promoter A SNP</th>
<th>Promoter B SNP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>MAF (%)</td>
<td>OR (95%CI)</td>
</tr>
<tr>
<td>CCF</td>
<td>596</td>
<td>3010</td>
<td>24.1/22.7</td>
<td>1.01 (0.86-1.18)</td>
</tr>
<tr>
<td>MGH</td>
<td>375</td>
<td>1101</td>
<td>22.4/22.1</td>
<td>1.01 (0.79-1.29)</td>
</tr>
<tr>
<td>ARIC</td>
<td>119</td>
<td>7395</td>
<td>21.8/23.7</td>
<td>0.90 (0.65-1.22)</td>
</tr>
<tr>
<td>Meta</td>
<td>1090</td>
<td>11506</td>
<td>23.3/23.3</td>
<td>0.99 (0.88-1.12)</td>
</tr>
</tbody>
</table>
Supplemental Figure 1. The promoter A genotype is not associated with Cx40 transcript A expression (P=0.66). Relative levels of Cx40 transcript A were quantified by real-time PCR in 31 atrial tissue samples.