Published in final edited form as:

Can J Cardiol. 2012 ; 28(2): 158–159. doi:10.1016/j.cjca.2012.01.010.

There Is Power in Numbers—Even/Especially in Genomic Medicine

David R. Van Wagoner, PhD and Mina K. Chung, MD

Cleveland Clinic Departments of Molecular Cardiology and Cardiovascular Medicine, Cleveland, Ohio, USA

Atrial fibrillation (AF) is a challenging global health problem that has proved difficult to treat and thus far impossible to prevent, due in part to our lack of a clear understanding of the numerous underlying mechanisms. While advanced age is the most common risk factor for AF, family history is also associated with significantly increased risk of AF.¹ Heritability suggests a significant contribution of genetic influences to the incidence of AF, but very few monogenic causes of AF have been identified, and the number of individuals affected by monogenic forms of AF is small. Genome-wide association studies (GWASs) have the potential to provide unbiased insights into novel genes and pathways that contribute to genetically determined disease risk. The first GWAS for AF, published in 2007, identified 2 independent single-nucleotide polymorphisms (SNPs) on chromosome 4q25 (rs2200733 and rs10033464) that were significantly associated with risk of AF in a discovery cohort of 550 Icelandic individuals with a history of AF, compared with 4476 unaffected individuals from the same population.² The initial study internally replicated this observation in additional cohorts in Sweden, the United States, and Hong Kong. Since the original report, the association of rs2200733 with AF has been replicated in additional populations^{3,4} and has been associated with risk of stroke (likely due to silent AF),⁵ with recurrence of AF after radiofrequency ablation,⁶ and with occurrence of AF after coronary artery bypass graft surgery.^{7,8} This is a robust association. The 4q25 locus is near a gene, *PITX2*, with more than plausible biologic relevance. *PITX2* has been shown to be critical to development of pulmonary veins, where AF may initiate.

Individuals with lone AF with the risk allele of rs2200733 have been shown to have a prolonged PR interval as an intermediate phenotype.³ While these studies document a robust association of rs2200733 with AF pathophysiology, fine mapping studies in larger numbers of individuals suggest that there are numerous additional SNPs in the 4q25 region that are also associated with AF risk.⁹ A GWAS of PR interval (in 28,517 individuals) has similarly identified SNPs in 5 additional loci that are associated with both PR interval and risk of AF (in 5741 AF cases).¹⁰ In addition to these loci, other large AF-focused GWASs have identified additional SNPs on chromosome 16q22 near the gene *ZFHX3*^{11,12} and on chromosome 1q21 in the *KCNN3* gene.¹

In this issue, Olesen and colleagues report on a study in which they sought to assess the relation of SNPs in the 8 loci previously identified in AF- and AF-related GWASs (*PITX2*, *SCN5A*, *SCN10A*, *CAVI*, *NKX2-5*, *SOX5*, *ZFHX3*, and *KCNN3*) to individuals with a diagnosis of lone AF before the age of 40 years.¹³ The primary strength of this study is the

population chosen for evaluation. In principle, detection of genetic influences on atrial arrhythmogenesis should be most straightforward in individuals with AF in the absence of confounding influences of coronary artery disease, valve disease, hypertension, heart failure, and underlying metabolic problems. In this respect, the study is similar to that of Ellinor and colleagues, who assessed the SNPs associated with lone AF in 1335 people with AF compared with 12,844 unaffected individuals.¹

Sample size is a critical determinant of the ability of GWASs to detect novel and reproducible associations. Larger sample sizes facilitate the identification of disease associations with less common SNPs. The corollary of this principle is that smaller studies have limited power to reliably detect genetic associations. In contrast to the genome-wide genotyping (and HapMap imputation) in the above lone-AF study,¹ Olesen et al. replicated earlier genome-wide studies in their early-onset lone-AF population using TaqMan-based polymerase chain reaction to genotype 1 SNP per locus (8 SNPs) in 209 people with lone AF and compared the genotype distribution with that of 534 ethnically matched middle-aged individuals with no history of AF or other cardiovascular disease or stroke, but with other risk factors for AF. Olesen et al. found that, of the SNPs tested, only 3 were significantly associated with AF.¹³ These included loci on chromosomes 4, 7, and 12, located near the genes *PITX2*, *CAVI*, and *SOX5*. After correction for multiple testing, only the SNPs near *PITX2* and *SOX5* remained significant.

How do we interpret the results of the Olesen et al. study in the context of the current literature? This study may suggest that the loci most strongly associated with lone, early-onset AF are those near *PITX2* and *SOX5*, suggesting that transcription factors are critical targets that either promote AF initiation or create a substrate for persistent AF. What might we make of the lack of significant association of *CAVI* and the other targets previously identified in other GWASs? While it is possible that the other targets are related less directly to AF and more to the secondary risk factors for AF (hypertension, heart rate, valve disease, etc), statistical issues seem as (or more) likely to dominate. The power to detect significant associations is determined in part by the frequency of the minor allele, and to detect rare variants, large sample sizes are critical. It is for this reason that large consortia (Cohorts for Heart and Aging Research in Genomic Epidemiology [CHARGE], etc) have been formed. Meta-analysis of multiple cohorts provides a statistically robust mechanism to discover novel associations. Use of genome-wide SNP data means that the possibility of detecting rare variants is enhanced. It is notable that, whereas rs2200733 was the most significant SNP in the 4q25 region in the initial study, a nearby but distinct SNP in this region (rs6843082) was even more strongly associated with AF in a large lone-AF meta-analysis.¹

With a sample size of 209 patients and a minor allele frequency of 0.1, the odds of being homozygous for the minor allele is $0.1 \times 0.1 = 0.01$; one would expect only 21 individuals to be heterozygous for the disease allele and 2 individuals to be homozygous for the risk allele in this patient cohort. Such small numbers increase the risk of random noise in the analysis. Olesen et al. previously evaluated the association of rs2200733 with lone AF in 196 early-onset lone-AF patients.¹⁴ When compared with 176 nonaffected age-matched individuals, no association of rs2200733 (4q25) with AF was detected.¹⁴ Sample size matters. Studies with < 1000 subjects have very limited power for SNP discovery and limited power for validation. Use of single SNPs per locus minimizes the penalty for multiple testing but also increases the odds that relevant associations within a locus may be missed. Limited statistical power remains the primary limitation of the Olesen study. It is clear that additional genetic determinants of AF remain to be discovered. In collaborative analyses currently under way, a focus on increasing sample size will help to discover additional loci and genes that contribute to AF risk. Consortia that create large datasets have been formed and will be required to achieve this goal.

The focus of current research is shifting toward bioinformatic studies that seek to identify the pathways in which the disease-related genes participate and the mechanisms by which these pathways contribute to the onset or persistence of AF. Studies that assess the relation of AF-associated SNPs to atrial tissue gene expression may help in this regard. Because many or most of the identified SNPs are located in intergenic or intronic regions, simple nearby gene knockout or overexpression studies may be insufficient to elucidate the mechanistic impact of AF-associated polymorphisms. Cellular studies (using atrial cell lines and stem-cell or fibroblast-derived cells transformed to an atrial phenotype) may be useful in evaluating the impact of SNPs on atrial cellular physiology and electrophysiology. Assays for promoter or repressor activity can help to assess the impact on expression of nearby genes. With the post-GWAS era nearly upon us, functional studies that make sense of identified genetic associations will be critical to translate genetic discoveries into useful knowledge that can affect clinical practice.¹⁵

Acknowledgments

Funding Sources Funding was provided by the Fondation Leducq European North American Atrial Fibrillation Alliance (CVD-07-03) and the National Institutes of Health/National Heart, Lung, and Blood Institute (R01-HL90620).

References

1. Ellinor PT, Lunetta KL, Glazer NL, et al. Common variants in KCNN3 are associated with lone atrial fibrillation. *Nat Genet.* 2010; 42:240–4. [PubMed: 20173747]
2. Gudbjartsson DF, Arnar DO, Helgadóttir A, et al. Variants conferring risk of atrial fibrillation on chromosome 4q25. *Nature.* 2007; 448:353–7. [PubMed: 17603472]
3. Goodloe AH, Herron KJ, Olson TM. Uncovering an intermediate phenotype associated with rs2200733 at 4q25 in lone atrial fibrillation. *Am J Cardiol.* 2011; 107:1802–5. [PubMed: 21481830]
4. Lee KT, Yeh HY, Tung CP, et al. Association of RS2200733 but not RS10033464 on 4q25 with atrial fibrillation based on the recessive model in a Taiwanese population. *Cardiology.* 2010; 116:151–6. [PubMed: 20606429]
5. Gretarsdóttir S, Thorleifsson G, Manolescu A, et al. Risk variants for atrial fibrillation on chromosome 4q25 associate with ischemic stroke. *Ann Neurol.* 2008; 64:402–9. [PubMed: 18991354]
6. Husser D, Adams V, Piorkowski C, Hindricks G, Bollmann A. Chromosome 4q25 variants and atrial fibrillation recurrence after catheter ablation. *J Am Coll Cardiol.* 2010; 55:747–53. [PubMed: 20170812]
7. Body SC, Collard CD, Shernan SK, et al. Variation in the 4q25 chromosomal locus predicts atrial fibrillation after coronary artery bypass graft surgery. *Circ Cardiovasc Genet.* 2009; 2:499–506. [PubMed: 20031626]
8. Virani SS, Brautbar A, Lee VV, et al. Usefulness of single nucleotide polymorphism in chromosome 4q25 to predict in-hospital and long-term development of atrial fibrillation and survival in patients undergoing coronary artery bypass grafting. *Am J Cardiol.* 2011; 107:1504–9. [PubMed: 21414601]
9. Lubitz SA, Sinner MF, Lunetta KL, et al. Independent susceptibility markers for atrial fibrillation on chromosome 4q25. *Circulation.* 2010; 122:976–84. [PubMed: 20733104]
10. Pfeufer A, van Noord C, Marcianti KD, et al. Genome-wide association study of PR interval. *Nat Genet.* 2010; 42:153–9. [PubMed: 20062060]
11. Gudbjartsson DF, Holm H, Gretarsdóttir S, et al. A sequence variant in ZFHX3 on 16q22 associates with atrial fibrillation and ischemic stroke. *Nat Genet.* 2009; 41:876–8. [PubMed: 19597491]
12. Benjamin EJ, Rice KM, Arking DE, et al. Variants in ZFHX3 are associated with atrial fibrillation in individuals of European ancestry. *Nat Genet.* 2009; 41:879–81. [PubMed: 19597492]

13. Olesen MS, Holst AG, Jabbari J, et al. Genetic loci on chromosomes 4q25, 7p31, and 12p12 are associated with onset of lone atrial fibrillation before the age of 40 years. *Can J Cardiol*. 2012; 28:191–5. [PubMed: 22336519]
14. Henningsen KM, Olesen MS, Haunsoe S, Svendsen JH. Association of rs2200733 at 4q25 with early onset of lone atrial fibrillation in young patients. *Scand Cardiovasc J*. 2011; 45:324–6. [PubMed: 21793630]
15. Nattel S. From guidelines to bench: implications of unresolved clinical issues for basic investigations of atrial fibrillation mechanisms. *Can J Cardiol*. 2011; 27:19–26. [PubMed: 21329858]