Taming Rare Variation With Known Biology in Long QT Syndrome

Marco V. Perez, MD; Euan A. Ashley, MRCP, DPhil

The tools necessary for linking novel genes to Mendelian diseases are now at the fingertips of most clinical scientists, thanks to the plummeting cost of whole genome and whole exome sequencing (WES).[1,2] Such easy access to sequence data would have been unimaginable 18 years ago, when *HERG*[3] and *SCN5A*[4] were first linked to long QT syndrome (LQTS) in 2 large families using time-consuming and expensive sequencing techniques. Since the completion of the human genome project[5] and development of next-generation sequencing technologies, there have been a handful of spectacular success stories, in which WES has led to timely diagnoses and life-saving therapies for previously uncharacterized Mendelian disorders.[6] Because WES became readily available, there was anticipation that the molecular nature of most Mendelian disease would be uncovered in very little time, and that a surge of publications would emerge with links between novel genes and Mendelian diseases that have for years remained mysteries. The 1000 Genomes Project[7] and the Exome Sequencing Project[8] showed us, however, that the number of rare variants in each individual's genome is much higher than we had anticipated. Each person carries thousands of rare, potentially unique, possibly disease-causing variants. Wading through this sea of rare and novel genomic variation has proven a formidable challenge. Thus, further establishing the burden of proof that a given variant causes disease is complex. Functional, rare variation in a plausible gene cosegregating with disease and absent in controls represents a minimum bar, but each of these features exists as a continuous function. Furthermore, even when causality seems clear by these metrics, the proportion of variance in disease expression explained needs independent assessment (oligogenic contribution to disease is inevitable). In this issue of *Circulation Cardiovascular Genetics*, Boczek et al[9] demonstrate, via WES and filtering for rare variation in genes with a predetermined high likelihood of association with disease, that *CACNA1C* is linked to LQTS in a family without Timothy Syndrome (nonsyndromic LQTS).

Article see p 279

Different techniques have been developed to narrow down or filter a large number of rare variants by excluding those that are unlikely to cause disease.[10] Excluding noncoding variants by focusing on exonic sequences, discounting synonymous variants, and eliminating variants cataloged in publicly available databases are standard filters that allowed Boczek et al[9] to narrow down the list of possible disease-causing variants to 263. The next filtering step was to require that the causative variant be present in affected family members and absent in unaffected members. Although family trios typically consist of mother/father/child, the decision to sequence an affected aunt instead of the mother allowed for a more powerful filter because only one fourth of the genome will be shared between the index and the aunt. In autosomal dominant Mendelian inheritance patterns, inclusion of the father for purposes of segregation does not add significantly to the filtering power; however, this does help eliminate variants that might have been erroneously called because of systematic sequencing errors. Although the expected number of variants left at this point was 65 (approximately one fourth of 263), there were actually 110 variants observed.

Additional exome sequencing of all phenotyped family members would have added power to this filtering step. However, accurate filtering assumes perfect phenotyping, which can be challenging in LQTS because the prolonged QT phenotype is variably expressed. Besides adding to the cost of WES of additional family members, inaccurate phenotyping can result in erroneously filtering out causative variants.

Sanger sequencing and functional validation is burdensome for 100 variants, so investigators have looked to predefined lists of candidate genes. Similar approaches, combining WES with functional genomic filters, have been used to prioritize variants and identify novel genes linked to endometrial cancer.[11] This method has the advantage of narrowing a large list of candidate variants substantially. The disadvantage of this approach is that if the causative variant resides in a gene from a completely novel pathway with no previous suspicion of involvement, then it would likely be filtered out. *CACNA1C* was fortunately included in the list of 1629 genes from the LQTS interactome,[12] which was used in the next filtering step. It is notable, however, that this list was generated from a protein–protein interaction network based on the set of genes that cause LQT1-12, which includes *CACNA1C*. Therefore, in this study, *CACNA1C* could be considered as a candidate gene.

The use of the LQTS interactome list allowed for a helpful filtering step, which reduced the number of candidate variants from 110 to 8. In fact, the only variant that cosegregated completely in all 12 family members, possibly most convincing evidence of causality, was the Pro857Arg *CACNA1C* variant. Complete cosegregation was achieved by classifying

borderline-affected family members phenotype-positive, a reasonable strategy in the investigation of causality in a condition with variable expressivity of the phenotype.

Of the 8 variants identified after the final filters were applied, all 3 bioinformatics/systems biology algorithms (Endeavour,[13] SUSPECTS,[14] and ToppGene[15]) ranked *CACNA1C* highest. In their ranking algorithms, 2 of these programs use resources such as semantic analysis of existing literature, in which *CACNA1C* is frequently included among the list of genes that are already linked to LQTS. The 3 LQTS-disease–causing genes (*KCNQ1*, *KCNH2*, and *SCN5A*) that were used as the algorithms' training genes are well characterized with extensive ontologic labels. These factors likely helped *CACNA1C* rise to the top of the ranked lists. Creating systems biology–based filters with sets of genes that are not as well characterized in Mendelian disorders for which several causative genes do not already exist may not result in such accurate predictions.

Impressively, these investigators went beyond human genetic evidence to explore the functional effect of this variant on the expressed channel. Functional analysis of the Pro857Arg *CACNA1C* variant in HEK293 cells demonstrated gain-of-function of the *CACNA1C* channel, likely mediated through impaired degradation of $Ca_v1.2$. These findings in addition to the cosegregation observation support the hypothesis that this variant in *CACNA1C* is LQTS-causing and suggest a novel mechanism by which *CACNA1C* trafficking disorders can lead to LQTS.

To further demonstrate that *CACNA1C* causes nonsyndromic LQTS and may play a more prominent role in LQT, Boczek et al[9] performed mutational analysis in 102 previously undiagnosed families with LQT syndrome and found 3 additional families with rare variation in *CACNA1C*. Limited cosegregation was evident in 1 of the families whose *CACNA1C* variant affected the same amino acid as the index family. Presence of rare variation alone is not sufficient to demonstrate pathogenicity.[16] However, the fact that 2 of these variants were located in the same domain (proline-, glutamic acid-, serine-, and threonine-rich doman [PEST]) responsible for degradation signaling is suggestive of common pathogenicity. Population variation reveals[17] rare missense variation in *CACNA1C* coding regions in 38 out of ≈4300 (0.9%) unrelated Caucasian individuals presumably free of Mendelian disease. This rate is lower than the rate of 3 out of 102 (2.9%) patients with nonsyndromic LQTS reported in the mutational analysis by Boczek et al[9] (χ^2 $P=0.031$). This further supports the hypothesis that *CACNA1C* is linked to nonsyndromic LQTS.

Although *CACNA1C* is known to cause prolongation of the QT interval as 1 of several clinical manifestations of Timothy Syndrome,[18] this is the first nonsyndromic LQT family whose disease has been linked to a *CACNA1C* mutation. Although variants in genes linked to LQT4-15 are responsible for <5% of LQTS,[19] this and other recent studies suggest that *CACNA1C* may play a more prominent role in nonsyndromic LQTS than was originally believed. Lieve et al[20] reported that, in patients with suspected LQTS who underwent sequencing of the genes responsible for LQTS1-10, rates of rare variation of unclear significance were highest in *CACNA1C*. This could be either because rates of benign rare variation are higher in this gene, or these variants were determined of unclear significance because they have not been as frequently documented in the literature.

We expect that as the cost of sequencing continues to plummet, WES in many more families with uncharacterized Mendelian diseases will ultimately be performed. However, finding large families with clear positive and negative phenotypes for conclusive cosegregation analysis will remain challenging. Because of this, filters that use systems biology algorithms, such as those presented by Boczek et al,[9] will become an important step in helping narrow down large lists of plausible variants. There will need to be a balance between using unbiased whole-genome approaches that are more likely to uncover genes in novel disease pathways and a candidate gene-based approach where effective filters link genes in well-established disease pathways.

Despite technical challenges in clinical grade coverage from current capture technologies and algorithmic challenges in calling small insertion/deletions and structural variation, exome sequencing continues to offer an exciting alternative for families where traditional panel–based sequencing has failed to identify a convincingly causal variant of disease. Boczek et al[9] have demonstrated with cosegregation, functional, and mutational analyses that *CACNA1C* is linked to nonsyndromic LQTS. Because commercially available LQT panels now routinely include sequencing of most exons in *CACNA1C*, not just exon 8 that is responsible for Timothy Syndrome, we anticipate a growing number of reports of *CACNA1C* variants linked to nonsyndromic LQTS.

Sources of Funding

Disclosures

E.A. Ashley is a cofounder of Personalis. Dr Perez reports no conflicts.

References

1. Ashley EA, Butte AJ, Wheeler MT, Chen R, Klein TE, Dewey FE, et al. Clinical assessment incorporating a personal genome. *Lancet*. 2010;375:1525–1535.
2. Dewey FE, Pan S, Wheeler MT, Quake SR, Ashley EA. DNA sequencing: clinical applications of new DNA sequencing technologies. *Circulation*. 2012;125:931–944.
3. Curran ME, Splawski I, Timothy KW, Vincent GM, Green ED, Keating MT. A molecular basis for cardiac arrhythmia: HERG mutations cause long QT syndrome. *Cell*. 1995;80:795–803.
4. Wang Q, Shen J, Splawski I, Atkinson D, Li Z, Robinson JL, et al. SCN5A mutations associated with an inherited cardiac arrhythmia, long QT syndrome. *Cell*. 1995;80:805–811.
5. Consortium IHGS. Finishing the euchromatic sequence of the human genome. *Nature*. 2004;431:931–945.
6. Worthey EA, Mayer AN, Syverson GD, Helbling D, Bonacci BB, Decker B, et al. Making a definitive diagnosis: successful clinical application of whole exome sequencing in a child with intractable inflammatory bowel disease. *Genet Med*. 2011;13:255–262.
7. Abecasis GR, Altshuler D, Auton A, Brooks LD, Durbin RM, Gibbs RA, et al. A map of human genome variation from population-scale sequencing. *Nature*. 2010;467:1061–1073.
8. Tennessen JA, Bigham AW, O'Connor TD, Fu W, Kenny EE, Gravel S, et al; Broad GO; Seattle GO; NHLBI Exome Sequencing Project. Evolution and functional impact of rare coding variation from deep sequencing of human exomes. *Science*. 2012;337:64–69.

9. Boczek NJ, Best JM, Tester DJ, Giudicessi JR, Middha S, Evans JM, et al. Exome sequencing and systems biology converge to identify novel mutations in the L-type calcium channel, *CACNA1C*, linked to autosomal dominant long QT syndrome. *Circ Cardiovasc Genet*. 2013;6:279–288.

10. Dewey FE, Chen R, Cordero SP, Ormond KE, Caleshu C, Karczewski KJ, et al. Phased whole-genome genetic risk in a family quartet using a major allele reference sequence. *PLoS Genet*. 2011;7:e1002280.

11. Liang H, Cheung LW, Li J, Ju Z, Yu S, Stemke-Hale K, et al. Whole-exome sequencing combined with functional genomics reveals novel candidate driver cancer genes in endometrial cancer. *Genome Res*. 2012;22:2120–2129.

12. Berger SI, Ma'ayan A, Iyengar R. Systems pharmacology of arrhythmias. *Sci Signal*. 2010;3:ra30.

13. Aerts S, Lambrechts D, Maity S, Van Loo P, Coessens B, De Smet F, et al. Gene prioritization through genomic data fusion. *Nat Biotechnol*. 2006;24:537–544.

14. Adie EA, Adams RR, Evans KL, Porteous DJ, Pickard BS. Speeding disease gene discovery by sequence based candidate prioritization. *BMC Bioinformatics*. 2005;6:55.

15. Chen J, Xu H, Aronow BJ, Jegga AG. Improved human disease candidate gene prioritization using mouse phenotype. *BMC Bioinformatics*. 2007;8:392.

16. Pan S, Caleshu CA, Dunn KE, Foti MJ, Moran MK, Soyinka O, et al. Cardiac structural and sarcomere genes associated with cardiomyopathy exhibit marked intolerance of genetic variation. *Circ Cardiovasc Genet*. 2012;5:602–610.

17. Exome Variant Server, NHLBI Go Exome Sequencing Project (ESP). Seattle, WA. Http://evs.Gs.Washington.Edu/evs/. Accessed May 2015.

18. Splawski I, Timothy KW, Sharpe LM, Decher N, Kumar P, Bloise R, et al. Ca(V)1.2 calcium channel dysfunction causes a multisystem disorder including arrhythmia and autism. *Cell*. 2004;119:19–31.

19. Ackerman MJ, Priori SG, Willems S, Berul C, Brugada R, Calkins H, et al. HRS/EHRA expert consensus statement on the state of genetic testing for the channelopathies and cardiomyopathies this document was developed as a partnership between the Heart Rhythm Society (HRS) and the European Heart Rhythm Association (EHRA). *Heart Rhythm*. 2011;8:1308–1339.

20. Lieve KV, Williams L, Daly A, Richard G, Bale S, Macaya D, et al. Results of genetic testing in 855 consecutive unrelated patients referred for long qt syndrome in a clinical laboratory. *Genet Test Mol Biomarkers*. 2013.

Taming Rare Variation With Known Biology in Long QT Syndrome

Marco V. Perez and Euan A. Ashley

The online version of this article, along with updated information and services, is located on the
World Wide Web at:
http://circgenetics.ahajournals.org/content/6/3/227