The Missing \textit{LINC} for Genetic Cardiovascular Disease?

Megan J. Puckelwartz, PhD

As of 2014, cardiovascular disease (CVD) was listed as the underlying cause for 1 of every 3 deaths in the United States.\textsuperscript{1} Many forms of CVD have a strong genetic component, and the number of causal genes elucidated greatly depends on the type of CVD under study. Understanding the molecular basis of CVD can guide medical decisions of the practitioner, patient, and family members. Currently, gene panels are used in the clinical setting to screen for disease-causing variation. Panels vary greatly in size depending on the number of genetic causes of the disease in question. There is often a great deal of phenotypic overlap between the inherited CVDs. To accommodate these variable phenotypes, many testing laboratories now use large panels that include >75 genes. Another route that is more commonly being used is the use of whole exome sequencing and whole genome sequencing to identify causal and contributory variants.\textsuperscript{2,3} Whole exome sequencing or whole genome sequencing can begin with focus on rare variants in genes most commonly associated with disease, and, if no attractive candidates are identified, the search is broadened to include genes not currently associated with disease. Such broad searches can lead to the identification of new disease genes; a great benefit in the research context, but a burden in the clinical-setting making disease risk assessment more difficult. It is also an ever-evolving target to determine the amount of validation necessary to deem variants pathogenic and therefore reportable, especially when variants are unique to individuals or to individual families.\textsuperscript{4} It can be especially troubling when the disease in question involves sudden death where there is significant risk for surviving family members. It is imperative that new, potential CVD genes continue to be validated using a variety of methods, including family studies, animal- and cell-based models, and in vitro techniques. Much is known about the role of sarcomere proteins and ion channels in inherited CVD. However, although proteins of the nuclear envelope (NE) are implicated in CVD, only a handful of the ≈80 proteins of the NE have been exhaustively screened, and more work needs to be done to understand the risk of variants in genes of the NE.

The opinions expressed in this article are not necessarily those of the editors or of the American Heart Association.

From the Department of Pharmacology, Center for Genetic Medicine, Northwestern University Feinberg School of Medicine, Chicago, IL.

Correspondence to Megan P. Puckelwartz, PhD, Center for Genetic Medicine, Northwestern University Feinberg School of Medicine, 303 E Superior St, Lurie Research Bldg, 7-109, Chicago, IL 60611. E-mail m.puckelwartz@northwestern.edu

\textit{Circ Cardiovasc Genet} is available at http://circgenetics.ahajournals.org

\textbf{DOI: 10.1161/CIRCGENETICS.117.001793.}

\textbf{See Article by Haskell et al}

In this issue, Haskell et al\textsuperscript{5} present whole exome sequencing data on 55 patients with suspected monogenic CVD, including cardiomyopathy, long QT, thoracic aortic aneurysm, and mitral valve prolapse. Variants were initially filtered using candidate gene panels specific to each subject’s phenotype. Analysis of established disease genes identified pathogenic or likely pathogenic variants in 12 subjects and variants of unknown significance in an additional 19 of 55 subjects resulting in identification of a potentially pathogenic variant in a known candidate gene in 56.4% of the cohort. This left >40% of the cohort negative for diagnostic genetic variants. To identify novel candidate disease genes in the remainder of the cohort, the authors examined variants across the entire captured exome. They examined rare (minor allele frequency <0.001), predicted loss-of-function variants (pLOF), including nonsense, splice site, and frameshifting variants. Four potential candidate variants were identified, all in genes of the NE. The NE separates the functions of the cytoplasm and nucleus and is composed of the inner and outer nuclear membrane separated by the perinuclear space and spanned by the nuclear pore complex (NPC). The nuclear lamina underlies the inner nuclear membrane and is composed of the A- and B-type lamins and their associated proteins. In addition to the mechanical support provided by the nuclear lamina, its other functions involve DNA replication, cell division, chromatin localization, and gene expression. Mutations in genes encoding proteins of the nuclear lamina cause several forms of striated muscle disease. Mutations in the A-type lamins, encoded by the \textit{LMNA} gene, are responsible for an entire class of diseases, termed the laminopathies, and affecting a diverse group of tissues, including striated muscle, nerves, and adipose tissue. The diseases of striated muscle include Emery–Dreifuss Muscular Dystrophy, a progressive muscle disease that can have a severe cardiac component and dilated cardiomyopathy (DCM) with and without conduction system disease.

Of the 4 pLOF variants identified by Haskell et al,\textsuperscript{5} one was identified in a proband with a severe DCM that required transplant at age 15. The proband’s father also had DCM; neither had a history of skeletal muscle weakness. The pLOF variant identified was in the \textit{SYNE1} gene. \textit{SYNE1} encodes the giant nesprin-1 protein, a vital component of the LINC complex. The LINC complex (Linker of Nucleoskeleton and Cytoskeleton) spans the NE and is a multiprotein complex composed of the laminas, nesprins, SUN proteins, and a host of other nuclear membrane components (Figure). The LINC complex is thought to transmit mechanical signals from the cytoplasm to the nucleus and plays a role in diverse functions, including moving and shaping the nucleus, signal transduction, and chromatin localization.\textsuperscript{4} Mutations in LINC complex components are found in both cardiomyopathy and skeletal...
muscle myopathy in both mice and humans.\textsuperscript{6} For example, mice lacking the C terminus of nesprin-1 have progressive muscle wasting and develop cardiomyopathy with conduction system disease and lack a functional LINC complex.\textsuperscript{7,8} In addition to the study by Haskell et al,\textsuperscript{1} other studies have also identified potentially deleterious variants in nesprin-1 in patients with either Emery–Dreifuss Muscular Dystrophy or isolated DCM.\textsuperscript{9} Cells derived from humans with nesprin-1 mutations show aberrant nuclear shape and mislocalization of LINC complex components. Despite evidence of nesprin-1’s role in cardiac disease, the SYNE1 gene is not routinely screened for variation, likely because of its large size. The open reading frame of SYNE1 spans over 27 kb and through alternative transcriptional initiation and splicing, results in many nesprin-1 isoforms that vary greatly in size, with the largest isoforms \textgtr 8700 amino acids in length.

The ExAC database (http://exac.broadinstitute.org) aggregates variant information on \textasciitilde 60000 exomes.\textsuperscript{10} These data provide insight not only into the rarity of particular variants, but also to the tolerance of particular genes to mutation, with the idea that genes with greater than the number of expected variants under neutrality are less likely to be deleterious.\textsuperscript{11} Aggregate data from the ExAC project indicates that SYNE1 is tolerant to variation, especially pLOF variants; therefore, it follows that pLOF variation may not cause disease. However, this interpretation may be flawed as evidenced by another giant gene, TTN. TTN’s coding region spans over 100 kb and produces a 33000 amino acid protein. Like SYNE1, TTN also produces multiple isoforms, only some of which are important in the heart. Herman et al\textsuperscript{12} performed large-scale sequencing of TTN in a cardiomyopathy cohort with \textgtr 590 subjects with either hypertrophic or dilated cardiomyopathy. They found rare TTN truncating variants in \textasciitilde 25% of the DCM cohort, but only 1% in the hypertrophic cardiomyopathy cohort. Three percent of the control group carried truncating variants in the TTN gene indicating that rare, pLOF variants are rather common in the giant gene. Although SYNE1 is only approximately a quarter of the size of TTN, it is still a massive protein and may also carry a higher than expected pLOF burden in the general population and still significantly contribute to CVD.

Further large-scale studies in patients with CVD are in order for this massive gene.

Haskell et al\textsuperscript{1} also identified 3 pLOF variants in NPC genes. The NPC is one of the largest protein complexes in the cell and consists of multiple copies of \textasciitilde 30 NUP (nucleoporins). Besides serving as conduits connecting the cytoplasm and nucleus, evidence suggests that NPCs have an active role in regulating gene expression and organizing chromatin, as well as DNA repair and RNA processing.\textsuperscript{13} There is evidence that the NPC associates with the SUN proteins, vital components of the LINC complex, and that the NPC relies on this interaction for assembly during interphase and for uniform distribution of NPCs across the nuclear surface.\textsuperscript{14,15} Although individual NUPs are not giant, the number of NUPs and their involvement in the sizable NPC makes studying their role in human disease challenging. To date, only 1 NUP gene, NUP155, was linked to inherited cardiac disease. A homozygous missense mutation was identified in a family with atrial fibrillation and sudden death in early childhood.\textsuperscript{16} Homozygous Nup155 null mice die before embryonic day 8.5, but heterozygous mice show a distinct atrial fibrillation phenotype. Haskell et al\textsuperscript{3} identified pLOF variants in NUP37, NUP43, and NUP188 in their subjects with CVD. These NUPs are all scaffolding NUPs, similar to NUP155 which may indicate a shared function. A recent study determined that both over- and underexpression of nup188 has a role in heart looping in Xenopus.\textsuperscript{17} Haskell et al\textsuperscript{3} used morpholino knockdown to test the function of NUP37 and NUP43 in zebrafish. Zebrafish lacking nup37 exhibited severe pericardial edema, a manifestation of heart failure, and possible arrhythmias. The zebrafish lacking nup43 also displayed pericardial edema and bradycardia. Although these studies do not conclusively prove that variation in the NUPs contribute to human CVD, they are suggestive of a role and indicate that further study is warranted.

In searching for new genes responsible for the inherited CVDs, it seems reasonable to examine genes enriched for cardiac expression. The LINC complex belies this thinking, as many components are rather ubiquitously expressed in nearly every cell type. One hypothesis put forth to explain this phenomenon is that cells under high mechanical load, such as heart and skeletal muscle, are uniquely sensitive to...
perturbation of the LINC complex. Further work needs to be done to examine the potentially deleterious variation in these genes and their role as a missing link to explain CVD.

Sources of Funding
Dr Puckelwartz is supported by National Institutes of Health grants HL128075 and U01HL131914.

Disclosures
None.

References

KEY WORDS: Editorials ◼ atrial fibrillation ◼ cardiovascular diseases ◼ genetics ◼ heart failure ◼ nuclear envelope ◼ nuclear pore
The Missing LINC for Genetic Cardiovascular Disease?
Megan J. Puckelwartz

_Circ Cardiovasc Genet._ 2017;10:
doi: 10.1161/CIRCGENETICS.117.001793
_Circulation: Cardiovascular Genetics_ is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2017 American Heart Association, Inc. All rights reserved.
Print ISSN: 1942-325X. Online ISSN: 1942-3268

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/10/3/e001793

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in _Circulation: Cardiovascular Genetics_ 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: Cardiovascular Genetics_ is online at:
http://circgenetics.ahajournals.org//subscriptions/