Editors

Modifying Mendel Redux
Unbiased Approaches Can Find Modifiers

Kim L. McBride, MD; Stephanie M. Ware, MD, PhD

Congenital heart disease (CHD) is one of the most common groups of birth defects, contributing to a major portion of mortality in early childhood and consuming large amounts of healthcare and family resources. They have a birth prevalence of 6 to 8/1000 live births, excluding late recognized defects, such as bicuspid aortic valve, which has a population frequency of 1% to 2%.1

See Article by Guo et al

Genetic epidemiology studies and reports of multiple recurrences of CHDs within families demonstrate a strong genetic component. Familial clustering of CHDs is particularly apparent when grouped by developmental mechanism.2 A recent large study using hundreds of these multiplex families confirmed the concept of grouping CHD by developmental mechanism, and supporting animal data suggest that these groupings are because of perturbations of genetic networks important in cardiogenesis.3 Indeed, of all risk factors for CHD, a family history of CHD has the highest relative risk, even over maternal diabetes mellitus or twinning.4 More formal segregation analyses have confirmed the strong genetic component, also noting that the inheritance pattern is likely complex and oligogenic.2

This evidence for the genetic basis of CHD spurred investigators to search for responsible loci and genes. A few early successes occurred using the traditional approach of linkage, identifying pathogenic variants in NKX2-5,5 NOTCH1,6 and GATA47 among multiplex families with CHDs. Unfortunately, further successes have been scarce, with the exception of CHDs occurring as part of a syndrome (such as the RASopathies). Genome-wide association studies for specific groups of CHD have added a few more loci but with limited replications in a second study.8,9 Copy number variant studies have identified novel genomic disorders in as many as 20% of syndromic cases, a few percent of nonsyndromic CHDs, including identification of individuals with multiple copy number variants.10-12 The advent of next-generation sequencing technology raised hopes that this bottleneck in gene discovery would be resolved. Early studies appeared exciting, with the identification of chromatin remodeling genes in CHD cases.13 Unfortunately, it seems at least some individuals may have had unrecognized or undefined syndromes that include CHDs rather than an isolated CHD. Subsequent studies have found de novo changes in only 1% to 2% of those cases,14 thus the cause remains elusive for isolated CHD. These studies also highlight the significant difficulty in case ascertainment, the need for careful (and longitudinal) phenotyping, and the challenges of variant interpretation.

Although the causes of syndromes that include CHD as part of the phenotype have been answered more successfully, an interesting question has arisen: Why does only a portion of individuals with the syndrome have a CHD? Conversely, individuals with specific syndromes offer a unique opportunity to study a population with susceptibility to CHDs. This topic was highlighted in our previous editorial on Modifying Mendel15 in which Li et al16 used a trisomy 21 (Down syndrome) mouse (Ts65dN) in an attempt to answer the question of CHD risk using this sensitized model. Their work implicated Creld1 and Gata4 in mouse, supported by data from human T21 subjects, providing one of the first studies on modification of a major susceptibility locus (in this case an extra chromosome 21) in the causation of CHD.

The article by Guo et al17 in this issue is among the first to use an unbiased approach to identifying a modifier locus for CHD in humans. Using genome-wide association studies in individuals with a deletion of 22q11.2 (22q11DS), the authors combined bioinformatics of the human locus with additional studies in mice to narrow down to a potential candidate gene. The 22q11DS is the most common genomic disorder that causes CHD. Individuals may have characteristic dysmorphic features, velopalatal abnormalities (cleft or insufficiency), hypocalcemia, immunodeficiency, and a variety of neurodevelopmental disorders. Most individuals have CHD related to second heart field developmental anomalies, including truncus arteriosus, interrupted aortic arch type B, and most commonly, tetralogy of Fallot.

Guo et al’s17 group was able to assemble an impressive cohort of ~1500 carefully phenotyped individuals with 22q11DS. Over 1200 samples were single nucleotide polymorphism genotyped on an Affymetrix 6.0 platform. The analysis was careful and thorough, including proper quality control and accounting for possible biases from population stratification. Comparing 22q11DS subjects with tetralogy of Fallot (n=326) to 22q11DS subjects with normal cardiac anatomy (n=566) in the cohort, they identified a locus on 5q14.3 with...
multiple single nucleotide polymorphisms (genotyped and imputed) exceeding the threshold for genome-wide significance. The highest associated single nucleotide polymorphism was rs12519770 \((P=2.98\times10^{-8})\) with an odds ratio of 1.69. This locus was further narrowed to an \(\approx100\) kb interval within an intron of \(GPRC7B\) using additional whole-genome sequencing of just under 400 22q11DS subjects. Bioinformatic analysis of chromatin conformation (Hi-C) showed that this region lay in a topological associated domain of 2.3 Mb size, containing 6 genes: \(GPR98\), \(MEF2C\), \(CETN3\), \(MBLAC2\), \(POLR3G\), and \(LYSMD3\). Expression studies in mouse were performed using in situ hybridization, and only \(Mef2c\) demonstrated expression in cardiac progenitor cells. \(MEF2C\) is a biologically plausible candidate gene because it is an important transcription factor during cardiogenesis, acting in a network that regulates second heart field development.

Several weaknesses in this study are important to note despite the huge effort by this group. Relatively speaking, this study is small by genome-wide association studies standards, thus it was powered to only be able to identify loci associated with large effect sizes. They did not have a second, replication cohort; however, assembling this size of initial cohort at all is a striking accomplishment. Finally, although the authors performed appropriate analyses to control for population stratification, inclusion of a variety of populations (European, Hispanic, American of African descent, Asian) does create some concern for false-positive results. Further work will also be required to identify whether the single nucleotide polymorphisms in the intron of \(GPRC7B\) are themselves functional or tag an as yet unidentified variant and to determine the mechanistic basis by which the functional variants modify or cause CHD in 22q11DS. Whether this study data can be generalized to other second heart field malformations not associated with 22q11DS will be interesting to see. Although further validation of these data will be important, the results provide initial data on a potential modifier that causes CHD in 22q11DS individuals.

The previous study using T21 as a susceptibility group and known mouse CHD genes provided proof of principle that variant burden threshold influences disease penetrance. This current study by Guo et al’s\(^{17}\) group advances the field in an important way by demonstrating that an unbiased approach in a human susceptibility population is able to identify a modifier locus. Other syndromes with CHD as part of the phenotype, such as Noonan syndrome and other RA-Sopathies, are additional candidates for this type of study.

Identifying major susceptibility loci, let alone modifiers, remains a challenge for isolated CHDs. Here, several approaches are promising. First, multiplex families are also a susceptibility population. Rates of identifying major susceptibility loci are much higher, with identification of pathogenic variants in 25\% to 30\% of families, compared with 1\% to 2\% of sporadic cases.\(^{14,18}\) Genome approaches in multiplex families also offer an unbiased screen for finding multiple loci. Second, brute force animal models, such as using mouse forward genetic screens (inducing mutation followed by high throughput phenotypic screening), can find not only major susceptibility loci but also digenic inheritance or modifier loci. This has already been successfully accomplished by Li et al.,\(^{19}\) who found 2 genes causing hypoplastic left heart syndrome in a forward mouse screen followed by identification of variants in these genes in human subjects with hypoplastic left heart syndrome.

Many difficulties lie ahead. Phenotyping remains a problem not only for anatomic features of a specific individual but also for the correct grouping of distinct, but related, malformations. Statistical modeling has yet to be successful given the small numbers of subjects with CHDs that can be collected. Even large groups, such as the Pediatric Cardiovascular Genetics Consortium, would be unable to collect the thousands of samples necessary to identify significant results in rare variant analyses by current statistical methods. Methods for multiplex families for narrowing the list of potential variants in oligogenic inheritance models are lacking. In addition, studying multiple variants in multiple genes in animal models is slow, laborious, and expensive. Newer techniques, such as use of CRISPR/Cas9 to more quickly create animal models, and use of fruit fly to rapidly and cheaply study variants in multiple loci via multiple crosses, may offer some hope.\(^{20}\)

The genetics of CHD seems to be more complex than previously thought. The field will need to move away from simple Mendelian genetic models to those with multiple loci, possibly a major locus with \(\geq1\) modifiers, or multiple loci acting together. This concept is not unique to the CHD genetics field. Other cardiovascular genetic diseases have well-established complex patterns of inheritance that are already altering how patients with these diseases are diagnosed and managed in clinical practice. Arrhythmogenic right ventricular cardiomyopathy, where up to 10\% of individuals have 2 disease-causing variants, and long-QT syndrome, where 2 pathogenic variants lead to more severe phenotype that includes deafness, are 2 examples.

Guo et al’s\(^{17}\) group shows that modifiers of Mendelian CHD traits can be found. These methods should become part of future study designs in CHD genetics.

Disclosures

None.

References

1. Hoffman JI, Kaplan S, Libethsson RR. Prevalence of congenital heart disease. Am Heart J. 2004;147:425–439. doi: 10.1016/j.ahj.2003.05.003.
2. McBride KL, Pignatelli R, Lewin M, Ho T, Fernbach S, Menesses A, et al. Inheritance analysis of congenital left ventricular outflow tract obstruction malformations: segregation, multiplex relative risk, and heritability. Am J Med Genet A. 2005;134A:180–186. doi: 10.1002/ajmg.a.30602.
3. Ellesøe SG, Workman CT, Bouvagnet P, Livardo CA, McBride KL, Hinton RB, et al. Familial co-occurrence of congenital heart defects follows distinct patterns. Eur Heart J. 2017. doi:10.1093/eurheartj/ehx314. https://academic.oup.com/eurheartj/article-lookup/doi/10.1093/eurheartj/ehx314. Accessed October 5, 2017.
4. Liu S, Joseph KS, Lisonkova S, Rouleau JL, Van den Hof M, Sauve R, et al; Canadian Perinatal Surveillance System (Public Health Agency of Canada). Association between maternal chronic conditions and congenital heart defects: a population-based cohort study. Circulation. 2013;128:583–589. doi: 10.1161/CIRCULATIONAHA.112.001054.
5. Schott JJ, Benson DW, Basson CT, Pease W, Silberbach GM, Moak JP, et al. Congenital heart disease caused by mutations in the transcription factor NKX2-5. Science. 1998;281:108–111.
6. Garg V, Muth AN, Ransom JF, Schluterman MK, Barnes R, King IN, et al. Mutations in NOTCH1 cause aortic valve disease. Nature. 2005;437:270–274. doi: 10.1038/nature03940.
7. Garg V, Kathiriya IS, Barnes R, Schluterman MK, King IN, Butler CA, et al. GATA4 mutations cause human congenital heart defects and reveal an interaction with TBX5. Nature. 2003;424:443–447. doi: 10.1038/ nature01827.
8. Cordell HJ, Töpf A, Mamasoula C, Postma AV, Bentham J, Zelenika D, et al. Genome-wide association study identifies loci on 12q24 and 13q32 associated with tetralogy of Fallot. *Hum Mol Genet*. 2013;22:1473–1481. doi: 10.1093/hmg/dds552.
9. Hanchard NA, Swaminathan S, Bucass K, Furthner D, Fernbach S, Azamian MS, et al. A genome-wide association study of congenital cardiovascular left-sided lesions shows association with a locus on chromosome 20. *Hum Mol Genet*. 2016;25:2331–2341. doi: 10.1093/hmg/ddw071.
10. Greenway SC, Pereira AC, Lin JC, DePalma SR, Israel SJ, Mesquita SM, et al. De novo copy number variants identify new genes and loci in isolated sporadic tetralogy of Fallot. *Nat Genet*. 2009;41:931–935. doi: 10.1038/ng.415.
11. Lander J, Ware SM. Copy number variation in congenital heart defects. *Current Genetic Medicine Reports*. 2014;2:168–178.
12. Hanchard NA, Umana LA, D’Alessandro L, Azamian M, Poopola M, Morris SA, et al. Assessment of large copy number variants in patients with apparently isolated congenital left-sided cardiac lesions reveals clinically relevant genomic events. *Am J Med Genet A*. 2017;173:2176–2188. doi: 10.1002/ajmg.a.38309.
13. Zaidi S, Choi M, Wakimoto H, Ma L, Jiang J, Overton JD, et al. De novo mutations in histone-modifying genes in congenital heart disease. *Nature*. 2013;498:220–223. doi: 10.1038/nature12141.
14. Sifrim A, Hitz MP, Wilson A, Breckpot J, Turki SH, Thienpont B, et al; INTERVAL Study; UK10K Consortium; Deciphering Developmental Disorders Study. Distinct genetic architectures for syndromic and non-syndromic congenital heart defects identified by exome sequencing. *Nat Genet*. 2016;48:1060–1065. doi: 10.1038/ng.3627.
15. McBride KL, Ware SM. Modifying Mendel: approaches for identification of susceptibility alleles for human cardiovascular malformations. *Circ Cardiovasc Genet*. 2012;5:274–276. doi: 10.1161/CIRCGENETICS.112.963579.
16. Li H, Cherry S, Klinedinst D, DeLeon V, Redig J, Reshey B, et al. Genetic modifiers predisposing to congenital heart disease in the sensitized Down syndrome population. *Circ Cardiovasc Genet*. 2012;5:301–308. doi: 10.1161/CIRCGENETICS.111.960872.
17. Guo T, Repetto GM, McDonald McGinn DM, Chung JH, Nomaru H, Campbell CL, et al. Genome-wide association study to find modifiers for tetralogy of Fallot in the 22q11.2 deletion syndrome identifies variants in the *GPR98* locus on 5q14.3. *Circ Cardiovasc Genet*. 2017;10:e001690. doi: 10.1161/CIRCGENETICS.116.001690.
18. LaHaye S, Corsmeier D, Basu M, Bowman JL, Fitzgerald-Butt S, Zender G, et al. Utilization of whole exome sequencing to identify causative mutations in familial congenital heart disease. *Circ Cardiovasc Genet*. 2016;9:320–329. doi: 10.1161/CIRCGENETICS.115.001324.
19. Liu X, Yagi H, Saeed S, Bais AS, Gabriel GC, Chen Z, et al. The complex genetics of hypoplastic left heart syndrome. *Nat Genet*. 2017;49:1152–1159. doi: 10.1038/ng.3870.
20. Zhu JY, Fu Y, Nettleton M, Richman A, Han Z. High throughput in vivo functional validation of candidate congenital heart disease genes in *Drosophila*. *Elife*. 2017;6:e22617. doi: 10.7554/eLife.22617.

Key Words: Editorials • genome-wide association study • genomics • heart defects, congenital • tetralogy of Fallot