Published in final edited form as:
Circ Cardiovasc Genet. 2012 June ; 5(3): . doi:10.1161/CIRCGENETICS.112.963215.

The Good SHP2 Association: A Porthole into the Genetics of Congenital Heart Disease

Bruce D. Gelb, MD[1] and **Christine E. Seidman, MD**[2]

[1]Child Health and Development Institute & Depts of Pediatrics and Genetics & Genomic Sciences, Mount Sinai School of Medicine, New York, NY

[2]Dept of Genetics, Harvard Medical School, Howard Hughes Medical Institute, and the Cardiovascular Division, Brigham and Women's Hospital, Boston, MA

Keywords

editorials; congenital; genetics; genome-wide analysis; genome-wide association study; tetralogy of Fallot

In this issue, Judith Goodship and colleagues report results from an association study for common genetic variants underlying tetralogy of Fallot (TOF).[1] By studying a limited number of single nucleotide polymorphisms (SNPs) tagging haplotypes at 22 candidate genes, they associated a single SNP at a chromosome 12 locus harboring *PTPN11*, which encodes the protein tyrosine phosphatase SHP2, and other genes with TOF. This SNP was calculated to have a per-allele risk of 1.34 and a population-attributable risk (PAR) of roughly 5%. If this study, undertaken with British Caucasian subjects, can be replicated in additional populations, it will represent a significant advance in our understanding of the genetics of congenital heart defects (CHD).

The genetic architecture of CHD has been the source of considerable debate and uncertainty. It has been known for some time that chromosomal defects and single-gene mutations can cause CHD, often in the context of a multisystem disease. For TOF, examples include trisomy 21 and *TBX5* point mutations underlying Down and Holt-Oram syndromes, respectively.[2] The discovery of 22q11 microdeletions, usually arising *de novo*, added significantly to our understanding of the genetic underpinnings of conotruncal forms of CHD as these account for substantial percentages of some cardiac lesions (e.g., 34% of truncus arteriosus and 16% of TOF).[3] Nonetheless, these known genetic causes of CHD are estimated to account for less than 20% of cases overall.

Epidemiologic studies of CHD have strongly pointed to genetic factors as the predominant cause, although environmental exposures are also relevant. Parental consanguinity significant increases the risk of CHD in offspring, likely due to the shared genetic background of unaffected parents.[4] In countries where consanguineous marriages are customary, the rates of consanguinity are 2–3 fold higher among couples with offspring with CHD than in the general population.[4,5] Moreover, a recent population-based study of CHD in Denmark, where a nation-wide medical registry enabled nearly complete ascertainment, revealed that the relative risk for any form of CHD in first-degree relatives was 3.2.[6] The

Correspondence to: Bruce D. Gelb, MD, Mount Sinai School of Medicine, One Gustave Levy Place, Box 1040, New York, NY 10029, Tel: 212-241-3302, Fax: 212-241-3310, bruce.gelb@mssm.edu.

Conflict of Interest Disclosures: None

recurrence risk of the same form of CHD varied by the lesion or lesion class; for conotruncal defects, of which TOF is one, that risk was 11.7. Finally, after exclusion of chromosomal defects, the population-associated risk given a positive family history of CHD was 4.2%.

In the 1960's, James Nora proposed a multifactorial model for CHD, in which polygenic inheritance combined with environmental factors would underlie cases not readily explained by aneuploidy, Mendelian genetic mutations or teratogens.[7] While the multifactorial model was (and is) broadly popular for complex genetic traits, it encountered substantial problems when applied to CHD. Specifically, most of the predictions that flow from polygenic inheritance proved untrue for nearly all forms of CHD, patent ductus arteriosus in full-term newborns being the singular exception. In the 1980's, the multifactorial model took a broadside hit from the groundbreaking work of Ruth Whittemore.[8] Studying women with CHD who had survived to adulthood and were having children, Whittemore observed a 16% CHD recurrence rate, often with the same heart lesion, in those offspring, which was wholly inconsistent with polygenic inheritance (expected offspring recurrence rate of 2–3%). Of interest, when the comparable issue was examined for fathers with CHD, recurrence risks were substantially lower. This sex-specific transmission risk remains unexplained to the present day.

With the advent of positional cloning in the 1980's, we and other investigators chose to elucidate gene mutations causing CHD using rare families inheriting these birth defects in Mendelian or near-Mendelian fashion, generally as part of a syndrome but sometimes in isolation. As the Human Genome Project progressed, this strategy proved robust- in the past 20 years, numerous genes causing CHD when mutated have been identified. While in aggregate, these mutations account for a relatively modest percentage of CHD, these discoveries provided the first evidence that altered dosages of cardiac developmental genes encoding transcription factor and signaling molecules cause CHD. Capitalizing on these observations, the study by Goodship and colleagues proposes that the chromosome 12 SNP associated with TOF may involve an important signaling molecule, the protein tyrosine phosphatase SHP2.

In the era after the sequencing of the human genome, a new type of genetic lesion was discovered: copy number variation (CNV). These gains or losses of DNA range in size from 1 kb to several Mb's and can now be assayed on a genome-wide basis using SNP microarrays or array comparative genomic hybridization. Significant challenges remain in differentiating pathologic CNVs from benign polymorphic one, which affect roughly 10% of the human genome. Nevertheless, it has become clear that pathologic CNVs contribute significantly to the pathogenesis of certain phenotypes such as autism and schizophrenia. Starting with the studies of Thienpont and colleagues in 2007,[9] the importance of pathologic CNVs for CHD is being elucidated. So far, it is clear that children with CHD plus involvement of other organ systems, particularly the central nervous system, are the most likely to harbor pathologic CNVs (upwards of 25%). For TOF specifically, one of us (C.E.S.) documented that pathologic CNVs are present in approximately 10% of patients, most often altering genes encoded at chromosome 1q21.1.[10] Of relevance in assessing the current association study by Goodship and colleagues, the subjects in their two TOF cohorts were screened for 22q11 deletions but not genome-wide for CNVs. That information about their cohorts would be interesting - aside from replicating the prior TOF study, it would be fascinating to know whether pathologic CNVs are mutually exclusive from the chromosomal 12 haplotype they associated with TOF or, alternatively, if the two interact to affect the cardiac phenotype.

Like CHD, many disorders of substantial importance to human health appear to be complex genetically. A major focus in human genetics in recent years has been the exploration of the

common disease/common variant hypothesis, generally through genome-wide association studies (GWAS). As most readers will be aware, GWAS has been enabled by the identification of large numbers of SNPs spanning the human genome, robust technologies for genotyping large cohorts of cases and controls, and the development of statistical genetic tools with which to analyze the resulting large datasets. The results have been a spectacular parade of GWAS that have successfully identified SNPs associated with a broad array of human disorders and traits. Along with the hoopla, some unpleasant realities have also become apparent. First, very large cohorts are required to power GWAS adequately for the typical effect size of many associated SNPs- many studies now use 10's of thousands or even more than a 100,000 cases. Second, the aggregate genetic variance accounted for after many GWAS's is quite modest and does not approach what might have been predicted based on heritability estimates, leading to the concept of the "missing heritability,"[11] which may reside in what Francis Collins has referred to as the dark matter of the genome.[12]

And CHD? To date, there has not been a published GWAS for CHD. There are probably two reasons that geneticists interested in CHD have lagged their colleagues studying other disorders. As noted, GWAS requires large cohorts in order to be powered properly (actually, two large cohorts as any positive results in the initial cohort need be replicated in a second, generally of comparable size). Since CHD prevalence (excluding bicuspic aortic valve) is just under 1%, the available subjects for such studies is restricted compared to common genetic traits. Moreover, CHD is comprised of numerous phenotypes (TOF, heterotaxy, ventricular septal defects, etc.). While it is already clear that certain heart lesions share genetic etiology (e.g., 22q11 deletions cause various conotruncal defects), it is unlikely that all forms of CHD share all of the same genetic causes. Hence, the cohorts to be studied for GWAS need to be assembled with care and can only draw from a subset of the entire CHD population. Taken as a whole, there are barriers to undertaking GWAS for CHD.

The second concern with respect to GWAS for CHD is theoretical. Such studies are designed to detect common genetic variants. When the size of cohorts is modest as dictated by the factors noted above, then the effect size of the common variant must be relative large in order to be detectable. For nearly all forms of CHD, reproductive fitness approached zero until quite recently as death during infancy or childhood was very high. Thus, purifying selection would have acted powerfully to eliminate variants associated with CHD whenever they arose. Thus, some would question whether there could be common variants causing CHD that GWAS could detect.

The counter argument favoring the existence of common variants associated with CHD with perhaps modest effect would be pleiotropy. That is, the common variant associated with CHD might have other, beneficial effects (e.g., a tradeoff between the occasional lethal CHD versus more frequent enhanced function of some other organ system). Alternatively, there could be hitchhiking- the weakly unfavorable variant associated with CHD could reside in complete linkage disequilibrium with another variant favorable for other function(s).

To this ongoing debate in the CHD genetics community, the current study of TOF provides a potentially illustrative example.[1] As the authors noted, the haplotype that they associated with TOF has previously been associated with a variety of autoimmune disorders including diabetes mellitus type I, celiac disease, and systemic lupus erythematosus. This haplotype shows strong evidence of selective sweep (i.e., positive selection) and has an estimated age of 3,400 years.[13] Whether the positive selection, postulated to be immunologic due to the associated autoimmune disorders, was driven by a variant altering *PTPN11* expression or instead one or more variants affecting other genes within the region of linkage disequilibrium such as *SH2B3*, which is important for T-cell signaling, is not clear. If the

authors are correct in surmising that the *PTPN11* variant is critical for TOF, then this association would provide the first clear example of pleiotropy or hitchhiking leading to a common variant for CHD.

While GWAS has not been completed for CHD, there are several publications in which association was sought for a limited number of SNPs. By reducing the number of hypotheses being posed by several orders of magnitude, such studies have lower nominal thresholds for statistical significance, in turn reducing the size of the cohorts required to detect associated SNPs of modest effect. Prior to the current study of TOF, Stevens and colleagues examined several SNPs capturing haplotypes of *ISL1*, a gene encoding a transcription factor critical for second heart field development.[14] Using cohorts with CHD putatively relevant for ISL1's function, they found significant associations of separate haplotypes for CHD in Caucasians and African-Americans with modest odds ratios (ORs; 1.27 and 1.57, respectively). To date, this study awaits replication.

Several association studies have focused on folate metabolism, particularly the C667T polymorphism of the methylenetetrahydrofolate reductase gene (*MTHFR*). These studies have been premised on epidemiologic studies of maternal folate supplementation, used to reduce neural tube defect incidence but also shown to reduce CHD incidence. The *MTHFR* C667T cSNP is nonsynonymous, substituting an alanine with a valine, and results in a thermolabile enzyme, which reduces blood folate levels. The findings of case-control and family-based studies of *MTHFR* C667T and CHD have been inconsistent. A recent meta-analysis revealed that the case-control studies in aggregate (n=20) achieved significant association of the TT genotype with CHD when present in fetuses and fathers (ORs of 1.55 and 1.84, respectively) but not in mothers.[15] Meta-analysis of three family-based studies did not achieve significance, perhaps due to poor statistical power. Taken as a whole, homozygosity of the *MTHFR* C667T allele may be relevant for CHD but additional studies, perhaps with larger cohorts or more focused on the relevant forms of CHD, are needed to determine this definitively.

Moving forward, the field of CHD genetics must continue the work begun already – especially given the increasing prevalence of children with corrected TOF and other CHD lesions,[16] who will soon reach reproductive ages. Collaboration amongst several large consortia now assembling cohorts of subjects with CHD, including the NHLBI-funded Pediatric Cardiovascular Genetics Consortium,[17] provide unprecedented opportunities to define the roles of pathologic CNVs and common variants through robust discovery and replications studies. In addition, with the power of high-throughput DNA sequencing, the hypothesis that rare variants are important to the etiologies for CHD can be readily tested. Hopefully, light will soon be streaming through the CHD genetics porthole!

Acknowledgments

Funding Sources: This work was supported in part by National Heart, Lung, and Blood Institute grants HL098123 (to Dr. Gelb) and HL098147 (to Dr. Seidman).

References

1. Goodship JA, Hall D, Topf A, Mamasoula C, Griffin H, Rahman TJ, Glen E, Tan H, Doza JP, Relton CL, Bentham J, Bhattacharya S, Cosgrove C, Brook D, Granados-Riveron J, Bu'Lock FA, O'Sullivan J, Stuart AG, Parsons J, Cordell HJ, Keavney B. Common variant in the *PTPN11* gene contributes to the risk of tetralogy of Fallot. Circ Cardiovasc Genet. 2012; 5:XXX–XXX.
2. Basson CT, Bachinsky DR, Lin RC, Levi T, Elkins JA, Soults J, Grayzel D, Kroumpouzou E, Traill TA, Leblanc-Straceski J, Renault B, Kucherlapati R, Seidman JG, Seidman CE. Mutations in

human *TBX5* cause limb and cardiac malformation in Holt-Oram syndrome. Nat Genet. 1997; 15:30–35. [PubMed: 8988165]

3. Goldmuntz E, Clark BJ, Mitchell LE, Jawad AF, Cuneo BF, Reed L, McDonald-McGinn D, Chien P, Feuer J, Zackai EH, Emanuel BS, Driscoll DA. Frequency of 22q11 deletions in patients with conotruncal defects. J Am Coll Cardiol. 1998; 32:492–498. [PubMed: 9708481]
4. Yunis K, Mumtaz G, Bitar F, Chamseddine F, Kassar M, Rashkidi J, Makhoul G, Tamim H. Consanguineous marriage and congenital heart defects: a case-control study in the neonatal period. Am J Med Genet A. 2006; 140:1524–1530. [PubMed: 16763961]
5. Nabulsi MM, Tamim H, Sabbagh M, Obeid MY, Yunis KA, Bitar FF. Parental consanguinity and congenital heart malformations in a developing country. Am J Med Genet A. 2003; 116A:342–347. [PubMed: 12522788]
6. Oyen N, Poulsen G, Boyd HA, Wohlfahrt J, Jensen PK, Melbye M. Recurrence of congenital heart defects in families. Circulation. 2009; 120:295–301. [PubMed: 19597048]
7. Nora JJ. Multifactorial inheritance hypothesis for the etiology of congenital heart diseases. The genetic-environmental interaction. Circulation. 1968; 38:604–617. [PubMed: 4876982]
8. Whittemore R, Hobbins JC, Engle MA. Pregnancy and its outcome in women with and without surgical treatment of congenital heart disease. Am J Cardiol. 1982; 50:641–651. [PubMed: 7113941]
9. Thienpont B, Mertens L, de Ravel T, Eyskens B, Boshoff D, Maas N, Fryns JP, Gewillig M, Vermeesch JR, Devriendt K. Submicroscopic chromosomal imbalances detected by array-CGH are a frequent cause of congenital heart defects in selected patients. Eur Heart J. 2007; 28:2778–2784. [PubMed: 17384091]
10. Greenway SC, Pereira AC, Lin JC, DePalma SR, Israel SJ, Mesquita SM, Ergul E, Conta JH, Korn JM, McCarroll SA, Gorham JM, Gabriel S, Altshuler DM, de Quintanilla-Dieck ML, Artunduaga MA, Eavey RD, Plenge RM, Shadick NA, Weinblatt ME, De Jager PL, Hafler DA, Breitbart RE, Seidman JG, Seidman CE. *De novo* copy number variants identify new genes and loci in isolated sporadic tetralogy of Fallot. Nat Genet. 2009; 41:931–935. [PubMed: 19597493]
11. Eichler EE, Flint J, Gibson G, Kong A, Leal SM, Moore JH, Nadeau JH. Missing heritability and strategies for finding the underlying causes of complex disease. Nat Rev Genet. 2010; 11:446–450. [PubMed: 20479774]
12. http://blogs.scientificamerican.com/guest-blog/2010/06/25/a-genome-story-10th-anniversary-commentary-by-francis-collins/; 2010
13. Soranzo N, Spector TD, Mangino M, Kuhnel B, Rendon A, Teumer A, Willenborg C, Wright B, Chen L, Li M, Salo P, Voight BF, Burns P, Laskowski RA, Xue Y, Menzel S, Altshuler D, Bradley JR, Bumpstead S, Burnett MS, Devaney J, Doring A, Elosua R, Epstein SE, Erber W, Falchi M, Garner SF, Ghori MJ, Goodall AH, Gwilliam R, Hakonarson HH, Hall AS, Hammond N, Hengstenberg C, Illig T, Konig IR, Knouff CW, McPherson R, Melander O, Mooser V, Nauck M, Nieminen MS, O'Donnell CJ, Peltonen L, Potter SC, Prokisch H, Rader DJ, Rice CM, Roberts R, Salomaa V, Sambrook J, Schreiber S, Schunkert H, Schwartz SM, Serbanovic-Canic J, Sinisalo J, Siscovick DS, Stark K, Surakka I, Stephens J, Thompson JR, Volker U, Volzke H, Watkins NA, Wells GA, Wichmann HE, Van Heel DA, Tyler-Smith C, Thein SL, Kathiresan S, Perola M, Reilly MP, Stewart AF, Erdmann J, Samani NJ, Meisinger C, Greinacher A, Deloukas P, Ouwehand WH, Gieger C. A genome-wide meta-analysis identifies 22 loci associated with eight hematological parameters in the HaemGen consortium. Nat Genet. 2009; 41:1182–1190. [PubMed: 19820697]
14. Stevens KN, Hakonarson H, Kim CE, Doevendans PA, Koeleman BP, Mital S, Raue J, Glessner JT, Coles JG, Moreno V, Granger A, Gruber SB, Gruber PJ. Common variation in ISL1 confers genetic susceptibility for human congenital heart disease. PLoS One. 2010; 5:e10855. [PubMed: 20520780]
15. Yin M, Dong L, Zheng J, Zhang H, Liu J, Xu Z. Meta analysis of the association between MTHFR C677T polymorphism and the risk of congenital heart defects. Ann Hum Genet. 2012; 76:9–16. [PubMed: 22175539]
16. Hoffman JI, Kaplan S, Liberthson RR. Prevalence of congenital heart disease. Am Heart J. 2004; 147:425–439. [PubMed: 14999190]

17. http://www.benchtobassinet.com/AboutPCGC.asp; 2012.