In this issue of the Journal, 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 limited number of single-nucleotide polymorphisms (SNPs) respectively.2 The discovery of 22q11 microdeletions, usually point mutations underlying Down and Holt-Oram syndromes, as the result of the shared genetic background of unaffected 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 (eg, 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.

Epidemiological studies of CHD have strongly pointed to genetic factors as the predominant cause, although environmental exposures are also relevant. Parental consanguinity significantly increases the risk of CHD in offspring, probably as the result of the shared genetic background of unaffected parents.4 In countries where consanguinous marriages are customary, the rates of consanguinity are 2- to 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 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-attributable risk given a positive family history of CHD was 4.2%.

In the 1960s, Nora7 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. Although 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, with patent ductus arteriosus in full-term newborns being the singular exception. In the 1980s, the multifactorial model took a broadside hit from the groundbreaking work of 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% to 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 1980s, 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 pathological CNVs from benign polymorphic ones, which affect roughly 10% of the human genome. Nevertheless, it has become clear that pathological CNVs contribute significantly to the pathogenesis of certain phenotypes such as autism and schizophrenia. Starting with the studies of Thiessen and colleagues in 2007,9 the importance of pathological 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 pathological CNVs (upwards of 25%). For TOF specifically, one of us (C.E.S.) documented that pathological 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 pathological 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 have 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 tens of thousands or even more than 100,000 cases. Second, the aggregate genetic variance accounted for after many GWAS is quite modest and does not approach what might have been predicted based on heritability estimates, leading to the concept of the “missing heritability,” which may reside in what Francis Collins has referred to as the dark matter of the genome.12

To date, there has not been a published GWAS for CHD. There are probably 2 reasons that geneticists interested in CHD have lagged their colleagues studying other disorders. As noted, GWAS require large cohorts to be powered properly (actually, 2 large cohorts as any positive results in the initial cohort need be replicated in a second, generally of comparable size). Since CHD prevalence (excluding bicuspid aortic valve) is just under 1%, the available subjects for such studies is restricted compared to common genetic traits. Moreover, CHD comprises numerous phenotypes (TOF, heterotaxy, ventricular septal defects, etc). Although it is already clear that certain heart lesions share genetic etiology (eg, 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 must 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 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 counterargument 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 (eg, 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 (ie, positive selection) and has an estimated age of 3400 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 1 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.

Although GWAS have 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 colleagues14 examined several SNPs capturing haplotypes of ISL1, a gene encoding a transcription factor critical for second heart field development. Using cohorts with CHD putatively relevant for ISL1 function, they found significant associations of separate haplotypes for CHD in whites and African-Americans with modest odds ratios (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 epidemiological 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 (odds ratios of 1.55 and 1.84, respectively) but not in mothers. 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, who will soon reach reproductive ages. Collaborative studies, perhaps with larger cohorts or more focused on the roles of pathological CNVs and common variants through the power of high-throughput DNA sequencing, the hypothesis that rare variants are important to the etiologies for CHD definitively.

**Sources of Funding**

This work was supported in part by National Heart, Lung, and Blood Institute grants HL098123 (to Dr Gelb) and HL098147 (to Dr Seidman), as well as the Howard Hughes Medical Institute (to Dr Seidman).

**Disclosures**

None.

**References**


