The Complex Genetic Basis of Congenital Heart Defects

Ehiolo Akhirome, BSc; Nephi A. Walton, MD; Julie M. Nogee, MD; Patrick Y. Jay, MD, PhD

Twenty years ago, chromosomal abnormalities were the only identifiable genetic causes of a small fraction of congenital heart defects (CHD). Today, a de novo or inherited genetic abnormality can be identified as pathogenic in one-third of cases. We refer to them here as monogenic causes, insofar as the genetic abnormality has a readily detectable, large effect. What explains the other two-thirds? This review considers a complex genetic basis. That is, a combination of genetic mutations or variants that individually may have little or no detectable effect contribute to the pathogenesis of a heart defect. Genes in the embryo that act directly in cardiac developmental pathways have received the most attention, but genes in the mother that establish the gestational milieu via pathways related to metabolism and aging also have an effect. A growing body of evidence highlights the pathogenic significance of genetic interactions in the embryo and maternal effects that have a genetic basis. The investigation of CHD as guided by a complex genetic model could help estimate risk more precisely and logically lead to a means of prevention.

Key Words: Congenital heart defects; Genetics; Maternal age; Maternal effects; Modifier genes

Gregor Mendel presented the laws of inheritance at a meeting of the Natural History Society of Brünn in 1865. The mathematics in his talk lost the audience, and the accompanying paper was soon forgotten. In 1866, Thomas Peacock described a child with tetralogy of Fallot whose sibling had “something the matter with its heart.” He speculated that “strong mental impressions or shocks which were sustained during pregnancy” caused cardiac malformation in some cases and a “hereditary predisposition” in others. Albeit vague, the notion that congenital heart defects (CHD) have a genetic basis preceded the rediscovery of Mendel’s laws in the 1900s.

Of course, most CHD occur sporadically and do not fit simple Mendelian patterns of inheritance, which led James Nora to propose in 1968 that a combination of genetic and environmental factors contribute to the development of a heart defect in an individual. The model appealed to him because it suggested that multiple factors could be targeted to prevent CHD. Nora’s reasoning was elegant, but the multifactorial hypothesis was difficult to validate when no etiology was known in the vast majority of cases. Investigators accordingly designed studies to discover causes that have readily detectable, large effects. Genetic abnormalities in the embryo comprise the vast majority of the known causes. For the sake of discussion, we refer to a genetic abnormality as monogenic if it can be considered the cause of a heart defect. The abnormality may involve a single gene, chromosomal interval or entire chromosome. As discussed later, other factors can modify the risk of a heart defect in the presence of the monogenic cause. The modifiers do not necessarily cause disease on their own. Currently, a monogenic basis can be identified in one-third of all cases of CHD (Figure). De novo genetic abnormalities can explain a significant fraction of CHD, but are not the entire explanation. Epidemiologic studies have consistently supported a significant role for inheritance. The largest study, encompassing >1.7 million persons born in Denmark between 1977 and 2005, showed that a family history of any CHD is a strong risk factor. Given an affected first-degree relative, the relative risk (RR) is more than 3-fold higher after excluding cases of chromosomal anomalies or extracardiac defects. The risk falls with the degree of distance from the proband but remains significant to third-degree relations. Few risk factors are as strong or consistent. A couple of notable exceptions include maternal phenylketonuria, which increases the risk 6-fold, and pregestational diabetes, which has comparable risk.

The fraction of cases that can be attributed to an inherited mutation is more difficult to estimate because of incomplete penetrance, unrecognized pathogenic mutations and undiscovered CHD genes. A recent whole-exome sequencing study of trios (i.e., the affected child and parents) offers one estimate. Non-syndromic patients with an isolated CHD are more likely to have inherited mutations of a known CHD gene or other genes. An inherited protein-truncating variant of a CHD gene was identified in a significant, but small fraction of non-syndromic patients (1.3%: 17/1,281; Figure). This fraction is a conservative estimate because it is limited to loss-of-function mutations of known CHD genes. The same study also found a significant excess (3,318) of protein-truncating variants of other genes that are likely to have undiscovered functions in cardiac development. Finally, there was a trend towards an excess of missense mutations of CHD genes (163 variants, P=0.0863). The number of patients who had these other mutations was not...
The vast majority of the known causes of congenital heart defects (CHD) are de novo or inherited genetic abnormalities. A monogenic basis can be identified in one-third of all cases. The Baltimore-Washington Infant Study and Metropolitan Atlanta Congenital Defects Program yielded similar estimates for the fraction attributable to chromosomal syndromes.\(^4\)\(^5\) De novo copy number variants (CNVs) include both the well-known, such as the 22q11.2 deletion, and more recently discovered ones.\(^6\) The fraction attributable to de novo mutations that affect protein-coding mutations were recently reported in 3 large, whole-exome sequencing studies of trios.\(^7\)\(^9\) The burden of inherited loss-of-function mutations of known CHD genes was estimated in one of these studies.\(^9\) Inherited CNVs and other forms of genetic mutation cause monogenic CHD, but their attributable fractions have not been quantified in large studies similar to the ones cited above. An oligogenic basis may explain a large fraction of the currently unknown causes.

**Monogenic CHD: One End of the Complex Genetic Continuum**

Viewing CHD as either monogenic or complex is as much a consequence of experimental design as biology. In a typical human study, cases are selected for the presence of a CHD, and the goal is to identify 1 highly penetrant genetic abnormality in each case. For instance, multiplex families are studied, or de novo mutations are sought in trios of unaffected parents and affected child. The high penetrance of a monogenic cause is demonstrated by its segregation with affected relatives or its enrichment in cases over controls. The identification of 1 genetic cause, however, does not exclude the contribution of other genes to the pathogenesis of each case.

When studies ascertain cases by genetic diagnosis rather than CHD, well-established monogenic causes consistently show high, but incomplete penetrance. The pattern is the same whether the cause is a chromosomal anomaly that affects more than 1 gene or the mutation of a single gene (Table 1).\(^16\)\(^25\) Additional factors modify the risk of CHD caused by a single genetic abnormality. Hearkening to Nora’s multifactorial hypothesis, investigators have sought to identify the other contributing environmental or genetic factors. Not for lack of effort, only a few environmental factors have been clearly and reproducibly demonstrated to cause disease or to affect the risk of CHD in general.\(^15\) Similarly, it has been difficult to establish the role of environmental factors in monogenic CHD. For example, maternal folate supplementation, which has received much attention in Down syndrome, has no or little detectable effect.\(^26\)\(^27\) In contrast, the evidence for modifier genes is growing in humans and well established in mouse models.

Consider the familial recurrence risk of CHD. The risk is commonly presumed to be from monogenic causes that are private to families, but it could also be caused by transmission of genetic modifiers that affect risk independently of the cause in a particular relative. Two groups have considered the latter hypothesis in the setting of CHD caused by a 22q11.2 deletion in the proband.\(^28\)\(^29\) Their studies yielded consistent results with a combined total of 212 probands and almost 2,000 first- and second-degree relatives who did not have the deletion. In the larger of the 2 studies, the incidence of CHD in the relatives was 4-fold greater if the proband had CHD than if not. Moreover, the incidence of severe CHD in the relatives of probands who had CHD was 6-fold greater than in the general popula-

<table>
<thead>
<tr>
<th>Species / Genetic abnormality</th>
<th>CHD incidence, %</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1q21.1 deletion</td>
<td>23</td>
<td>16</td>
</tr>
<tr>
<td>Turner syndrome (45, XO karyotype)</td>
<td>30</td>
<td>17</td>
</tr>
<tr>
<td>Cornelia de Lange (NIPBL, SMC1A, SMC3)</td>
<td>31</td>
<td>18</td>
</tr>
<tr>
<td>Down syndrome (trisomy 21)</td>
<td>44</td>
<td>19</td>
</tr>
<tr>
<td>Holt-Oram syndrome (TBX5)</td>
<td>62</td>
<td>20</td>
</tr>
<tr>
<td>DiGeorge syndrome (22q11.2 deletion)</td>
<td>75</td>
<td>21</td>
</tr>
<tr>
<td>NKX2–5 mutation</td>
<td>88</td>
<td>22</td>
</tr>
<tr>
<td>Mouse</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gata4(^+/−)</td>
<td>12–76</td>
<td>23</td>
</tr>
<tr>
<td>Nkx2–5(^+/−)</td>
<td>5–50</td>
<td>24</td>
</tr>
<tr>
<td>Tbx5(^+/−)</td>
<td>40–80</td>
<td>25</td>
</tr>
</tbody>
</table>

Genetic causes of CHD show high, but never complete penetrance. Examples of human copy number variants, chromosomal, and genetic abnormalities are shown in which cases were ascertained by the presence of the genetic abnormality. CHD incidence in mouse mutant models depends on the genetic background, which can be systematically varied via inbred strain crosses. The variability is a sign of interactions between the mutated CHD gene and polymorphic modifier genes. CHD, congenital heart defects.

![Figure. The vast majority of the known causes of congenital heart defects (CHD) are de novo or inherited genetic abnormalities. A monogenic basis can be identified in one-third of all cases.](image-url)
tion. More studies are necessary to determine whether the familial recurrence risk is independent of a monogenic cause in the proband.

Common genetic variation, either common copy number variants (CNVs) or single nucleotide polymorphisms, has not been found to have a detectable risk on monogenic CHD, with one telling exception. A 12p13.31 CNV duplication is associated with the risk of conotruncal heart defects and left-sided lesions in DiGeorge and Turner syndromes, respectively. The interval includes 3 genes, SLC2A3, SLC2A14, and NANOGP1. The functions of SLC2A14 and NANOGP1 are unknown. SLC2A14 encodes GLUT14, one of a large class of proteins with homology to facilitative glucose transporters. NANOGP1 is a pseudogene of NANOG, the transcription factor that maintains pluripotency. SLC2A3 encodes GLUT3, a facilitative glucose transporter that is expressed by trophoblasts in the placenta. GLUT3 facilitate the transport of glucose across cell membranes and down a concentration gradient. Haploinsufficiency of SLC2A3 causes a quantitative reduction in placental glucose transport and embryonic glucose concentration. One may speculate that maternal diabetes and a SLC2A3 duplication similarly increase the risk of CHD by causing embryonic hyperglycemia. Maternal diabetes would do so by increasing the glucose concentration gradient between the mother and embryo, while the SLC2A3 duplication would increase glucose transport into the embryo. The RR associated with maternal diabetes in a recent Danish epidemiologic study (RR 4.00, 95% confidence interval (CI) 3.51–4.53) was modestly higher than the risk associated with the SLC2A3 duplication in DiGeorge (RR 1.4, 95% CI 1.3–1.6) and Turner syndromes (RR 1.8, 95% CI 1.3–2.5). The RRs for each syndrome were recalculated from the published data for comparison. The risk associated with maternal diabetes may be higher because maternal hyperglycemia can induce greater embryonic hyperglycemia than can a SLC2A3 duplication when the mother is normoglycemic.

Rare genetic variation of cardiac developmental genes has been found to affect the risk of monogenic CHD. For example, mutations of CRELD1 and HEY2 are known to cause CHD in humans and mice, respectively. Rare variants of these 2 genes have also been associated with atrioventricular septal defects (AVSD) in Down syndrome. Genetic experiments in a mouse model of human trisomy 21 support the findings. Similarly, mutations of histone-modifying genes are an important class of monogenic causes of CHD. In a whole-exome sequencing study of DiGeorge syndrome patients, rare variants of several histone-modifying genes were associated with an increased risk of conotruncal heart defects. Conversely, rare variants of genes that have the opposite biochemical effect on histone modification may be associated with a reduced risk.

The major role of modifier genes in monogenic CHD has been demonstrated in mouse models via a few experimental strategies. The most common is to examine the interaction between a CHD gene and a candidate modifier gene. Typically, 2 mutant lines that each carry a knockout mutation of the CHD or candidate gene are crossed. A difference in the incidences of CHD between the 2 single and the double mutant progeny indicates a genetic interaction, as has been shown for several monogenic CHD models (Table 2). Many more pathologically significant genetic interactions undoubtedly remain to be discovered.

The second strategy demonstrates the effects of naturally occurring genetic polymorphisms that interact with the disease-causing mutation. Mouse mutants are usually studied in a homogeneous, inbred strain background so that phenotypes can be clearly attributed to the mutated gene. In contrast, a few groups have characterized monogenic CHD models in heterogeneous genetic backgrounds. Introduction of the mutation into systematic inbred strain crosses permits the analysis of genetic interactions with polymorphic modifier genes. Regardless of the particular mutation, the incidence of CHD varies widely with the genetic background (Table 1). In a study of >3,000 Nkx2–5 mice from 5 different crosses, the incidence of specific malformations, such as membranous ventricular septal defect (VSD) or AVSD, also varied with the genetic background. Alleles of modifier genes either increase or decrease the susceptibility of a cardiac developmental pathway to the causative mutation. Combinations of low- and high-risk genotypes at quantitative trait loci (QTL) determine the risk of specific cardiac malformations in the Nkx2–5 mouse. Several QTLs for VSDs have been described from a cross between 2 inbred strains. An ongoing effort in our group has yielded at least a dozen QTLs for simple and complex heart defects.

### Evidence for the Oligogenic Basis of Human CHD

Although a monogenic cause of CHD is defined by its high phenotypic penetrance in a population, its interaction with modifier genes can significantly promote or suppress risk in an individual. Genetic interactions can likewise potentiate the deleteriousness of a mutation that has little or no effect on its own. Oligogenic combinations of inherited genetic variants could explain the majority of CHD cases that lack a detectable monogenic basis.

Early studies noted that affected individuals not uncommonly carry rare, heterozygous mutations of 2 different CHD genes that are predicted to be deleterious. The studies were limited by DNA-sequencing capacity, so only a few genes were examined in a small number of individuals. Two recent studies suggest that oligogenic combinations of mutations are a common pathogenic phenomenon.

Using whole-exome sequencing, the first study examined individuals with AVSDs caused by mutations of known AVSD genes. The cases carried more, presumably pathogenic mutations of AVSD genes per person than the general population. On average, each case and control respectively carried 3.62 and 2.40 rare, nonsynonymous mutations and
Mutations of approximately 200 genes are associated with human CHD. The mutations mainly affect transcription factors, epigenetic regulators, and signaling pathways. Tremendous progress has been made in understanding how the mutations affect cardiac development. The results have found relevance to clinical problems such as arrhythmia and cardiomyopathy, but how the knowledge can be applied to prevent CHD is far less clear. Nevertheless, we believe that CHD prevention should remain a driving force for research. With that in mind we outline 3 research areas that could yield novel methods to quantify and reduce the risk of CHD.

Contributions to the Incomplete Penetrance of Monogenic CHD Mutations
Mutations that cause severe CHD should be subject to negative selection. Indeed, our analysis of data from the Exome Aggregation Consortium (ExAC) indicates that humans are more intolerant of loss-of-function mutations of CHD genes than all other genes (pLI scores 0.59±0.03 vs. 0.30±0.003, mean±SEM; P=3×10−11). The supplementary table 20 in Sifrim et al. lists the CHD genes.) On the other hand, 1.9% of adults in the ExAC study, who did not have any severe pediatric disease, carry a loss-of-function mutation of a CHD gene. Given that the incidence of moderate-to-severe CHD at birth is 0.6%, there may be genetic mechanisms that suppress the deleteriousness of mutations, as has been observed in mouse models.24,25 Delineation of the genetic modifiers could help to provide more personalized estimates of familial recurrence risks that are currently based on epidemiologic data.

Genetic Interactions that Contribute to the Oligogenic Basis of CHD
The interactions between genetic variants may be just as important as individual genes in the pathogenesis of a large fraction of cases. Guided by the monogenic model, large whole-exome sequencing studies of trios have focused on high-heart expressed or known CHD genes and loss-of-function mutations. In contrast, statistical and bioinformatic analyses of human genomic studies are much more difficult when there are exponentially more interactions than genes to consider and the interacting genetic variants cannot be as simply or narrowly defined as a loss-of-function mutation. Systems genetic analysis in animal models could help to narrow the search space for human studies. If synergistic interactions contribute significantly to the pathogenesis of severe CHD, targeting them could have a tremendous effect on the most challenging clinical cases.

The Genetic Basis of Maternal Age-Associated Risk of CHD
The maternal age-associated risk of CHD in the Nkx2–5+/− mouse model is a quantitative genetic trait. If the same maternal pathway operates in humans, some offspring may face a higher or lower than chronologically predicted mater-
nal age-associated risk depending on the mother's genetics. Consideration of biological pathways in the mother that affect the gestational milieu offers potential opportunities to prevent CHD that might not otherwise present themselves by a focus on the embryo.

Insights from each of these areas into the complex genetic basis of CHD could lead to the identification of high-risk populations who would benefit from future prevention strategies. Elucidation of the genetic variants and interactions that affect risk, especially ones that reduce risk, could logically suggest therapies that a focus on monogenic causes has not. One can imagine therapies that target the embryo or the mother depending on where the genes act. The outcomes for newborns who have severe heart defects have improved dramatically in the past few decades, but there is still room for improvement. A prevention strategy, even if only modestly effective, would generate tremendous benefits. In the USA, >30,000 affected children are born each year. The number of adults who have severe CHD is growing by almost 10,000 per year. They exceed the number of children with severe CHD. By conservative estimates, the prevention of one case would save around US$100,000 in direct patient care costs and provide society an additional estimated $1,000,000 of economic productivity over a lifetime ($25,000/year/40 years). As little as a 1% reduction or 300 additional, healthy children per year in perpetuity would produce a huge return on investment in research to prevent CHD. Of course, the return on a healthy child for the family is priceless.

Acknowledgments

We thank Drs. Paul Hruz and David Wilson for their comments. E.A. is supported by a NIH predoctoral training grant (T32 HL007873) and a Ruth L. Kirschstein National Research Service Award (F30 HL136077). J.M.N. is supported by an NIH training grant (T32 HD043010, Training of the Pediatric Physician-Scientist). P.Y.J. is an Established Investigator of the American Heart Association and the Pediatric Physician-Scientist). P.Y.J. is supported by a NIH predoctoral training grant (T32 HL007873) and is supported by a NIH predoctoral training grant (T32 HL007873) and is supported by a NIH predoctoral training grant (T32 HL007873) and is supported by a NIH predoctoral training grant (T32 HL007873). We thank Drs. Paul Hruz and David Wilson for their comments. E.A. is supported by a NIH predoctoral training grant (T32 HL007873) and is supported by a NIH predoctoral training grant (T32 HL007873) and is supported by a NIH predoctoral training grant (T32 HL007873). We thank Drs. Paul Hruz and David Wilson for their comments. E.A. is supported by a NIH predoctoral training grant (T32 HL007873) and is supported by a NIH predoctoral training grant (T32 HL007873).

References

29. Swaby JA, Silversides CK, Bekeshus SC, Piran S, Oechslin EN,


