Developments in Our Understanding of the Genetic Basis of Birth Defects

Daniel M. Webber[1], Stewart L. MacLeod*[1], Michael J. Bamshad[2], Gary M. Shaw[3], Richard H. Finnell[4], Sanjay S. Shete[5], John S. Witte[6], Stephen W. Erickson[7], Linda D. Murphy[1], and Charlotte Hobbs[1]

Birth defects are a major cause of morbidity and mortality worldwide. There has been much progress in understanding the genetic basis of familial and syndromic forms of birth defects. However, the etiology of nonsydromic birth defects is not well-understood. Although there is still much work to be done, we have many of the tools needed to accomplish the task. Advances in next-generation sequencing have introduced a sea of possibilities, from disease-gene discovery to clinical screening and diagnosis. These advances have been fruitful in identifying a host of candidate disease genes, spanning the spectrum of birth defects. With the advent of CRISPR-Cas9 gene editing, researchers now have a precise tool for characterizing this genetic variation in model systems. Work in model organisms has also illustrated the importance of epigenetics in human development and birth defects etiology. Here we review past and current knowledge in birth defects genetics. We describe genotyping and sequencing methods for the detection and analysis of rare and common variants. We remark on the utility of model organisms and explore epigenetics in the context of structural malformation. We conclude by highlighting approaches that may provide insight into the complex genetics of birth defects.

Birth Defects Research (Part A) 103:680–691, 2015.

Key words: Birth defects; congenital heart defects; cleft lip and palate; epigenetics; genetics; hypospadias; neural tube defects; next-generation sequencing

Introduction

Nearly 8 million children are born in the world each year with a serious birth defect (Christianson et al., 2005). In the United States, birth defects affect at least 1 in every 33 newborns and result in considerable mortality and long-term disability (Centers for Disease Control and Prevention, 2008). Progress has been made in identifying environmental risk factors in nonsyndromic birth defects.* However, ample work remains in terms of characterizing the genetic basis for most of these conditions. Here we review the genetic basis of nonsyndromic structural birth defects, with a focus on the four most common structural birth defects: congenital heart defects (CHD), neural tube defects (NTD), clefts of the lip and/or palate (CLP), and hypospadias. We provide a historical perspective and describe current microarray- and sequencing-based approaches for identifying common and rare variants underlying structural birth defects. We discuss the strengths and limitations of each technique and provide examples of the successful implementation of each approach to identify genetic factors influencing the risk of nonsyndromic birth defects.

CHD, NTD, CLP, and hypospadias account for nearly half of the birth defects that occur in the United States (Porter et al., 2005; Parker et al., 2010). CHDs are abnormalities of the heart or great vessels that are present at birth. They are the most common type of birth defect. These malformations occur in approximately 8 of every 1000 live births, and approximately 40% of babies born with the most serious CHDs die in infancy (Moller et al., 1993; Yoon et al., 2001; Hoffman and Kaplan, 2002; Pierpont et al., 2007; Gilboa et al., 2010; Mathews et al.,

[1]Division of Birth Defects Research, Department of Pediatrics, College of Medicine, University of Arkansas for Medical Sciences, Little Rock, Arkansas
[2]Division of Genetic Medicine, Department of Pediatrics, University of Washington, Seattle, Washington
[3]Stanford University School of Medicine, Stanford, California
[4]Dell Pediatric Research Institute, Department of Nutritional Sciences, The University of Texas at Austin, Austin, Texas
[5]Department of Epidemiology, MD Anderson Cancer Center, Houston, Texas
[6]Department of Epidemiology and Biostatistics, University of California San Francisco, San Francisco, California
[7]Department of Biostatistics, College of Medicine, University of Arkansas for Medical Sciences, Little Rock, Arkansas

Supported the National Institute of Child Health and Human Development (5R01HD039054-12) and the National Center on Birth Defects and Developmental Disabilities (U01DD000491-05).

*Correspondence to: Stewart MacLeod, University of Arkansas for Medical Sciences, College of Medicine, Department of Pediatrics, 13 Children's Way, Mail Slot 512-40, Little Rock, AR 72202. E-mail: macleodstewartl@uams.edu

Published online 28 May 2015 in Wiley Online Library (wileyonlinelibrary.com). Doi: 10.1002/bdra.23385

*Nonsyndromic birth defects are defined in the context of this study as congenital malformations that are not associated with a known or identifiable syndrome.

2013). Affected infants who survive often require repeated surgeries and lengthy hospitalization. Similarly, neural tube defects are often severe and debilitating. These malformations result from improper closure of the skull or vertebrae, leaving the brain or spinal cord exposed. In the United States, NTDs affect 0.6 in every 1000 births (Parker et al., 2010). Rates of NTDs are higher in some developing countries (Castilla et al., 2003). CLP is a congenital malformation in which facial development is disrupted. It affects 2 of every 1000 births in the United States. Although it is not a major cause of infant mortality, children with craniofacial malformations require surgery to repair the cleft lip or cleft palate and may encounter problems with feeding, speaking, hearing, or social stigmatism. Hypospadias is a structural malformation in which the opening of the urethra is located on the underside of the penis rather than on the tip. It affects approximately 3 per 1000 births (Dolk et al., 2004; Porter et al., 2005; Fisch et al., 2009).

Genetic Landscape: Past

Several lines of evidence, in both animal and human studies, indicate that most nonsyndromic defects have a genetic component. Existing evidence from human studies includes increased concordance among monozygotic twins compared with dizygotic twins, increased recurrence among full siblings compared with half siblings, and increased recurrence among first-degree relatives compared with second- and third-degree relatives. Such studies point to a genetic basis for CHD (Oyen et al., 2009), NTD (Janerich and Piper, 1978), CLP (Christensen and Mitchell, 1996), and hypospadias (Schnack et al., 2008).

CANDIDATE GENE STUDIES

Early genetic studies of nonsyndromic birth defects focused on testing the association of a small number of candidate single-nucleotide polymorphisms (SNPs) with common birth defects, including CLP, NTDs, and CHDs. For many of these studies, candidate genes were selected based upon mouse models of normal and abnormal development. For example, animal studies initially highlighted growth factors involved in development of the palate. Several of these developmental genes were later included in a small case-control study of nonsyndromic cleft lip with or without cleft palate (Ardinger et al., 1989). This work revealed an association between transforming growth factor-alpha (*TGFA*) and CLP – an association that has been replicated in subsequent studies (Lu et al., 2014). Characterization of NTDs in two mouse models (*spin cycle* and *crash*) led to the discovery of *CELSR1* (Curtin et al., 2003) and associated proteins within the planar cell polarity pathway. Genes within this pathway (e.g., *CELSR1, FUZ*, *VANGL1*, *VANGL2*, and *SCRIB*) have since been linked to NTDs among humans. In the case of nonsyndromic CHDs, many of the critical cardiac transcription factors (e.g., *NKX2–5*, *GATA4*) were first characterized in the mouse and then included as targets in candidate gene studies (Molkentin et al., 1994; Lyons et al., 1995). Discovery of mutations in the transcription factors, *ZIC3*, *GATA4*, and *NKX2–5* in CHD has since highlighted the critical role of these proteins in cardiac development (McCulley and Black, 2012).

Candidate gene studies of birth defects not only built upon findings from developmental biology but also gained insight from epidemiologic studies. It was clear by 1992 that periconceptional folate intake reduced the risk of NTD. (In 1992, the United States Public Health Service made the recommendation that women of childbearing age consume 0.4 mg of folic acid per day to reduce the risk of having a pregnancy affected by a neural tube defect [Houk et al., 1992].) Subsequent work established a link between maternal periconceptional multivitamin use and reduced risk of conotruncal heart defects, limb deficiencies, and CLP (Shaw et al., 1995a,b). Frosst et al. (1995) described a polymorphism in methylenetetrahydrofolate reductase (*MTHFR* 677C>T) that encodes a thermolabile enzyme with diminished activity. Individuals with this form of *MTHFR* have a decreased concentration of serum folate and an increased concentration of homocysteine. Recognizing the potential implications, researchers soon tested for associations between *MTHFR* 677C>T and common birth defects. By the end of the decade, *MTHFR* 677C>T was established as an important risk factor in NTD (van der Put et al., 1998) and conotruncal heart defects (Junker et al., 2001; van Beynum et al., 2006; Yin et al., 2012).

As custom genotyping microarrays became more available in the mid-2000s, researchers began to genotype entire pathways rather than individual genes. Despite changes in technology, folate-related genes continued to be an important focus of study (Shaw et al., 2009; Zhu et al., 2012; Hobbs et al., 2014). This approach cast a broader net in search of common variants affecting disease risk. It did so by examining tens to hundreds of polymorphisms, within the context of gene-environment (Zhu et al., 2012; Hobbs et al., 2014) and maternal-fetal interactions (Li et al., 2014). An increasingly large number of study participants were necessary to account for multiple testing. Therefore, common birth defects (CHD, NTD, and CLP) received much of the initial focus.

Candidate gene studies are especially well-suited to situations where there is strong evidence for involvement of a pathway in disease. However, this method may be vulnerable to inadvertent bias in the selection of candidate genes. It is also possible that a significant association is detected by chance alone. Therefore, it is important that results be replicated. In the past, there have been several questionable genotype-phenotype associations (Hirschhorn et al., 2002), which might have been clarified by rigorous validation efforts. To address this issue, the National Human Genome Research Institute (NHGRI) working group has outlined several best practices for replicating genotype-phenotype associations (Chanock et al., 2007). The guidelines have become requisite standards for GWAS, but they are equally applicable to candidate gene studies. The report recommends that a

comparable phenotype and population should be analyzed in both the initial study and the replication. It also stresses that the replication study should be large enough to identify the initial association. Ideally, the replication sample should be at least as large in number as the discovery sample.

Genetic Landscape: Present

Birth defects sometimes cluster within families and have a higher recurrence rate among full-siblings compared with half-siblings. This is especially true of syndromic birth defects, which often segregate as autosomal dominant, autosomal recessive, or X-linked traits (Fahed et al., 2013). In contrast, genotype may play only a minor role in birth defects caused by maternal exposure to a teratogen such as isotretinoin (Rosa, 1983).

Genetic variation among humans is often classified as either rare or common. Common variants are arbitrarily defined as those with a minor allele frequency (MAF) of at least 5% (1000 Genomes Project Consortium, 2012); whereas, rare variants are often characterized as having a MAF less than 1%. There is a longstanding debate about the respective contributions of common and rare variants to complex diseases (Bodmer and Bonilla, 2008; Gibson, 2012). In reality, both common and rare variants are thought to contribute to risk of nonsyndromic birth defects. In the following pages, we will describe both categories and will review methods for variant detection and analysis.

COMMON VARIANTS

Genome wide association studies (GWAS) interrogate hundreds of thousands to millions of SNPs to identify associations between a genotype and complex disease. Nearly 2000 GWAS have been reported since the initial publication of Ozaki et al. (2002). These studies have identified numerous common variants that are risk factors for disease (Welter et al., 2014). The majority of GWAS of birth defects have focused on CLP, CHD, and hypospadias (Table 1). To our knowledge, none have included NTDs. GWAS of CHDs have identified five SNPs reaching genome-wide significance ($p < 5 \times 10^{-8}$). The CHD-associated SNPs have modest odds ratios (OR) ranging from 1.2 to 1.5 and have global minor allele frequencies of 0.2 to 0.3 (Table 1). Meanwhile, GWAS on CLP have reported SNPs with greater effects size (OR = 1.4–2.6). Two GWAS have been conducted on hypospadias. The first identified a common variant in the gene *DGKK* that is strongly associated with risk of hypospadias and has a global minor allele frequency of 0.44 (van der Zanden et al., 2011). The second, much larger study identified 18 SNPs with genome-wide significance, accounting for a total of 8.7% of disease liability (Geller et al., 2014). A total of 56.9% disease liability was explained when all SNPs in the study were considered. This lends support to the infinitesimal model, in which hundreds or thousands of loci contribute to disease risk. In this model, statistically significant GWAS results represent only the largest of effects drawn from a normal distribution (Gibson, 2012).

GENOTYPING TECHNOLOGY

Genotyping methods have evolved rapidly since the first GWAS was completed in 2002. The number of variants assayed by GWAS has increased from around 10,000 SNPs in the early 2000s to 5 million SNPs at present (Hopper et al., 2012). The additional SNPs provide increased resolution of haplotypes, increased coverage of low frequency variants, and improved ability to infer genomic structural variation (Alkan et al., 2011). Structural variation, including insertions, deletions, and inversions, is broadly associated with nonsyndromic birth defects (Southard et al., 2012). SNP microarray and array comparative genomic hybridization (CGH) have historically been the workhorses for detecting insertions and deletions, referred to as copy number variations. However, SNP microarrays may be gaining an edge because of their versatility. Increased density of SNP microarrays has made it possible to detect smaller structural variation and accurately resolve their breakpoints. Although the high density SNP microarray is a powerful tool, it is biased by a lower sensitivity in detecting single copy gains compared with single copy deletions. Even with high density chips, it can be challenging to consistently detect small events. Next-generation sequencing, a high-throughput method for sequencing millions or billions of DNA strands in parallel, overcomes many of these limitations and is arguably better suited at detecting small structural variations. There are multiple bioinformatics and statistical considerations in detecting copy number variations (Alkan et al., 2011) that are beyond the scope of this study.

GWASs of CHD, CLP, and hypospadias have been successful in identifying common variants that influence disease risk (Birnbaum et al., 2009; Beaty et al., 2010; van der Zanden et al., 2011; Ludwig et al., 2012; Hu et al., 2013; Cordell et al., 2013a,b; Geller et al., 2014). These SNPs (Table 1) occur frequently in the population (af = 0.17–.49) and have modest effects size (OR = 1.3–2.6) (Fig. 1). They may be directly involved in the disease etiology (i.e., functional or causal variant) or may "tag" a nearby functional variant that is involved in development of the disease. In most cases, the SNP being "tagged" is common within the population and has an effect size that is similar to, if not slightly larger than the original SNP. In some instances, the GWASs point to genomic regions that are not only susceptible to perturbations caused by common variants but are also vulnerable to rare variants. Ventral anterior homeobox 1 (*VAX1*) proves an example of a gene that contains both common SNPs and rare functional variants. A GWAS first identified SNPs within *VAX1* that were strongly associated with CLP. Sequencing of the gene among affected family trios later replicated the risk

TABLE 1. *Genome Wide Association Studies of Congenital Heart Defects, Clefts of the Lip and/or Palate, and Hypospadias*

Phenotype	SNP	Allele[a]	GMAF	Chr.	Gene	*P*	OR het (95% CI)	OR hom (95% CI)	Author
Atrial septal defect	rs870142	**T**/C	0.207	4p16.2	*MSX1 - STX18* (between)	2.60×10^{-10}	1.46 (NR)		Cordell HJ et al.
CHD	rs1531070	**A**/G	0.251	4q31.1	*MAML3*	5.0×10^{-12}	1.40 (1.27–1.54)		Hu Z et al.
CHD	rs2474937	**C**/T	0.211	1p12	*TBX15* (closest)	8.0×10^{-10}	1.40 (1.26–1.56)		Hu Z et al.
TOF	rs11065987	**G**/A	0.195	12q24.12	*ATXN2–BRAP* (between)	7.7×10^{-11}	1.34 (1.21–1.50)		Cordell HJ et al.
TOF	rs7982677	**G**/A	0.297	13q31.3	*GPC5*	3.03×10^{-11}	1.29 (1.15–1.44)		Cordell HJ et al.
CLP	rs560426	**G**/A	0.399	1p22.1	*ABCA4* (closest)	5.01×10^{-12}	1.42 (1.29–1.59)		Beaty TH et al.
CLP	rs987525	**A**/C	0.254	8q24	*PVT1 - GSDMC* (between)	3.34×10^{-24}	2.57 (2.02–3.26)	6.05 (3.88–9.43)	Birnbaum S et al.
CLP	rs7078160	**A**/G	0.266	10q25	*KIAA1598 - VAX1*	1.92×10^{-8}	1.36 (1.21–1.53)	2.50 (1.95–3.21)	Mangold E et al
CLP	rs227731	**C**/A	0.365	17q22	*NOG* (closest)	1.07×10^{-8}	1.38 (1.21–1.56)	1.91 (1.63–2.24)	Mangold E et al
CLP	rs13041247	C/**T**	0.488	20q12	*MAFB* (closest)	1.44×10^{-11}	.70 (0.64–0.78)		Beaty TH et al.
CLP	rs861020	**A**/G	0.182	1q32.2	*IRF6*	3.24×10^{-12}	1.44 (1.27–1.64)	2.04 (1.60–2.60)	Ludwig KU et al.
CLP	rs742071	**T**/G	0.304	1p36	*PAX7*	7.02×10^{-9}	1.32 (1.13–1.54)	1.88 (1.52–2.32)	Ludwig KU et al.
CLP	rs7590268	**G**/T	0.167	2p21	*THADA*	1.25×10^{-8}	1.42 (1.23–1.64)	1.98 (1.47–2.66)	Ludwig KU et al.
CLP	rs7632427	C/**T**	0.388	3p11.1	*EPHA3* (closest)	3.90×10^{-8}	0.73 (0.64–0.83)	0.61 (0.49–0.76)	Ludwig KU et al.
CLP	rs12543318	**C**/A	0.392	8q21.3	*DCAF4L2* (closest)	1.90×10^{-8}	1.27 (1.11–1.46)	1.68 (1.40–2.01)	Ludwig KU et al.
CLP	rs8001641	**A**/G	0.292	13q31.1	*SPRY2* (closest)	2.62×10^{-10}	1.31 (1.13–1.51)	1.86 (1.54–2.26)	Ludwig KU et al.
Hypospadias	rs1934179	**T**/C	0.438	Xp11.22	*DGKK*	2.80×10^{-21}	2.60 (2.10–3.10)		Van der Zanden LFM et al.
Hypospadias	18 SNPs[b]								Geller F et al.

Odds ratios (ORs) are given with the major allele set as baseline.

[a]The minor allele is given first. The risk allele is shown in bold.

[b]A total of 18 loci were significantly associated with hypospadias.

Chr., chromosome; het, heterozygous; hom, homozygous; CHD, congenital heart defect; TOF, tetralogy of Fallot, CLP, clefts of the lip and/or palate.

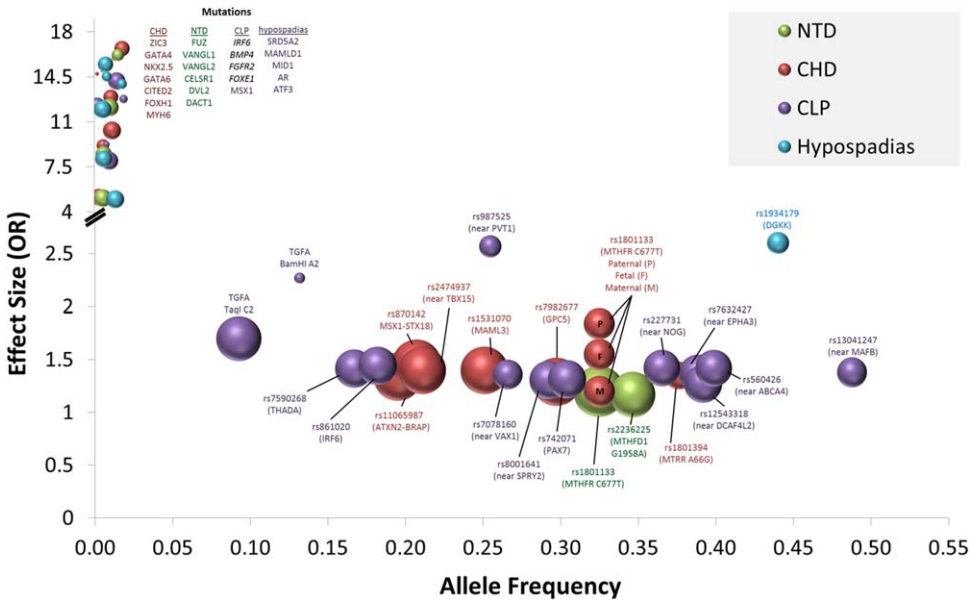

FIGURE 1. *Allele frequency, effect size, and selected birth defects. Common single nucleotide polymorphisms (SNPs) with modest effect size (middle-center) have been identified by association studies of sporadic cases of congenital heart defects (CHD), neural tube defects (NTD), clefts of the lip and/or palate (CLP), and hypospadias. In contrast, certain rare mutations appear to have high penetrance/effect size for the above-listed birth defects. The genes listed in the upper left are rare causal mutations that have been substantiated among patient groups and characterized in model systems. There are few if any adequately powered studies for these mutations; therefore, the OR for mutations is estimated based upon published literature and is intended to illustrate the general trend. Disease-associated SNPs are identified in the lower middle part of the figure by spheres, which are scaled in size based upon the total number of cases and controls in combined discovery and validation studies, from 479 individuals in the smallest study to 10,091 individuals in the largest study. Allele frequency represents global minor allele frequency (GMAF) from dbSNP. Effect sizes in the figure are based upon GWASs in Table 1 and meta-analyses of candidate-gene association studies (Yin et al., 2012; Cai et al., 2014; Feng et al., 2014; Jiang et al., 2014; Lu et al., 2014; Yadav et al., 2015).*

markers from the GWAS and identified both common and rare variants associated within CLP (Butali et al., 2013). Interferon regulatory factor 6 (*IRF6*) provides a similar example. Mutations in this gene are known to cause Van der Woude syndrome, an autosomal dominant form of cleft lip and palate. Not only do rare variants in IRF6 cause syndromic cleft lip and palate, but a common variant within an *IRF6* enhancer influences non-syndromic disease risk by disrupting transcription factor binding (Rahimov et al., 2008). Genomic variants that have a major impact on health or reproductive fitness are more likely to be under selective pressures, limiting allele transmission and reducing allele frequencies within the population, as shown in the upper left-hand quadrant of Figure 1. These mutations often have large effects but occur at such low frequency that even in large GWASs they fail to meet genome-wide significance.

Risk loci identified by GWASs account for only a small fraction of the observed heritability of any particular birth defect (Gibson, 2012). Rare variants are seen to play an increasingly important role in the etiology of birth defects, and it is likely that rare alleles explain some of the missing heritability of complex traits. However, next-generation sequencing provides an effective means of identifying rare variants. Whole-exome sequencing (WES), for example, is an especially efficient approach for functional variant discovery. WES uses probe hybridization enrichment to capture 50 to 60 Mb of genomic DNA, including protein coding sequences, micro RNA, and in some cases untranslated regions flanking each gene. The exome comprises only 2% of the human genome; however, it contains the majority of known, disease-causing mutations. WES has been widely successful in identifying mutations responsible for inherited Mendelian diseases (Rabbani et al., 2012) and familial forms of birth defects (Arrington et al., 2012; Yu et al., 2013). As the cost of WES continues to fall, it becomes a more attractive tool for studying complex disorders, including birth defects. For example, exome sequencing of 362 case-parent trios with CHD and 264 control-parent trios recently identified histone-modifying genes that are involved in CHD (Zaidi et al., 2013). Because CHD is under strong selective pressure, the investigators in this example focused on de novo mutations that might account for the sporadic pattern of occurrence among their cases.

They found a very similar number of de novo protein-altering mutations among cases and controls, but interestingly the mutations among cases were more likely to occur in genes required for heart development. Based on the number of de novo mutations in heart developmental genes, the authors estimated that such mutations have a role in 10% of severe nonsyndromic CHD. The results suggest that risk for isolated CHD is influenced by mutations that affect any one of a broad range of developmental genes. This mirrors a study among trios with autism, which showed that predicted damaging, de novo mutations among cases were more likely to affect genes expressed in the developing brain (Sanders et al., 2012).

Whole exome sequencing has recently been used alongside SNP arrays to identify de novo copy number variation in CHD (Glessner et al., 2014). In a study conducted by the Pediatric Cardiac Genomics Consortium, two complementary technologies, WES and microarray, were used to detect de novo mutations among 538 CHD case-parent trios and 1301 healthy controls. SNP microarray is very effective at identifying large copy number variants that reside throughout the genome; however, it is unable to identify small copy number variants, has a bias toward detecting single copy losses, and cannot map the location of copy number gains. In contrast, exome sequencing is often inaccurate at identifying large structural variation, but is well-suited at detecting small copy number variants, which take the form of insertions or deletions (indels). By pairing WES with microarray, the investigators enhanced the effectiveness of both technologies and were able to identify recurrent de novo copy number variants at 15q11.2 and detect variants in genes that interact with key CHD proteins, NKX2–5 and GATA4 (Glessner et al., 2014).

WES has additional limitations. First, it covers only a small fraction of the genome. Regulatory elements play a critical role in development and disease risk (Wamstad et al., 2014); however, only a fraction of these elements are assessed by WES. A second limitation of WES is that it does not provide uniform coverage. Some regions of the exome have high coverage, whereas others receive limited coverage and suffer from low sensitivity in SNP detection. Uniformity of coverage for WES has improved in recent years, but there is still much room for improvement (Meynert et al., 2014). WES also has limited accuracy in identifying structural variation greater than 30 bp. Synthetic long-read approaches may improve accuracy of indel detection from WES datasets (McCoy et al., 2014), but microarray and whole genome sequencing are currently more accurate alternatives for detecting long indels (Fang et al., 2014).

Whole-genome sequencing overcomes many of the limitations of WES. This method uses next-generation sequencing technology to determine the DNA sequence of the human genome. Unlike WES, whole-genome sequencing does not depend upon targeted enrichment. As a result, it is more uniform in genomic coverage and is less biased in detecting nonreference alleles (Meynert et al., 2014). Due to improved coverage uniformity, whole-genome sequencing requires a mean depth of only 14 reads to achieve 95% sensitivity, whereas WES requires a mean on-target depth of 40 reads to reach this threshold (Meynert et al., 2014). Thus, at a given read depth, whole genome sequencing is more likely to detect a genomic variant than WES. Whole-genome sequencing has been used to study autism, CHD, and other complex diseases (Michaelson et al., 2012; Chaiyasap et al., 2014). It is able to identify potential disease-causing variants both within and between genes and is more accurate than WES in detecting structural variation, such as insertions or deletions (Fang et al., 2014). With the release of the Illumina HiSeq X Ten system in 2014, the cost of sequencing the whole genome at 30× depth has dropped to approximately $1000 per sample (Meynert et al., 2014). An economic analysis by Meynert et al. (2014) recently demonstrated that the cost of sequencing the whole exome and the whole genome are roughly equivalent for institutions with access to a HiSeq X Ten system. Because access to this sequencing platform is limited, WES still remains in most cases a more affordable option for sequencing protein coding regions. A major consideration with whole genome sequencing is that it generates an enormous amount of data, which must be stored and analyzed. Analysis requires bioinformatics and statistical expertise that is not broadly available. There are also challenges in predicting the functional consequences of variants, especially those within the intragenic region (Kircher et al., 2014).

Identifying de novo mutations is another strategy for discovering rare causal variants. Samocha et al. (2014) recently described a framework for interpreting de novo mutations in which the number of de novo mutations within a gene or gene-set is compared with the expected number of mutations. In this framework, expected mutation rates are estimated based on local sequence context and selective constraint. Evolutionary constraint is not only informative within this context, but is also an important factor in prioritizing rare sequence variants. Several algorithms are available for estimating sequence conservation among species: GERP++, PhyloP, SiPhy (Liu et al., 2013). Sequence variants may also be prioritized using functional prediction algorithms, such as PolyPhen2, SIFT, MutationTaster, MutationAssessor, FATHMM, or LRT (Liu et al., 2013). Functional prediction methods are individually prone to false positives; therefore, it is prudent to evaluate the consensus among models (Tennessen et al., 2012). Software is also available to help predict the effects of genomic variation on RNA splicing; Human Splicing Finder represents one such tool (Desmet et al., 2009). Conservation- and functional-prediction tools are useful for prioritizing single nucleotide variation within the exon.

However, there are still major challenges to interpreting single nucleotide variation within noncoding regions (Khurana et al., 2013). Prediction models are also not well suited to evaluate the consequences of insertions and deletions.

FUNCTIONAL VALIDATION

Despite limitations, functional prediction models can be very helpful for population-based studies. Studies often identify multiple disease-associated variants. Functional predictions can help prioritize variants for further study. Likely causal variants may then be characterized in cell culture or in animal models. In studies where there is appropriate consent, lymphoblastoid cell lines can be generated from peripheral B lymphocytes. Cell lines possessing the candidate variant can be screened to detect phenotypic changes at a cellular level. In cases where a cell line from the proband is unavailable, gene editing may be used to introduce the candidate variant into an appropriate cell type. Several options are available for gene editing, including clustered regularly interspaced short palindromic repeats (CRISPRs), zinc-finger nucleases (ZFNs), and transcription activator–like effector nucleases (TALENs) (Gaj et al., 2013). The most recent addition to this group, CRISPR/Cas, has proven to be a powerful tool for sequence specific gene editing. CRISPR/Cas systems can be used to efficiently introduce a putative causal variant into model systems (Cong et al., 2013). CRISPR/Cas systems have been used to generate mice (Heckl et al., 2014), zebrafish (Hwang et al., 2013), and *Xenopus tropicalis* (Nakayama et al., 2013) models with targeted mutations.

Gene knockdown provides an alternative method of studying a candidate disease gene. Fakhro et al. (2011) used a *Xenopus tropicalis* gene knockdown system to identify human gene orthologs that are responsible for heterotaxy, which is a type of congenital heart disease caused by defects in left–right body patterning. A large excess of copy number variants in 61 genes had been discovered in human heterotaxy subjects compared with unaffected subjects. Twenty-two of these genes had Xenopus orthologs, and 7 of these were found to be expressed in the ciliated left–right organizer. Gene knockdown experiments with 5 of the genes resulted in left–right heart morphological anomalies, thereby validating their function in cardiac left–right patterning. Animal models provide much of the foundation for what we know about developmental processes. However, these models are not without their shortcomings. Animals and humans often have differences in the rate and production of birth defects related to environment, teratogens, and modifying factors. In addition, genes are not uniformly conserved between humans and models organisms. These differences can hinder translation of findings between animals and humans. Nonetheless, model systems are invaluable in the study of birth defects and will play an increasingly vital role in characterizing candidate disease-genes related to human birth defects.

EPIGENETICS AND NONSYNDROMIC BIRTH DEFECTS

Epigenetic modifications include DNA methylation, modification of histones, and DNA interactions with noncoding RNA. DNA methylation, the most studied epigenetic modification, constitutes an epigenetic mechanism whereby a methyl group is covalently bound to a cytosine base in the context of CpG dinucleotides. Methylation of cytosines in DNA has been implicated in the mechanism of silencing of gene expression, genomic imprinting, chromosomal stability and protection against repetitive element expression.

During pregnancy, folate-dependent nucleotide synthesis and DNA methylation are increased (Oommen et al., 2005), and altered DNA methylation may be an underlying mechanism in the development of birth defects (Okano et al., 1999; Li et al., 2005; Blom et al., 2006). Indeed, recent studies showed that altered DNA methylation is associated with NTD and CHD (Chen et al., 2010; Wang et al., 2010; Chowdhury et al., 2011a,b). DNA methylation is an attractive therapeutic target for congenital defects because maternal dietary supplementation may restore DNA methylation patterns and negate the hypomethylating effects of harmful maternal exposures such as bisphenol-A (Dolinoy et al., 2007). Various maternal factors (Waterland and Jirtle, 2003) implicated in abnormal fetal development have been shown to affect DNA methylation patterns (Cooney et al., 2002; Baccarelli et al., 2009; Candiloro and Dobrovic, 2009). Alterations in epigenetic phenomena, such as DNA methylation, are likely to play a crucial role in determining the fetal phenotype (Wolff et al., 1998). The combined effects of genetics and epigenetics in the intrauterine environment and subsequent fetal development are not well understood and warrant further investigation.

Animal studies have confirmed the importance of folic acid and folate metabolism in normal fetal growth and development. A homomorphic mutation in the mouse *MTRR* gene, which is necessary for usage of methyl groups from folate metabolism, resulted in developmental delay and congenital malformations, including neural tube and heart defects (Padmanabhan et al., 2013). Transgenerational effects of *MTRR* deficiency were also observed. When maternal grandparents were *MTRR* deficient, wild-type female grand progeny exhibited wide-spread genomic instability and altered placental gene expression, as well as an increased level of congenital malformations. These malformations persisted in wild-type progeny for five generations and were independent of maternal environment, suggesting transgenerational epigenetic inheritance, triggered by defects in *MTRR*.

In human studies, McKay et al. (2012) found that inter-individual differences in DNA methylation patterns at birth are influenced by environmental factors, such as maternal vitamin B12 levels, genetic factors such as infant *MTRR*

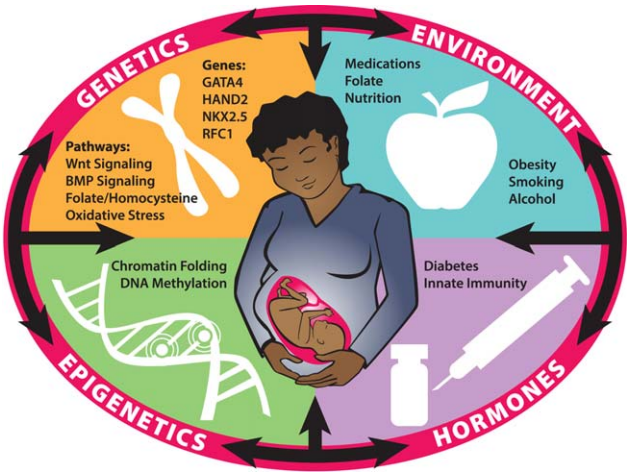

FIGURE 2. *Interactions between genetics, epigenetics, maternal hormonal levels, and environmental exposures during embryonic development. Disruptions or anomalies in these interrelated systems may perturb the precise developmental program, leading to an increased risk of structural birth defects. Environmental and lifestyle factors, such as obesity, cigarette smoke, alcohol, nutrient intake, and folate supplementation influence folate and homocysteine metabolism and may impact oxidative stress. Oxidative stress and increased homocysteine can affect signaling pathways that include critical developmental genes, such as GATA4, HAND2, NKX2.5, and RFC1. Aberrant folate or homocysteine metabolism can also drain the pool of methyl groups that is critical to maintaining gene expression levels through DNA- and histone-methylation. This reflects the multifactorial origin of most nonsyndromic birth defects.*

and maternal *MTHFR* genotypes, as well as length of gestation. These factors influence folate metabolism and affect the pool of methyl groups available for DNA methylation. Differences in methylation at the IGF2 locus, important in intrauterine growth (Börzsönyi et al., 2012), were associated with maternal MTHFR 677C>T polymorphism. They concluded that both global and gene-specific DNA methylation patterns in the developing fetus are dependent on genetic factors in both fetus and mother that influence folate metabolism, as well as the intrauterine environment (vitamin B12 levels). In a study of one pair of monozygotic twins who were discordant for renal agenesis, Jin et al. (2014) found no differences in SNPs or indels between the twins. They did, however, find 514 differentially methylated regions that were localized to 10 signaling pathways and 25 genes, including 6 genes that are known to be involved in organ development. These data implicate DNA methylation in the mechanism of organogenesis as well as in congenital malformations.

Recent work by Zaidi et al. (2013) suggests that aberrant histone modification may play a role in certain birth defects. This study compared the frequency of harmful de novo mutations among infants affected by CHD to healthy controls. Infants with CHD had a 7.5-fold excess of protein-altering de novo mutations (premature termination, frameshift, or splice site) in genes that are important in cardiac development. Notably, cases also had a significant excess of mutations in genes involved in histone modifications. Histone modifications influence gene transcription throughout life and are especially important in regulating developmental genes.

Most research concerning congenital malformations has focused on maternal or fetal genetics and/or epigenetics. The paternal genetic component of human birth defects has not been sufficiently studied. Some recent evidence has emerged from experimental work to suggest paternal genetics contribute to risk of birth defects in offspring. For example, Lambrot et al. (2013) determined that paternal dietary folate deficiency increases birth defects in offspring. It was demonstrated that a folate deficient diet alters the sperm DNA methylation at loci associated with genes responsible for normal development and disease. Also, sperm histone methylation was altered in folate deficient males, suggesting that dietary insufficiency of folate could alter gene expression. Folate deficient males had a significantly reduced pregnancy rate when mated to control females due to increased postimplantation resorption. An increase in the frequency of malformations was observed in offspring of folate deficient males, including hydrocephalus, limb defects and muscle or skeletal defects. These data suggest that male folate levels may also be important in the prevention of structural birth defects. Therefore, it may be useful to consider both maternal and paternal folate deficiency in future studies of birth defects.

Genetic Landscape: Future

As Figure 2 illustrates, most birth defects are multifactorial in origin. Specifically, maternal environment (e.g., medications, folate, nutrition, obesity, smoking, alcohol, etc.) interacts with maternal genetics, epigenetics, and hormones to influence various metabolic processes and signaling pathways. These factors shape the intrauterine environment and can interact with fetal genetics and epigenetics to either facilitate or disrupt embryogenesis. Recent developments have helped to better understand such risk factors. In short, epidemiology has identified environmental risk factors; genomics has provided an unprecedented tool for identifying genetic variants within an affected population; model systems and gene editing have been invaluable in studying development and disease; and new analytical approaches have provided a means of interpreting complex datasets. Aided by these tools, research has shifted from a gene-centered analysis to a systems-based analysis that examines the interactions among metabolic pathways and gene networks within the context of maternal and intrauterine environments. Much progress has been made; however, there is still much work to be done.

There have been successful reductions in some birth defects through research and population-based initiatives,

for example, folate supplementation and fortification; yet, birth defects continue to be a global public health problem. Developing countries may not have the resources to follow through with nutritional or environmental preventive measures when malnutrition remains all too common. Birth defects persist in developed countries because nutritional, genetic, epigenetic, environmental and unknown factors have not been fully investigated. Continued research in the areas of epidemiology, genetics, and epigenetics will lessen the global burden of birth defects through personalized and population-based preventative strategies.

Acknowledgments

The authors thank Ashley S. Block for assistance in the preparation of this manuscript. All authors have stated that they have no conflict of interest regarding this manuscript.

References

1000 Genomes Project Consortium. 2012. An integrated map of genetic variation from 1,092 human genomes. Nature 491:56–65.

Alkan C, Coe BP, Eichler EE. 2011. Genome structural variation discovery and genotyping. Nat Rev Genet 12:363–376.

Ardinger HH, Buetow KH, Bell GI, et al. 1989. Association of genetic variation of the transforming growth factor-alpha gene with cleft lip and palate. Am J Hum Genet 45:348–353.

Arrington CB, Bleyl SB, Matsunami N, et al. 2012. Exome analysis of a family with pleiotropic congenital heart disease. Circ Cardiovasc Genet 5:175–182.

Baccarelli A, Wright RO, Bollati V, et al. 2009. Rapid DNA methylation changes after exposure to traffic particles. Am J Respir Crit Care Med 179:572–578.

Beaty TH, Murray JC, Marazita ML, et al. 2010. A genome-wide association study of cleft lip with and without cleft palate identifies risk variants near MAFB and ABCA4. Nat Genet 42:525–529.

Birnbaum S, Ludwig KU, Reutter H, et al. 2009. Key susceptibility locus for nonsyndromic cleft lip with or without cleft palate on chromosome 8q24. Nat Genet 41:473–477.

Blom HJ, Shaw GM, den Heijer M, Finnell RH. 2006. Neural tube defects and folate: case far from closed. Nat Rev Neurosci 7:724–731.

Bodmer W, Bonilla C. 2008. Common and rare variants in multifactorial susceptibility to common diseases. Nat Genet 40:695–701.

Börzsönyi B, Demendi C, Pajor A, et al. 2012. Gene expression patterns of the 11β-hydroxysteroid dehydrogenase 2 enzyme in human placenta from intrauterine growth restriction: the role of impaired feto-maternal glucocorticoid metabolism. Eur J Obstet Gynecol Reprod Biol 161:12–17.

Butali A, Suzuki S, Cooper ME, et al. 2013. Replication of genome wide association identified candidate genes confirm the role of common and rare variants in PAX7 and VAX1 in the etiology of nonsyndromic CL (P). Am J Med Genet A 161:965–972.

Cai B, Zhang T, Zhong R, et al. 2014. Genetic variant in MTRR, but not MTR, is associated with risk of congenital heart disease: an integrated meta-analysis. PloS One 9:e89609.

Candiloro IL, Dobrovic A. 2009. Detection of MGMT promoter methylation in normal individuals is strongly associated with the T allele of the rs16906252 MGMT promoter single nucleotide polymorphism. Cancer Prev Res (Phila) 2:862–867.

Castilla EE, Orioli IM, Lopez-Camelo JS, et al. 2003. Preliminary data on changes in neural tube defect prevalence rates after folic acid fortification in South America. Am J Med Genet A 123:123–128.

Centers for Disease Control and Prevention. 2008. Update on overall prevalence of major birth defects–Atlanta, Georgia, 1978–2005. MMWR Morb Mortal Wkly Rep 57:1–5.

Chaiyasap P, Kulawonganunchai S, Srichomthong C, et al. 2014. Whole genome and exome sequencing of monozygotic twins with trisomy 21, discordant for a congenital heart defect and epilepsy. PloS One 9:e100191.

Chanock SJ, Manolio T, Boehnke M, et al. 2007. Replicating genotype–phenotype associations. Nature 447:655–660.

Chen X, Guo J, Lei Y, et al. 2010. Global DNA hypomethylation is associated with NTD-affected pregnancy: a case-control study. Birth Defects Res A Clin Mol Teratol 88:575–581.

Chowdhury S, Cleves MA, MacLeod SL, et al. 2011a. Maternal DNA hypomethylation and congenital heart defects. Birth Defects Res A Clin Mol Teratol 91:69–76.

Chowdhury S, Erickson SW, MacLeod SL, et al. 2011b. Maternal genome-wide DNA methylation patterns and congenital heart defects. PloS One 6:e16506.

Christensen K, Mitchell LE. 1996. Familial recurrence-pattern analysis of nonsyndromic isolated cleft palate–a Danish Registry study. Am J Hum Genet 58:182–190.

Christianson A, Howson CP, Modell B. 2005. March of Dimes: global report on birth defects, the hidden toll of dying and disabled children. March of dimes: global report on birth defects, the hidden toll of dying and disabled children.

Cong L, Ran FA, Cox D, et al. 2013. Multiplex genome engineering using CRISPR/Cas systems. Science 339:819–823.

Cooney CA, Dave AA, Wolff GL. 2002. Maternal methyl supplements in mice affect epigenetic variation and DNA methylation of offspring. J Nutr 132:2393S–2400S.

Cordell HJ, Bentham J, Topf A, et al. 2013a. Genome-wide association study of multiple congenital heart disease phenotypes identifies a susceptibility locus for atrial septal defect at chromosome 4p16. Nat Genet 45:822–824.

Cordell HJ, Topf A, Mamasoula C, et al. 2013b. Genome-wide association study identifies loci on 12q24 and 13q32 associated with tetralogy of Fallot. Hum Mol Genet 22:1473–1481.

Curtin JA, Quint E, Tsipouri V, et al. 2003. Mutation of *Celsr1* disrupts planar polarity of inner ear hair cells and causes severe neural tube defects in the mouse. Curr Biol 13:1129–1133.

Desmet FO, Hamroun D, Lalande M, et al. 2009. Human splicing finder: an online bioinformatics tool to predict splicing signals. Nucleic Acids Res 37:e67.

Dolinoy DC, Huang D, Jirtle RL. 2007. Maternal nutrient supplementation counteracts bisphenol A-induced DNA hypomethylation in early development. Proc Natl Acad Sci U S A 104:13056.

Dolk H, Vrijheid M, Scott JE, et al. 2004. Toward the effective surveillance of hypospadias. Environ Health Perspect 112:398–402.

Fahed AC, Gelb BD, Seidman JG, Seidman CE. 2013. Genetics of congenital heart disease: the glass half empty. Circ Res 112:707–720.

Fakhro KA, Choi M, Ware SM, et al. 2011. Rare copy number variations in congenital heart disease patients identify unique genes in left-right patterning. Proc Natl Acad Sci U S A 108:2915–2920.

Fang H, Wu Y, Narzisi G, et al. 2014. Reducing INDEL calling errors in whole genome and exome sequencing data. Genome Med 6:89.

Feng C, Zhang E, Duan W, et al. 2014. Association between polymorphism of TGFA Taq I and cleft Lip and/or palate: a meta-analysis. BMC Oral Health 14:88.

Fisch H, Lambert SM, Hensle TW, Hyun G. 2009. Hypospadias rates in New York State are not increasing. J Urol 181:2291–2294.

Frosst P, Blom HJ, Milos R, et al. 1995. A candidate genetic risk factor for vascular disease: a common mutation in methylenetetrahydrofolate reductase. 10:111–113.

Gaj T, Gersbach CA, Barbas CF III. 2013. ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. Trends Biotechnol 31:397–405.

Geller F, Feenstra B, Carstensen L, et al. 2014. Genome-wide association analyses identify variants in developmental genes associated with hypospadias. Nat Genet 46:957–963.

Gibson G. 2012. Rare and common variants: twenty arguments. Nat Rev Genet 13:135–145.

Gilboa SM, Salemi JL, Nembhard WN, et al. 2010. Mortality resulting from congenital heart disease among children and adults in the United States, 1999 to 2006. Clinical perspective. Circulation 122:2254–2263.

Glessner JT, Bick AG, Ito K, et al. 2014. Increased frequency of de novo copy number variants in congenital heart disease by integrative analysis of single nucleotide polymorphism array and exome sequence data. Circ Res 115:884–896.

Heckl D, Kowalczyk MS, Yudovich D, et al. 2014. Generation of mouse models of myeloid malignancy with combinatorial genetic lesions using CRISPR-Cas9 genome editing. Nat Biotechnol 32:941–946.

Hirschhorn JN, Lohmueller K, Byrne E, Hirschhorn K. 2002. A comprehensive review of genetic association studies. Genet Med 4:45–61.

Hobbs CA, Cleves MA, MacLeod SL, et al. 2014. Conotruncal heart defects and common variants in maternal and fetal genes in folate, homocysteine, and transsulfuration pathways. Birth Defects Res A Clin Mol Teratol 100:116–126.

Hoffman JIE, Kaplan S. 2002. The incidence of congenital heart disease. J Am Coll Cardiol 39:1890–1900.

Hopper J, Makalic E, Schmidt D, et al. 2012. 'Next-generation' genome wide association studies. Hered Cancer Clin Pract 10:A48.

Houk V, Oakley G Jr, Erickson JD, et al. 1992. Recommendations for the use of folic acid to reduce the number of cases of spina bifida and other neural tube defects. MMWR Morb Mortal Weekly Rep 41:1–7.

Hu Z, Shi Y, Mo X, et al. 2013. A genome-wide association study identifies two risk loci for congenital heart malformations in Han Chinese populations. Nat Genet 45:818–821.

Hwang WY, Fu Y, Reyon D, et al. 2013. Efficient genome editing in zebrafish using a CRISPR-Cas system. Nat Biotechnol 31:227–229.

Janerich DT, Piper J. 1978. Shifting genetic patterns in anencephaly and spina bifida. J Med Genet 15:101–105.

Jiang J, Zhang Y, Wei L, et al. 2014. Association between MTHFD1 G1958A polymorphism and neural tube defects susceptibility: a meta-analysis. PloS One 9:e101169.

Jin M, Zhu S, Hu P, et al. 2014. Genomic and epigenomic analyses of monozygotic twins discordant for congenital renal agenesis. Am J Kidney Dis. 64:119–122.

Junker R, Kotthoff S, Vielhaber H, et al. 2001. Infant methylenetetrahydrofolate reductase 677TT genotype is a risk factor for congenital heart disease. Cardiovasc Res 51:251–254.

Khurana E, Fu Y, Colonna V, et al. 2013. Integrative annotation of variants from 1092 humans: application to cancer genomics. Science 342:1235587.

Kircher M, Witten DM, Jain P, et al. 2014. A general framework for estimating the relative pathogenicity of human genetic variants. Nat Genet 46:310–315.

Lambrot R, Xu C, Saint-Phar S, et al. 2013. Low paternal dietary folate alters the mouse sperm epigenome and is associated with negative pregnancy outcomes. Nat Commun 4:2889.

Li M, Erickson SW, Hobbs CA, et al. 2014. Detecting maternal-fetal genotype interactions associated with conotruncal heart defects: a haplotype-based analysis with penalized logistic regression. Genet Epidemiol 38:198–208.

Li D, Pickell L, Liu Y, et al. 2005. Maternal methylenetetrahydrofolate reductase deficiency and low dietary folate lead to adverse reproductive outcomes and congenital heart defects in mice. Am J Clin Nutr 82:188–195.

Liu X, Jian X, Boerwinkle E. 2013. dbNSFP v2. 0: a database of human non-synonymous SNVs and their functional predictions and annotations. Hum Mutat 34:E2393–E2402.

Lu XC, Yu W, Tao Y, et al. 2014. Contribution of transforming growth factor alpha polymorphisms to nonsyndromic orofacial clefts: a HuGE review and meta-analysis. Am J Epidemiol 179: 267–281.

Ludwig KU, Mangold E, Herms S, et al. 2012. Genome-wide meta-analyses of nonsyndromic cleft lip with or without cleft palate identify six new risk loci. Nat Genet 44:968–971.

Lyons I, Parsons LM, Hartley L, et al. 1995. Myogenic and morphogenetic defects in the heart tubes of murine embryos lacking the homeo box gene Nkx2-5. Genes Dev 9:1654–1666.

Mathews TJ, MacDorman MF. 2013. Infant mortality statistics from the 2010 period linked birth/infant death data set. Vatl Vital Stat Rep 62:1–28.

McCoy RC, Taylor RW, Blauwkamp TA, et al. 2014. Illumina TruSeq synthetic long-reads empower de novo assembly and resolve complex, highly repetitive transposable elements. PLoS One 9: e106689.

McCulley DJ, Black BL. 2012. Transcription factor pathways and congenital heart disease. Curr Top Dev Biol 100:253–277.

McKay JA, Groom A, Potter C, et al. 2012. Genetic and non-genetic influences during pregnancy on infant global and site specific DNA methylation: role for folate gene variants and vitamin B12. PLoS One 7:e33290.

Meynert AM, Ansari M, FitzPatrick DR, Taylor MS. 2014. Variant detection sensitivity and biases in whole genome and exome sequencing. BMC Bioinformatics 15:247.

Michaelson JJ, Shi Y, Gujral M, et al. 2012. Whole-genome sequencing in autism identifies hot spots for de novo germline mutation. Cell 151:1431–1442.

Molkentin JD, Kalvakolanu DV, Markham BE. 1994. Transcription factor GATA-4 regulates cardiac muscle-specific expression of the alpha-myosin heavy-chain gene. Mol Cell Biol 14:4947–4957.

Moller J, Allen H, Clark E, et al. 1993. Report of the task force on children and youth. Circulation 88:2479–2486.

Nakayama T, Fish MB, Fisher M, et al. 2013. Simple and efficient CRISPR/Cas9-mediated targeted mutagenesis in Xenopus tropicalis. Genesis 51:835–843.

Okano M, Bell DW, Haber DA, Li E. 1999. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. Cell 99:247–257.

Oommen AM, Griffin JB, Sarath G, Zempleni J. 2005. Roles for nutrients in epigenetic events. J Nutr Biochem 16:74–77.

Oyen N, Poulsen G, Boyd HA, et al. 2009. Recurrence of congenital heart defects in families. Circulation 120:295–301.

Ozaki K, Ohnishi Y, Iida A, et al. 2002. Functional SNPs in the lymphotoxin-α gene that are associated with susceptibility to myocardial infarction. Nat Genet 32:650–654.

Padmanabhan N, Jia D, Geary-Joo C, et al. 2013. Mutation in folate metabolism causes epigenetic instability and transgenerational effects on development. Cell 155:81–93.

Parker SE, Mai CT, Canfield MA, et al. 2010. Updated national birth prevalence estimates for selected birth defects in the United States, 2004–2006. Am J Med Genet A 88:1008–1016.

Pierpont ME, Basson CT, Benson DW, et al. 2007. Genetic basis for congenital heart defects: current knowledge. Circulation 115: 3015–3038.

Porter MP, Faizan MK, Grady RW, Mueller BA. 2005. Hypospadias in Washington State: maternal risk factors and prevalence trends. Pediatrics 115:e495–9.

Rabbani B, Mahdieh N, Hosomichi K, et al. 2012. Next-generation sequencing: impact of exome sequencing in characterizing Mendelian disorders. J Hum Genet 57:621–632.

Rahimov F, Marazita ML, Visel A, et al. 2008. Disruption of an AP-2α binding site in an IRF6 enhancer is associated with cleft lip. Nat Genet 40:1341–1347.

Rosa F. 1983. Teratogenicity of isotretinoin. Lancet 322:513.

Samocha KE, Robinson EB, Sanders SJ, et al. 2014. A framework for the interpretation of de novo mutation in human disease. Nat Genet 46:944–950.

Sanders SJ, Murtha MT, Gupta AR, et al. 2012. De novo mutations revealed by whole-exome sequencing are strongly associated with autism. Nature 485:237–241.

Schnack TH, Zdravkovic S, Myrup C, et al. 2008. Familial aggregation of hypospadias: a cohort study. Am J Epidemiol 167:251–256.

Shaw GM, Lu W, Zhu H, et al. 2009. 118 SNPs of folate-related genes and risks of spina bifida and conotruncal heart defects. BMC Med Genet 10:49.

Shaw GM, O'Malley CD, Wasserman CR, et al. 1995a. Maternal periconceptional use of multivitamins and reduced risk for conotruncal heart defects and limb deficiencies among offspring. Am J Med Genet 59:536–545.

Shaw GM, Wasserman C, O'Malley C, et al. 1995b. Risks of orofacial clefts in children born to women using multivitamins containing folic acid periconceptionally. Lancet 346:393–396.

Southard AE, Edelmann LJ, Gelb BD. 2012. Role of copy number variants in structural birth defects. Pediatrics 129:755–763.

Tennessen JA, Bigham AW, O'Connor TD, et al. 2012. Evolution and functional impact of rare coding variation from deep sequencing of human exomes. Science 337:64–69.

van Beynum IM, Kouwenberg M, Kapusta L, et al. 2006. MTRR 66A> G polymorphism in relation to congenital heart defects. Clin Chem Lab Med 44:1317–1323.

van der Put, Nathalie MJ, Gabreëls F, et al. 1998. A second common mutation in the methylenetetrahydrofolate reductase gene: an additional risk factor for neural-tube defects?. Am J Hum Genet 62:1044–1051.

van der Zanden, Loes FM, van Rooij IA, et al. 2011. Common variants in DGKK are strongly associated with risk of hypospadias. Nat Genet 43:48–50.

Wamstad JA, Wang X, Demuren OO, Boyer LA. 2014. Distal enhancers: new insights into heart development and disease. Trends Cell Biol 24:294–302.

Wang L, Wang F, Guan J, et al. 2010. Relation between hypomethylation of long interspersed nucleotide elements and risk of neural tube defects. Am J Clin Nutr 91:1359.

Waterland RA, Jirtle RL. 2003. Transposable elements: targets for early nutritional effects on epigenetic gene regulation. Mol Cell Biol 23:5293.

Welter D, MacArthur J, Morales J, et al. 2014. The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. Nucleic Acids Res 42:D1001–6.

Wolff GL, Kodell RL, Moore SR, Cooney CA. 1998. Maternal epigenetics and methyl supplements affect agouti gene expression in Avy/a mice. FASEB J 12:949–957.

Yadav U, Kumar P, Yadav SK, et al. 2015. Polymorphisms in folate metabolism genes as maternal risk factor for Neural Tube Defects: an updated meta-analysis. Metab Brain Dis:1–18.

Yin M, Dong L, Zheng J, et al. 2012. Meta analysis of the association between MTHFR C677T polymorphism and the risk of congenital heart defects. Ann Hum Genet 76:9–16.

Yoon PW, Rasmussen SA, Lynberg MC, et al. 2001. The National Birth Defects Prevention Study. Public Health Rep 116:32.

Yu L, Wynn J, Cheung YH, et al. 2013. Variants in GATA4 are a rare cause of familial and sporadic congenital diaphragmatic hernia. Hum Genet 132:285–292.

Zaidi S, Choi M, Wakimoto H, et al. 2013. De novo mutations in histone-modifying genes in congenital heart disease. Nature 498: 220–223.

Zhu H, Yang W, Lu W, et al. 2012. Gene variants in the folate-mediated one-carbon metabolism (FOCM) pathway as risk factors for conotruncal heart defects. Am J Hum Genet 158:1124–1134.