Title
Genetic Epidemiology and Nonsyndromic Structural Birth Defects: From Candidate Genes to Epigenetics

Authors
Hobbs, Charlotte A
Chowdhury, Shimul
Cleves, Mario A
et al.

Peer reviewed

Genetic Epidemiology and Nonsyndromic Structural Birth Defects: From Candidate Genes to Epigenetics

Dr. Charlotte A. Hobbs, MD, PhD, **Dr. Shimul Chowdhury, PhD**, **Dr. Mario A. Cleves, PhD**, **Dr. Stephen Erickson, PhD**, **Dr. Stewart L. MacLeod, PhD**, **Dr. Gary M. Shaw, DPH**, **Dr. Sanjay J. Shete, PhD**, **Dr. John S. Witte, PhD**, and **Dr. Benjamin Tycko, MD, PhD**

Department of Pediatrics, University of Arkansas for Medical Sciences, Little Rock, AR (Drs. Hobbs, Chowdhury, Cleves, Erickson, MacLeod); Department of Pediatrics, Stanford University School of Medicine, Stanford, CA (Dr. Shaw); Department of Epidemiology, MD Anderson Cancer Center, Houston, TX (Dr. Shete); Institute for Cancer Genetics and Department of Pathology and Cell Biology, Columbia University Medical Center, New York, NY (Dr. Tycko); Department of Epidemiology and Biostatistics, University of California in San Francisco, San Francisco, CA (Dr. Witte)

Abstract

Birth defects are a leading cause of infant morbidity and mortality worldwide. The vast majority of birth defects are nonsyndromic, and although their etiologies remain mostly unknown, evidence supports the hypothesis that they result from the complex interaction of genetic, epigenetic, environmental, and lifestyle factors. Since our last review published in 2002 describing the basic tools of genetic epidemiology used to study nonsyndromic structural birth defects, many new approaches have become available and have been used with varying success. Through rapid advances in genomic technologies, investigators are now able to interrogate large portions of the genome at a fraction of previous costs. With next generation sequencing (NGS), research has progressed from assessing a small percentage of single nucleotide polymorphisms (SNPs) to assessing the entire human protein-coding repertoire (exome) – an approach that is starting to uncover rare but informative mutations associated with nonsyndromic birth defects. Here we report on the current state of genetic epidemiology of birth defects and comment on future challenges and opportunities. We consider issues of study design, and we discuss common variant approaches including candidate gene studies and genome-wide association studies (GWAS). We also discuss the complexities embedded in exploring gene-environment interactions. We complete our review by describing new and promising NGS technologies and examining how the study of

epigenetic mechanisms could become the key to unraveling the complex etiologies of nonsyndromic structural birth defects.

Research on birth defects in the post-genomic era

On April 14th, 2003, the Human Genome Project (HGP) was completed. To celebrate the 10th year anniversary, JAMA published a thematic issue on genomics on April 10th, 2013. In 2002, in a commentary published in the *Archives of Pediatrics and Adolescent Medicine*[1], we described the basic tools of genetic epidemiology as applied to the study of nonsyndromic structural birth defects. Since that time, many new approaches have become available to discover genetic factors leading to birth defects, and epigenetics has also come to the forefront as an important contributing factor. Here we present an updated review of genetic epidemiology and birth defects, focusing on these new approaches.

The prevalence of structural birth defects varies globally[2], ranging from approximately 3% to 6% of all live births. These birth defects are a leading cause of infant mortality, and are more prevalent than most chronic diseases of childhood such as autism[3], pediatric cancers[4], and Type I diabetes.[5] Most structural birth defects develop early in embryogenesis, during the first 10 weeks of pregnancy, and the majority of these defects occur in isolation affecting only one organ system. Despite recent progress in finding highly informative mutations in some cases a majority of nonsyndromic birth defects still does not appear to be accounted for by a single gene or chromosomal abnormality. The most prevalent defects are orofacial clefts, heart, neural tube, and limb defects.[6] When birth defects are not associated with known multi-organ syndromes; they are referred to as *nonsyndromic* defects. The etiologies of most nonsyndromic structural birth defects remain incompletely understood, and most are thought to result from a complex interplay between genetic, epigenetic, environmental, and lifestyle factors.[7] Maternal lifestyle factors, such as smoking, can alter developmental processes and expression of key developmental genes, such as *GATA4*, and the impact of environmental exposures and lifestyle on the developing fetus can in turn be influenced by maternal and fetal genetic susceptibilities.

While prior research on nonsyndromic birth defects has focused largely on the independent roles of environmental and/or lifestyle exposures and genetics[8], epigenetic causes have now begun to receive increased attention in this and many other human diseases. As early as 1940, Waddington, a British embryologist, geneticist, and philosopher of science, defined *epigenetics* as "…the interactions of genes with their environment which bring the phenotype into being".[9] Although studies on epigenetics and birth defects are still limited, they are very important as they may help to establish the molecular basis for gene-environment interactions.[8]

Embryogenesis requires an intricate coordination of cell migration, proliferation, and death that ultimately determines three-dimensional events in embryo formation and development. The complexity of embryogenic processes requires that multiple genes and biological pathways are involved in an intricate series of events that are susceptible to perturbations due to environmental exposures or maternal conditions[10]. In the following sections we will (a) review current approaches to identify genetic factors associated with birth defects, (b)

introduce potential epigenetic approaches, and (c) provide a perspective on the challenges and opportunities for future studies.

Candidate Gene Studies

Studies to identify etiologies of nonsyndromic birth defects in humans are for the most part limited to demonstrating association between genetic variants and birth defect phenotypes. Such association studies are most readily compared with traditional case-control epidemiologic designs. Population-based samples of cases and controls or small nuclear families (mother, father and child) or both are enrolled in a study in which the investigator compares the frequencies of specific genetic variants, polymorphisms, in affected and unaffected individuals. Such comparisons could include the index children as well as one or both parents.

Once it has been established that genetic variations impact the occurrence of a particular birth defect, studies are designed to identify and evaluate candidate genes. Candidate genes are those that have been proven or are believed to be associated with the malformation from animal studies, known or suspected developmental and signaling pathways, or from studying a small number of human pedigrees, and thus need to be confirmed in large, carefully selected, and well characterized human population samples. The selection of appropriate candidate genes is a key step in such studies. Knowledge of the pathogenesis of the malformation and the function of one or more proteins implicated in the disease can facilitate the identification of a suitable candidate gene. Genes that are known to be associated with variations in biologically relevant metabolic processes also may be selected as candidate genes. For example, several studies have demonstrated the protective effect of maternal periconceptional intake of folic acid on the occurrence of orofacial clefts[11], neural tube defects[12], and cardiac defects.[13] Subsequently several studies have investigated the association between genetic polymorphisms in the folate metabolic pathway and the risk of these three defect groups.[14–16]

Single nucleotide polymorphisms (SNPs) are the most common sources of variation in human genomes.[17] Each SNP is a difference in a single nucleotide at a specific site within the genome. For example, a SNP may form by substituting the nucleotide cytosine with the nucleotide thymine at a specific genomic location, annotated as C>T. If both of these variants (alleles) are compatible with life, then they can be present at detectable frequencies in the general population. In fact, we now know that there are approximately 7 million SNPs with a minor allele frequency (MAF) of over 5%.[18] Until completion of the International HapMap Project in 2009, population- or family-based studies of candidate genes were limited to the detection of associations between phenotypes of interest and only a small number of functional SNPs within candidate genes. By identifying a dense map of common variants and their correlations with each other (i.e. linkage disequilibrium), investigations such as the International HapMap Project and the 1000 Genomes Project (Figure 1) have made it possible to test for associations between birth defects and common (MAF>5% in the population) or even somewhat rarer genetic variants within each gene.[19,20] In the past 5 years, multiple reviews of candidate gene studies and birth defects have been published. Key pathways that have been implicated in the development of orofacial clefts and neural tube,

heart, and kidney defects include the Wnt signaling pathway, the BMP signaling pathway, the Hedgehog signaling pathway and variants in genes coding for key enzymes in the folate/homocysteine and oxidative stress pathways [21–26].

Genome Wide Association Studies (GWAS) and Copy Number Variants (CNVs)

Candidate gene or pathway approaches are limited by their reliance on preexisting knowledge of relevant pathways and/or mechanisms that may adversely impact embryogenesis. With the completion of the HapMap Project and the advances in technology that provide platforms to genotype large number of SNPs efficiently, GWAS was introduced as a popular method for studying common but genetically complex human diseases.[27] In contrast to candidate-gene studies, GWAS are not limited by prior knowledge but rather take an agnostic approach in which no SNP is considered, *a priori*, to have a higher likelihood of being associated with the relevant phenotype than any other SNP in the genome.[28] Between January 2005 and December 2012, over 1350 GWAS studies were reported. With varying success, GWAS have identified common SNPs associated with risks for specific pediatric or adult diseases, but only a few have been completed on birth defects, including one on hypospadias[29] and several on nonsyndromic cleft lip with or without cleft palate (NSCL/P). Mangold et al. identified two loci, in chromosome bands 17q22 and 10q25.3, which were associated with this type of birth defect,[30] while the other three studies identified strong associations between NSCL/P and a locus in 8q24.21.[31–33]

These findings support the utility of GWAS for identifying novel chromosomal regions associated with birth defects. However, most genetic studies of nonsyndromic birth defects continue to rely on candidate SNPs, leaving most of the genome unexplored.[34,35] Thus, there is a need to comprehensively explore the genome to identify new regions harboring genes associated with birth defects.[36] In addition to SNPs, which affect only a single nucleotide base, multiple lines of evidence indicate that copy number variants (CNVs) can play an important role in the etiology of some cases of birth defects. CNVs are defined as DNA sequences, ranging from kilobases to megabases in length, which are present in variable copy number in comparison to a reference genome.[37] In the past decade, molecular techniques such as array-based comparative genomic hybridization (aCGH), genotyping microarrays, and high throughput DNA sequencing have given a much richer picture of this form of genetic variation, and CNVs have increasingly been discovered to be associated with birth defects. For example, a survey of CNVs in 114 subjects with tetralogy of Fallot[38] and their unaffected parents identified 11 *de novo* CNVs that were absent or extremely rare in >2,000 controls, and pathogenic CNVs affecting the *GATA4* and *NODAL* genes have been found in more than one study of congenital heart disease.[39] Rare and/or *de novo* CNVs have also been implicated in microphthalmia,[40] congenital diaphragmatic hernia,[41] cleft lip and/or palate,[42] other craniofacial defects[43] and renal defects.[44]

For most GWAS, replication and validation of findings is necessary to separate true relationships from chance findings (Type 1 statistical errors). Genetic variants identified by GWAS are theoretically in linkage disequilibrium with functional variants that may be causal. To identify causal variants, targeted resequencing of genomic regions in close

proximity to selected SNPs, and whole- exome sequencing are new technological advances that hold considerable promise. In contrast to SNPs, which are often simply regional markers, CNV associations with birth defects are more likely to be directly causal of the phenotype being studied - via gene dosage increase or decrease, direct gene disruption, and cis-acting effects via disruption of gene regulatory sequences.[45]

Gene-environment Interactions

Genetic predispositions in conjunction with environmental influences are thought to be implicated in most birth defects. However, investigations into birth defects and interactions between functional or marker SNPs and environmental or lifestyle exposures are few.[8] This paucity of studies may in part reflect the difficulty of obtaining robust genetic and environmental measures on the same samples, and the increased statistical power needed to detect gene-environment interactions. Nonetheless, there have been some indications of an interaction between functional SNPs in the methyl donor pathway and the beneficial effects of folic acid supplementation against NTDs and other birth defects.[35]

Next Generation Sequencing (NGS) Technologies

The first human genome was sequenced over a period of 10 years using classical Sanger (dideoxy terminator) sequencing, at a cost of almost $3 billion.[46] "Next-generation" massively parallel sequencing (NGS) has exploded into the research and clinical genetics arena since its availability in 2005, with major advantages including markedly reduced sequencing time, reduced cost per nucleotide base, and substantial increases in data output.[47] With new technologies, a genome can be sequenced within days at a current cost less than $10,000. Reviews detailing NGS chemistries and the practical advantages of specific platforms have been published.[48–50] Data storage and protection, data analysis, clinical interpretation, and ethical issues, such as whether to report incidental, potentially adverse, findings to the patient or family, remain major challenges in the incorporation of NGS into the clinical arena.[51,52]

For the discovery and clinical testing of genes implicated in common birth defects such as CHDs, neural tube defects (NTDs), and cleft lip and palate (CL/P), different NGS-based approaches may be utilized. One approach is to select candidate genes for targeted deep resequencing. For conditions in which multiple genes with strong evidence of disease pathogenicity exist, such as those identified through GWAS and CNV studies, a targeted resequencing approach may be warranted. Although targeted NGS gene panels are offered clinically for certain conditions such as mitochondrial disorders,[51] they are not yet offered clinically for the major types of non-syndromic birth defects. Another approach is the interrogation of the exome (all gene exons, that is, all protein coding regions) of the genome. Although the human exome consists of only 1–2% of the entire genome, it is estimated that up to 85% of disease-causing mutations are harbored within it. The merits of sequencing the exome as a diagnostic tool have been delineated and exome sequencing is starting to aid in the diagnosis and treatment of inherited conditions.[48,52] Whole genome sequencing is also available, but to date the sequencing cost and the difficulties associated with data warehousing, analysis and interpretation remain prohibitive.

NGS technology holds promise for non-syndromic birth defects research and genetic diagnosis because 1) it allows the simultaneous analysis of the many candidate genes that have been identified so far; and 2) the technology is capable of detecting rare genetic variation. Although GWAS have identified many loci associated with complex traits, common genetic variation accounts for only a small percentage of heritability.[53] Thus, identification of rare genetic variation may yield larger effect sizes in complex diseases, including isolated birth defects.

The use of NGS for discovering causal mutations in non-syndromic birth defects is still in its early stages. NGS was used recently to show that de-novo mutations contribute to approximately 10% of severe CHD.[54] Strikingly, several genes encoding histone modifying and chromatin remodeling enzymes, such as *MLL2*, *CHD7*, and KDM5-A and –B, were found mutated in that study, implicating epigenetic alterations in CHD pathogenesis. Likewise, NGS approaches may become fruitful for finding mutations, such as those already known in the *HNF1B* and *PAX2* genes, which can underlie congenital kidney malformations.[55] Genome scanning methods for CNVs continue to identify lesions underlying CHD, some predisposing to it generally, and others with lesion specificity[56] and it can be expected that additional recurrent CNVs will detected in nonsyndromic birth defect cases using NGS.

Clinical Testing using NGS

NGS targeted resequencing of candidate genes, whole exome sequencing, and even whole genome sequencing are being offered clinically in a growing number of laboratories. The GeneTests website[57] and the NIH genetic testing registry[58] provide valuable resources for clinicians to identify clinical genetic testing laboratories. Recently, the merits, considerations, and challenges of using NGS technologies in a clinical diagnostic setting have been discussed in publications through the American College of Medical Genetics (ACMG) and the Association for Molecular Pathology (AMP).[59,60] Although resequencing assays are clinically available for complex diseases including cardiomyopathy,[61] autism spectrum disorder[62] and intellectual disability,[63] targeted assays using NGS are not yet available for non-syndromic birth defects such as CHDs, NTDs and CL/P. Whole exome sequencing is a potential genetic testing option for these conditions, but large research studies are only now being conducted in non-syndromic birth defects. Nonetheless, the apparent multigenic nature of non-syndromic birth defects makes NGS experimental approaches ideal for gene discovery and further pathogenic mutation characterization. Overall, the evidence accumulated through many dedicated research studies suggests that non-syndromic CHDs, NTDs, and CL/P may be caused by multiple rare, familial, genetic mutations, interacting with maternal genotype and exposures. Large population studies utilizing NGS may delineate genetic causes of common birth defects, and may aid in preconception risk and guide future therapeutic interventions.

Epigenetic Alterations in Non-syndromic Birth Defects

Despite the tremendous advances in human genetics enabled by the HGP and brought to fruition with GWAS and NGS, many aspects of human embryology and biology still cannot

be adequately explained by genetics alone. Normal embryogenesis requires the specification of a multitude of cell types/organs that depend on transcriptional regulation programmed by epigenetic mechanisms, namely modifications to DNA and its associated proteins that define the distinct gene expression profiles for individual cell types at specific developmental stages. Disruption of such control mechanisms is associated with a variety of diseases with behavioral, endocrine or neurologic manifestations, and quite strikingly with disorders of tissue growth, which will in all likelihood include structural birth defects. As a precedent, several well-studied syndromic birth defects, including Prader-Willi syndrome, Angelman syndrome, Beckwith-Wiedemann syndrome and Russell-Silver syndrome are known to be caused by loss of imprinting, uniparental disomy, or deletion/mutation of epigenetically regulated genes.[64]

An epigenetic trait can be defined as a "stably (somatically) heritable phenotype resulting from changes in a chromosome without alterations in the DNA sequence".[65] Several classes of epigenetic phenomena have been identified and recently reviewed.[8,66] Epigenetic patterns, essential for controlling gene expression in normal growth and development, are established by a number of mechanisms including DNA methylation at cytosine residues in CpG dinucleotides and covalent modifications of histone proteins, as well as by less well understood mechanisms controlling long-range chromatin architecture within the cell nucleus. DNA methylation involves the transfer of a methyl group to cytosine in a CpG dinucleotide, catalyzed by DNA methyltransferase enzymes that establish and maintain these patterns through cell division. Importantly DNA methylation has been shown to be essential for normal development.[66,67] DNMT1, the major maintenance methylase, has a high affinity for hemimethylated DNA[68] and it therefore acts to propagate methylation patterns in somatic cell divisions, while other enzymes, such as DNMT3A-DNMT3L, are responsible for initiating the epigenetic patterns. Importantly, there is already some evidence for an interaction of environment with epigenetics. Research on mice with a mutation in the *Agouti* gene has provided an excellent example of how maternal diet and epigenetics may affect fetal phenotypes.[69] This model has shown that variations in maternal dietary constituents affecting the methyl donor pool, such as folic acid, can result in alterations in coat color in the offspring, as a result of differential CpG methylation.[69,70] As another example, with clinical implications, a high folate diet given to H. felis-infected gastric cancer-prone mice at weaning prevented the development of gastric dysplasia and cancer.[71]

Many techniques, which we reviewed recently,[64] have been developed to study DNA methylation. The gold standard for comprehensive analysis of the methylation status of CpG sites, that is, DNA methylation patterns, is sodium bisulfite chemical conversion of DNA.[72] This procedure deaminates non-methylated deoxycytidine (dC) to deoxyuracyl (dU) residues; during subsequent PCR amplification, the latter are converted to A/T base pairs. However, if the C is methylated, the DNA sequence obtained after PCR does not change. Methylation-specific PCR or Pyrosequencing using bisulfite converted DNA provide quantitative measurements of DNA methylation levels, while another approach, which has now been made high throughput via NGS, involves amplification of bisulfite PCR products followed by sequencing of clones. This more thorough approach permits DNA methylation levels of a large number of contiguous CpG sites to be quantified, and the precise patterns of

methylation, including clonal heterogeneity and allele-specificity, to be displayed. By combining sodium bisulfite conversion and microarrays or NGS, genome-wide DNA methylation patterns can be determined essentially genome-wide. A number of initiatives have now been implemented to define human epigenetic patterns at high resolution with complete genomic coverage, with the goal of integrating epigenetics into the study of common but complex human diseases.[73] These new technologies and approaches may provide keys to unravel genetic and environmental factors that impinge on epigenomes to affect normal processes in embryological development and lead to human malformations when these processes go awry. In searching for evidence of "epigenome x environment" interactions in datasets from such studies, it will be crucial to take the cell types being sampled and the age of the subjects into account as methylation patterns are highly cell type-specific and, for some genes, can change with age.[74] It will also be important to remember that the genetic makeup of an individual exerts a strong influence on his/her epigenome: specifically, multiple studies have now shown a strong influence of human haplotypes, that is, clusters of SNPs in a given chromosomal region, on the patterns of DNA methylation in that region.[75]

The Future: linking genome, epigenome and environment in non-syndromic birth defects

We concluded our 2002 article by stating that identification of genes associated with birth defects does not lead to an immediate understanding of the relation between the gene and the birth defect with which it is associated. Identification "is only the first step in a long path to understanding the cause of the condition and ultimately to find preventive or corrective strategies.[1] As new technologies are made available, genetic epidemiologists are quick to utilize these new platforms to generate databases that seem to be ever-increasing in size. Our increased understanding of the importance of epigenetics in the development of birth defects suggests that an approach that simultaneously investigates genome-wide genetic and epigenetic variation in participants for whom environmental exposure data has been obtained may be a major step forward. Such studies may help to establish the mechanistic link between genetic variants and environmental exposures.

Acknowledgments

References

1. Hobbs CA, Cleves MA, Simmons CJ. Genetic epidemiology and congenital malformations: from the chromosome to the crib. Arch Pediatr Adolesc Med. 2002; 156(4):315–320. [PubMed: 11929362]

2. Christianson, A.; Howson, CBM. March of Dimes Global Report on Birth Defects: The hidden toll of dying and disabled children. White Plains; NY: 2006.

3. Prevalence of Autism Spectrum Disorders-Autism and Developmental Disabilities Monitoring Network, 14 Sites, United States, 2008. MMWR. 2012; 61(SS03):1–19.

4. United States Cancer Statistics: 1999–2005 Incidence and Mortality Web-based Report. Atlanta (GA): Department of Health and Human Services, Centers for Disease Control and Prevention, and National Cancer Institute; 2009.

5. The Burden of Diabetes Mellitus Among US Youth. Prevalence Estimates From the SEARCH for Diabetes in Youth Study. Pediatrics. 2006; 118(4):1510–1518. [PubMed: 17015542]

6. Parker SE, Mai CT, Canfield MA, et al. Updated National Birth Prevalence estimates for selected birth defects in the United States, 2004–2006. Birth Defects Res A Clin Mol Teratol. 2010; 88(12): 1008–1016. [PubMed: 20878909]

7. Hobbs CA, Cleves MA, Melnyk S, Zhao W, James SJ. Congenital heart defects and abnormal maternal biomarkers of methionine and homocysteine metabolism. Am J Clin Nutr. 2005; 81(1): 147–153. [PubMed: 15640474]

8. Friedman JM. Using genomics for birth defects epidemiology: can epigenetics cut the GxE Gordian knot? Birth Defects Res A Clin Mol Teratol. 2011; 91(12):986–989. [PubMed: 22140073]

9. Waddington, CH. Organisers and Genes. Cambridge: Cambridge University Press; 1940.

10. Srivastava D. Genetic regulation of cardiogenesis and congenital heart disease. Annu Rev Pathol. 2006; 1:199–213. [PubMed: 18039113]

11. Itikala PR, Watkins ML, Mulinare J, Moore CA, Liu Y. Maternal multivitamin use and orofacial clefts in offspring. Teratology. 2001; 63(2):79–86. [PubMed: 11241430]

12. van Beynum IM, den HM, Blom HJ, Kapusta L. The MTHFR 677C->T polymorphism and the risk of congenital heart defects: a literature review and meta-analysis. QJM. 2007; 100(12):743–753. [PubMed: 17965089]

13. Botto LD, Mulinare J, Erickson JD. Occurrence of congenital heart defects in relation to maternal mulitivitamin use. Am J Epidemiol. 2000; 151(9):878–884. [PubMed: 10791560]

14. Shaw GM, Lu W, Zhu H, et al. 118 SNPs of folate-related genes and risks of spina bifida and conotruncal heart defects. BMC Med Genet. 2009; 10:49. [PubMed: 19493349]

15. Pan Y, Zhang W, Ma J, et al. Infants' MTHFR polymorphisms and nonsyndromic orofacial clefts susceptibility: a meta-analysis based on 17 case-control studies. Am J Med Genet A. 2012; 158A(9):2162–2169. [PubMed: 22847888]

16. Yin M, Dong L, Zheng J, Zhang H, Liu J, Xu Z. Meta analysis of the association between MTHFR C677T polymorphism and the risk of congenital heart defects. Ann Hum Genet. 2012; 76(1):9–16. [PubMed: 22175539]

17. Hinds DA, Stuve LL, Nilsen GB, et al. Whole-genome patterns of common DNA variation in three human populations. Science. 2005; 307(5712):1072–1079. [PubMed: 15718463]

18. Kruglyak L, Nickerson DA. Variation is the spice of life. Nat Genet. 2001; 27(3):234–236. [PubMed: 11242096]

19. The International HapMap Project. Nature. 2003; 426(6968):789–796. [PubMed: 14685227]

20. Couzin J. Genomics. Consensus emerges on HapMap strategy. Science. 2004; 304(5671):671–673. [PubMed: 15118138]

21. Leslie E, Murray J. Evaluating rare coding variants as contributing causes to non-syndromic cleft lip and palate. Clin Genet. 2012

22. Greene ND, Stanier P, Copp AJ. Genetics of human neural tube defects. Hum Mol Genet. 2009; 18(R2):R113–129. [PubMed: 19808787]

23. Gelb B, Brueckner M, Chung W, et al. The Congenital Heart Disease Genetic Network Study: rationale, design, and early results. Circ Res. 2013; 112(4):698–706. [PubMed: 23410879]

24. Dixon MJ, Marazita ML, Beaty TH, Murray JC. Cleft lip and palate: understanding genetic and environmental influences. Nat Rev Genet. 2011; 12(3):167–178. [PubMed: 21331089]

25. Rana MS, Christoffels VM, Moorman AF. A molecular and genetic outline of cardiac morphogenesis. Acta Physiol (Oxf). 2013; 207(4):588–615. [PubMed: 23297764]

26. Tan, WH.; Gilmore, EC.; Baris, HN. Human Developmental Genetics. In: Rimoin, DL.; Pyeritz, RE.; Korf, BR., editors. Emery & Rimoin's Principles and Practice of Medical Genetics. 6. Philadelphia, PA: Elsevier Science; 2013.

27. National Human Genome Research Institute. [Accessed June 17, 2013.] Genome-Wide Association Studies. http://www.genome.gov/20019523

28. Hunter DJ, Chanock SJ. Genome-wide association studies and "the art of the soluble". J Natl Cancer Inst. 2010; 102(12):836–837. [PubMed: 20505151]

29. van der Zanden LF, van Rooij IA, Feitz WF, et al. Common variants in DGKK are strongly associated with risk of hypospadias. Nat Genet. 2011; 43(1):48–50. [PubMed: 21113153]

30. Mangold E, Ludwig KU, Birnbaum S, et al. Genome-wide association study identifies two susceptibility loci for nonsyndromic cleft lip with or without cleft palate. Nat Genet. 2010; 42(1): 24–26. [PubMed: 20023658]

31. Birnbaum S, Ludwig KU, Reutter H, et al. Key susceptibility locus for nonsyndromic cleft lip with or without cleft palate on chromosome 8q24. Nat Genet. 2009; 41(4):473–477. [PubMed: 19270707]

32. Beaty TH, Murray JC, Marazita ML, et al. A genome-wide association study of cleft lip with and without cleft palate identifies risk variants near MAFB and ABCA4. Nat Genet. 2010; 42(6):525–529. [PubMed: 20436469]

33. Grant SF, Wang K, Zhang H, et al. A genome-wide association study identifies a locus for nonsyndromic cleft lip with or without cleft palate on 8q24. J Pediatr. 2009; 155(6):909–913. [PubMed: 19656524]

34. Hobbs CA, Cleves MA, Keith C, Ghaffar S, James SJ. NKX2.5 and congenital heart defects: A population-based study. Am J Med Genet A. 2005; 134a(2):223–225. [PubMed: 15633194]

35. Hobbs CA, James SJ, Parsian A, et al. Congenital heart defects and genetic variants in the methylenetetrahydroflate reductase gene. J Med Genet. 2006; 43(2):162–166. [PubMed: 15951337]

36. Olshan AF, Hobbs CA, Shaw GM. Discovery of genetic susceptibility factors for human birth defects: an opportunity for a National Agenda. Am J Med Genet A. 2011; 155A(8):1794–1797. [PubMed: 21739590]

37. Redon R, Ishikawa S, Fitch KR, et al. Global variation in copy number in the human genome. Nature. 2006; 444(7118):444–454. [PubMed: 17122850]

38. Greenway SC, Pereira AC, Lin JC, et al. De novo copy number variants identify new genes and loci in isolated sporadic tetralogy of Fallot. Nat Genet. 2009; 41(8):931–935. [PubMed: 19597493]

39. Warburton D, Ronemus M, Kline J, et al. The contribution of de novo and rare inherited copy number changes to congenital heart disease in an unselected sample of children with conotruncal defects or hypoplastic left heart disease. Hum Genet. 2013

40. Bardakjian TM, Kwok S, Slavotinek AM, Schneider AS. Clinical report of microphthalmia and optic nerve coloboma associated with a de novo microdeletion of chromosome 16p11.2. Am J Med Genet A. 2010; 152A(12):3120–3123. [PubMed: 21082658]

41. Srisupundit K, Brady PD, Devriendt K, et al. Targeted array comparative genomic hybridisation (array CGH) identifies genomic imbalances associated with isolated congenital diaphragmatic hernia (CDH). Prenat Diagn. 2010; 30(12–13):1198–1206. [PubMed: 21064195]

42. Shi M, Mostowska A, Jugessur A, et al. Identification of microdeletions in candidate genes for cleft lip and/or palate. Birth Defects Res A Clin Mol Teratol. 2009; 85(1):42–51. [PubMed: 19137569]

43. Mefford HC, Shafer N, Antonacci F, et al. Copy number variation analysis in single-suture craniosynostosis: multiple rare variants including RUNX2 duplication in two cousins with metopic craniosynostosis. Am J Med Genet A. 2010; 152a(9):2203–2210. [PubMed: 20683987]

44. Weber S, Landwehr C, Renkert M, et al. Mapping candidate regions and genes for congenital anomalies of the kidneys and urinary tract (CAKUT) by array-based comparative genomic hybridization. Nephrol Dial Transplant. 2011; 26(1):136–143. [PubMed: 20605837]

45. Zhang F, Gu W, Hurles ME, Lupski JR. Copy number variation in human health, disease, and evolution. Annu Rev Genomics Hum Genet. 2009; 10:451–481. [PubMed: 19715442]

46. Finishing the euchromatic sequence of the human genome. Nature. 2004; 431(7011):931–945. [PubMed: 15496913]

47. Metzker ML. Sequencing technologies - the next generation. Nat Rev Genet. 2010; 11(1):31–46. [PubMed: 19997069]

48. Korf BR, Rehm HL. New approaches to molecular diagnosis. JAMA. 2013; 309(14):1511–1521. [PubMed: 23571590]

49. Liu L, Li Y, Li S, et al. Comparison of next-generation sequencing systems. J Biomed Biotechnol. 2012; 2012:251364. [PubMed: 22829749]

50. Loman NJ, Misra RV, Dallman TJ, et al. Performance comparison of benchtop high-throughput sequencing platforms. Nat Biotechnol. 2012; 30(5):434–439. [PubMed: 22522955]

51. Vasta V, Merritt JL 2nd, Saneto RP, Hahn SH. Next-generation sequencing for mitochondrial diseases: a wide diagnostic spectrum. Pediatr Int. 2012; 54(5):585–601. [PubMed: 22494076]

52. Bamshad MJ, Ng SB, Bigham AW, et al. Exome sequencing as a tool for Mendelian disease gene discovery. Nat Rev Genet. 2011; 12(11):745–755. [PubMed: 21946919]

53. Manolio TA, Collins FS, Cox NJ, et al. Finding the missing heritability of complex diseases. Nature. 2009; 461(7265):747–753. [PubMed: 19812666]

54. Zaidi S, Choi M, Wakimoto H, et al. De novo mutations in histone-modifying genes in congenital heart disease. Nature. 2013; 498(7453):220–223. [PubMed: 23665959]

55. Thomas R, Sanna-Cherchi S, Warady BA, Furth SL, Kaskel FJ, Gharavi AG. HNF1B and PAX2 mutations are a common cause of renal hypodysplasia in the CKiD cohort. Pediatr Nephrol (Berlin, Germany). 2011; 26(6):897–903.

56. Gelb BD. Recent advances in understanding the genetics of congenital heart defects. Curr Opin Pediatr. 2013

57. GeneTests. National Center for Biotechnology Information; http://www.genetests.org [Accessed May 14,2013.]

58. Rubinstein WS, Maglott DR, Lee JM, et al. The NIH genetic testing registry: a new, centralized database of genetic tests to enable access to comprehensive information and improve transparency. Nucleic Acids Res. 2013; 41(Database issue):D925–935. [PubMed: 23193275]

59. Genomics ACoMGa. ACMG Recommendations for Reporting of Incidental Findings in Clinical Exome and Genome Sequencing.

60. Schrijver I, Aziz N, Farkas DH, et al. Opportunities and challenges associated with clinical diagnostic genome sequencing: a report of the Association for Molecular Pathology. J Mol Diagn. 2012; 14(6):525–540. [PubMed: 22918138]

61. Norton N, Li D, Hershberger RE. Next-generation sequencing to identify genetic causes of cardiomyopathies. Curr Opin Cardiol. 2012; 27(3):214–220. [PubMed: 22421630]

62. Iossifov I, Ronemus M, Levy D, et al. De novo gene disruptions in children on the autistic spectrum. Neuron. 2012; 74(2):285–299. [PubMed: 22542183]

63. de Ligt J, Willemsen MH, van Bon BW, et al. Diagnostic exome sequencing in persons with severe intellectual disability. N Engl J Med. 2012; 367(20):1921–1929. [PubMed: 23033978]

64. Weksberg, R.; Butcher, DT.; Grafodatskaya, D.; Choufani, S.; Tycko, B. Epigenetics. In: Rimoin, DL.; Pyeritz, RE.; Korf, BR., editors. Emery & Rimoin's Principles and Practice of Medical Genetics. 6. Philadelphia, PA: Elsevier Sciences; 2013.

65. Berger SL, Kouzarides T, Shiekhattar R, Shilatifard A. An operational definition of epigenetics. Genes Dev. 2009; 23(7):781–783. [PubMed: 19339683]

66. Li E, Bestor TH, Jaenisch R. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. Cell. 1992; 69(6):915–926. [PubMed: 1606615]

67. Okano M, Bell DW, Haber DA, Li E. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. Cell. 1999; 99(3):247–257. [PubMed: 10555141]

68. Pradhan S, Bacolla A, Wells RD, Roberts RJ. Recombinant human DNA (cytosine-5) methyltransferase. I. Expression, purification, and comparison of de novo and maintenance methylation. J Biol Chem. 1999; 274(46):33002–33010. [PubMed: 10551868]

69. Wolff GL, Kodell RL, Moore SR, Cooney CA. Maternal epigenetics and methyl supplements affect agouti gene expression in Avy/a mice. FASEB J. 1998; 12(11):949–957. [PubMed: 9707167]

70. Cooney CA, Dave AA, Wolff GL. Maternal methyl supplements in mice affect epigenetic variation and DNA methylation of offspring. J Nutr. 2002; 132(8 Suppl):2393S–2400S. [PubMed: 12163699]

71. Gonda TA, Kim YI, Salas MC, et al. Folic acid increases global DNA methylation and reduces inflammation to prevent Helicobacter-associated gastric cancer in mice. Gastroenterology. 2012; 142(4):824–833.e827. [PubMed: 22248660]

72. Frommer M, McDonald LE, Millar DS, et al. A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. Proc Natl Acad Sci U S A. 1992; 89(5):1827–1831. [PubMed: 1542678]

73. Rakyan VK, Down TA, Balding DJ, Beck S. Epigenome-wide association studies for common human diseases. Nat Rev Genet. 2011; 12(8):529–541. [PubMed: 21747404]

74. Maegawa S, Hinkal G, Kim HS, et al. Widespread and tissue specific age-related DNA methylation changes in mice. Genome Res. 2010; 20(3):332–340. [PubMed: 20107151]

75. Tycko B. Allele-specific DNA methylation: beyond imprinting. Hum Mol Genet. 2010; 19(R2):R210–220. [PubMed: 20855472]

Genome-wide studies of genetic risk factors for birth defects have only recently been made possible due to the vast increase in genomic data first provided by the sequencing of the human genome by the Human Genome Project and subsequent studies that identified and measured genetic diversity among human subjects.

The Human Genome Project:
http://www.ornl.gov/sci/techresources/Human_Genome/project/about.shtml
  • **Goal:** To sequence the entire 3 billion bases in the human genome over a 13 year period. Project was a major international effort including 18 countries.
  • Private company, Celera Genomics sequenced human genome and published first draft sequence simultaneously with NIH effort in Science and Nature respectively.

The International HapMap Project:
http://hapmap.ncbi.nlm.nih.gov/
  • **Goal:** To identify and catalog patterns of human genetic variation. Project was a collaboration between institutions in US, Canada, Japan, China, Nigeria and the UK.

The 1000 Genomes Project:
http://www.1000genomes.org/
  • **Goal:** To identify genetic variants with frequencies of 1% or more in populations studied, which include individuals and families from diverse backgrounds around the world.

ENCODE (Encyclopedia of DNA Elements):
http://www.genome.gov/10005107
  • **Goal:** To identify all functional elements in the human genome.

CardioGenomics:
http://cardiogenomics.med.harvard.edu/pga-overview
  • **Goal:** To advance our understanding of congenital heart defects and genomics research by linking genes to structure, function, dysfunction and abnormalities.

FaceBase Consortium: https://www.facebase.org/
  • **Goal:** To provide craniofacial researchers with data repositories, technology and bioinformatics through a central data management and bioinformatics hub.

National Birth Defects Prevention Study:
http://www.nbdps.org/
  • **Goal:** To find genetic and environmental causes of nonsyndromic birth defects and identify prevention strategies. Over 45,000 case-/control-parental triads enrolled.

GENEVA: Gene Environment Association Studies:
http://www.genome.gov/27541319
  • **Goal:** to identify genetic variants related to complex disease, identify gene-trait variations associated with environmental exposures.

Figure 1.
Stepping stones: Projects that made genome-wide association studies (GWAS) possible.