

An opportunity to address the genetic causes of birth defects

Mustafa K. Khokha^{1,2}, Laura E. Mitchell³ and John B. Wallingford⁴

The following White Paper was submitted to the National Institutes of Health (NIH) to highlight the importance of birth defects on infant and child health as well as the need to address the challenges of birth defects research by fostering multidisciplinary approaches. A number of discussions including those at the January 2014 workshop, “Developing an Interdisciplinary Research Agenda for Genetics of Birth Defects,” and the December 2014 10th Structural Birth Defects Meeting culminated in this White Paper. These discussions combined the advice of many members of the birth defects research community including clinicians, epidemiologists, geneticists, and basic developmental biologists (see Acknowledgments and **Supplementary File S1** online). Subsequently, the Society for Pediatric Research Council, American Society for Human Genetics Executive Council, and the Teratology Society Council endorsed this White Paper as well as past and current Presidents of the American Pediatric Society and the Society of Developmental Biology. The following is this White Paper in its entirety including recommendations moving forward.

BIRTH DEFECTS HAVE A DISPROPORTIONATE IMPACT ON THE YOUNGEST MEMBERS OF SOCIETY

Approximately 8 million children are born with serious birth defects each year—roughly 6% of births worldwide (1). Importantly, an additional, untold number of birth defects lead to stillbirths and miscarriages, which can have a devastating effect on parents hoping to start a family (2,3). The impact of birth defects for live-born children is profound. In the United States, during the first year of life, birth defects are the leading cause of pediatric hospitalizations (4), medical expenditures (5), and death (6). Further, birth defects rank as a leading cause of death among children aged 1–4 y (#2 cause of death), 5–14 y (#3), and 15–24 y (#6) (7). Hence, for the youngest members of our population, birth defects are a leading healthcare concern. While advances in surgical outcomes are improving the survival of these children, the impact of long-term sequelae remains significant. For example, brain abnormalities result in intellectual disability; structural heart defects are detrimental to exercise tolerance; abnormal limbs dramatically reduce fine motor function in the workplace; and craniofacial malformations affect individual identity and can lead to substantial psychosocial issues.

BIRTH DEFECTS ARE ASSOCIATED WITH INCREASED RATES OF CHILDHOOD CANCER

Although many birth defect-related morbidities are phenotype specific (e.g., children with spina bifida are at risk for hydrocephalus, while those with congenital heart defects are at risk for cardiac fibrosis), some morbidities (e.g., neurodevelopmental deficits) are common across a number of different birth defect phenotypes. Notably, there is a strong association between birth defects and childhood cancers. Children with chromosomal birth defects appear to be at the highest risk (10–15-fold) for developing childhood cancer, but even those with nonchromosomal birth defects seem to have a higher rate of childhood cancer (two- to threefold) than children without birth defects (8–10). In addition, many developmental pathways that can cause birth defects also have clear roles in the development of cancer. Hence, knowledge gained about the causes of birth defects may also elucidate the etiology of childhood cancers, which are rarer than birth defects and, therefore, generally more difficult to study. The NIH-funded Gabriella Miller Kids First Pediatric Research Program is an exciting opportunity to explore this relationship between birth defects and cancer.

THE GENETIC CAUSES OF BIRTH DEFECTS ARE DISCOVERABLE BUT CURRENTLY MOSTLY UNKNOWN

If we are going to make an impact on birth defects, research efforts **MUST** identify the underlying causes of these conditions. Efforts to identify environmental factors associated with birth defects (such as toxin exposure and nutrition) have had some successes (e.g., folic acid) and warrant further investigation. It is clear, however, that genetic factors play a major role in birth defect etiology, either as a primary cause or as a modifier to environmental exposures. The current explosion in DNA sequencing technologies offers an unprecedented opportunity to discover these genetic factors. Unfortunately, the application of these technologies to birth defects has lagged significantly behind its application to other disease processes, despite the extraordinary potential for these approaches to reveal the genetic etiology of birth defects and thereby impact treatment. Hence, there is a pressing need to engage large-scale efforts to identify the causes of birth defects and develop prevention and treatment strategies.

¹Program in Vertebrate Developmental Biology, Department of Pediatrics, Yale University School of Medicine, New Haven, Connecticut; ²Program in Vertebrate Developmental Biology, Department of Genetics, Yale University School of Medicine, New Haven, Connecticut; ³Division of Epidemiology, Human Genetics and Environmental Sciences, University of Texas School of Public Health, Houston, Texas; ⁴Department of Molecular Biosciences, Center for Systems and Synthetic Biology and Institute for Cellular and Molecular Biology, University of Texas at Austin, Austin, Texas. Correspondence: Mustafa K. Khokha (Mustafa.khokha@yale.edu)

Received 6 April 2016; accepted 31 August 2016; advance online publication 7 December 2016. doi:10.1038/pr.2016.229

UNDERSTANDING THE GENETICS OF HUMAN BIRTH DEFECTS WILL DIRECTLY BENEFIT PATIENTS

Despite their personal and public health impact, the causes of most birth defects are unknown. Understanding the genetic or epigenetic variants causing birth defects will have a number of benefits. We can improve the quality of genetic counseling and prognostic information. Because recurrence risks for many conditions are currently based on empiric estimates (e.g., the risk of recurrence of spina bifida is approximately 3%), identification of the genetic cause will allow for a more precise estimation of recurrence risk (e.g., low risk of recurrence for a *de novo* mutation; a predictable high risk of recurrence for a recessive condition).

Identification of specific birth defect-related genetic variations would also provide the basis for functional analyses that may improve potential therapies for these patients. Currently, most specific birth defects (i.e., spina bifida) are lumped together as single diseases. However, clinically we know that the outcomes of these phenotypically-similar patients can be dramatically different. Whether such differences are genetically-mediated remains a critical open question. By identifying birth defect-related genetic variations, and improving our understanding of the function of these sequence variations, we may identify subsets of patients who are at high risk for specific adverse outcomes. Such knowledge would allow us to *personalize* the clinical management of patients with birth defects and improve the *precision* of our care based on gene function. In the end, without an understanding of birth defect related genetic variations, improving birth defect therapy is inevitably hampered.

CHALLENGES TO BIRTH DEFECTS RESEARCH THAT NEED TO BE OVERCOME

To realize the benefits of genomic approaches to the study of birth defects, there are a number of challenges to overcome:

Individual Birth Defect Phenotypes are Rare

Although birth defects as a group are quite common, this group is comprised of many different conditions that are etiologically distinct and individually rare. Consequently, enrollment of cohorts of sufficient size to identify common genetic variants, which generally have small effects on disease risk, and identify variants of larger effect, which tend to be rare, can be challenging even for large clinical centers. For this reason, we need to develop a collaborative system to enroll patients with birth defects; collect, record, and store phenotypic information in a standardized database; and, collect DNA samples from case–parent trios and unaffected siblings as well as any available tissue samples.

While there are many existing birth defects-specific cohorts, the majority are small ($n < 1,000$) and, thus, under-powered for sequence-based studies. Further, many of the existing cohorts were recruited years ago and/or outside of the United States and may not be representative of current cases (e.g., the relative contribution of genetic risk factors may have shifted after folic acid fortification in the United States). In addition, many of

the existing cohorts are not optimal for birth defects research (based on case–control designs rather than case–parent trios) and/or lack detailed phenotype information collected using a uniform nomenclature. Finally, many of these cohorts were collected under consents that did not include current genome sequence sharing rules making them ineligible for many NIH-sponsored research initiatives. Therefore, the prospective recruitment of large cohorts with systematically detailed phenotypic information including prenatal exposure information and optimal study designs (trios/quads) are necessary for birth defects research to advance. Importantly, with these large cohorts in hand, we can identify second alleles which when combined with functional studies (see below) can provide critical evidence for causality.

Timely collection of appropriately sized study cohorts will require a national recruitment network, with select Clinical Centers coordinating the activities of a larger number of regional patient enrollment centers that will ascertain study subjects and systematically collect detailed phenotype descriptions and other relevant data. Patient enrollment will differ from the typical approach, which is often condition-specific (e.g., cases with heart defects enrolled in the Pediatric Cardiac Genomics Consortium (11)), in that each center will enroll the full spectrum of birth defect phenotypes. Hence, the network will build a resource of the complete array of birth defects, subsets of which could be analyzed by investigators with the necessary interest and expertise.

Critical components for patient enrollment include: (i) a consent process that allows for sharing of de-identified clinical and genomic data; (ii) dedicated nurse/counselors to recruit patients and family members, conduct interviews for prenatal exposures and medical/family history and collect blood samples as well as any available tissue samples including surgical samples; (iii) clinicians with the necessary expertise to evaluate phenotypic features; (iv) a standardized nomenclature for describing each birth defect phenotype; (v) a secure database to house phenotype data that can be shared across different types of computer platforms; (vi) a central biorepository for sample processing, storage and distribution; and (vii) community engagement by large and small centers to recruit patients. Further, in anticipation of falling costs, we must plan for patient and family enrollment to exceed our ability to sequence cases at current rates. Creating a comprehensive infrastructure that can readily recruit patients and family members is essential. Fortunately, the NIH already has models for such infrastructure in the Rare Disorders Consortia, Clinical and Translational Science Award program as well as the Pediatric Cardiac Genomics Consortium.

Genetic Analysis of Birth Defects Will Require the Latest, Most Cost-Effective Sequencing Technologies

The revolution in human genomics offers many exciting opportunities for genetic variant discovery in patients with birth defects. However, to be effective, sequencing quality must be high and comparable data must be available for appropriately matched controls. Therefore, centralized exome/genome

sequencing may have substantial advantages both in terms of throughput and quality. In addition, since DNA from patients is limited, sequence information needs to be publicly available, to avoid resequencing the same patient without justifiable improvements in sequence quality or coverage. The Centers for Mendelian Genomics offers models for the necessary genomic resources. Finally, a wealth of data can be generated from patient samples including epigenetic modifications, transcriptome, and proteomic data that may be informative in the disease process; strategies to generate and share these data are essential to elucidate birth defect disease mechanisms.

Functional Studies in Animal Models are Essential to Test Variants and Discover Underlying Etiologies

Birth defects are fundamentally a problem in embryonic development. Surprisingly, at least in the case of congenital heart disease, most of the genes with identified variants have no known role in cardiac development, no known role in embryonic development, and many are completely novel (11–14). In addition, very few putative disease-related alleles currently have been identified in more than one case and the functional impact of many such variants has not been established. Therefore, functional testing in model systems is needed: (i) to bolster the evidence that these genes/genetic variations may be disease causing; and (ii) to investigate the embryological mechanisms underlying the disease-association. In particular, low cost, high-throughput, *in vivo*, animal models of human disease are necessary to screen variants for functional effects. Additionally, cell culture systems, including iPSCs from patients and control family members, can prove highly informative. High-throughput strategies are necessary to screen both genic and intergenic sequences for disease phenotypes as well as initiate projects to discover underlying developmental mechanisms. Screening novel birth defect candidate genes in animal models is relatively risky and challenging; therefore, funding mechanisms to foster this work are essential to generate the necessary preliminary data for further functional studies.

We Must Create a Research Infrastructure That Will Immediately Benefit Patients

The potential to provide research results to patients and families affected by birth defects is an especially exciting outcome of genomic studies. As these studies will necessarily be focused on children, discoveries have the potential to make a significant impact on the lives of these children and their families. Therefore, for maximal impact, we need to create the infrastructure to return results to patients, their families, and their physicians, especially in cases where disease causality can be established. This needs to be built into the consent process. We recognize that phenotypes in model systems may be insufficient proof to alter treatment or even to assign disease causality. However, many patients may find this information useful, and it will provide a critical opportunity for the research community to “give back” to the patients and families who inspire these studies. For patient and families that crave

as much information about their disease process as possible, frank discussions about the available evidence and their relative strength for disease causality and management are critical.

Birth Defects Research Requires Building Collaborative Interdisciplinary Teams

For birth defects research to succeed, cementing collaborative science between clinicians, genomicists, geneticists, epidemiologists, biostatisticians/bioinformaticians, and basic developmental/cell biologists is essential (Figure 1). Unfortunately, in many institutions, these different groups operate in different “silos” that slow research progress. By creating a collaborative environment, birth defects research has considerable potential for accelerating research especially in the current human genomics era.

In addition to investigator collaboration, we need collaboration between different Institutes at the NIH. Birth defects affect organ systems that span the NIH mission. Many of the resources proposed here can only be effectively implemented if each Institute participates to create the resources needed. In so doing, the Institutes ensure that these birth defect resources will go on to benefit the mission of each Institute in the future. Therefore, trans-NIH initiatives are essential for birth defects research to grow. The Trans-NIH Structural Birth Defects Working Group established in 2014 should be tasked with taking the lead on coordinating funding opportunity announcements that will foster the efforts proposed here.

Finally, we need to train the next generation of birth defects researchers to become adept with human phenotypes, big data, and functional experiments. This training must incorporate connections between patients, genomic analyses, and cellular/

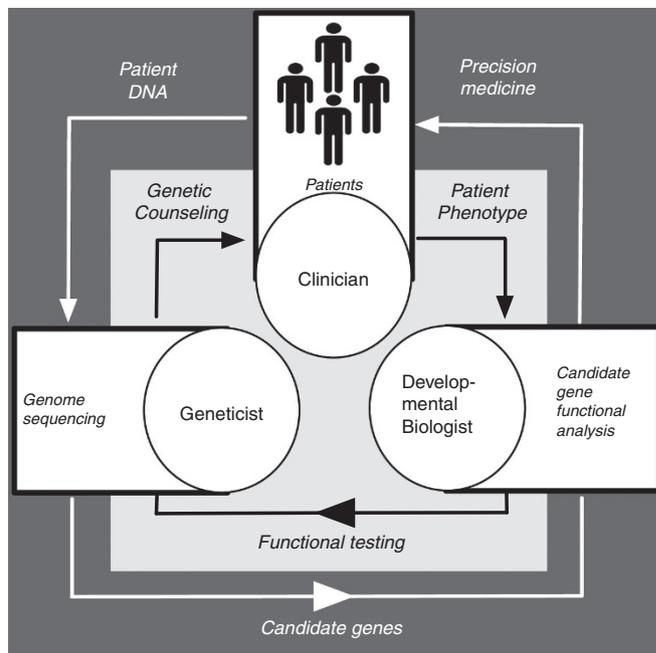


Figure 1. Multidisciplinary approach in birth defects research. Interrelationships between clinicians, geneticists, and basic biologists are outlined schematically.

developmental biology to foster the necessary collaborations that will be essential for success in the future. This support must come at all levels—training in graduate and postdoctoral levels, mentored clinical scientist awards, as well as transition awards towards independence. Opportunities for institutions to create training grants that foster a multi-disciplinary training environment are essential as well as course work that spans disciplines.

What is needed:

1. Trans-NIH collaborative initiatives—birth defects affect the breadth of the NIH Institutes' missions.
2. Funding opportunity announcements with allocated resources to recruit birth defect patients, record and store detailed objective, quantitative, and comprehensive phenotype information, and create a biorepository for DNA and tissues. Recruiting Centers should span the nation in order to be comprehensive and continue to collect outcome data, which can then be integrated with genotype information.
3. Funding opportunity announcements creating sequencing centers to generate sequence, distribute sequence data, provide bioinformatic support, and develop analysis methods that are more suited towards children (i.e., trio-based methods including analysis of maternal genetic effects, parent of origin effects, gene-environment interactions).
4. Infrastructure to make sequence information publicly available that is particularly geared towards children and birth defects research. Infrastructure needs to accommodate "big data" (genome, epigenome, transcriptome, and proteome data) to enable insights into birth defects and pediatric cancers.
5. Funding opportunity announcements for functional testing of human variants, either genic or nongenetic (e.g., gene regulatory elements). They should foster collaborative projects where genetic analysis and functional testing are tightly integrated within a project in order to maximize research potential. Funding mechanisms must include pilot projects for high-throughput screening of genetic variants for functional effects with further support for preliminary characterization of novel genes and nongenetic variants in order to discover their developmental mechanisms.
6. Mechanisms to return research results to patients, families, and clinical providers.
7. Training mechanisms with a specific focus on interdisciplinary training. Specifically focus on breaking down barriers (e.g., developmental biology and genomics training for pediatricians, human genetic and epidemiology training for PhD developmental biologists, etc.) and foster multidisciplinary interactions and training (conferences and coursework).

SUPPLEMENTARY MATERIAL

Supplementary material is linked to the online version of the paper at <http://www.nature.com/pr>

ACKNOWLEDGMENTS

The authors wish to acknowledge helpful discussion that occurred at the January 2014 workshop, "Developing an Interdisciplinary Research Agenda for Genetics of Birth Defects" and the December 2014 10th Structural Birth Defects Meeting as well as comments from the following members of the 11th Structural Birth Defects Steering Committee: Nadav Ahituv, Daniel Hart, Ethylin Jabs, Ian Krantz, Catherine Keegan, Joseph Levy, Bernice Morrow, and Wendy Nembhard.

STATEMENT OF FINANCIAL SUPPORT

No financial assistance was received to support this study.

Disclosure: The authors declare no competing financial interests.

REFERENCES

1. Christianson A, Howson CP, Modell B. March of Dimes global report on birth defects. White Plains, New York: March of Dimes Research Foundation; 2006.
2. Gregory EC, MacDorman MF. Fetal and Perinatal Mortality: United States, 2013. National vital statistics reports: from the Centers for Disease Control and Prevention, National Center for Health Statistics, National Vital Statistics System. 2015;64:1–24.
3. Stillbirth Collaborative Research Network Writing Group. Causes of death among stillbirths. *JAMA* 2011;306:2459–68.
4. Yoon PW, Olney RS, Khoury MJ, Shappanfield WM, Chavez GF, Taylor D. Contribution of birth defects and genetic diseases to pediatric hospitalizations. A population-based study. *Arch Pediatr Adolesc Med* 1997;151:1096–103.
5. Waitzman NJ, Romano PS, Scheffler RM. Estimates of the economic costs of birth defects. *Inquiry* 1994;31:188–205.
6. Martin JA, Kochanek KD, Strobino DM, Guyer B, MacDorman MF. Annual summary of vital statistics—2003. *Pediatrics* 2005;115:619–34.
7. United States Department of Health and Human Services. Deaths: Final Data for 2013. (Table 10). (http://www.cdc.gov/nchs/data/nvsr/nvsr64/nvsr64_02.pdf).
8. Fisher PG, Reynolds P, Von Behren J, Carmichael SL, Rasmussen SA, Shaw GM. Cancer in children with nonchromosomal birth defects. *J Pediatr* 2012;160:978–83.
9. Botto LD, Flood T, Little J, et al. Cancer risk in children and adolescents with birth defects: a population-based cohort study. *PLoS One* 2013;8:e69077.
10. Carozza SE, Langlois PH, Miller EA, Canfield M. Are children with birth defects at higher risk of childhood cancers? *Am J Epidemiol* 2012;175:1217–24.
11. Gelb B, Brueckner M, Chung W, et al.; Pediatric Cardiac Genomics Consortium. The Congenital Heart Disease Genetic Network Study: rationale, design, and early results. *Circ Res* 2013;112:698–706.
12. Fakhro KA, Choi M, Ware SM, et al. Rare copy number variations in congenital heart disease patients identify unique genes in left-right patterning. *Proc Natl Acad Sci USA* 2011;108:2915–20.
13. Glessner JT, Bick AG, Ito K, et al. 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* 2014;115:884–96.
14. Zaidi S, Choi M, Wakimoto H, et al. De novo mutations in histone-modifying genes in congenital heart disease. *Nature* 2013;498:220–3.