Applying Deep DNA Sequencing to Common, Complex Pediatric Traits

AUTHORS: John M. Dagle, MD, PhD,^{a,b} and Jeffrey C. Murray, MD^{a,c,d,e}

Departments of ^aPediatrics, ^bBiochemistry, ^cBiology, ^dNursing, and ^eEpidemiology, University of Iowa, Iowa City, Iowa

ABBREVIATION

RDS—respiratory distress syndrome

Opinions expressed in these commentaries are those of the author and not necessarily those of the American Academy of Pediatrics or its Committees.

www.pediatrics.org/cgi/doi/10.1542/peds.2012-2870

doi:10.1542/peds.2012-2870

Accepted for publication Sep 25, 2012

Address correspondence to Jeff Murray, MD, 2182 MedLabs, Department of Pediatrics, University of Iowa, Iowa City, IA 52242. E-mail: jeff-murray@uiowa.edu

PEDIATRICS (ISSN Numbers: Print, 0031-4005; Online, 1098-4275).

Copyright © 2012 by the American Academy of Pediatrics

FINANCIAL DISCLOSURE: The authors have indicated they have no financial relationships relevant to this article to disclose.

FUNDING: Supported by National Institutes of Health, March of Dimes, and Burroughs Wellcome Foundation. Funded by the National Institutes of Health (NIH).

COMPANION PAPER: A companion to this article can be found on page e1575, and online at www.pediatrics.org/cgi/doi/10.1542/peds.2012-0918.

The specific genetic factors contributing to the causes of complex pediatric diseases are gradually being identified through a combination of technological advances, advanced statistical methods, and large biorepositories with appropriate samples and data. These developments have the capacity to personalize diagnosis and treatment of the common complications faced by preterm infants. The article in this issue of *Pediatrics* by Wambach et al¹ is illustrative of this transformation and was applied to respiratory distress syndrome (RDS), the most common respiratory morbidity associated with prematurity. The authors considered RDS as a complex disease with both genetic and environmental/developmental risk factors. By excluding infants of <34 weeks' gestation, a sample cohort was studied that was enriched for genetic causes of RDS and depleted of some developmental risk factors. They examined several genes previously associated with Mendelian single-gene (recessive usually) contributors to RDS that may be severe even in the term neonate. They searched for heterozygotes of known mutations or more minor variants that would result in a less-extreme phenotype but still be significant contributors to RDS in the near-term infant. Their finding that *ABCA3* mutations are overrepresented in preterm infants with RDS compared with those without RDS is consistent with a growing body of literature demonstrating that a non-Mendelian complex disease may be caused by alterations in the same gene responsible for a more severe syndromic, single-gene effect, presentation. Although complex traits are usually thought of as arising from the interplay of many genetic and environmental factors, each of small-effect, current work suggests that recently arising rare variants may individually have a substantial impact on a disease phenotype.² These effects are driven by what is a presumably lower dose of a "genetic abnormality." Although they focused on coding sequence variants to assist in determining causality by using functional modeling, the study strongly suggests that future studies should also focus on regulatory regions for *ABCA3* where the protein structure may be intact but dosage altered. If these results are confirmed, it will allow for more accurate diagnosis and estimation of recurrences in subsequent pregnancies with perhaps altering plans for delivery hospital or intensity of monitoring of term or near-term infants for RDS after birth.

Germline contributions to the inherited components of disease can involve searches for common single-nucleotide variants using genome-wide association, copy number variants using array technologies, or rare sequence variants using sequencing.³ The introduction of massively parallel sequencing technologies has enabled investigators to move from sequencing a single human genome to sequencing substantial portions of the human genomes of thousands or even tens of thousands of individuals. This approach allows for the detection of

rare variants causing Mendelian disorders and rare variants contributing to complex traits.⁴ These same technologies are now being applied to analysis of the human microbiome with similar broad implications for pediatric health and disease.⁵

The work by Wambach et al¹ takes advantage of a very deep biorepository of samples and illustrates the value of these resources for contributing to our knowledge of the causes of pediatric diseases. The use of such biorepositories, particularly those derived from residual samples from newborn screening programs, raises many complex ethical and social issues.⁶ Care must be taken with respect to the rights of individuals (and their progeny) to determine if and when their genetic identity is to be explored. On the other hand, the use of de-identified biological samples to further disease-specific research can foster important public health initiatives if carried out so there is minimal risk for exposing an individual or there are untoward

consequences if such an exposure did occur. The results of Wambach et al¹ demonstrate that biorepository samples can be used appropriately when careful attention has been paid to protecting the identities of specific individuals and working closely with families and state and federal agencies to ensure sound ethical use.

The recent publications using cord blood or maternal serum obtained noninvasively to sequence the entire genome of a neonate fetus opens further doors for the application of sequencing technology.^{7,8,9} For example, prenatal screening for disorders now tested for in the newborn could be done for the entire range of Mendelian disorders and not just limited to the few dozen currently tested in which biochemical assays are available. This would provide for prebirth anticipatory informing of parents and therapy initiated immediately at birth. Specific environmental exposures could be limited in susceptible

children. Such sequencing is imbued with ethical challenges as well.¹⁰ A paramount concern is the capacity of the scientists and clinicians to interpret the data in a way that is both understandable and useful to the patient/family and to find mechanisms to deal with the enormous amount of data that arise from deep-sequencing studies.¹¹ Another major concern is how to manage and report data that are outside of what a primary goal of a sequencing effort might have been designed to accomplish. For example, unsuspected chromosomal anomalies or point be readily identified in genome-wide sequencing approaches. When (and how) to report such findings are currently topics of wide debate. Thus, substantial challenges remain in bringing these exciting technologies truly to the bedside. The future can be bright but only when coupled with existing resources, novel technologies, ethical approaches to the investigations, and bioinformatic approaches that enable the best and most appropriate interpretation of the data.

REFERENCES

1. Wambach JA, Wegner DJ, DePass K, et al. Single *ABCA3* mutations increase risk for neonatal respiratory distress syndrome. *Pediatrics*. 2012;130(6). Available at: www.pediatrics.org/cgi/content/full/130/6/e1575
2. Lupski JR, Belmont JW, Boerwinkle E, Gibbs RA. Clan genomics and the complex architecture of human disease. *Cell*. 2011;147(1):32–43
3. Connolly JJ, Hakonarson H. The impact of genomics on pediatric research and medicine. *Pediatrics*. 2012;129(6):1150–1160
4. Bamshad MJ, Shendure JA, Valle D, et al; Centers for Mendelian Genomics. The Centers for Mendelian Genomics: a new large-scale initiative to identify the genes underlying rare Mendelian conditions. *Am J Med Genet A*. 2012;158A(7):1523–1525
5. Johnson CL, Versalovic J. The human microbiome and its potential importance to pediatrics. *Pediatrics*. 2012;129(5):950–960
6. Rothwell E, Anderson R, Goldenberg A, et al. Assessing public attitudes on the retention and use of residual newborn screening blood samples: a focus group study. *Soc Sci Med*. 2012;74(8):1305–1309
7. Kitzman JO, Snyder MW, Ventura M, et al. Noninvasive whole-genome sequencing of a human fetus. *Sci Transl Med*. 2012;4(137):137ra76
8. Fan HC, Gu W, Wang J, Blumenfeld YJ, El-Sayed YY, Quake SR. Non-invasive prenatal measurement of the fetal genome. *Nature*. 2012;487(7407):320–324
9. Saunders CJ, Miller NA, Soden SE, et al. Rapid whole-genome sequencing for genetic disease diagnosis in neonatal intensive care units. *Sci Transl Med*. 2012 Oct 3;4(154):154ra135
10. Tabor HK, Murray JC, Gammill HS, et al. Non-invasive fetal genome sequencing: opportunities and challenges. *Am J Med Genet A*. 2012;158A(10):2382–2384
11. Solomon BD, Hadley DW, Pineda-Alvarez DE; NISC Comparative Sequencing Program, Kamat A, Teer JK, Cherukuri PF, et al. Incidental medical information in whole-exome sequencing. *Pediatrics*. 2012;129(6). Available at: www.pediatrics.org/cgi/content/full/129/6/e1605