

BRIEF REPORT

Clinical Diagnosis by Whole-Genome Sequencing of a Prenatal Sample

Michael E. Talkowski, Ph.D., Zehra Ordulu, M.D., Vamsee Pillalamarri, M.S., Carol B. Benson, M.D., Ian Blumenthal, B.E., Susan Connolly, M.D., Carrie Hanscom, M.S., Naveed Hussain, M.D., Shahrin Pereira, B.S., Jonathan Picker, M.B., Ch.B., Ph.D., Jill A. Rosenfeld, M.S., Lisa G. Shaffer, Ph.D., Louise E. Wilkins-Haug, M.D., James F. Gusella, Ph.D., and Cynthia C. Morton, Ph.D.

SUMMARY

From the Center for Human Genetic Research, Massachusetts General Hospital, Boston (M.E.T., V.P., I.B., C.H., J.F.G.), the Program in Medical and Population Genetics, Broad Institute, Cambridge (M.E.T., J.F.G., C.C.M.), the Departments of Neurology (M.E.T., J.F.G.) and Genetics (J.F.G.), Harvard Medical School, Boston, the Departments of Obstetrics, Gynecology and Reproductive Biology (Z.O., S.P., L.E.W.-H., C.C.M.), Radiology (C.B.B.), and Pathology (C.C.M.), Brigham and Women's Hospital and Harvard Medical School, Boston, and the Department of Radiology (S.C.) and the Division of Genetics, Department of Psychiatry (J.P.), Children's Hospital Boston and Harvard Medical School, Boston — all in Massachusetts; the Department of Pediatrics, University of Connecticut, Farmington (N.H.); and Signature Genomic Laboratories, PerkinElmer, Spokane, WA (J.A.R., L.G.S.). Address reprint requests to Dr. Morton at Harvard Medical School, 77 Ave. Louis Pasteur, Boston, MA 02115, or at cmorton@partners.org.

Drs. Talkowski and Ordulu contributed equally to this article.

N Engl J Med 2012;367:2226-32.
DOI: 10.1056/NEJMoa1208594

Copyright © 2012 Massachusetts Medical Society.

Conventional cytogenetic testing offers low-resolution detection of balanced karyotypic abnormalities but cannot provide the precise, gene-level knowledge required to predict outcomes. The use of high-resolution whole-genome deep sequencing is currently impractical for the purpose of routine clinical care. We show here that whole-genome “jumping libraries” can offer an immediately applicable, nucleotide-level complement to conventional genetic diagnostics within a time frame that allows for clinical action. We performed large-insert sequencing of DNA extracted from amniotic-fluid cells with a balanced *de novo* translocation. The amniotic-fluid sample was from a patient in the third trimester of pregnancy who underwent amniocentesis because of severe polyhydramnios after multiple fetal anomalies had been detected on ultrasonography. Using a 13-day sequence and analysis pipeline, we discovered direct disruption of *CHD7*, a causal locus in the CHARGE syndrome (coloboma of the eye, heart anomaly, atresia of the choanae, retardation, and genital and ear anomalies). Clinical findings at birth were consistent with the CHARGE syndrome, a diagnosis that could not have been reliably inferred from the cytogenetic breakpoint. This case study illustrates the potential power of customized whole-genome jumping libraries when used to augment prenatal karyotyping.

DEEP SEQUENCING OF THE WHOLE GENOME HOLDS DIAGNOSTIC PROMISE but is currently thought to be impractical for routine prenatal care. In contrast, large-insert mate-pair, or jumping-library, sequencing provides a tractable approach for immediate clinical application and could complement conventional prenatal diagnostics. The risk of major structural birth defects among live births in the United States is approximately 3%¹ and is associated with inherited or *de novo* genetic rearrangements and mutations as well as with maternal factors, such as advanced age, certain clinical conditions, and exposure to teratogenic factors. Approximately 1 in 2000 prenatal cases analyzed with conventional karyotyping has a *de novo*, apparently balanced reciprocal translocation that carries a 6.1% risk of congenital malformation.² Ultrasound examination between 18 and 20 weeks of gestation allows detection of major malformations and is offered routinely, since 90% of infants with congenital anomalies are born to women without predisposing risk factors.³ An abnormal finding on fetal ultrasonography necessitates counseling and a discussion of a diagnostic procedure that can be used to assess the possibility that the abnormality has a genetic basis.

Conventional karyotyping, which is the standard method used for prenatal cytogenetic diagnosis, can detect numerical abnormalities as well as unbalanced and

apparently balanced rearrangements within microscopical resolution (range, 3 to 10 Mb). Fluorescence in situ hybridization analyses can be used to detect chromosomal abnormalities smaller than 3 Mb, but this method is not suitable for high-throughput analyses because only a limited number of probes can be screened simultaneously. Array-based comparative genomic hybridization (CGH) has been introduced in prenatal diagnosis to detect genomewide gains and losses with higher resolution,⁴ but its use is typically limited to dosage imbalances on the order of tens to hundreds of kilobases. For example, a recent study of more than 36,000 persons revealed karyotypic abnormalities in 0.78% of persons with intellectual disabilities in whom array-based CGH tests were unremarkable.⁵ The ability to rapidly localize breakpoints of cytogenetically balanced chromosomal rearrangements to individual genes could substantially improve the prediction of phenotypic outcomes and inform postnatal medical care. Here we describe such an approach in a clinical setting. We used massively parallel paired-end sequencing of customized large-insert jumping libraries to define the precise consequences of a balanced *de novo* translocation in DNA extracted from amniotic-fluid cells after the detection of multiple fetal anomalies.

CASE REPORT

A pregnant 37-year-old woman with a history of infertility and spontaneous abortion and no previous full-term pregnancy presented after ultrasonography performed at 18.8 weeks of gestation revealed fetal abnormalities, including a hypoplastic right ventricle and tricuspid atresia (Fig. 1A). Conception had occurred after the fourth cycle of *in vitro* fertilization, and the results of ultrasonography and genetic screening performed during the first trimester had been normal. (All prenatal care and testing had been performed in the United States.) A pediatric cardiology review included the consideration of two or three surgeries for possible palliation, as well as pregnancy termination. Follow-up fetal surveys revealed a level of amniotic fluid that was elevated but within the normal range at 27.3 weeks and polyhydramnios at 30.4 weeks; a small, intermittently undetected stomach was also noted (Fig. 1B). Esophageal atresia was considered along with the possibility of surgical repair after delivery. At 33.3 weeks' gestation, additional findings on fetal ultrasonography indicated the

possibility of micrognathia, flexed extremities, and severe polyhydramnios, suggesting a differential diagnosis that included arthrogryposis, the Stickler syndrome, and trisomy 18 (Fig. 1C, 1D, and 1E). Therapeutic amnioreduction was performed, and 20 ml of fluid was submitted for cytogenetic analysis, with a portion of the fluid saved for possible array-based CGH testing. Karyotyping with Giemsa (GTG) banding revealed an apparently balanced *de novo* translocation, 46,XY,t(6;8)(q13;q13)dn (see Fig. S1 in the Supplementary Appendix, available with the full text of this article at NEJM.org). Abnormal findings detected on fetal magnetic resonance imaging at 34.4 weeks' gestation included moderately severe polyhydramnios, the absence of a fluid-filled stomach, a nondilated esophagus to the level of the carina, microstomia, an enlarged protruding superior lip, intermittent abnormal swallowing motion with mild protrusion of the tongue, and abnormal fetal position, with flexed elbows and knees, abducted hips, and clenched hands.

A medical genetics consultation at 34.4 weeks included discussion of the possibility of a syndrome resulting from the disruption of one or more genes, microdeletions or microduplications in the breakpoint regions created during an apparently balanced chromosome rearrangement, or a combination of these abnormalities. An array-based CGH analysis revealed no clinically significant loss or gain of genetic material (Table S1 in the Supplementary Appendix). The results of ultrasonography performed at 35.3 weeks suggested the possibility of an undescended right testicle. Absence of fetal movement during ultrasonography at 36.2 weeks led to an immediate cesarean section. At the time of the cesarean section, polyhydramnios was observed and the baby was found to have neurologic and respiratory depression, prompting intubation after delivery. On the basis of clinical features, the infant received a postnatal diagnosis of the CHARGE syndrome (Online Mendelian Inheritance in Man [OMIM] number, 214800) (Table S2 in the Supplementary Appendix). Plans for any immediate surgeries were postponed until after stabilization, but the infant's clinical condition worsened, and he died at 10 days of age.

METHODS

STUDY OVERSIGHT

We obtained written informed consent from the parents in accordance with the Developmental Genome Anatomy Project protocol, approved by

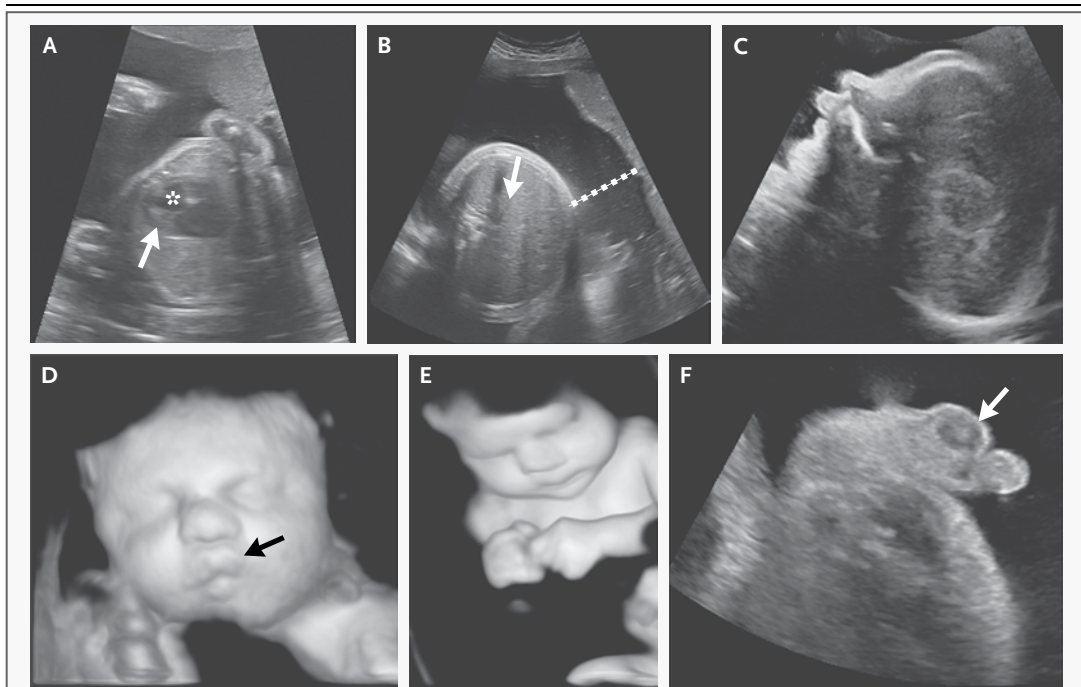


Figure 1. Clinical Findings Detected with Prenatal Imaging.

A transaxial ultrasonogram with a four-chamber view of the heart, obtained at 27.3 weeks of gestation (Panel A), shows a small right ventricle (arrow), as compared with the left ventricle (star), which was first detected at 18.8 weeks; tricuspid atresia was also detected on earlier imaging. A transaxial ultrasonogram obtained at 35.3 weeks of gestation (Panel B) shows polyhydramnios (dashed line), first detected at 30.4 weeks; also noteworthy is the absence of a fluid-filled stomach in the upper abdomen (arrow). An ultrasonogram of the fetal profile (Panel C) and a three-dimensional ultrasonogram of the fetal face (Panel D), both obtained at 34.4 weeks of gestation, show microstomia and protrusion of the upper lip (Panel D, arrow), and a three-dimensional ultrasonogram obtained at 33.3 weeks of gestation (Panel E) shows abnormally clenched hands and flexed arms. A transaxial ultrasonogram of the perineum in a phenotypic male fetus, obtained at 35.3 weeks of gestation, shows only one testicle in the scrotum (Panel F, arrow).

the Partners HealthCare System Institutional Review Board. The study was initiated in January 2012 and concluded in March 2012.

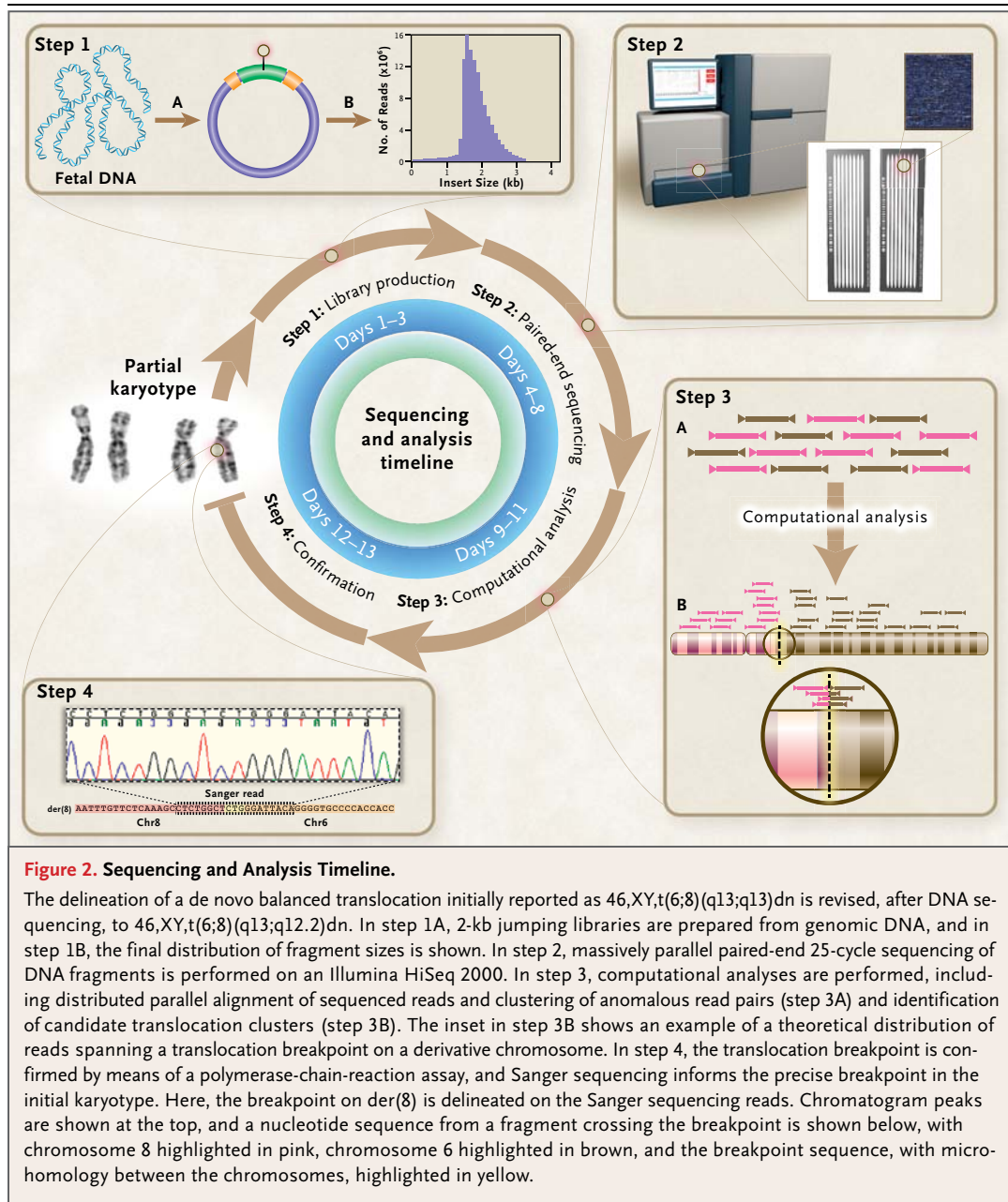
SEQUENCING AND ANALYSIS

We sequenced the paired ends of approximately 220-bp DNA fragments separated by approximately 2 kb of contiguous genomic DNA.^{6,7} The entire four-step, 13-day process is shown in Figure 2.

On days 1 through 3, genomic DNA was sheared and selected according to size such that the majority of DNA fragments were approximately 2 kb. These fragments were circularized with adapters containing an *EcoP15I* recognition site and a biotinylated thymine. The circularized DNA was processed into fragments by means of a restriction digest, and the fragments at the circularization junction were retained by binding the biotinylated thymine to streptavidin beads. Genomic libraries suitable for next-generation sequencing on an Illumina platform were creat-

ed from these fragments (which spanned the circularization junction) while they were bound to the streptavidin beads, yielding a library of DNA fragments with ends separated by a genomic distance equal to the size of the circularized fragments (Fig. 2).⁷ On days 4 through 8, paired-end, 25-cycle sequencing was performed on a single lane of an Illumina HiSeq 2000.

During computational and statistical analysis, on days 9 through 11, reads were aligned with the use of the Burrows–Wheeler alignment tool,⁸ and BAM files were then processed with a C++ program (BamStat) to tabulate mapping statistics and output lists of anomalous read pairs (i.e., ends that map to two different chromosomes, abnormally sized inserts, or unexpected strand orientations).⁷ Mapping and assembly artifacts were excluded on the basis of our previous sequencing experiments^{9–11} to elucidate chimeric read pairs. These chimeric read pairs suggested possible candidate translocation “clusters” throughout the



genome (see the Methods section in the Supplementary Appendix for the definition of candidate clusters and information about filtering criteria and sequence-analysis metrics). Read-pair clustering and translocation discovery were performed by two independent analysts with knowledge only of the chromosomes involved in the translocation.

On days 12 and 13, DNA was amplified from cells obtained from the amniotic fluid with the use of polymerase-chain-reaction (PCR) primers designed according to the sequence reads

supporting the translocation junction. The amplified products were then sequenced (Tables S4 and S5 in the Supplementary Appendix).

We tested the significance of the structural rearrangements disrupting this locus using copy-number variant data.¹¹ The data were derived from 33,573 cases referred to a clinical diagnostic laboratory for array-based CGH testing (Signature Genomic Laboratories, PerkinElmer) and from 13,991 unaffected controls in previous genomewide association studies (Table S3 in the Supplementary Appendix).

RESULTS

Large-insert, paired-end sequencing of DNA from cells in the amniotic fluid generated 282,294,280 individual reads (141 million pairs). Each aligned pair allowed assessment of a chromosomal region corresponding to the original fragment size (median, 1914 bp; standard deviation, 369 bp). Consequently, the inserts between aligned pairs covered each base in the genome 68 times on average, despite a mean coverage of only two reads spanning each nucleotide of the genome. We identified only one cluster of reads with ends mapping to chromosomes 6 and 8 (Fig. S2 in the Supplementary Appendix). This cluster contained 35 read pairs with high mapping quality (Fig. S3 in the Supplementary Appendix). The translocation breakpoint in chromosome 8 directly disrupted *CHD7*, and the chromosome 6 breakpoint disrupted *LMBRD1* (Fig. 3). The transcriptional orientation of each gene was incompatible with the generation of a fusion transcript involving *CHD7* and *LMBRD1*. PCR amplification and capillary sequencing of the breakpoints resulted in a revised karyotype of 46,XY,t(6;8)(q13;q12.2)dn (for additional breakpoint information, see the Results section and Table S5 in the Supplementary Appendix).

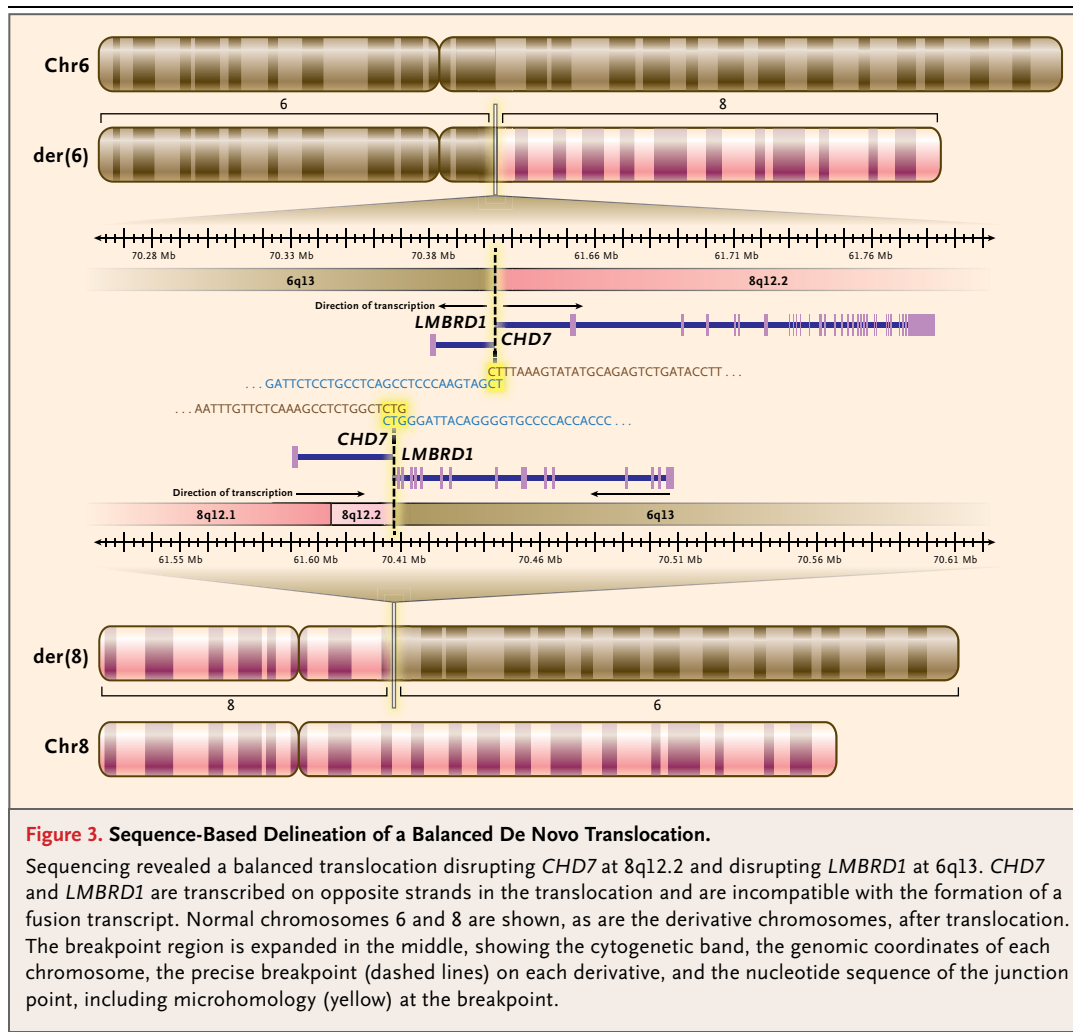
Point mutations of *CHD7* cause the CHARGE syndrome, which has also been attributed to functional hemizygoty arising from deletion of one copy of the gene.¹² We analyzed copy-number variant data on more than 47,000 persons and identified two gene-specific deletions of *CHD7*, both of which were found in persons with features consistent with the CHARGE syndrome. We found no *LMBRD1*-specific variations among cases referred to a clinical diagnostic laboratory for array CGH testing, and no disruption of either locus among controls (Table S3 in the Supplementary Appendix). Taken together, these findings suggest that functional mutations and the disruption of a single copy of *CHD7* by means of structural variation can cause the CHARGE syndrome.

DISCUSSION

We report the identification of a 46,XY,t(6;8)(q13;q13)dn karyotype in a fetus with an isolated heart defect at 18.8 weeks of gestation and additional abnormalities revealed on imaging studies performed throughout the third trimester (Fig. 1).

After delivery, the neonate received a clinical diagnosis of the CHARGE syndrome, a result that could not have been unequivocally diagnosed on the basis of ultrasonography, original karyotyping, or subsequent array-based CGH testing. Following an optimized 13-day protocol, we used large-insert sequencing of the prenatal DNA sample to identify precise translocation breakpoints that directly disrupted *CHD7* at 8q12.2, a pathogenic locus in the CHARGE syndrome,¹² and *LMBRD1* at 6q13, a pathogenic locus in a recessive disorder of vitamin B₁₂ metabolism (cobalamin F type)¹³ (Fig. 3). We thus identified a pathogenic gene disruption by sequencing the DNA obtained from a prenatal sample with a balanced translocation providing a definitive sequence-based prenatal diagnosis that was consistent with the diagnosis based on postnatal clinical findings.

Our study suggests that innovations in genome sequencing aimed specifically at detecting structural variations can offer a rapid adjunct to cytogenetic techniques. Sequencing enables precise definition of individual disrupted genes, thereby adding to the information available for outcome prediction, medical planning, and genetic counseling. In the case described here, results obtained with cytogenetic testing and array-based CGH were consistent with a balanced de novo translocation, but these tests did not identify the gene or genes responsible either for the isolated cardiac defect or for the additional fetal abnormalities that were subsequently detected. Designation of the 8q13 breakpoint through karyotyping neither supported a prediction of a disruption in *CHD7* at 8q12.2 nor provided sufficient resolution to consider specific genes in a differential diagnosis (Fig. S4 in the Supplementary Appendix). Indeed, were GTG-banded breakpoints to be misinterpreted by a visible band in each direction (which is not an uncommon occurrence according to our sequencing of such balanced rearrangements^{8,11}), this would entail consideration, on chromosome 8 alone, of approximately 38 Mb of DNA containing 288 potential phenotype-contributing genes, of which 39 have been associated with disease, according to the OMIM database, and at least 4 have been associated with cardiac defects. In addition, our previous analyses have shown that rearrangements appearing to be balanced at karyotypic resolution can be complex at nucleotide resolution (with complex rearrangements accounting for approximately 20% of all



events).⁸ In the case we describe here, sequence-based revision of the karyotype permitted a definitive description of the causal syndromic locus. Such diagnostic precision and consequent phenotypic prediction cannot currently be obtained with the use of other methods, and the results were obtained within a time frame similar to that required for conventional prenatal cytogenetic methods.

The CHARGE syndrome is a rare, usually sporadic disease that may include cranial-nerve abnormalities and tracheoesophageal fistula in addition to the features listed in the Summary section.^{14,15} Previous studies have implicated *CHD7* alterations in 90% of patients meeting the diagnostic criteria for the CHARGE syndrome.¹² *CHD7* is a highly conserved member of the chromodomain helicase family; it alters gene expression by remodeling chromatin.¹⁶ *CHD7* mutations thus have potentially wide-ranging phenotypic effects.

Disruption of *CHD7* represents a strong genetic risk factor for the CHARGE syndrome, although not all disruptions of *CHD7* are fully penetrant, since deletions affecting a portion of the upstream or coding region of *CHD7* have been identified in phenotypically normal persons of Asian and African ancestry.^{17,18} Nonetheless, if a *CHD7* mutation is detected, clinical follow-up and genetic counseling are recommended.¹⁸

The chromosome 6 breakpoint disrupted *LMBRD1*, which encodes a lysosomal membrane protein involved in the transport and metabolism of cobalamin. Frameshift mutations leading to loss of function of *LMBRD1* are associated with the recessive disorder methylmalonic aciduria and homocystinuria (cobalamin F type) (OMIM number, 277380).¹³ Disruption of a single copy of the locus is unlikely to result in the disorder, and our postnatal metabolic workup of the present case ruled out a metabolic syndrome.

Delineation of a *CHD7* disruption and consequent diagnosis of the CHARGE syndrome would probably have influenced genetic counseling, subsequent discussions of management of the pregnancy, and preparation of the health care team and parents for the possibility of multiple life-threatening medical conditions requiring immediate management of breathing and feeding difficulties on delivery.¹⁹ Had we not detected a clearly predictive causal locus, we would have assessed the rearrangement for its likelihood of representing a benign alteration or its designation as a variant of unknown significance, using analyses of the clinical diagnostic and population-based copy-number variant data, along with available findings of standing genetic variation from exome-sequencing studies, resources such as the 1000 Genomes Project, and genomewide association studies in clinical cohorts and controls. Although the merits, limitations, and interpretation of such additional data sets warrant careful discussion and appropriate caution in view of our still-limited understanding of the functional consequences of disrupting specific sequences in the human genome, it is important to consider that medical decisions are usually based on the presence of the chromosome rearrangement, without additional

predictive information. At best, a secondary array-based CGH test is performed, which in this subject and most subjects with apparently balanced abnormalities will yield a normal result. Our study thus shows the predictive power of pangenomic paired-end sequencing and points toward the complexity of interpretation likely to confront the enterprise of ultra-high-resolution diagnostics.

Although scientific, medical, and ethical issues should be evaluated carefully, this strategy, when used in the prenatal setting, can detect genomic alterations that may change the obstetrical course and outcome, providing a basis for decisions regarding termination, fetal therapy, mode of delivery, and postnatal referral to a tertiary-care center with advanced expertise in management.

Supported by grants from the National Institutes of Health (GM061354, to Drs. Morton and Gusella; HD065286, to Dr. Gusella; and MH087123 and MH095867, to Dr. Talkowski). Drs. Talkowski and Gusella are supported by the Simons Foundation for Autism Research and the Nancy Lurie Marks Family Foundation.

Ms. Rosenfeld and Dr. Shaffer report being employed by Signature Genomic Laboratories, PerkinElmer.

Disclosure forms provided by the authors are available with the full text of this article at NEJM.org.

We thank the family for their participation in this research while enduring difficult personal circumstances; Dr. Stanisława Weremowicz for karyotype interpretation; Dr. Harvey Levy for his helpful discussions; and Svetlana Kozlova and Marc L. Listewnik for their excellent technical assistance.

REFERENCES

- Update on overall prevalence of major birth defects — Atlanta, Georgia 1978–2005. *MMWR Morb Mortal Wkly Rep* 2008; 57:1-5.
- Warburton D. De novo balanced chromosome rearrangements and extra marker chromosomes identified at prenatal diagnosis: clinical significance and distribution of breakpoints. *Am J Hum Genet* 1991; 49:995-1013.
- Long G, Sprigg A. A comparative study of routine versus selective fetal anomaly ultrasound scanning. *J Med Screen* 1998;5:6-10.
- ACOG Committee Opinion No. 446: array comparative genomic hybridization in prenatal diagnosis. *Obstet Gynecol* 2009; 114:1161-3.
- Hochstenbach R, van Binsbergen E, Engelen J, et al. Array analysis and karyotyping: workflow consequences based on a retrospective study of 36,325 patients with idiopathic developmental delay in the Netherlands. *Eur J Med Genet* 2009;52:161-9.
- Korbel JO, Urban AE, Affourtit JP, et al. Paired-end mapping reveals extensive structural variation in the human genome. *Science* 2007;318:420-6.
- Talkowski ME, Ernst C, Heilbut A, et al. Next-generation sequencing strategies enable routine detection of balanced chromosome rearrangements for clinical diagnostics and genetic research. *Am J Hum Genet* 2011;88:469-81.
- Li H, Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* 2010;26:589-95.
- Chiang C, Jacobsen JC, Ernst C, et al. Complex reorganization and predominant non-homologous repair following chromosomal breakage in karyotypically balanced germline rearrangements and transgenic integration. *Nat Genet* 2012;44:390-7.
- Talkowski ME, Mullegama SV, Rosenfeld JA, et al. Assessment of 2q23.1 microdeletion syndrome implicates *MBD5* as a single causal locus of intellectual disability, epilepsy, and autism spectrum disorder. *Am J Hum Genet* 2011;89:551-63.
- Talkowski ME, Rosenfeld JA, Blumenthal I, et al. Sequencing chromosomal abnormalities reveals neurodevelopmental loci that confer risk across diagnostic boundaries. *Cell* 2012;149:525-37.
- Janssen N, Bergman JE, Swertz MA, et al. Mutation update on the *CHD7* gene involved in CHARGE syndrome. *Hum Mutat* 2012;33:1149-60.
- Rutsch F, Gailus S, Miousse IR, et al. Identification of a putative lysosomal cobalamin exporter altered in the *cbf* defect of vitamin B12 metabolism. *Nat Genet* 2009;41:234-9.
- Blake KD, Davenport SL, Hall BD, et al. CHARGE association: an update and review for the primary pediatrician. *Clin Pediatr (Phila)* 1998;37:159-73.
- Pagon RA, Graham JM Jr, Zonana J, Yong SL. Coloboma, congenital heart disease, and choanal atresia with multiple anomalies: CHARGE association. *J Pediatr* 1981;99:223-7.
- Schnetz MP, Bartels CF, Shastri K, et al. Genomic distribution of *CHD7* on chromatin tracks H3K4 methylation patterns. *Genome Res* 2009;19:590-601.
- Park H, Kim JI, Ju YS, et al. Discovery of common Asian copy number variants using integrated high-resolution array CGH and massively parallel DNA sequencing. *Nat Genet* 2010;42:400-5.
- Shaikh TH, Gai X, Perin JC, et al. High-resolution mapping and analysis of copy number variations in the human genome: a data resource for clinical and research applications. *Genome Res* 2009;19:1682-90.
- Hara Y, Hirota K, Fukuda K. Successful airway management with use of a laryngeal mask airway in a patient with CHARGE syndrome. *J Anesth* 2009;23:630-2.

Copyright © 2012 Massachusetts Medical Society.