BRIEF REPORT | Genetics inMedicine

Identifying gene disruptions in novel balanced de novo constitutional translocations in childhood cancer patients by whole-genome sequencing

Deborah I. Ritter, PhD[1], Katherine Haines[2,3], Hannah Cheung, PhD, PMP[3], Caleb F. Davis, PhD[1], Ching C. Lau, MD, PhD[3,4], Jonathan S. Berg, MD, PhD[5], Chester W. Brown, MD[2,3], Patrick A. Thompson, MD[3], Richard Gibbs, PhD[1,2,4], David A. Wheeler, PhD[1,2,4] and Sharon E. Plon, MD, PhD[1–4]

Purpose: We applied whole-genome sequencing (WGS) to children diagnosed with neoplasms and found to carry apparently balanced constitutional translocations to discover novel genic disruptions.

Methods: We applied the structural variation (SV) calling programs CREST, BreakDancer, SV-STAT, and CGAP-CNV, and we developed an annotative filtering strategy to achieve nucleotide resolution at the translocations.

Results: We identified the breakpoints for t(6;12)(p21.1;q24.31), disrupting *HNF1A* in a patient diagnosed with hepatic adenomas and maturity-onset diabetes of the young (MODY). Translocation as the disruptive event of *HNF1A*, a gene known to be involved in MODY3, has not been previously reported. In a subject with Hodgkin lymphoma and subsequent low-grade glioma, we identified t(5;18)(q35.1;q21.2), disrupting both *SLIT3* and *DCC*, genes previously implicated in both glioma and lymphoma.

Conclusion: These examples suggest that implementing clinical WGS in the diagnostic workup of patients with novel but apparently balanced translocations may reveal unanticipated disruption of disease-associated genes and aid in prediction of the clinical phenotype.

Genet Med advance online publication 8 January 2015

Key Words: cancer; next-generation cytogenetics; structural variation; translocation; whole-genome sequencing

Although potentially harmless, balanced translocations can alter gene expression or function if the breakpoints occur within genes or regulatory regions. Balanced de novo translocations (not inherited from a parent) were seen with a prevalence of ~0.08% in a large cohort of prenatal cytogenetic results.[1] Translocations can result in gene fusions or disruptions, and both can be identified through whole-genome sequencing (WGS) with structural variation (SV) programs reporting translocation events. We used paired-end Illumina WGS on two subjects with constitutional balanced translocations, applied three SV calling algorithms (CREST, BreakDancer, and SV-STAT), and developed an annotative filtering strategy to identify the precise breakpoints of novel genic disruptions.[2–4] Both patients and parents (when available) were entered into a human subjects protocol approved by the institutional review board of Baylor College of Medicine.

The first patient, FCP637, is a 12-year-old girl with dysmorphic craniofacial features, seizure disorder, vesicoureteral reflux, patent foramen ovale and mildly dilated aortic root, Hashimoto thyroiditis, moderate speech delay (receptive language superior to expressive), poor speech articulation with normal hearing, friendly demeanor, subclinical seizures, and hypotonia. Diffuse hepatic adenomas were discovered incidentally and confirmed by needle biopsy. Perturbation of *HNF1A* (*TCF1*) is associated with familial hepatic adenomas and with maturity-onset diabetes of the young type 3 (MODY3) (OMIM 142410 and OMIM 600496, respectively).[5] MODY3 was diagnosed ~2 years after adenoma discovery. Peripheral blood karyotype revealed an apparently balanced translocation between 6p11.1 and 12q24.2. Karyotype analyses of both parents were normal, confirming the translocation was *de novo.* Clinical Sanger sequence analysis of *HNF1A* was normal, and an oligonucleotide chromosomal microarray analysis (Molecular Genetics Laboratory, Baylor College of Medicine) with exon coverage of *HNF1A* showed no copy-number variation, or gain or loss of material. However, the microarray results included a microdeletion of chromosome 17q21.31, encompassing 0.341–1.328 Mb. This results in a well-described microdeletion syndrome (Koolen–de Vries syndrome, OMIM 610443), which includes many of the developmental problems seen in the patient.

The second patient, FCP672, is a male who was diagnosed at 5 years of age by lymph node biopsy with classic Hodgkin lymphoma of mixed cellularity with interfollicular growth

[1]Human Genome Sequencing Center, Baylor College of Medicine, Houston, Texas, USA; [2]Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, Texas, USA [3]Department of Pediatrics, Texas Children's Cancer Center, Baylor College of Medicine, Houston, Texas, USA; [4]Dan L. Duncan Cancer Center, Baylor College of Medicine, Houston, Texas, USA; [5]Department of Genetics, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA. Correspondence: Sharon E. Plon (splon@bcm.edu)

Submitted 21 September 2014; accepted 16 November 2014; advance online publication 8 January 2015. doi:10.1038/gim.2014.189

pattern (OMIM 236000). The patient displayed unexplained growth deceleration and obesity but no apparent developmental delay. Bone marrow at the time of diagnosis demonstrated an apparently balanced translocation t(5;18)(q35.1;q21.2) in all lymphoma cells analyzed. Subsequent peripheral blood and skin fibroblast karyotype analyses demonstrated the translocation in all cells, confirming it was constitutional. A focused chromosomal analysis of both parents found no rearrangements for chr5 or chr18. Two chromosomal microarrays were performed (version V7.2CMA, repeated in 2013 as CMA-HR+SNP(V9.1.1)); these failed to detect loss or gain of genomic material in the region of the translocation. After completing treatment for lymphoma, the patient was diagnosed by magnetic resonance imaging with a brain lesion suggestive of low-grade glioma and is being followed with serial imaging without biopsy.

METHODS

Genomic DNA from both patients and one maternal parental sample (FCP664, the mother of FCP672) underwent WGS (Illumina, San Diego, CA;, 100-bp paired ends, 400-bp insert size, 30× coverage) at the Human Genome Sequencing Center, Baylor College of Medicine, following previously described protocols and quality metrics with alignment to HG19.[6]

Structural variation. We applied CREST (http://www.stjuderesearch.org/site/lab/zhang), SV-STAT (https://gitorious.org/svstat), and BreakDancer (http://breakdancer.sourceforge.net/) (using default parameters, apart from BreakDancer $q \geq 50$) and intersected outputs using 600-bp windows. Filtering metrics were as follows: CREST, ≥4 right/left soft clips and ≥20 reads coverage; BreakDancer, ≥20 anomalous read pairs; and SV-STAT, "PASS" filter calls only. We excluded translocations between autosomal and X/Y chromosomes due to mismapping. As controls, we used the BAB195 48X whole genome[7] and an additional set of 69 samples from normal tissues of adults from The Cancer Genome Atlas. We removed any proband translocation within 600 bp of control translocations. The analyst was aware of t(6;12) (p11.1;q24.2) reported in FCP637 but was blinded to the cytogenetic breakpoints for FCP672.

Copy-number variation. We applied CGAP-CNV, an in-house whole-genome copy-number program[8], to obtain microdeletion breakpoint windows. These were manually refined using Integrative Genomics Viewer (http://www.broadinstitute.org/igv/).

Single-nucleotide variation and insertions and deletions. For samples and a large set of controls (1,079 randomly selected nontumor whole exomes from the Atherosclerosis Risk in Communities cohort, run on the same design), variants were called with ATLAS Suite (ATLAS-SNP: http://sourceforge.net/p/atlas2/wiki/Atlas-SNP/ and ATLAS-INDEL: http://sourceforge.net/p/atlas2/wiki/Atlas-Indel/) and Pindel (http://gmt.genome.wustl.edu/packages/pindel/). (≤200-bp insertions and deletions) and annotated with ANNOVAR, dbSNP, COSMIC, and Exome Variant Server. Variants were filtered for ≥20/30% variant allele fraction in whole-exome sequencing and WGS, respectively, ≤1% dbSNP minor allele frequency, and ≤1.5% presence in the control set. Variants were manually reviewed in Integrative Genomics Viewer, with the BAB195 40X whole genome as a visual control.

Breakpoint reporting. We utilized the recently described Next-Gen Cytogenetics nomenclature in reporting breakpoints from cytogenetic karyotype and WGS.[9]

Breakpoint databases queried, software, and annotation resources. Mittelman Database, Atlas of Genetics and Cytogenetics in Oncology and Haematology, TICdb, and DACRO. Calling, annotation, and database sources can be found by an Internet search of listed names and/or URL; for brevity, we could not list individual references.

RESULTS

The breakpoints for the proband with MODY3 and hepatic adenomas (FCP637), using the newly proposed genomic nomenclature, are: 46,XX,t(6;12)(p11.1;q24.2)dn.seq[GRCh37/hg19]t(6;12)(12qter→12q24.31(121,420,346)::6p21.1(44,758,577)→6qter;12pter→12q24.31(121,420,33{1–2})::6p21.1(44,758,57{3-2})→6pter)dn.

On chr6, this is a nongenic region ~35 kb away from *SUPT3H* and ~336 kb from *CDC5L* in the telomeric direction. On chr12, the breakpoint occurs in the first intron of *HNF1A*, likely disrupting gene function, as the open reading frame is broken into two segments (**Figure 1a** and **Supplementary Table S1** online).

For FCP672, the proband with Hodgkin lymphoma and subsequent putative low-grade glioma, the breakpoints are: 46,XY,t(5;18)(q35.1;q21.2)dn.seq[GRCh37/hg19] t(5;18)(5pter→5q34(168,236,81{0–3})::18q21.2(50,099,{299–302})→18qter;18pter→18q21.2(50,099,30{2–3})::5q34(168,236,81{5–6})→5qter)dn.

On chr5, the breakpoint occurs in intron 8 of *SLIT3* and the first intron of *DCC* on chr18. *DCC* is forward-transcribed, whereas *SLIT3* is transcribed in the negative direction. Again, from reconstructing the derivative chromosomes (**Figure 2a** and **Supplementary Table S1** online), we reason both the *DCC* and *SLIT3* open reading frames are disrupted.

We validated translocations using polymerase chain reaction (PCR) assays (**Figures 1** and **2**), followed by Sanger sequencing. As expected, all samples from the patients, mother, and control lines contained PCR products from intact copies of the involved chromosomes. PCR reactions spanning the putative breakpoints reveal sequencing that aligns to reconstructed derivatives in the probands. For t(6;12)(p21.1;q24.31), Sanger sequencing of the der6 product revealed an overlap of only 2 bp between chr6 and chr12: AGTATAAAAA***CA***GAGCTAGGATTAGGATG with a 5-bp deletion of chr6 (**Figure 1b**). For the der12 product, Sanger

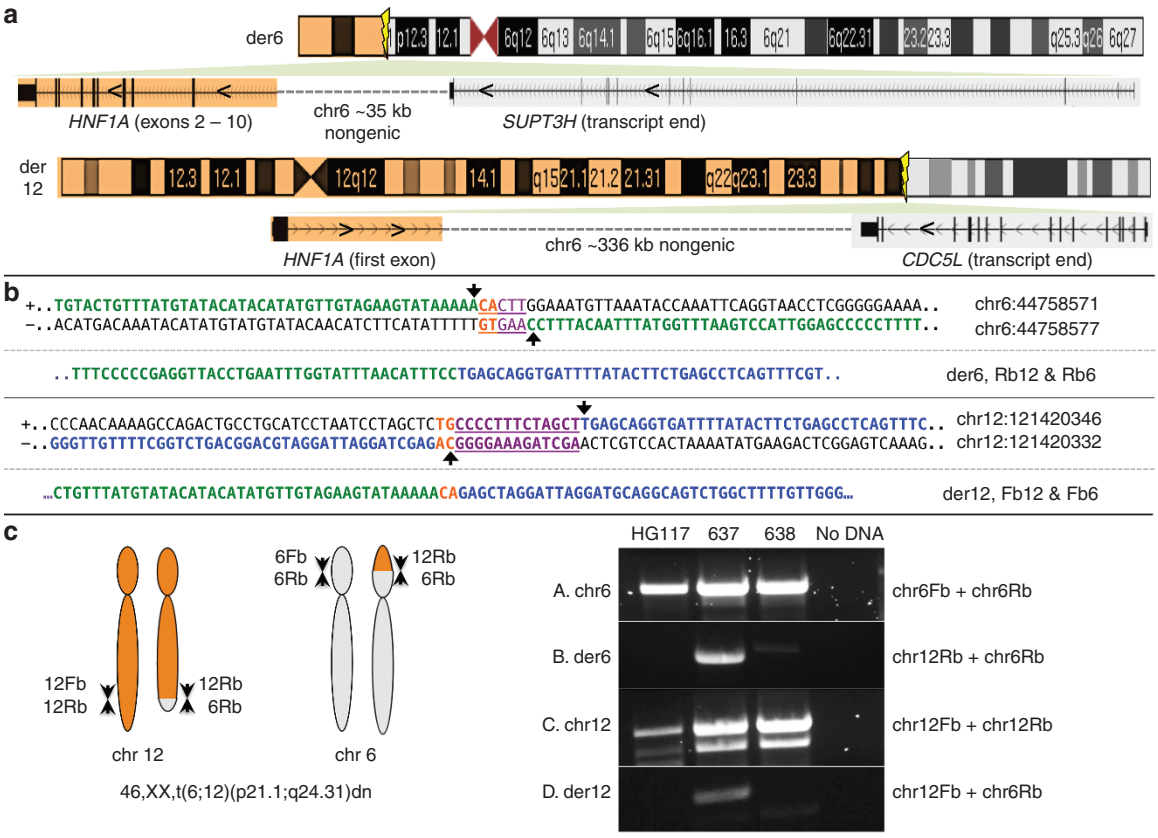

Figure 1 Translocation characterization and polymerase chain reaction (PCR) confirmation of t(6;12)(p21.1;q24.31). (**a**) Depiction of the derivative chromosomes and gene rearrangements. The der6 allele has *HNF1A* exons 2 to 10 in the reverse direction and ~35 kb of nongenic sequence from the terminus of *SUPT3H*. The der12 allele displays forward transcription of *HNF1A*, ending in a large noncoding region of chr6, prior to the *CDC5L* gene, which is transcribed on the opposite strand. (**b**) Breakpoint characterization: there is a 2-bp microhomology at the breakpoint (CA, orange), deleted in a 5-bp deletion on chr6 (purple and orange underline). A 13-bp deletion occurs on chr12 (purple underline), maintaining the CA microhomology. Arrows detail the breakpoint bases, with the GRCh37 reference sequence on the side. (**c**) PCR validation: primers (6Fb and 6Rb; 12Fb and 12Rb) amplified the wild-type chromosomes 6 and 12. The proband's mother (FCP638) and a 1000 Genomes sample (HG117) were also tested. Pairing primers 6Fb with 12Fb and 6Rb with 12Rb amplifies the expected translocation products in the proband.

sequencing results show a deletion of 13 bp in the breakpoint region and a 12-bp region of homology to both chr6 and chr12, adjacent to the translocation breakpoint. This indicates a balanced translocation with minimal sequence loss. Similarly, for t(5;18)(q34;q21) (**Figure 2b**), we found a 4-bp region of homology (CACA) between chr5 and chr18 at the region of the breakpoint. There is a 2-bp deletion on chr5 and no loss on chr18. These WGS results showing nearly precise breakpoints are consistent with the normal microarray results.

As mentioned, FCP637 carries a well-described heterozygous microdeletion detected within a megabase range by clinical array comparative genomic hybridization. We determined the microdeletion breakpoints as: 46,XX.arr[hg19]17q21.31.seq[GRCh37/hg19]del(17)(pter→q21.31(43691189)::q21.31(44354365)→qter) (**Supplementary Figure S1** online). This is an ~663-kb span, which is typical of previously reported microdeletions, and it deletes the reported causal gene, *KANSL1*.

To determine if the clinical phenotypes might result from mutations near the translocation breakpoints, we reviewed the single-nucleotide variations and insertions and deletions within 1 kb of transcription start/stop from whole-exome and genome for *DCC*, *SLIT3*, and *HNF1A*. No rare coding or intronic variants that passed visual inspection were identified in *HNF1A*, and the noncoding region ~35 kb downstream from *SUPT3H* was not further analyzed for variants. Interestingly, *RUNX2* is ~537 kb from the chr6 breakpoint, and *CDC5L* is ~326 kb upstream. *RUNX2* is a global transcriptional regulator (OMIM 600211); *CDC5L* is similar in sequence to cell-cycle regulatory genes (OMIM 602868). Position effects, such as distal gene disruptions ablating or inducing long-range *cis*-regulation, have been reported in a t(6;7)(p21.1;q36) breakpoint ~735 kb upstream of *RUNX2* (although the proband phenotype is markedly dissimilar to ours).[10] Position effects present a challenge due to large numbers of possible affected genes, and we did not have RNA-sequencing data for further analysis. For *SLIT3*, only one rare (rs34260167, minor allele frequency = 0.006) maternal (FCP664) missense variant in exon 18 of *SLIT3* (chr5, 168180047 C>T, S629N) was found in whole exome, and no rare variants were identified in the proband FCP672. Whole-genome variants produced far more results (>2k per individual), but after filtering and visualization,

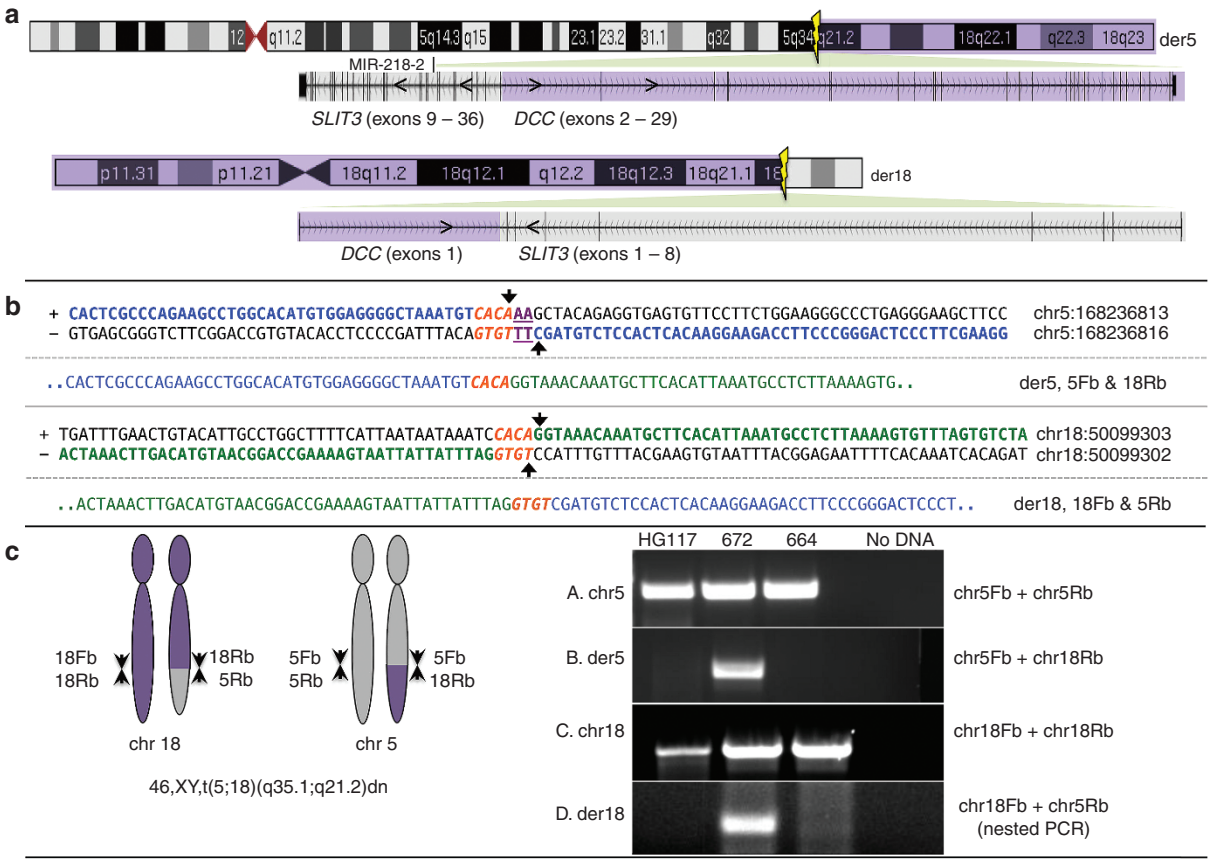

Figure 2 Translocation characterization and polymerase chain reaction (PCR) confirmation of t(5;18)(q35.1;q21.2). (**a**) The der5 allele displays exons 9–36 through the terminus of *SLIT3*, joined in intron 8 of *SLIT3* and intron 1 of *DCC*, with *DCC* exons 2 through 29 in the opposite direction. The der18 allele shows the first exon of *DCC* in the forward direction, ending in intron 8 of *SLIT3*, which is transcribed in the opposite direction. (**b**) Breakpoint characterization: there is a 4-bp region of homology between chr5 and chr18 at the breakpoint (CACA, orange), with a 2-bp deletion on chr5 (purple underline). On chr18, there is no loss of genomic material. Arrows and coordinates detail the breakpoint bases from the GRCh37 reference sequence. (**c**) PCR validation: primers (5Fb and 5Rb; 18Fb and 18Rb) amplified the wild-type chr5 and chr18. The proband's mother (FCP664) and a 1000 Genomes sample (HG117) were also tested. Pairing primers 5Fb with 18Rb and 5Rb with 18Fb amplifies the expected translocation product. Nested PCR was used to confirm the der18 product and eliminate nonspecific bands.

we found only the above maternal coding variant and a novel intronic variant at the site of the translocation in *SLIT3* (chr5, IVS8, 168236814 A>C). All other variants were novel intronic events of unknown significance (~140 per individual).

DISCUSSION

We identified two novel disruptive constitutional translocations occurring in introns of *HNF1A*, *SLIT3*, and *DCC*. We found no previous reports of a constitutional balanced translocation involving *HNF1A*. The prior knowledge of *HNF1A* underlying MODY3 with hepatic adenomas strongly supports our novel finding of *HNF1A* disruption caused by a translocation as the underlying genetic defect in FCP637.[5] We found a single report of an RNA-seq gene fusion in sarcoma tumors involving the *HNF1A* and *CMKLR1* genes, both on chr12, (ref. 11) and two prior reports identified constitutional balanced translocations in other forms of MODY-diagnosed individuals, one affecting *HNF4A* [t(3;20)(p21.2;q12)][12] and one involving *MPP7*, which shares functional overlap with *HNF4A* [t(7;10)(q22;p12)].[13] Our finding emphasizes the benefits of obtaining intronic and nongenic sequences through WGS, because in this case clinical exon tests of *HNF1A* did not find coding mutations and array comparative genomic hybridization did not indicate any material loss of *HNF1A*.

In addition to reports of germ-line variants of *HNF1A* (point mutations and small frameshift mutations) in MODY3 patients, Bluteau et al.[14] reported biallelic mutations of *HNF1A* in adenomas of affected individuals, indicative of a two-hit predisposition model. Unfortunately, there was insufficient tissue from the diagnostic biopsy to look for alterations of *HNF1A* in the hepatomas.

The WGS data fully resolved breakpoints of the 17q21.31 microdeletion, consistent with size and genomic region in other reports. We found no other published cases of MODY3 and hepatic adenomas concurrent with Koolen–de Vries syndrome; these two conditions appear to be the result of one patient carrying two unrelated genomic events, similar to recent reports from exome sequencing tests.[15]

We found no previous report of a constitutional balanced translocation disrupting both *DCC* and *SLIT3*. However, we

have identified reports with translocations in the same or similar cytobands (5q34 and 18q21) in hematological malignancies (**Supplementary Table S2** online). The analyses focused on *BCL2* in 18q21; however, the translocations were not characterized to gene resolution. *DCC* (deleted in colorectal cancer) at 18q21 has been extensively studied as a tumor suppressor in colorectal cancers but has also shown loss in lymphoid malignancies[16] and specifically in Hodgkin lymphoma.[17] Likewise, *SLIT3* has been reported to undergo epigenetic regulation in cancers and is downregulated in gliomas,[18] which is of relevance to the evidence of low-grade glioma on magnetic resonance imaging. Dickinson et al.[18] showed *SLIT3* inactivation via 5′ hypermethylation in 21/60 (~35%) glioma tumors. In addition, a miRNA (mir-218-2) was recently reported in *SLIT3* intron 22, and mir-218-2 expression decreased with concurrent reduction of *SLIT3* transcript.[19] Our t(5;18)(q34;q21) breakpoint is upstream of mir-218-2 and thus may disrupt miRNA transcription. The implications for cancer and *DCC* and the *UNC5*s/*SLIT*s/*ROBO*s gene families have been recently reviewed.[20] Our work suggests further analysis of germ-line changes in *SLIT3* should be considered in glioma patients.

In clinical genetics, the counseling and prediction of future clinical manifestations for a child with a previously uncharacterized balanced translocation has been challenging. Although both of the patients described had de novo translocations, previous work identified balanced translocations involving chr3 transmitted through multiple generations, resulting in autosomal dominant familial renal cell cancer.[21] The integration of clinical WGS, with appropriate data analysis for precise mapping of rearrangements, into diagnostic testing may provide important knowledge of the genes disrupted by balanced translocations and their potential role in disease manifestation. Current pediatric practice does not routinely include WGS, particularly for balanced (as based on normal microarray analysis) or de novo balanced translocations lacking overt phenotypes. Our findings emphasize the need to resolve to the base-pair level the breakpoints of constitutional balanced translocations by WGS to provide optimal care.

SUPPLEMENTARY MATERIAL

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

ACKNOWLEDGMENTS

The authors acknowledge the following funding sources: Cancer Prevention Research Institute of Texas RP10189 and National Institutes of Health R01-CA138836 to S.E.P.; National Institute of General Medical Sciences, Institutional Research and Academic Career Development K12 GM084897-06 (to D.I.R.); Keck Center for Interdisciplinary Bioscience Training of the Gulf Coast Consortia T15 LM007093 (to C.F.D.); and National Institute of General Medical Sciences T32 GM08307-22 (in support of K.M.H.).

DISCLOSURE

The authors declare no conflict of interest.

REFERENCES

1. Peng HH, Chao AS, Wang TH, Chang YL, Chang SD. Prenatally diagnosed balanced chromosome rearrangements: eight years' experience. *J Reprod Med* 2006;51:699–703.
2. Wang J, Mullighan CG, Easton J, et al. CREST maps somatic structural variation in cancer genomes with base-pair resolution. *Nat Methods* 2011;8:652–654.
3. Chen K, Wallis JW, McLellan MD, et al. BreakDancer: an algorithm for high-resolution mapping of genomic structural variation. *Nat Methods* 2009;6:677–681.
4. Davis CF. SV-STAT [computer program]. https://gitorious.org/svstat. 2013.
5. Willson JS, Godwin TD, Wiggins GA, Guilford PJ, McCall JL. Primary hepatocellular neoplasms in a MODY3 family with a novel HNF1A germline mutation. *J Hepatol* 2013;59:904–907.
6. Cancer Genome Atlas Network. Comprehensive molecular characterization of human colon and rectal cancer. *Nature* 2012;487:330–337.
7. Lupski JR, Reid JG, Gonzaga-Jauregui C, et al. Whole-genome sequencing in a patient with Charcot-Marie-Tooth neuropathy. *N Engl J Med* 2010;362:1181–1191.
8. Daines B, Wang H, Li Y, Han Y, Gibbs R, Chen R. High-throughput multiplex sequencing to discover copy number variants in *Drosophila*. *Genetics* 2009;182:935–941.
9. Ordulu Z, Wong KE, Currall BB, et al. Describing sequencing results of structural chromosome rearrangements with a suggested next-generation cytogenetic nomenclature. *Am J Hum Genet* 2014;94:695–709.
10. Fernandez BA, Siegel-Bartelt J, Herbrick JA, Teshima I, Scherer SW. Holoprosencephaly and cleidocranial dysplasia in a patient due to two position-effect mutations: case report and review of the literature. *Clin Genet* 2005;68:349–359.
11. McPherson A, Hormozdiari F, Zayed A, et al. deFuse: an algorithm for gene fusion discovery in tumor RNA-Seq data. *PLoS Comput Biol* 2011;7:e1001138.
12. Gloyn AL, Ellard S, Shepherd M, et al. Maturity-onset diabetes of the young caused by a balanced translocation where the 20q12 break point results in disruption upstream of the coding region of hepatocyte nuclear factor-4alpha (HNF4A) gene. *Diabetes* 2002;51:2329–2333.
13. Bhoj EJ, Romeo S, Baroni MG, Bartov G, Schultz RA, Zinn AR. MODY-like diabetes associated with an apparently balanced translocation: possible involvement of MPP7 gene and cell polarity in the pathogenesis of diabetes. *Mol Cytogenet* 2009;2:5.
14. Bluteau O, Jeannot E, Bioulac-Sage P, et al. Bi-allelic inactivation of TCF1 in hepatic adenomas. *Nat Genet* 2002;32:312–315.
15. Yang Y, Muzny DM, Reid JG, et al. Clinical whole-exome sequencing for the diagnosis of mendelian disorders. *N Engl J Med* 2013;369:1502–1511.
16. Castets M, Broutier L, Molin Y, et al. DCC constrains tumour progression via its dependence receptor activity. *Nature* 2012;482:534–537.
17. Devilard E, Bertucci F, Trempat P, et al. Gene expression profiling defines molecular subtypes of classical Hodgkin's disease. *Oncogene* 2002;21:3095–3102.
18. Dickinson RE, Dallol A, Bieche I, et al. Epigenetic inactivation of SLIT3 and SLIT1 genes in human cancers. *Br J Cancer* 2004;91:2071–2078.
19. Tie J, Pan Y, Zhao L, et al. MiR-218 inhibits invasion and metastasis of gastric cancer by targeting the Robo1 receptor. *PLoS Genet* 2010;6:e1000879.
20. Mehlen P, Delloye-Bourgeois C, Chédotal A. Novel roles for Slits and netrins: axon guidance cues as anticancer targets? *Nat Rev Cancer* 2011;11:188–197.
21. Poland KS, Azim M, Folsom M, et al. A constitutional balanced t(3;8)(p14;q24.1) translocation results in disruption of the TRC8 gene and predisposition to clear cell renal cell carcinoma. *Genes Chromosomes Cancer* 2007;46:805–812.