HHS Public Access

Author manuscript

Genet Med. Author manuscript; available in PMC 2016 April 01.

Published in final edited form as:

Genet Med. 2015 October ; 17(10): 831–835. doi:10.1038/gim.2014.189.

Identifying Gene Disruptions in Novel Balanced *de novo* Constitutional Translocations in Childhood Cancer Patients by Whole Genome Sequencing

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

[1]Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX 77030

[2]Department of Pediatrics, Division of Hematology-Oncology, Texas Children's Cancer Center, Baylor College of Medicine, Houston, TX 77030

[3]Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX 77030

[4]Department of Genetics, University of North Carolina, Chapel Hill, NC 27599

[5]Dan L. Duncan Cancer Center, Baylor College of Medicine, Houston, TX 77030

Abstract

Purpose—We applied whole genome sequencing to children diagnosed with neoplasms and found to carry apparently balanced constitutional translocations, to discover novel genic disruptions.

Methods—We applied SV calling programs CREST, Break Dancer, SV-STAT and CGAP-CNV, and 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's 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.

Conclusions—These examples suggest that implementing clinical whole genome sequencing in the diagnostic work-up of patients with novel but apparently balanced translocations may reveal unanticipated disruption of disease-associated genes and aid in prediction of the clinical phenotype.

*Corresponding Author: Sharon E. Plon (splon@bcm.edu).

Supplemental Data:
Supplemental data include one figure and two tables.

Keywords

Whole-Genome Sequencing; Structural Variation; Translocation; Cancer; Next-Generation Cytogenetics

INTRODUCTION

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 with structural variation (SV) programs reporting translocation events. We used paired-end Illumina whole genome sequencing (WGS) on two subjects with constitutional balanced translocations, applied three SV calling algorithms (CREST, Break Dancer 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 (BCM).

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 (MODY 3, MIM600496, MIM142410)[5]. MODY 3 was diagnosed approximately two years after adenoma discovery. Peripheral blood karyotype revealed an apparently balanced translocation between 6p11.1 and 12q24.2. Karyotype analysis of both parents was normal, confirming the translocation was *de novo.* Clinical Sanger sequence analysis of *HNF1A* was normal, and an oligonucleotide chromosomal microarray analysis (BCM, Molecular Genetics Laboratory) with exon coverage of *HNF1A* showed no copy number variation (CNV), gain or loss of material. However, the microarray results included a microdeletion of chromosome 17q21.31, encompassing 0.341 Mb – 1.328 Mb. This results in a well-described microdeletion syndrome, (Koolen-de Vries syndrome, MIM610443) which includes many of the developmental problems seen in the patient.

The second patient, FCP672, is a male, diagnosed at age 5 by lymph node biopsy, with classical Hodgkin's Lymphoma of mixed cellularity with interfollicular growth pattern (MIM236000). 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 chromosomes 5 or 18. Two chromosomal

microarrays were performed (version V7.2CMA, repeated in 2013 as CMA-HR +SNP(V9.1.1)) and 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 MRI with a brain lesion suggestive of low-grade glioma and is being followed by serial imaging without biopsy.

METHODS

Genomic DNA from both patients and one maternal parental sample (FCP664, the mother of FCP672) underwent WGS (Illumina, 100bp paired-ends, 400bp insert size, 30× coverage) at BCM Human Genome Sequencing Center, following previously described protocols and quality metrics, with alignment to HG19[6]. *SV*: We applied CREST, SV-STAT and Break Dancer (using default parameters, apart from Break Dancer q>=50) and intersected outputs using 600bp windows. Filtering metrics were as follows: CREST: >=4 right/left soft clips and >=20 reads coverage. Break Dancer: >=20 anomalous read pairs, SV-STAT: "PASS" filter calls only. We excluded translocations between autosomal and X/Y chromosomes due to mis-mapping. As controls, we used the BAB195 48× 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 600bp 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. *CNV*: We applied CGAP-CNV, a local whole genome copy number program, to obtain microdeletion breakpoint windows. These were manually refined using Integrative Genomics Viewer (IGV). *SNV and INDELS*: For samples and a large set of controls (1079 randomly selected non-tumor whole exomes from the Atherosclerosis Risk in Communities (ARIC) cohort, run on the same design), variants were called with ATLAS Suite and Pindel (<=200bp indels), and annotated with ANNOVAR, dbSNP, COSMIC and Exome Variant Server. Variants were filtered for >=20%/30% variant allele fraction in WE and WGS, respectively, <=1% dbSNP minor allele frequency and <=1.5% presence in control set. Variants were manually reviewed in IGV, with the BAB195 40× whole genome as a visual control. *Breakpoint Reporting*: We utilize the recently described Next-Gen Cytogenetics nomenclature in reporting breakpoints from cytogenetic karyotype and whole genome sequencing[8]. *Breakpoint Databases Queried*: 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; 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 non-genic region approximately 35 kilobases (kb) away from *SUPT3H*, and approximately 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, Table S1).

For FCP672, the proband with Hodgkin's 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 chromosome 5, the breakpoint occurs in intron 8 of *SLIT3*, and the first intron of *DCC* on chromosome 18. *DCC* is forward transcribed, while *SLIT3* is transcribed in the negative direction. Again, from reconstructing the derivative chromosomes (Figure 2A, Table S1), we reason both *DCC* and *SLIT3* open reading frames are disrupted.

We validated translocations using PCR assays (Figures 1 and 2), followed by Sanger sequencing. As expected, all samples from the patients, mother and control lines contain PCR products from intact copies of the involved chromosomes. PCR reactions spanning the putative breakpoints reveal sequencing that aligns to re-constructed 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 der12 product, Sanger sequencing results show a deletion of 13bp 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 find 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 arrayCGH. 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) (Figure S1). This is a ~663kb span, which is typical of previously reported microdeletions, and 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 variation (SNVs) and insertions and deletions (INDELs) within 1kb 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 approximately 35kb downstream *SUPT3H* was not further analyzed for variants. Interestingly, *RUNX2* is ~537kb from the chromosome 6 breakpoint, and *CDC5L* is ~326kb upstream. *RUNX2* is a global transcriptional regulator (MIM600211); *CDC5L* is similar in sequence to cell-cycle regulatory genes (MIM602868). 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 ~735kb upstream of *RUNX2* (although the proband phenotype is markedly dissimilar to

ours)[9]. 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, MAF=.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, 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 find 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 *HNF1A* and *CMKLR1* genes, both genes on chr12[10], and two prior reports identified constitutional balanced translocations in other forms of MODY-diagnosed individuals, one affecting *HNF4A* [t(3;20) (p21.2;q12)][11] and one involving *MPP7*, which shares functional overlap with *HNF4A* [t(7;10) (q22;p12)][12]. Our finding emphasizes the benefits of obtaining intronic and nongenic sequences through whole genome sequencing, as in this case clinical exon tests of *HNF1A* did not find coding mutations, and arrayCGH did not indicate any material loss of *HNF1A*.

In addition to reports of germline variants of *HNF1A* (point mutations and small frame shift mutations) in MODY3 patients, Bluteau et. al. reports bi-allelic mutations of *HNF1A* in adenomas of affected individuals, indicative of a two-hit predisposition model[13]. Unfortunately there was insufficient tissue from the diagnostic biopsy to look for alterations of *HNF1A* in the hepatomas.

The WGS sequencing data fully resolved breakpoints of the 17q21.31microdeletion, consistent in size and genomic region to other reports. We find 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[14].

We find 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 (Table S2). 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[15] and specifically in Hodgkin's Lymphoma[16]. Likewise, *SLIT3* has been reported to undergo epigenetic regulation in cancers and is down-regulated in gliomas[17], which is of relevance to the evidence of low-grade glioma on MRI imaging. Dickinson et al. showed *SLIT3*

inactivation via 5' hypermethylation in 21/60 (~35%) glioma tumors[17]. Additionally, a miRNA (mir-218-2) was recently reported in *SLIT3* intron 22, and mir-218-2 expression decreased with concurrent reduction of *SLIT3* transcript[18]. 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[19]. Our work suggests further analysis of germline 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 chromosome 3 transmitted through multiple generations, resulting in autosomal dominant familial renal cell cancer[20]. The integration of clinical whole genome sequencing, 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 whole genome sequencing, 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 base-pair resolution the breakpoints of constitutional balanced translocations by whole genome sequencing to provide optimal care.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the following funding sources: (1) Cancer Prevention Research Institute of Texas (CPRIT) RP10189 and National Institutes of Health (NIH) R01-CA138836 to SEP, (2) National Institute of General Medical Sciences (NIGMS), Institutional Research and Academic Career Development (IRACDA) K12 GM084897-06 to DIR and (3) Keck Center for Interdisciplinary Bioscience Training of the Gulf Coast Consortia (T15 LM007093) to CFD.

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(9):699–703. [PubMed: 17039698]
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(8):652–654. 10.1038/nmeth.1628. [PubMed: 21666668]
3. Chen K, Wallis JW, McLellan MD, et al. Break Dancer: An algorithm for high-resolution mapping of genomic structural variation. Nat Methods. 2009; 6(9):677–681. 10.1038/nmeth.1363. [PubMed: 19668202]
4. Davis CF. SV-STAT [computer program]. 2013 https://gitorious.org/svstat.
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(4):904–907. 10.1016/j.jhep.2013.05.024. [PubMed: 23707370]
6. Cancer Genome Atlas Network. Comprehensive molecular characterization of human colon and rectal cancer. Nature. 2012; 487(7407):330–337. [doi]. [PubMed: 22810696]

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(13):1181–1191. [doi]. [PubMed: 20220177]
8. 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 [pii].
9. 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(4):349–359. doi: CGE498 [pii]. [PubMed: 16143022]
10. 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(5):e1001138. [doi]. [PubMed: 21625565]
11. 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(7):2329–2333. [PubMed: 12086970]
12. 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 [doi]. 5-8166-2-5.
13. Bluteau O, Jeannot E, Bioulac-Sage P, et al. Bi-allelic inactivation of TCF1 in hepatic adenomas. Nat Genet. 2002; 32(2):312–315. [PubMed: 12355088]
14. Yang Y, Muzny DM, Reid JG, et al. Clinical whole-exome sequencing for the diagnosis of mendelian disorders. N Engl J Med. 2013; 369(16):1502–1511. [doi]. [PubMed: 24088041]
15. Castets M, Broutier L, Molin Y, et al. DCC constrains tumour progression via its dependence receptor activity. Nature. 2011; 482(7386):534–537. 10.1038/nature10708. [PubMed: 22158121]
16. Devilard E, Bertucci F, Trempat P, et al. Gene expression profiling defines molecular subtypes of classical hodgkin's disease. Oncogene. 2002; 21(19):3095–3102. [doi]. [PubMed: 12082542]
17. Dickinson RE, Dallol A, Bieche I, et al. Epigenetic inactivation of SLIT3 and SLIT1 genes in human cancers. Br J Cancer. 2004; 91(12):2071–2078. [PubMed: 15534609]
18. 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(3):e1000879. [doi]. [PubMed: 20300657]
19. Mehlen P, Delloye-Bourgeois C, Chedotal A. Novel roles for slits and netrins: Axon guidance cues as anticancer targets? Nat Rev Cancer. 2011; 11(3):188–197. 10.1038/nrc3005. [PubMed: 21326323]
20. 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(9):805–812. [doi]. [PubMed: 17539022]

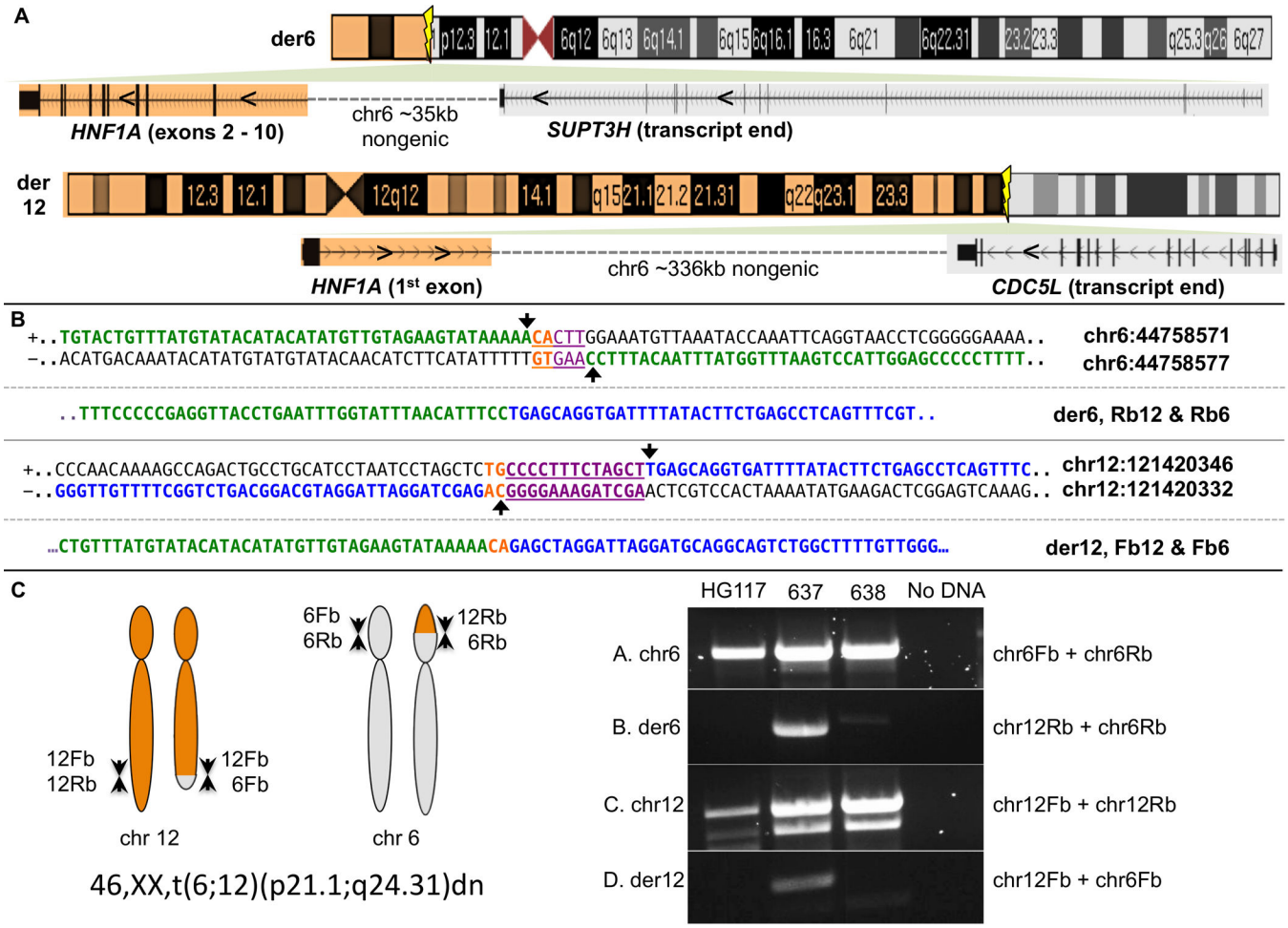

Figure 1. Translocation Characterization and PCR Confirmation of t(6;12) (p21.1;q24.31)

(A) Depicting the derivative chromosomes and gene re-arrangements. The der6 allele has *HNF1A* exons 2 to 10 in the reverse direction and ~ 35kb 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 *CDCL5* gene, which is transcribed on the opposite strand. **(B) Breakpoint Characterization**: There is a 2bp microhomology at the breakpoint (CA, orange), deleted in a 5bp deletion on chr6 (purple and orange underline). A 13bp deletion occurs on chr12 (purple underline), maintaining the CA microhomology. Arrows detail the breakpoint bases, with from 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 Thousand Genomes sample (HG117) were also tested. Pairing primers 6Fb with 12Fb and 6Rb with 12Rb amplifies the expected translocation products in the proband.

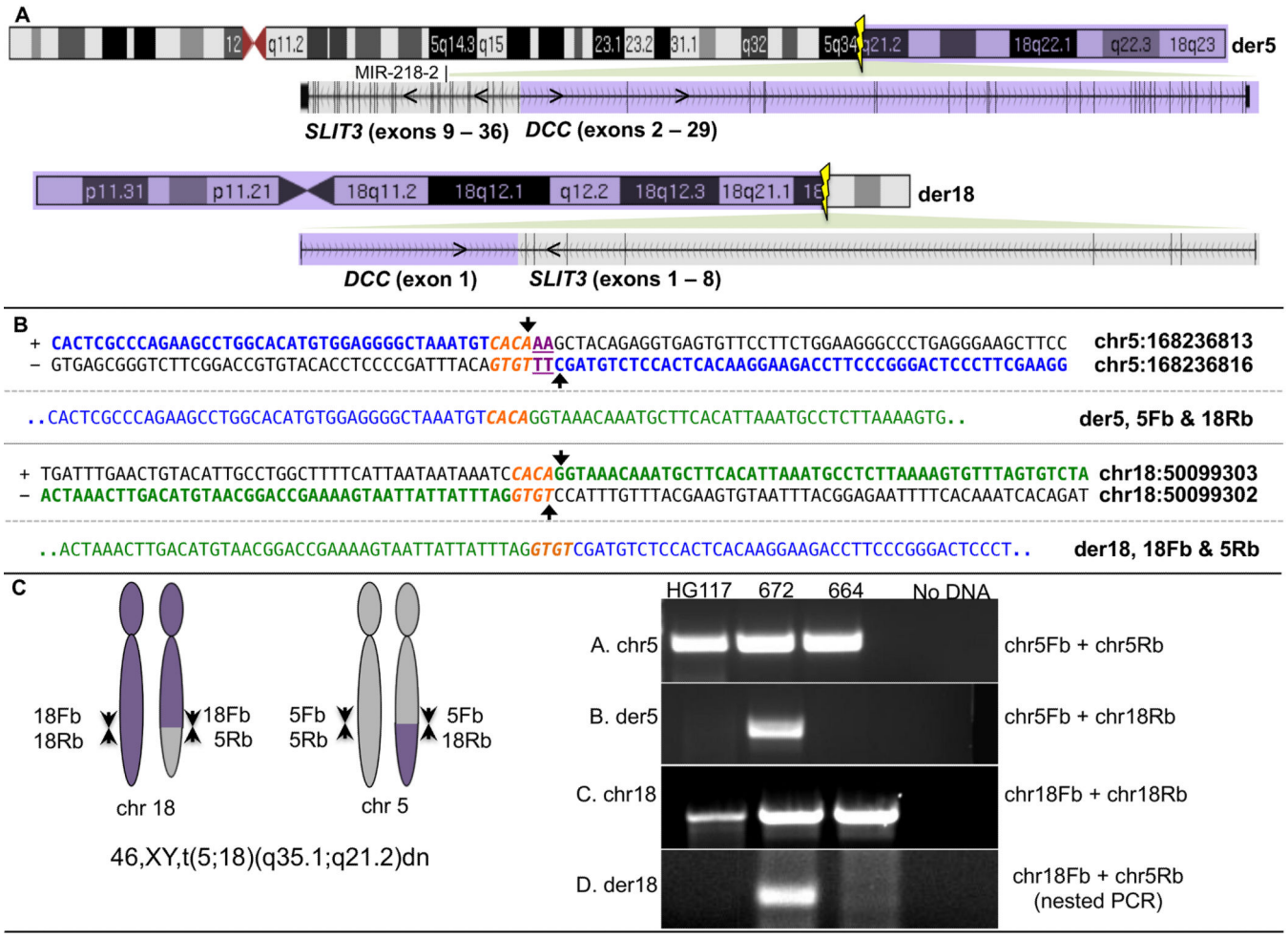

Figure 2. Translocation Characterization and PCR Confirmation of t(5;18) (q35.1;q21.2)

(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 1st 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 4bp region of homology between chr5 and chr18 at the breakpoint (CACA, orange), with a 2bp 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 wildtype chromosomes 5 and 18. The proband's mother (FCP664) and a Thousand 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 non-specific bands.