

Whole-genome sequencing in patients with ciliopathies uncovers a novel recurrent tandem duplication in IFT140

Véronique Geoffroy, Corinne Stoetzel, Sophie Scheidecker, Elise Schaefer, Isabelle Perrault, Séverine Bär, Ariane Kröll, Marion Delbarre, Manuela Antin, Anne-Sophie Leuvrey, et al.

▶ **To cite this version:**

Véronique Geoffroy, Corinne Stoetzel, Sophie Scheidecker, Elise Schaefer, Isabelle Perrault, et al.. Whole-genome sequencing in patients with ciliopathies uncovers a novel recurrent tandem duplication in IFT140. Human Mutation, 2018, 39 (7), pp.983-992. 10.1002/humu.23539 . hal-02371583

HAL Id: hal-02371583
https://hal.science/hal-02371583

Submitted on 19 Nov 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Human Mutation

Whole Genome Sequencing in patients with ciliopathies uncovers a novel recurrent tandem duplication in IFT140

Journal:	*Human Mutation*
Manuscript ID	humu-2018-0045
Wiley - Manuscript type:	Research Article
Date Submitted by the Author:	29-Jan-2018
Complete List of Authors:	geoffroy, veronique; Laboratoire de Génétique Médicale, Institut de Génétique Médicale d'Alsace, INSERM U1112, Fédération de Médecine Translationnelle de Strasbourg (FMTS), Université de Strasbourg Stoetzel, Corinne; UMR_S INSERM U1112, IGMA, Laboratoire de Génétique médicale, Scheidecker, Sophie; UMR_S INSERM U1112, IGMA, Laboratoire de Génétique médicale, SCHAEFER, Elise; Hôpital de Hautepierre, GENETICS PERRAULT, Isabelle; INSERM U781, GENETICS Bär, Séverine; Department of Molecular and Cellular Genetics, UMR7156, Centre National de Recherche Scientifique (CNRS), Université de Strasbourg, Strasbourg, France Kröll, Ariane; UMR_S INSERM U1112, IGMA, Laboratoire de Génétique médicale, Delbarre, Marion; Hôpitaux universitaire de Strasbourg, Laboratoires de diagnostic génétique Antin, Manuela; Nouvel Hopital Civi, Laboratoire de Diagnostic Génétique Jaeger, Anne-Sophie; Hôpitaux universitaire de Strasbourg, Laboratoires de diagnostic génétique Henry, Charline; Inserm U1163-institut Imagine Blanché, Hélène; Fondation Jean-Dausset-Centre d'Etude du Polymorphisme Humain (CEPH), Institut de Génétique Moléculaire, Decker, Eva; Bioscentia, Kloth, Katja; Institut für Humangenetik, Universitätsklinikum Hamburg-Eppendorf, Hamburg, Germany Klaus, Günter; University Marburg, KfH-Nierenzentrum für Kinder und Jugendliche, Marburg, Germany Mache, Christopher; Graz, Martin-Coignard, dominique; CH le Mans, Service de Génétique Médicale McGinn, Steven; CNRGH, Institut de Biologie François Jacob, DRF, CEA, Evry, France Boland, Anne; CNG Deleuze, Jean-François; CNG Friant, Sylvie; Department of Molecular and Cellular Genetics, UMR7156, Centre National de Recherche Scientifique (CNRS), Université de Strasbourg, Strasbourg, France SAUNIER, Sophie; Université René Descartes, ; INSERM U-574, Hopital Necker, Enfants Malades

	ROZET, Jean-Michel; INSERM U781, GENETICS Bergmann, Carsten; Aachen University, Department of Human Genetics Dollfus, Helene; UMR_S INSERM U1112, IGMA, Laboratoire de Génétique médicale, Muller, Jean; Hôpitaux universitaire de Strasbourg, Laboratoires de diagnostic génétique; Laboratoire de Génétique Médicale, Institut de Génétique Médicale d'Alsace, INSERM U1112, Fédération de Médecine Translationnelle de Strasbourg (FMTS), Université de Strasbourg,
Key Words:	Mainzer-Saldino syndrome, IFT140, Structural variation, tandem duplication, whole-genome sequencing, Alu-mediated recombination, copy number variation

SCHOLARONE™
Manuscripts

For Peer Review

Whole Genome Sequencing in patients with ciliopathies uncovers a novel recurrent tandem duplication in *IFT140*

Véronique Geoffroy[1,*], Corinne Stoetzel[1,*], Sophie Scheidecker[1,2], Elise Schaefer[1,3], Isabelle Perrault[4], Séverine Bär[5], Ariane Kröll[1], Marion Delbarre[2], Manuela Antin[2], Anne-Sophie Leuvrey[2], Charline Henry[6], Hélène Blanché[7], Eva Decker[8], Katja Kloth[9], Günter Klaus[10], Christoph Mache[11], Dominique Martin-Coignard[12], Steven McGinn[13], Anne Boland[13], Jean-François Deleuze[7,13], Sylvie Friant[5], Sophie Saunier[6], Jean-Michel Rozet[4], Carsten Bergmann[8,14], Hélène Dollfus[1,15], Jean Muller[1,2]

[1]Laboratoire de Génétique médicale, UMR_S INSERM U1112, IGMA, Faculté de Médecine FMTS, Université de Strasbourg, Strasbourg, France

[2]Laboratoires de Diagnostic Génétique, Hôpitaux Universitaires de Strasbourg, Strasbourg Cedex, France

[3]Service de Génétique Médicale, Hôpitaux Universitaires de Strasbourg, Strasbourg, France.

[4]Laboratory of Genetics in Ophthalmology (LGO), INSERM UMR1163, Institute of Genetic Diseases, Imagine, Paris Descartes University, 75015 Paris, France.

[5]Department of Molecular and Cellular Genetics, UMR7156, Centre National de Recherche Scientifique (CNRS), Université de Strasbourg, Strasbourg, France.

[6]INSERM, U983, Paris Descartes University, Paris, France

[7]Centre d'études du polymorphisme humain-Fondation Jean Dausset, Paris, France

[8]Center for Human Genetics, Bioscientia, Ingelheim, Germany

[9]Institut für Humangenetik, Universitätsklinikum Hamburg-Eppendorf, Hamburg, Germany

[10]University Marburg, KfH-Nierenzentrum für Kinder und Jugendliche, Marburg, Germany

[11]Department of Pediatrics, Medical University of Graz, Graz, Austria.

For Peer Review

[12]Service de Génétique, Centre Hospitalier, CCLAD, Le Mans, France

[13]CNRGH, Institut de Biologie François Jacob, DRF, CEA, Evry, France

[14]Department of Medicine, University Hospital Freiburg, Freiburg, Germany

[15]Centre de Référence pour les affections rares en génétique ophtalmologique, CARGO, Filière SENSGENE, Hôpitaux Universitaires de Strasbourg, 67091 Strasbourg, France.

*equal contributor

Corresponding Author: Jean Muller

Email address: jeanmuller@unistra.fr

For Peer Review

Abstract

Ciliopathies represent a wide spectrum of rare diseases with overlapping phenotypes and a high genetic heterogeneity. Among those, *IFT140* is implicated in a variety of phenotypes ranging from isolated retinis pigmentosa to more syndromic cases such as the Bardet-Biedl syndrome. Using whole genome sequencing in patients with uncharacterized ciliopathies, we identified a novel recurrent tandem duplication of exon 27 to 30 (6.7 kb) in *IFT140,* c.3454-488_4182+2588dup p.(Tyr1152_Thr1394dup), missed by whole exome sequencing. Pathogenicity of the mutation was assessed on the patients' skin fibroblasts. Several hundreds of patients with a ciliopathy phenotype were screened and biallelic mutations were identified in 11 families representing 12 pathogenic variants of which 7 are novel. Among those unrelated families especially with a Mainzer-Saldino syndrome, 8 carried the same tandem duplication (2 at the homozygous state and 6 at the heterozygous state).

In conclusion, we demonstrated the implication of structural variations in *IFT140* related diseases expanding its mutation spectrum. We also provide evidences for a unique genomic event mediated by an *Alu-Alu* recombination occurring on a shared haplotype. We confirm that whole genome sequencing can be instrumental in the ability to detect structural variants for genomic disorders.

Keywords: *IFT140*, Mainzer-Saldino syndrome, structural variation, copy number variation, tandem duplication, *Alu*-mediated recombination, whole-genome sequencing.

For Peer Review

Background

Mainzer-Saldino syndrome (MSS, MIM 266920) is a rare (< 1/1,000,000) autosomal recessive ciliopathy characterized by severe early-onset retinal dystrophy, phalangeal cone-shaped epiphyses, chronic renal failure, and mild radiographic abnormality of the proximal femur (Perrault, et al., 2012), known to be caused by *IFT140* (Khan, et al., 2014; Perrault, et al., 2012; Schmidts, et al., 2013) and *IFT172* mutations (Halbritter, et al., 2013). Both genes have been implicated in several other ciliopathies ranging from isolated Retinis Pigmentosa (RP) to more syndromic cases such as Jeune or Bardet-Biedl syndromes (Bifari, et al., 2016; Bujakowska, et al., 2015; Schaefer, et al., 2016).

Intraflagellar transport (*IFT*) genes are involved in a bidirectional (anterograde and retrograde) transport process essential for the assembly and the maintenance of the cilia through the redistribution of ciliary proteins between the cell body and the cilium. In particular, IFT140 a component of the IFT complex A (IFT-A) responsible of the retrograde IFT, is vital for both the development and the maintenance of outer segments of photoreceptors and has a specific role in opsin transport across the connecting cilium (Blacque, et al., 2006; Crouse, et al., 2014) of the photoreceptors. Previously, 46 different mutations in 56 families have been reported in *IFT140* (Additional File 1: Table S1), encompassing missenses, essential splice sites, stop and frameshifts mutations but no structural variations (SV) such as copy number variant (CNV). Interestingly, 10 mono allelic variants cases have also been reported with no second pathogenic allele detected to date (Additional File 1: Table S1).

In this study, we report for the first time a structural variation (tandem duplication) in *IFT140* identified by Whole Genome Sequencing (WGS) and missed by Whole Exome Sequencing (WES). The impact of this mutation was further assessed in patients' skin fibroblasts. Given

For Peer Review

the difficulty to identify this mutation and the genomic context surrounding the breakpoints, we speculated that it might have been missed by other genetic screenings and that several other families might carry this mutation. We finally report 8 unrelated families carrying the mutation either at the homozygous or the heterozygous state, out of 11 families identified with *IFT140* biallelic mutations in total. Moreover, the characterization of the breakpoints allowed to delineate a potential molecular mechanism and to design a specific duplex PCR that will help screening further patients (including our cohorts).

Materials and Methods

Subjects

Study protocols used in each cohort have been approved by the corresponding Institutional Review Board or equivalent committees (as an example in Strasbourg, "Comité Protection des Personnes" EST IV, N°DC-20142222), and written informed consent was given from each participant or parents. Our research complies with the Declaration of Helsinki. Written informed consent for open-access publication was provided by the participants or their parents. DNA of additional affected and unaffected family members was requested whenever it was considered informative. Skin fibroblasts were obtained for family A. Clinical data for all 11 families are presented in Additional File 1: Table S2.

Whole Genome Sequencing

WGS was performed for the two affected siblings from family A (A-II.1 and A-II.2) and their parents (A-I.1 and A-I.2) by the Centre National de Génotypage (Institut de Génomique, CEA). Genomic DNA was used to prepare a library for whole genome sequencing, using the Illumina TruSeq DNA PCR-Free Library Preparation Kit, according to the manufacturer's

For Peer Review

instructions. After normalization and quality control, qualified libraries have been sequenced on a HiSeq2000 platform from Illumina (Illumina Inc., CA, USA), as paired-end 100 bp reads. At least 3 lanes of HiSeq2000 flow cell have been produced for each sample, in order to reach an average sequencing depth of 30x for each sample. Sequence quality parameters have been assessed throughout the sequencing run and standard bioinformatics analysis of sequencing data was based on the Illumina pipeline to generate FASTQ file for each sample. The sequence reads were aligned to the reference sequence of the human genome (GRCh37) using the Burrows-Wheeler Aligner (BWA V7.12) (Li and Durbin, 2010). The UnifiedGenotyper and HaplotypeCaller modules of the Genome Analysis ToolKit (GATK) (DePristo, et al., 2011), Platypus (http://www.well.ox.ac.uk/platypus) and Samtools (Li, et al., 2009) were used for calling both single nucleotide variations (SNV) and small insertion/deletion (indel).

Bioinformatics analysis

Annotation and ranking of SNV and indel were performed by VaRank (Geoffroy, et al., 2015) in combination with the Alamut Batch software (Interactive Biosoftware, Rouen, France). Very stringent filtering criteria were used for excluding non-pathogenic variants, in particular: (1) variants represented with an allele frequency of more than 1% in public variation databases either dbSNP 138 (Sherry, et al., 2001), the Exome Variant Server (NHLBI GO Exome Sequencing Project, http://evs.gs.washington.edu/EVS/), the 1000Genomes (Genomes Project, et al., 2015), the ExAC browser database(Lek, et al., 2016) or our internal exome database, (2) variants in 5′ UTR, 3′ UTR, downstream, upstream or intronic locations without pathogenic prediction of local splice effect, and (3) synonymous variants without prediction of local splice effect. Variant effect on the nearest splice site was predicted using MaxEntScan (Yeo and Burge, 2004), NNSplice (Reese, et al., 1997) and Splice Site Finder (Shapiro and

For Peer Review

Senapathy, 1987). Our analysis was focused on compound heterozygous and homozygous variants (SNV/indel/SV) consistent with a recessive mode of transmission. Structural variants were predicted using by default the CANOES program (Backenroth, et al., 2014) and annotated thanks to our in house script AnnotSV (manuscript in preparation, http://www.lbgi.fr/AnnotSV/) based on the classical annotations such as the Database of Genomic Variants (DGV) (MacDonald, et al., 2014). The *IFT140* nomenclature is based on the accession number NM_014714.3 from the RefSeq database (O'Leary, et al., 2016). Genomic coordinates are defined according to GRCh37/hg19 assembly downloaded from the University of California Santa Cruz (UCSC) genome browser (Tyner, et al., 2017).

Sanger validation and segregation

Sanger sequencing was performed by PCR amplification with 50 ng of genomic DNA template. The primers were designed with Primer 3 (http://frodo.wi.mit.edu/primer3) and are detailed in Additional File 1: Table S3. Bidirectional sequencing of the purified PCR products was performed by GATC Sequencing Facilities (Konstanz, Germany).

qPCR quantification

Absolute quantification was performed using the SyberGreen Mastermix Quantitect (Qiagen) measured on the LightCycler 480 (Roche). Amplicons were designed in exon 30 of *IFT140* and compared to two reference genes (*HBB* and *HMBS*). Each samples have been done in duplicate and standard deviation was <0.2. Standard curve have been established using 5 different DNA quantity (50ng, 25ng, 10ng, 5ng and 1ng) and PCR efficiency was assessed for each amplicon (1.9<PCR efficiency<2.0). Ratio between the mean concentrations from the amplicon of interest compared to reference gene was collected. The mean values for both

For Peer Review

control genes was calculated for each patient and plotted. The primers were designed with Primer 3 and are detailed in Additional File 1: Table S3

RNA analysis

RNA was extracted from skin fibroblasts of individual II.1 and a healthy unrelated control using Rneasy RNA kit (Qiagen) then we performed reverse transcription using the iScriptTM cDNA Synthesis Kit (BioRad, Hercules, CA).

Cell culture

Fibroblasts of patients and control individuals were obtained by skin biopsy as previously described (Scheidecker, et al., 2014). To induce primary cilium formation, cells were deprived of serum by growth for 24 hrs in DMEM with 1% PSG but only 0.1% FCS (conditions -FCS) as previously described (Stoetzel, et al., 2016).

Immunofluorescence

Primary fibroblasts from patients and control individuals were grown in Nunc Lab-Tek chamber slides (Thermo Scientific) and ciliogenesis was done as described above. Primary cilia were labelled with an antibody directed against acetylated a-tubulin highlighting the axoneme41. Pictures were taken either on a fluorescence microscope (Figure 2A) or a confocal microscope (Figure 2C). Primary and secondary antibodies used in this study as well as their dilution are indicated in Additional File 1: Table S4.

Cohort screening

A duplex PCR was designed to specifically detect the tandem duplication. Primers and example conditions are provided in Additional File 1: Table S3 and Figure S1. Additional

For Peer Review

targeted exome sequencing datasets have been reanalyzed (see Additional File 1: supplementary Methods).

Alu pair analysis

We obtained the annotations and sequences for the two *Alu* repeat in the duplication junction region from the RepeatMasker track of the UCSC genome browser (Tyner, et al., 2017). We subsequently aligned the *Alu* sequences and computed the identity using the online tool EMBOSS Matcher (Rice, et al., 2000).

Results

Identification of biallelic mutations in *IFT140*

Two affected individuals born from healthy non consanguineous parents (family A, Figure 1A), named II.1 and II.2, were referred to our lab with retinitis pigmentosa, short stature (152cm and 158cm respectively at 16 and 15yo), brachydactyly, moderate renal failure and overweight for II.1 (Additional File 1: Table S2). This Bardet-Biedl like phenotype prompted us to perform an extensive first genetic analysis including Sanger sequencing of recurrent BBS mutations (Muller, et al., 2010), targeted exome sequencing (Redin, et al., 2012) and WES without success. Indeed no biallelic mutations including copy number variant could be identified (see Additional File 1: Table S5). Hence, considering that affected individuals from family A may harbor mutations in a region not (well) covered by the WES (i.e. deep intronic, promotor…), we applied WGS to the 2 affected individuals and their parents. A combined analysis including both SNV/indel and structural variant revealed biallelic mutations in *IFT140,* a gene known to be responsible of several ciliopathies (ranging from isolated RP to more syndromic cases such as Jeune or Mainzer-Saldino syndromes) compatible with the phenotype of the affected patients.

For Peer Review

First, the two affected individuals were found heterozygous for a possible splice mutation in intron 20, c.2577+25G>A, predicted to have a local splice effect as a new strong donor site. This variant was reported previously and is not present in any variation database (e.g. ExAC or 1000G). Co-segregation analysis revealed the paternal inheritance (Figure 1A, primers used are detailed in Additional File 1: Table S3). Patient's RNA analysis confirmed that a novel donor splice site within the intron 20 resulted in the incorporation of 21 additional bp (r.2577_2578insGTGAGGGGCGCCCGCCATGGG) that are predicted to add 7 new amino acids to the protein sequence (p.Leu859_Glu860insValArgGlyAlaArgHisGly) (Figure 1B).

Second, in *trans* to this mutation, structural variant analysis revealed a maternally inherited (Figure 1A) duplication of exons 27 to 30 (of sizes 6.7 kb). Analysis of the mapped reads revealed that the duplication is occurring in tandem in direct orientation within the gene (Additional File 1: Figure S2) with breakpoints located in intron 26 and intron 30. This tandem duplication has not been reported previously in patients and is not present in DGV. The results were confirmed by qPCR (Figure 1C) delineating the expected maternal inheritance. The duplication is predicted to be in frame and to add 243 amino acids within the tetratricopeptide repeat (TPR) domain of the IFT140 protein (c.3454-488_4182+2588dup, p.Tyr1152_Thr1394dup). RNA analysis revealed that the duplicated exons are transcribed (Figure 1D). Western blot analysis revealed only a single band of the approximate size of the wild type protein suggesting that the allele carrying the duplication is not detected as a protein (Figure 1E).

Immunofluorescence analysis in patients fibroblasts revealed that IFT140 is mislocalized compared to sex/age matched control cells. Indeed, when grown in ciliary conditions, patient skin fibroblasts developed primary cilia less often (Figure 2**A**). Thus while 75% of control cells developed a primary cilium, only about 50-55% of patient cells did so (Figure 2**B**). It is to notice that primary cilia did not seem to be altered since when cells were deprived of

For Peer Review

serum, fixed for immunofluorescence and labeled with an anti-acetylated tubulin antibody, their cilium as observed by fluorescence microscope and measured with image J (Schneider, et al., 2012) was about 2-2,5 µm both in control and patient cells (Figure 2**B**). Nevertheless, a difference in organisation could be noticed when co-labelling was performed with an anti-IFT140 antibodies (Figure 2**C**). Then, localisation of IFT140 at the base of the cilium could be observed using confocal microscopy in 82% of ciliated control cells but only in about 14% of ciliated patient fibroblasts (Figure 2**D**). Thus, the mutations found in the patients lead to a significant decrease of cilium formation and a loss of IFT140 localisation at the base of the cilium.

Mutation screening in a large cohort

To explore additional cases of this duplication possibly missed by prior analysis, we retrospectively screened our cohorts and reanalyzed available high throughput sequencing panels including the *IFT140* gene. As mutations in *IFT140* are known to cause isolated to syndromic retinal degeneration (Bifari, et al., 2016), this included 126 patients using the Leber panel of the Imagine institute, 117 patients using the RP panel of Strasbourg Hospital , and 104 patients using the Ciliome panel of the Imagine institute (see Methods and Additional File 1: methods online). We also setup a duplex PCR around the breakpoints (Additional File 1: Figure S1) and screened all our negative cases, including 207 from our Bardet-Biedl cohort, 150 isolated RP and 40 *IFT140* heterozygous patients. All together we identified 11 additional families positive for biallelic mutations in *IFT140* (Figure 3 and Additional File 1: Table S1) among which 10 families have a phenotype compatible with a Mainzer Saldino syndrome and one patient presents with isolated retinitis pigmentosa. Among the MSS patients, three patients presented a phenotype overlapping with another ciliopathy: two with Jeune syndrome and one with Sensenbrenner syndrome. Eight of the 11 families carried the exact same tandem duplication either at the heterozygous state (3 copies) or at the

For Peer Review

homozygous state (4 copies) (Figure 3A, Additional File 1: Figure S3). Among the novel mutations identified, we observed 2 missenses at the heterozygous state c.1319T>C (p.Leu440Pro), c.2177C>T (p.Pro726Leu) that are both predicted deleterious (Additional File 1: supplementary Methods). Among the novel mutations we also identified another CNV, a large deletion encompassing exon 27 to 29 (c.3454-1005_4040+737delinsCCC, Additional File 1: Figure S4) that does not share the same breakpoints compared to the tandem duplication. Thus to date, 120 pathogenic alleles have been reported for *IFT140* of which the tandem duplication represents 10 alleles. Interestingly this tandem duplication is the second most frequent mutation for *IFT140*-affected patients.

Founder effect or recurrent mechanism?

Given the relatively high frequency (8/65 families) of this duplication in the patients described with *IFT140* mutations, we performed 3 additional explorations to assess whether the duplication was an independent recurrent mechanism or the same mutational event derived from a common ancestry suggesting a founder effect. First, we screened a French cohort of 354 WGS of healthy individuals (FREX dataset), for which CNV analysis (CANOES) as well as visual inspections of the bam files (IGV batch scripting (Thorvaldsdottir, et al., 2013)) did not reveal any corresponding allele, as expected due to the *IFT140* diseases prevalence (e.g. MSS: 1/1,000,000). Second, we compared the haplotype at the *IFT140* locus from the different patients for which NGS data was available. In total 15 haplotypes, including 6 carrying the duplication, from 3 different families, were analyzed (see Additional File 1: Table S6). We observed a shared haplotype between the duplication carriers at the *IFT140* locus, suggesting a founder effect. Third, we characterized the genomic localization around the breakpoint junctions. Sanger sequencing confirmed the breakpoint's coordinates (Figure 1A and Additional File 1: Figure S2) and demonstrated that the exact same breakpoint is shared among all patients tested (Family C was not investigated). The

For Peer Review

breakpoints overlapped two distinct repeated elements from the *Alu* family, *AluJb* and *AluJr* (Figure 4A). Pairwise alignment of the two *Alu* sequences (Additional File 1: Figure S5) revealed a conservation of 68% of identity that reached 81% in the junction region (Figure 4B). The tandem duplication is resulting in an *Alu* hybrid (Figure 4C) at both sides. *Alu-Alu* recombination is a prominent mechanism underlying the formation of pathogenic SV associated with distinct diseases (Boone, et al., 2014; Stankiewicz, et al., 2009). Altogether, this led us to speculate that this tandem duplication is a rare event, which arises from a common ancestor due to an *Alu-Alu* mediated genomic recombination.

Discussion

The advent of next generation sequencing, especially Whole Exome Sequencing (WES), in the past 10 years was a tremendous step that led to the identification of many mutations accelerating the discovery of novel genes involved in human genetic diseases (Chong, et al., 2015). Nevertheless, the diagnostic yield of WES is plateauing between 25% to 50% depending on the disease (Chong, et al., 2015; Taylor, et al., 2015) leaving many patients still with no molecular diagnosis. The use of RNA sequencing has recently extended the percentage by 10% to 66% (depending on the context: with or without a strong candidate by prior DNA sequencing) by helping identifying unseen defects or interpreting variations found in the WES (Cummings, et al., 2017; Kremer, et al., 2017). Lastly, Whole Genome Sequencing (WGS) of human patients has proven to be an instrumental tool for identifying the whole spectrum of genetic anomalies (Gilissen, et al., 2014) and will replace in a close future other genetic screening.

In line with this, we applied a range of genetic screening including targeted sequencing, WES and then finally WGS to unsolved patients affected with a ciliopathy. The combination of

For Peer Review

WGS and CNV analysis was able to uncover a tandem duplication unseen by the WES and to reconsider a distant splice site mutation.

Interestingly, the tandem duplication could be identified in family A only by using the WGS and not the WES data. Indeed, CANOES failed to identify the event in one of the individuals (Additional File 1: Table S7), which prevented us from considering this as a mutation of interest but made us rather discard it as a false positive of the method. Thereupon, it is well known that CNV detection from next generation sequencing (NGS) datasets is feasible and many publication reported positive results either on gene panels or WES (de Ligt, et al., 2013; Redin, et al., 2012). However, this remains challenging (Tan, et al., 2014). WGS provides a much more uniform distribution of sequencing-quality parameters and by nature does not have the restriction of noncontiguous regions of interests (captured exons). Therefore, it is more suited for SV identification (Belkadi, et al., 2015). Having that in mind, we further tested the CNV detection using 2 other popular programs (see Additional File 1: Table S7) which failed to detect the tandem duplication in family A patients (Additional File 1: Table S7). These results confirmed many efforts are still required to ensure a high quality CNV detection from NGS data, even with the large number of programs and methods available. Nevertheless, thanks to the split reads from WGS data, once a CNV is detected one can easily define the CNV breakpoints and further characterize the mechanism, as done in this study (Additional File 1: Figure S2).

Analysis of the breakpoint positions and in particular the split reads allowed us to identify overlapping *Alu* elements at the breakpoint junction. Alignment of the *Alu* regions surrounding the junction revealed that both elements share 81% identity and exhibit a microhomology region of 6 nucleotides. Because the recombination occurred between homologous sequences (e.g. imperfect match substrates), it is most likely mediated by mechanisms other than nonallelic homologous recombination, which are mostly recurrent

For Peer Review

events (NAHR, for review see (Carvalho and Lupski, 2016)). Among the possible molecular mechanisms, that include non-homologous end joining (NHEJ) and replication-based mechanisms such as break-induced replication (BIR), microhomology-mediated BIR (MMBIR), serial replication slippage (SRS) and fork stalling and template switching (FoSTeS), we believe that the presence of a microhomology region points most likely to a MMBIR driven recombination. Given the nonrecurrent nature of those mechanisms, the common haplotype and the exact same breakpoint junction observed in most of our patients, we hypothesized that the genomic event that gave rise to this tandem duplication occurred once on a common *IFT140* haplotype.

IFT140 is a 1462-amino-acid protein encoded by 31 exons and composed of 5 WD repeats and 9 tetratricopeptide repeats (TPR) both known to act as protein-protein interaction domains that might be involved in IFT140 interactions within the IFT-A complex (Zhu, et al., 2017). To date, 45 different pathogenic alleles have already been described throughout the *IFT140* gene with no clustering or domain preference (Figure 4A). Most of them are private mutations but several are found multiple times like the c.634G>A (p.Gly212Arg) (7 alleles) which has been proven to affect the splicing of exon 6 (Helm, et al., 2017) or the c.2399+1G>T (5 alleles) (Additional File 1: Table S1). The most frequent mutation (31 alleles) is a missense (c.1990G>A, p.Glu664Lys) that has been observed in multiple studies and especially in 11 consanguineous families from the Arabian Peninsula sharing a common ancestor (Bifari, et al., 2016) which might bias the allele count. Nevertheless, the tandem duplication described in our study is the second most frequent cause of mutation in *IFT140* representing 10 alleles.

In family A, the two mutations identified required extensive and careful analysis to be considered. One mutation was predicted to affect the splicing of exon 20 (c.2577+25G>A) which was confirmed at the RNA level (r.2577_2578insGTGAGGGGCGCCCGCCATGGG), and suggested to add 7 amino acid to the protein sequence

For Peer Review

(p.Leu859_Glu860insValArgGlyAlaArgHisGly). The tandem duplication of exons 27 to 30 (c.3454-488_4182+2588dup, p.Tyr1152_Thr1394dup) is predicted to add 243 amino acids within the TPR repeat domain. Skin fibroblast analysis revealed RNA synthesis in the patient cells but no protein could be identified by Western Blot at the predicted size. We hypothesize that the protein was not stable enough to maintain its 3D structure. Functional analysis on the patients' cells further revealed a reduced level of ciliated cells and mislocalization of the IFT140 mutant away from the cilia base.

The Mainzer Saldino syndrome is a syndrome characterized by skeletal phenotype with phalangeal cone-shaped epiphyses, chronic renal disease, and retinal dystrophy. Overlapping phenotypes have been described with other skeletal ciliopathies such as the Jeune or Sensenbrenner syndrome. In our cohort the diagnosis of MSS was not made in all patients before the molecular analysis. However, all of them except one with an isolated retinitis pigmentosa, fulfilled the criteria for this diagnosis after a new careful clinical and radiological examination. Indeed, 11 patients presented digits anomalies, including 8 with phalangeal cone-shaped epiphyses. Nine patients developed renal failure among which 7 have a severe and early onset renal disease (1.5 years old to 23 years old) as previously described (Schmidts, et al., 2013). The retinal dystrophy is constantly reported with a highly variable age of onset. Among our cohort, a single patient presented an isolated retinitis pigmentosa (as previously reported by others (Hull, et al., 2016)). Two patients have thoracic dystrophy which is a clinical criteria of Jeune syndrome and one patient has a craniosynostosis compatible with a Sensenbrenner syndrome. These descriptions illustrate well the overlap existing between the different ciliopathies associated with mutation in *IFT140*. On a molecular level, no genotype-phenotype correlation could be established in our cohort, especially regarding the patients carrying the duplication at the homozygous or the heterozygous state.

For Peer Review

Conclusions

In summary, we report here 11 novel unrelated families with mutations in *IFT140*. Among them, 8 families carry a recurrent tandem duplication of 4 exons either at the heterozygous state or the homozygous state, for which we have assessed the pathogenicity in the pateints cells. This is the first time that a structural variation is reported in *IFT140* expanding the mutation spectrum for this gene. Notably, this large duplication was missed by the WES analysis but uncoverd thanks to the whole genome, pointing out the power of such analysis.

Availability of data and material

Data generated or analyzed during this study are included in the published article and the corresponding supplementary data. The raw sequencing data generated in the course of this study are not publicly available due to the protocol and the corresponding consents used that did not include such information. All variants have been submitted to ClinVar (https://www.ncbi.nlm.nih.gov/clinvar/). Anonymised NGS data and genomic variant data files will be made available upon request from qualified investigators studying the molecular basis of genomic disorders. Datasets can be obtained via the corresponding author on reasonable request.

For Peer Review

Acknowledgements

We would like to thank the patients and their family for their participation. We thank Daniel Backenroth and Olivier Quenez for their help in the bioinformatics setup of CANOES, Arnaud Kress and Antony Le Béchec for informatic support, Amélie Piton, Bénédicte Gérard and Ilia Humbert for scientific discussion. We also thank Emmanuelle Génin, Pierre Lindenbaum and Richard Redon for giving us access to the genome dataset from the FREX

project and the Genomic and Bioinformatic Platform of Institut Imagine (Patrick Nitschké, and Cécile Masson).

Disclosure Statment

Eva Decker and Carsten Bergmann are employees of Bioscientia/Sonic Healthcare. Günter Klaus benefited from a travel grants from The Apheresis Research Institute Cologne, the DGFF (Die Deutsche Gesellschaft zur Bekämpfung von Fettstoffwechselstörungen und ihren Folgeerkrankungen) and is consultant to Vifor pharma, Switzerland concerning a pediatric phosphat binder trial.

Funding Information

The WGS research was supported by the Laboratory of Excellence GENMED (Medical Genomics) grant no. ANR-10-LABX-0013 managed by the National Research Agency (ANR) part of the Investment for the Future program. Whole exome sequencing was performed by the IGBMC Microarray and Sequencing platform, a member of the 'France Génomique' consortium (ANR-10-INBS-0009) and funded by "La Fondation Maladie Rare". The RP panel is supported by the French program PHRC I 2013 HUS N° 5724 and by grants from the Retina France to IP, UNADEV- AVIESAN ITMO MNP to JMR. In addition, CB holds a part-time faculty appointment at the University of Freiburg. His research lab receives support from the Deutsche Forschungsgemeinschaft (DFG) Collaborative Research Centre (SFB) KIDGEM 1140 and the Federal Ministry of Education and Research (BMBF, 01GM1515C).

For Peer Review

References

Backenroth D, Homsy J, Murillo LR, Glessner J, Lin E, Brueckner M, Lifton R, Goldmuntz E, Chung WK, Shen Y. 2014. CANOES: detecting rare copy number variants from whole exome sequencing data. Nucleic Acids Res 42(12):e97.

Belkadi A, Bolze A, Itan Y, Cobat A, Vincent QB, Antipenko A, Shang L, Boisson B, Casanova JL, Abel L. 2015. Whole-genome sequencing is more powerful than whole-exome sequencing for detecting exome variants. Proc Natl Acad Sci U S A 112(17):5473-8.

Bifari IN, Elkhamary SM, Bolz HJ, Khan AO. 2016. The ophthalmic phenotype of IFT140-related ciliopathy ranges from isolated to syndromic congenital retinal dystrophy. Br J Ophthalmol 100(6):829-33.

Blacque OE, Li C, Inglis PN, Esmail MA, Ou G, Mah AK, Baillie DL, Scholey JM, Leroux MR. 2006. The WD repeat-containing protein IFTA-1 is required for retrograde intraflagellar transport. Mol Biol Cell 17(12):5053-62.

Boone PM, Yuan B, Campbell IM, Scull JC, Withers MA, Baggett BC, Beck CR, Shaw CJ, Stankiewicz P, Moretti P and others. 2014. The Alu-rich genomic architecture of SPAST predisposes to diverse and functionally distinct disease-associated CNV alleles. Am J Hum Genet 95(2):143-61.

Bujakowska KM, Zhang Q, Siemiatkowska AM, Liu Q, Place E, Falk MJ, Consugar M, Lancelot ME, Antonio A, Lonjou C and others. 2015. Mutations in IFT172 cause isolated retinal degeneration and Bardet-Biedl syndrome. Hum Mol Genet 24(1):230-42.

Carvalho CM, Lupski JR. 2016. Mechanisms underlying structural variant formation in genomic disorders. Nat Rev Genet 17(4):224-38.

Chong JX, Buckingham KJ, Jhangiani SN, Boehm C, Sobreira N, Smith JD, Harrell TM, McMillin MJ, Wiszniewski W, Gambin T and others. 2015. The Genetic Basis of Mendelian Phenotypes: Discoveries, Challenges, and Opportunities. Am J Hum Genet 97(2):199-215.

Crouse JA, Lopes VS, Sanagustin JT, Keady BT, Williams DS, Pazour GJ. 2014. Distinct functions for IFT140 and IFT20 in opsin transport. Cytoskeleton (Hoboken) 71(5):302-10.

Cummings BB, Marshall JL, Tukiainen T, Lek M, Donkervoort S, Foley AR, Bolduc V, Waddell LB, Sandaradura SA, O'Grady GL and others. 2017. Improving genetic diagnosis in Mendelian disease with transcriptome sequencing. Sci Transl Med 9(386).

de Ligt J, Boone PM, Pfundt R, Vissers LE, Richmond T, Geoghegan J, O'Moore K, de Leeuw N, Shaw C, Brunner HG and others. 2013. Detection of clinically relevant copy number variants with whole-exome sequencing. Hum Mutat 34(10):1439-48.

DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, Philippakis AA, del Angel G, Rivas MA, Hanna M and others. 2011. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet 43(5):491-8.

Genomes Project C, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, Korbel JO, Marchini JL, McCarthy S, McVean GA and others. 2015. A global reference for human genetic variation. Nature 526(7571):68-74.

Geoffroy V, Pizot C, Redin C, Piton A, Vasli N, Stoetzel C, Blavier A, Laporte J, Muller J. 2015. VaRank: a simple and powerful tool for ranking genetic variants. PeerJ 3:e796.

Gilissen C, Hehir-Kwa JY, Thung DT, van de Vorst M, van Bon BW, Willemsen MH, Kwint M, Janssen IM, Hoischen A, Schenck A and others. 2014. Genome sequencing identifies major causes of severe intellectual disability. Nature 511(7509):344-7.

Halbritter J, Bizet AA, Schmidts M, Porath JD, Braun DA, Gee HY, McInerney-Leo AM, Krug P, Filhol E, Davis EE and others. 2013. Defects in the IFT-B component IFT172 cause Jeune and Mainzer-Saldino syndromes in humans. Am J Hum Genet 93(5):915-25.

For Peer Review

Helm BM, Willer JR, Sadeghpour A, Golzio C, Crouch E, Vergano SS, Katsanis N, Davis EE. 2017. Partial uniparental isodisomy of chromosome 16 unmasks a deleterious biallelic mutation in IFT140 that causes Mainzer-Saldino syndrome. Hum Genomics 11(1):16.
Hull S, Owen N, Islam F, Tracey-White D, Plagnol V, Holder GE, Michaelides M, Carss K, Raymond FL, Rozet JM and others. 2016. Nonsyndromic Retinal Dystrophy due to Bi-Allelic Mutations in the Ciliary Transport Gene IFT140. Invest Ophthalmol Vis Sci 57(3):1053-62.
Khan AO, Bolz HJ, Bergmann C. 2014. Early-onset severe retinal dystrophy as the initial presentation of IFT140-related skeletal ciliopathy. J AAPOS 18(2):203-5.
Kremer LS, Bader DM, Mertes C, Kopajtich R, Pichler G, Iuso A, Haack TB, Graf E, Schwarzmayr T, Terrile C and others. 2017. Genetic diagnosis of Mendelian disorders via RNA sequencing. Nat Commun 8:15824.
Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, O'Donnell-Luria AH, Ware JS, Hill AJ, Cummings BB and others. 2016. Analysis of protein-coding genetic variation in 60,706 humans. Nature 536(7616):285-91.
Li H, Durbin R. 2010. Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics 26(5):589-95.
Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, Genome Project Data Processing S. 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25(16):2078-9.
MacDonald JR, Ziman R, Yuen RK, Feuk L, Scherer SW. 2014. The Database of Genomic Variants: a curated collection of structural variation in the human genome. Nucleic Acids Res 42(Database issue):D986-92.
Muller J, Stoetzel C, Vincent MC, Leitch CC, Laurier V, Danse JM, Helle S, Marion V, Bennouna-Greene V, Vicaire S and others. 2010. Identification of 28 novel mutations in the Bardet-Biedl syndrome genes: the burden of private mutations in an extensively heterogeneous disease. Hum Genet 127(5):583-93.
O'Leary NA, Wright MW, Brister JR, Ciufo S, Haddad D, McVeigh R, Rajput B, Robbertse B, Smith-White B, Ako-Adjei D and others. 2016. Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. Nucleic Acids Res 44(D1):D733-45.
Perrault I, Saunier S, Hanein S, Filhol E, Bizet AA, Collins F, Salih MA, Gerber S, Delphin N, Bigot K and others. 2012. Mainzer-Saldino syndrome is a ciliopathy caused by IFT140 mutations. Am J Hum Genet 90(5):864-70.
Redin C, Le Gras S, Mhamdi O, Geoffroy V, Stoetzel C, Vincent MC, Chiurazzi P, Lacombe D, Ouertani I, Petit F and others. 2012. Targeted high-throughput sequencing for diagnosis of genetically heterogeneous diseases: efficient mutation detection in Bardet-Biedl and Alstrom syndromes. J Med Genet 49(8):502-12.
Reese MG, Eeckman FH, Kulp D, Haussler D. 1997. Improved splice site detection in Genie. J Comput Biol 4(3):311-23.
Rice P, Longden I, Bleasby A. 2000. EMBOSS: the European Molecular Biology Open Software Suite. Trends Genet 16(6):276-7.
Schaefer E, Stoetzel C, Scheidecker S, Geoffroy V, Prasad MK, Redin C, Missotte I, Lacombe D, Mandel JL, Muller J and others. 2016. Identification of a novel mutation confirms the implication of IFT172 (BBS20) in Bardet-Biedl syndrome. J Hum Genet 61(5):447-50.
Scheidecker S, Etard C, Pierce NW, Geoffroy V, Schaefer E, Muller J, Chennen K, Flori E, Pelletier V, Poch O and others. 2014. Exome sequencing of Bardet-Biedl syndrome patient identifies a null mutation in the BBSome subunit BBIP1 (BBS18). J Med Genet 51(2):132-6.
Schmidts M, Frank V, Eisenberger T, Al Turki S, Bizet AA, Antony D, Rix S, Decker C, Bachmann N, Bald M and others. 2013. Combined NGS approaches identify mutations in the

For Peer Review

intraflagellar transport gene IFT140 in skeletal ciliopathies with early progressive kidney Disease. Hum Mutat 34(5):714-24.
Schneider CA, Rasband WS, Eliceiri KW. 2012. NIH Image to ImageJ: 25 years of image analysis. Nat Methods 9(7):671-5.
Shapiro MB, Senapathy P. 1987. RNA splice junctions of different classes of eukaryotes: sequence statistics and functional implications in gene expression. Nucleic Acids Res 15(17):7155-74.
Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM, Sirotkin K. 2001. dbSNP: the NCBI database of genetic variation. Nucleic Acids Res 29(1):308-11.
Stankiewicz P, Sen P, Bhatt SS, Storer M, Xia Z, Bejjani BA, Ou Z, Wiszniewska J, Driscoll DJ, Maisenbacher MK and others. 2009. Genomic and genic deletions of the FOX gene cluster on 16q24.1 and inactivating mutations of FOXF1 cause alveolar capillary dysplasia and other malformations. Am J Hum Genet 84(6):780-91.
Stoetzel C, Bar S, De Craene JO, Scheidecker S, Etard C, Chicher J, Reck JR, Perrault I, Geoffroy V, Chennen K and others. 2016. A mutation in VPS15 (PIK3R4) causes a ciliopathy and affects IFT20 release from the cis-Golgi. Nat Commun 7:13586.
Tan R, Wang Y, Kleinstein SE, Liu Y, Zhu X, Guo H, Jiang Q, Allen AS, Zhu M. 2014. An evaluation of copy number variation detection tools from whole-exome sequencing data. Hum Mutat 35(7):899-907.
Taylor JC, Martin HC, Lise S, Broxholme J, Cazier JB, Rimmer A, Kanapin A, Lunter G, Fiddy S, Allan C and others. 2015. Factors influencing success of clinical genome sequencing across a broad spectrum of disorders. Nat Genet 47(7):717-726.
Thorvaldsdottir H, Robinson JT, Mesirov JP. 2013. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief Bioinform 14(2):178-92.
Tyner C, Barber GP, Casper J, Clawson H, Diekhans M, Eisenhart C, Fischer CM, Gibson D, Gonzalez JN, Guruvadoo L and others. 2017. The UCSC Genome Browser database: 2017 update. Nucleic Acids Res 45(D1):D626-D634.
Yeo G, Burge CB. 2004. Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. J Comput Biol 11(2-3):377-94.
Zhu B, Zhu X, Wang L, Liang Y, Feng Q, Pan J. 2017. Functional exploration of the IFT-A complex in intraflagellar transport and ciliogenesis. PLoS Genet 13(2):e1006627.

For Peer Review

Figure legends

Figure 1. Identification of 2 mutations in *IFT140* for a family explored by WGS.

(A) Pedigree of family A, which has 2 affected individuals. Segregation analysis of the 2 mutations noted M1: 2577+25G>A and M2: 3454-488_4182+2588dup. Example of Sanger sequencing profiles for individual II.1. The breakpoint junction between exon 30 and intron 26 are shown.

(B) M1 analysis. Sanger sequencing was performed on RNA extracted from fibroblasts of individual II.1 and a healthy unrelated control amplified between exon 18 and exon 22.

(C) M2 analysis. Quantitative real-time PCR was performed on DNA from exon 30 in all individuals from family A and one unrelated control. DNA quantity from exon 30 of *IFT140* was compared to 2 reference genes (*HBB* and *HMBS*) using the absolute quantification method.

(D) M2 RNA expression in patient's II.1 skin fibroblasts is demonstrated by PCR amplification (*PRMT9* as a control) and by Sanger sequencing.

(E) IFT140 expression in skin fibroblasts was revealed by western blot using anti-IFT140 antibody.

Figure 2. Patients' fibroblasts have reduced number of ciliated cells and mislocalized IFT140

(A) Number of ciliated cells were counted on serum deprived (24H) control and patient's skin fibroblasts fixed for immunofluorescence and stained with anti-IFT140 (red) and anti-acetylated tubulin (green) antibodies. Nuclei were colored in blue (Dapi).

For Peer Review

(B) Based on 15 fields in three independent experiments (100-200 cells per experiment), mean percentages of ciliated cells are shown in a histogram together with corresponding standard deviation and p-values (n.s.: non-significant, *: p< 0,0001).

(C) IFT140 localisation was assessed on serum deprived (24H) control and patient's skin fibroblasts fixed for immunofluorescence and stained with anti-IFT140 (red) and anti-acetylated tubulin (green) antibodies. Nuclei were colored in blue (Dapi).

(D) Primary cilia containing IFT140 were counted based on three independent experiments (60-200 ciliated cells per experiment), mean percentages are shown in a histogram together with corresponding standard deviation and p-values (n.s.: non-significant, *: p<0,0005).

Figure 3. Mutations in *IFT140*

(A) Pedigrees of 7 additional families with affected individuals carrying the tandem duplication.

(B) Pedigrees of additional families with *IFT140* pathogenic variants.

Figure 4. Schematic representation of *IFT140* and of the tandem duplication

(A) *IFT140* extends over 105.6 kb in chromosomal region 16p13.3 and contains 31 exons encoding a 1462 amino acids protein containing two repeated domains, 5 WD repeats and 9 tetratricopeptide (TPR) repeats. Previously reported mutations in *IFT140* are represented using a black "asterisk" whereas novel mutations are colored in red. In particular, the two mutations from family A are also shown in introns 20 (c.2577+25G>A) and 30 (c.3454-490_4182+2588dup). The tandem duplication (black horizontal bar) spans from exon 27 to 30 (6.7 kb). The breakpoints are located within *Alu* elements (*AluJb* and *AluJr*).

For Peer Review

(B) Breakpoint junction sequence of affected individuals with tandem duplication. Breakpoint junction sequence is aligned to the proximal and distal genomic references, respectively *AluJr* (chr16:1565570-1565697) and *AluJb* (chr16:1571239-1571365), and color-matched. Microhomology at the breakpoint is indicated in red. The aligned genomic region is 128 bp long showing 81% of sequence identity.

(C) Proposed rearrangement of *Alu-Alu* mediated duplication in affected individuals.

For Peer Review

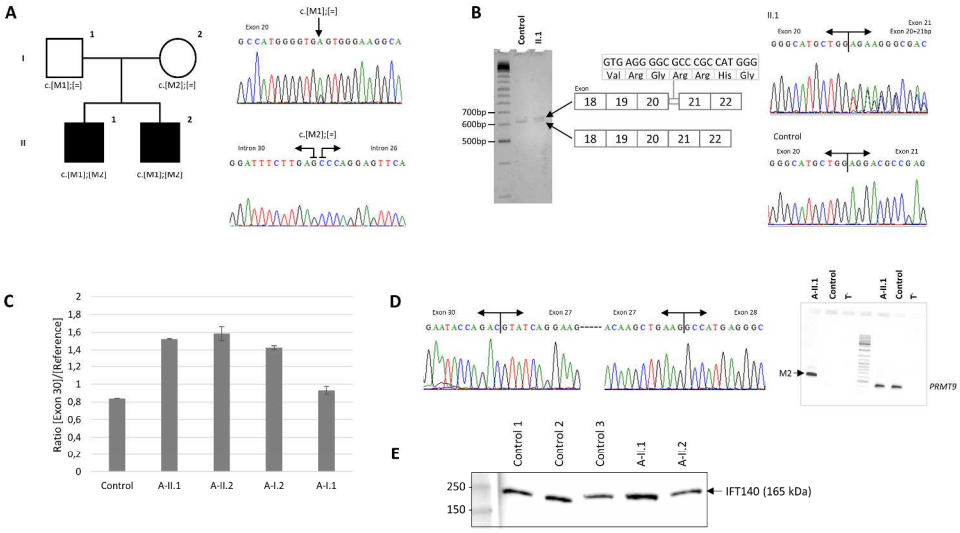


Figure 1. Identification of 2 mutations in IFT140 for a family explored by WGS.
(A) Pedigree of family A, which has 2 affected individuals. Segregation analysis of the 2 mutations noted M1: 2577+25G>A and M2: 3454-488_4182+2588dup. Example of Sanger sequencing profiles for individual II.1. The breakpoint junction between exon 30 and intron 26 are shown.
(B) M1 analysis. Sanger sequencing was performed on RNA extracted from fibroblasts of individual II.1 and a healthy unrelated control amplified between exon 18 and exon 22.
(C) M2 analysis. Quantitative real-time PCR was performed on DNA from exon 30 in all individuals from family A and one unrelated control. DNA quantity from exon 30 of IFT140 was compared to 2 reference genes (HBB and HMBS) using the absolute quantification method.
(D) M2 RNA expression in patient's II.1 skin fibroblasts is demonstrated by PCR amplification (PRMT9 as a control) and by Sanger sequencing.
(E) IFT140 expression in skin fibroblasts was revealed by western blot using anti-IFT140 antibody.

297x166mm (300 x 300 DPI)

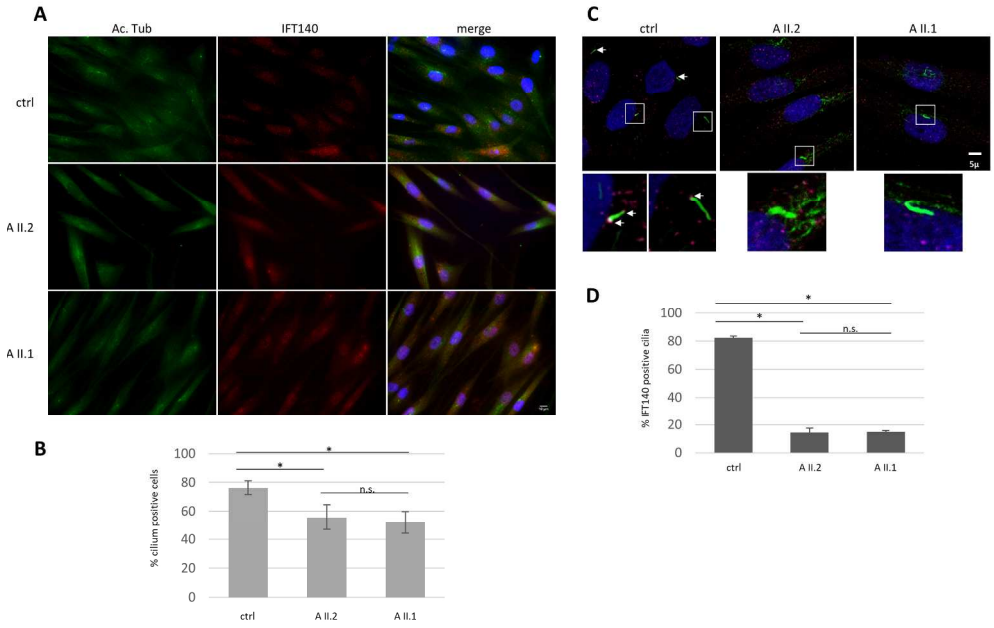


Figure 2. Patients' fibroblasts have reduced number of ciliated cells and mislocalized IFT140
(A) Number of ciliated cells were counted on serum deprived (24H) control and patient's skin fibroblasts fixed for immunofluorescence and stained with anti-IFT140 (red) and anti-acetylated tubulin (green) antibodies. Nuclei were colored in blue (Dapi).
(B) Based on 15 fields in three independent experiments (100-200 cells per experiment), mean percentages of ciliated cells are shown in a histogram together with corresponding standard deviation and p-values (n.s.: non-significant, *: p< 0,0001).
(C) IFT140 localisation was assessed on serum deprived (24H) control and patient's skin fibroblasts fixed for immunofluorescence and stained with anti-IFT140 (red) and anti-acetylated tubulin (green) antibodies. Nuclei were colored in blue (Dapi).
(D) Primary cilia containing IFT140 were counted based on three independent experiments (60-200 ciliated cells per experiment), mean percentages are shown in a histogram together with corresponding standard deviation and p-values (n.s.: non-significant, *: p<0,0005).

220x141mm (300 x 300 DPI)

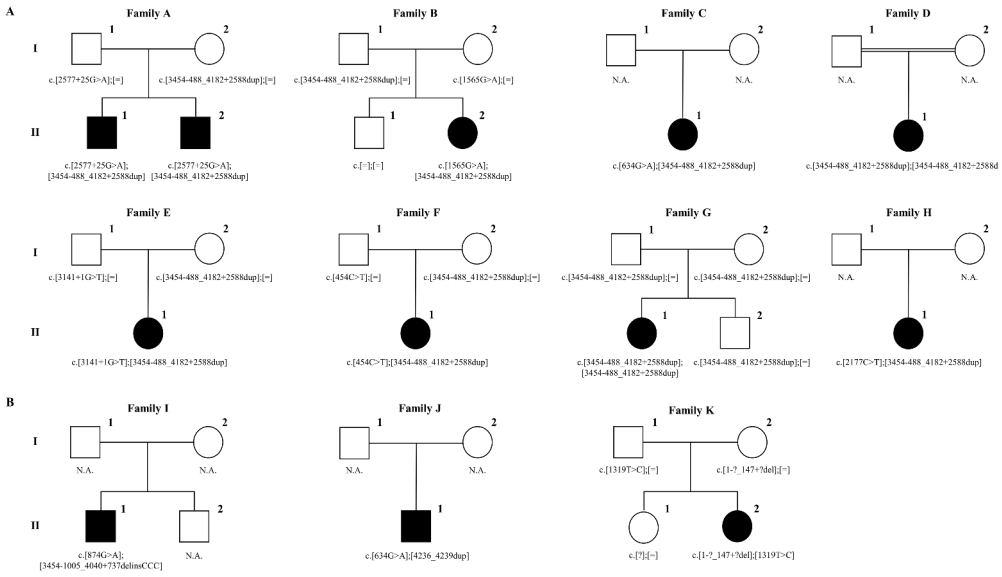


Figure 3. Mutations in IFT140
(A) Pedigrees of 7 additional families with affected individuals carrying the tandem duplication.
(B) Pedigrees of additional families with IFT140 pathogenic variants.

558x325mm (300 x 300 DPI)

Review

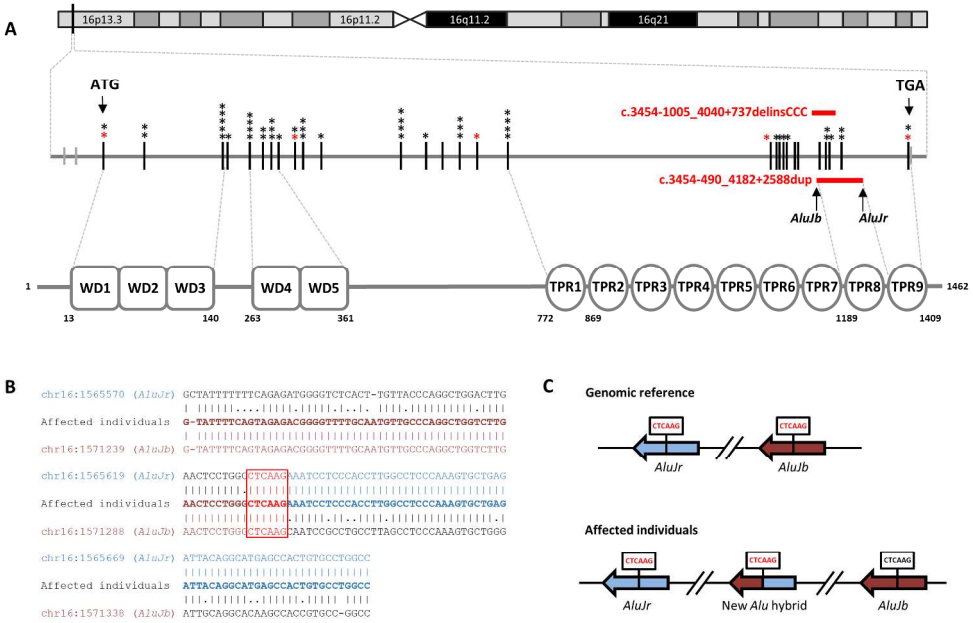


Figure 4. Schematic representation of IFT140 and of the tandem duplication
(A) IFT140 extends over 105.6 kb in chromosomal region 16p13.3 and contains 31 exons encoding a 1462 amino acids protein containing two repeated domains, 5 WD repeats and 9 tetratricopeptide (TPR) repeats. Previously reported mutations in IFT140 are represented using a black "asterisk" whereas novel mutations are colored in red. In particular, the two mutations from family A are also shown in introns 20 (c.2577+25G>A) and 30 (c.3454-490_4182+2588dup). The tandem duplication (black horizontal bar) spans from exon 27 to 30 (6.7 kb). The breakpoints are located within Alu elements (AluJb and AluJr).
(B) Breakpoint junction sequence of affected individuals with tandem duplication. Breakpoint junction sequence is aligned to the proximal and distal genomic references, respectively AluJr (chr16:1565570-1565697) and AluJb (chr16:1571239-1571365), and color-matched. Microhomology at the breakpoint is indicated in red. The aligned genomic region is 128 bp long showing 81% of sequence identity.
(C) Proposed rearrangement of Alu-Alu mediated duplication in affected individuals.

232x153mm (300 x 300 DPI)

SUPPLEMENTARY METHODS

Whole Exome Sequencing

Whole Exome Sequencing (WES) was performed for the two affected siblings from family A (A-II.1 and A-II.2) by the IGBMC (Institute of Genetics and Molecular and Cellular Biology) Microarray and Sequencing platform. Exons of DNA samples were captured with in-solution enrichment methodology (Agilent SureSelect All Exon XT2 50 Mb Kit) and sequenced with an Illumina HiSeq 2500 instrument as paired-end 100 bp reads, in order to reach an average sequencing depth of 80x for each sample. The sequence reads were aligned to the reference sequence of the human genome (GRCh37) using the Burrows-Wheeler Aligner (BWA V7.12) (Li and Durbin, 2010). The HaplotypeCaller module of the Genome Analysis ToolKit (GATK, v.3.4.46) (DePristo, et al., 2011) was used for calling both SNV and indel.

Targeted Exome Sequencing

Leber panel

Targeted exome sequencing of 14 full genes related to Leber congenital amaurosis was performed for 126 patients (including family B, C, D, E and I). Illumina compatible precapture barcoded genomic DNA libraries were constructed according to the manufacturer's sample preparation protocol (Ovation Ultralow, Nugen Technologies). Exons of DNA samples were captured with in-solution enrichment methodology (Agilent SureSelect custom panel) and sequenced with an Illumina HiSeq2500 (Paired-End sequencing 2x125 bases, Rapid Mode, 80 libraries per lane). Bioinformatics analysis was performed as described in Gerber *et al (Gerber, et al., 2016)* using an in house pipeline (POLYWEB).

For Peer Review

RP panel

Targeted exome sequencing of 267 genes related to retinitis pigmentosa was performed for 117 patients (including family J). Exons of DNA samples were captured with in-solution enrichment methodology (Agilent QXT SureSelect custom panel) and sequenced with an Illumina NextSeq 550 instrument (Paired-End sequencing 2x150 bases, 48 libraries per lane). SNVs and indels were called with the Genome Analysis Toolkit v.3.4.46 thanks to our in house pipeline (STARK) and following the GATK best practice.

Ciliome panel

Targeted exome sequencing of 1221 ciliary candidate genes including genes related to ciliopathies like Jeune asphyxiating thoracic dystrophy, Mainzer-Saldino syndrome, Bardet-Biedl syndrome, Joubert syndrome and Senior Loken was performed for 364 patients (including families C, D, I and K). Sequencing was conducted using a custom SureSelect capture kit (Agilent Technologies) targeting 4.5 Mb of 20,168 exons, and performed on SOLiD5500XL (Life Technologies) or HiSeq (Illumina). Bioinformatics analysis was performed as described in Grampa *et al* (Grampa, et al., 2016) using an in house pipeline (POLYWEB).The mean depth of coverage obtained was greater than 165x, with ≥89% of the bases covered at least 15x. In order to evaluate duplication and large deletion events, for each individual the relative read count for each targeted region was determined as the ratio of the read count for that region divided by the total absolute read counts of all targeted regions of the design. The ratio of the relative read count of a region in a given individual over the average relative read counts in other individuals of the run resulted in the estimated copy number for that region in that individual.

For Peer Review

Missense pathogenicity analysis

For the 2 novel missenses identified in this study, conservation has been assessed using the multiple sequence alignment available from the eggnog database (KOG3717, (Huerta-Cepas, et al., 2016)) and they were analyzed thanks to the PolyPhen-2, SIFT and MutationTaster softwares (Adzhubei, et al., 2010; Kumar, et al., 2009; Schwarz, et al., 2010)

.

For Peer Review

Disease	Allele 1	Location	Allelic count	Allele 2	Location	Allelic count	Family
Perrault *et al* 2012 (Perrault, et al., 2012) and this study							
MSS	c.2399+1G>T (p.?)	Intron 19	5/120	c.1990G>A (p.Glu664Lys)	Exon 17	31/120	
MSS	c.932A>G (p.Tyr311Cys)	Exon 9	1/120	c.857_860del (p.Ile286Lysfs*6)	Exon 8	1/120	
MSS	c.1990G>A (p.Glu664Lys)	Exon 17	31/120	c.1990G>A (p.Glu664Lys)	Exon 17	31/120	
MSS	c.1990G>A (p.Glu664Lys)	Exon 17	31/120	c.1990G>A (p.Glu664Lys)	Exon 17	31/120	
MSS	c.634G>A (p.Gly212Arg)	Exon 7	7/120	c.3916dupG (p.Ala1306Glyfs*56)	Exon 29	1/120	
MSS	c.699T>G (p.Ile233Met)	Exon 7	2/120	c.699T>G (p.Ile233Met)	Exon 7	2/120	
JATD	c.2399+1G>T (p.?)	Intron 19	5/120	c.634G>A (p.Gly212Arg)	Exon 7	7/120	
MSS	c.1565G>A (p.Gly522Glu)	Exon 14	4/120	***c.3454-488_4182+2588dup p.(Tyr1152_Thr1394dup)***	Introns 26 30	10/120	Family B
JATD/MSS	c.874G>A (p.Val292Met)	Exon 8	4/120	***c.3454-1005_4040+737delinsCCC, (p.Tyr1152Aspfs*14)***	Exons 27-30	1/120	Family I
MSS	c.1727G>A (p.Arg576Gln)	Exon 15	1/120	2nd allele not found			
MSS	c.489C>T (p.Gly163Gly), and c.488_491del (p.Glu164Thrfs*10) from creation of additional donor splice site	Exon 5	1/120	2nd allele not found			
Schmidts *et al* 2013 (Schmidts, et al., 2013)							
JATD	c.1380delC (p.Asn460Lysfs*28)	Exon 12	1/120	c.874G>A (p.Val292Met)	Exon 8	4/120	
JATD	c.1565G>A (p.Gly522Glu)	Exon 14	4/120	c.874G>A (p.Val292Met)	Exon 8	4/120	
JATD	c.1565G>A (p.Gly522Glu)	Exon 14	4/120	c.874G>A (p.Val292Met)	Exon 8	4/120	
JATD	c.454C>T (p.Leu152Phe)	Exon 5	3/120	c.454C>T (p.Leu152Phe)	Exon 5	3/120	
JATD	c.2278C>T (p.Arg759*)	Exon 19	2/120	2nd allele not found			
MSS	c.2399+1G>T (p.?)	Intron 19	5/120	c.4078T>C (p.Cys1360Arg)	Exon 30	1/120	
MSS	c.418G>A (p.Gly140Arg) and c.800A>G (p.Glu267Gly)	Exons 5/7	1/120	c.490G>T (p.Glu164*)	Exon 6	1/120	
JATD	c.2399+1G>T (p.?)	Intron 19	5/120	2nd allele not found			
JATD	c.4058 G>C (p.Pro1353Arg)	Exon 30	1/120	2nd allele not found			
JATD	c.2360 T>C (p.Asp787Gly)	Exon 19	1/120	2nd allele not found			
JATD	c.1541 A>T (p.Leu514His)	Exon 14	1/120	2nd allele not found			
JATD	c.329G>A (p.Arg110His)	Exon 4	1/120	2nd allele not found			
JATD	c.481C>A (p.Pro161Thr)	Exon 5	2/120	2nd allele not found			

For Peer Review

JATD	c.481C>A (p.Pro161Thr)	Exon 5	2/120	2nd allele not found			
Khan *et al* 2014 (Khan, et al., 2014)							
MSS	c.1990G>A (p.Glu664Lys)	Exon 17	31/120	c.1990G>A (p.Glu664Lys)	Exon 17	31/120	
MSS	c.1990G>A (p.Glu664Lys)	Exon 17	31/120	c.1990G>A (p.Glu664Lys)	Exon 17	31/120	
Bifari *et al* 2015 (Bifari, et al., 2016)							
EORD	c.1990G>A (p.Glu664Lys)	Exon 17	31/120	c.1990G>A (p.Glu664Lys)	Exon 17	31/120	
EORD	c.1990G>A (p.Glu664Lys)	Exon 17	31/120	c.1990G>A (p.Glu664Lys)	Exon 17	31/120	
EORD	c.1990G>A (p.Glu664Lys)	Exon 17	31/120	c.1990G>A (p.Glu664Lys)	Exon 17	31/120	
EORD	c.1990G>A (p.Glu664Lys)	Exon 17	31/120	c.1990G>A (p.Glu664Lys)	Exon 17	31/120	
EORD	c.1990G>A (p.Glu664Lys)	Exon 17	31/120	c.1990G>A (p.Glu664Lys)	Exon 17	31/120	
EORD	c.1990G>A (p.Glu664Lys)	Exon 17	31/120	c.1990G>A (p.Glu664Lys)	Exon 17	31/120	
EORD	c.1990G>A (p.Glu664Lys)	Exon 17	31/120	c.1990G>A (p.Glu664Lys)	Exon 17	31/120	
EORD	c.1990G>A (p.Glu664Lys)	Exon 17	31/120	c.1990G>A (p.Glu664Lys)	Exon 17	31/120	
EORD	c.1990G>A (p.Glu664Lys)	Exon 17	31/120	c.1990G>A (p.Glu664Lys)	Exon 17	31/120	
EORD	c.1990G>A (p.Glu664Lys)	Exon 17	31/120	c.1990G>A (p.Glu664Lys)	Exon 17	31/120	
EORD	c.1990G>A (p.Glu664Lys)	Exon 17	31/120	c.1990G>A (p.Glu664Lys)	Exon 17	31/120	
EORD	c.1541_1542delinsAA (p.Leu514Gln)	Exon 14	2/120	c.1541_1542delinsAA (p.Leu514Gln)	Exon 14	2/120	
Xu *et al* 2015 (Xu, et al., 2015)							
RP	c.4196T>C (p.Leu1399Pro)	Exon 31	31/120	c.1898_1901delATAA (p.Asn633Serfs*10)	Exon 17	1/120	
RP	c.3826G>A (p.Gly1276Arg)	Exon 28	1/120	c.1989C>G (p.Cys663Trp)	Exon 17	1/120	
RP	c.650_651delTG (p.Val217Glyfs*2)	Exon 7	1/120	c.212C>T (p.Pro71Leu)	Exon 4	1/120	
RP	c.2611C>T (p.Arg871Cys)	Exon 21	1/120	c.1377G>A (p.Trp459*)	Exon 12	1/120	
RP	c.1252G>C (p.Ala418Pro)	Exon 11	1/120	c.2921C>T (p.Ala974Val) presumed in trans	Exon 23	1/120	
LCA	c.1451C>T (p.Thr484Met)	Exon 13	3/120	c.985T>C (p.Cys329Arg)	Exon 9	1/120	
LCA	c.2368G>A (p.Glu790Lys)	Exon 19	1/120	c.1655_1656delAG (p.Glu522Glyfs*6)	Exon 14	1/120	
Beheshtian *et al* 2015 (Beheshtian, et al., 2015)							
RP	c.3827G>A (p.Gly1276Glu)	Exon 28	2/120	c.3827G>A (p.Gly1276Glu)	Exon 28	2/120	
Weisschuh *et al* 2016 (Weisschuh, et al., 2016)							
RP	c.472C>T (p.Arg158Trp)	Exon 5	1/120	c.1565G>A (p.Gly522Glu)	Exon 14	4/120	
Hull *et al* 2016 (Hull, et al., 2016)							

For Peer Review

RP	c.1451C>T (p.Thr484Met)	Exon 13	3/120	c.1451C>T (p.Thr484Met)	Exon 13	3/120	
RP	c.2399+1G>T (p.?)	Intron 19	5/120	c.2815T>C (p.Ser939Pro)	Exon 22	1/120	
RP	c.998G>A (p.Cys333Tyr)	Exon 9	4/120	c.998G>A (p.Cys333Tyr)	Exon 9	4/120	
RP	c.998G>A (p.Cys333Tyr)	Exon 9	4/120	c.998G>A (p.Cys333Tyr)	Exon 9	4/120	
RP	c.1021G>A (p.Ala341Thr)	Exon 10	1/120	c.1422_23insAA (p.Arg475Asnfs*14)	Exon 12	1/120	
Pena-Padilla *et al* 2016 (Pena-Padilla, et al., 2016)							
OTCS	c.3141+1G>T (p.?)	Exon 24	2/120	c.-11_6del (p.?)	Exon 3	1/120	
Bayat *et al* 2017 (Bayat, et al., 2017)							
SB	c.634G>A (p.Gly212Arg)	Exon 7	7/120	c.2278C>T (p.Arg759*)	Exon 19	2/120	
Helm *et al* 2017 (Helm, et al., 2017)							
MSS	c.634G>A (p.Gly212Arg)	Exon 7	7/120	c.634G>A (p.Gly212Arg)	Exon 7	7/120	
This study							
MSS	***c.2577+25G>A (p.Leu859_Glu860insValArgGlyAlaArgHisGly)***	Intron 20	1/120	***c.3454-488_4182+2588dup (p.Tyr1152_Thr1394dup)***	Introns 26-30	10/120	Family A
MSS	c.634G>A (p.Gly212Arg)	Exon 7	7/120	***c.3454-488_4182+2588dup (p.Tyr1152_Thr1394dup)***	Introns 26-30	10/120	Family C
MSS	***c.3454-488_4182+2588dup (p.Tyr1152_Thr1394dup)***	Introns 26-30	10/120	***c.3454-488_4182+2588dup (p.Tyr1152_Thr1394dup)***	Introns 26-30	10/120	Family D
JATD/MSS	c.3141+1G>T (p.?)	Exon 24	2/120	***c.3454-488_4182+2588dup (p.Tyr1152_Thr1394dup)***	Introns 26-30	10/120	Family E
MSS	c.454C>T (p.Leu152Phe)	Exon 5	3/120	***c.3454-488_4182+2588dup (p.Tyr1152_Thr1394dup)***	Introns 26-30	10/120	Family F
MSS	***c.3454-488_4182+2588dup (p.Tyr1152_Thr1394dup)***	Introns 26-30	10/120	***c.3454-488_4182+2588dup (p.Tyr1152_Thr1394dup)***	Introns 26-30	10/120	Family G
MSS/SB	***c.2177C>T (p.Pro726Leu)***	Exon 18	1/120	***c.3454-488_4182+2588dup (p.Tyr1152_Thr1394dup)***	Introns 26-30	10/120	Family H
RP	c.634G>A (p.Gly212Arg)	Exon 7	7/120	***c.4236_4239dup (p.Tyr1414Leufs*48)***	Exon 31	1/120	Family J
MSS (mild)	***c.1-?_147+?del, start loss***	Exon 3	1/120	***c.1319T>C (p.Leu440Pro)***	Exon 11	1/120	Family K

Table S1. Previously reported and novel mutations in *IFT140*.

Mutations are reported only once per family. Those reported in this study are in bold and those described for the first time are also in italic. Mutation allelic counts are reported for only one affected individual per family. It is to notice that some studies (i.e. Bifari *et al*) have analyzed consanguineous families from the same region increasing the global allelic count. EORD: early-onset retinal dystrophy; JATD: Jeune asphyxiating thoracic dystrophy; MSS: Mainzer-Saldino Syndrome; OCTS: Opitz trigonocephaly C syndrome; RP: retinitis pigmentosa; SB: Sensenbrenner syndrome.

Individual	A-II.1	A-II.2	B-II.1	C-II.2	D-II.1	E	F	G	H	I-II.1	J-II.1	K-II.2
Birth	10/01/1998 France, Italy	31/08/1999 France, Italy	17/03/1978 Serbia, Yugoslavia	26/08/1976 France	27/12/1968 France	20/09/2010 Caucasian	11/02/2007 Austria	21/05/1993 Austria		16/04/1986	05/06/1993	14/02/2008
Sex	M	M	F	F	F	F	F	F	F	M	M	F
Suspected diagnostic	MSS	MSS	MSS	MSS	MSS	Jeune/MSS	MSS	MSS	MSS/Sensenbrenner	Jeune/MSS	Isolated RP	MSS (mild)
Onset	14 yo	15 yo	3 yo	6 yo	45 yo	Neonatal	4 yo	1st yo	Neonatal			7 yo
Short stature (SD)	+ (-3 SD)	+ (-2 SD)	+ (< - 5 SD)	+ (- 3 SD)	+ (< - 5 SD)	+ disproportionate (-3 SD)	(-1 SD)	(+ 1,5 SD)			(- 1 SD)	
Skeletal features	Broad long bones	Hip anomalies (narrow pelvis with narrow iliac wings, lack of coverage of superior femoral epiphyses)	Hip anomalies Metaphyseal defect (forearms) Micromelia Scoliosis	Metaphyseal defect (long bones)	Short and broad bones Bilateral ulnar valgus Marked curved radius Hip anomalies (large pelvis, short femoral heads) Varus of the feet	Hypoplastic thorax Short, slightly broader ribs Hip anomalies (hip dysplasia, broad epiphyses, steep acetabular roofs) Genu varum Hyperlordosis Brachymelia No premature craniosynostosis	Moderate hips' anomalies	No	Craniostenosis Hip anomalies not known	Thoracic dystrophy		
Digit anomalies	Brachymetacarpia (1, 3, 4 digits) Brachymetatarsia	Short hands Brachydactyly and brachymetacarpia (1, 4 and 5 digits) brachymetatarsia	Short hands Cone-shaped phalangeal epiphyses (hands)	Cone-shaped phalangeal epiphyses (hands)	Brachytelephalangy Brachymesophalangy with clinodactyly of the 5th digits (feet and hands) Enlarged metaphyses (feet) No cone-shaped epiphyses due to the age of the patient (52 yo)	Short hands Brachyphalangy Cone-shaped phalangeal epiphyses (hands) Pes planovalgus Incomplete syndactyly of digits 3-4 (toes) Brachydactyly of 3rd toe Delayed bone age	Short hands Brachymesophalangy Cone-shaped phalangeal epiphyses (feet, hands)	Short hands Brachymesophalangy Cone-shaped phalangeal epiphyses (feet, hands)	Short hands Brachymesophalangy Cone-shaped phalangeal epiphyses No polydactyly	Cone-shaped phalangeal epiphyses		Cone-shaped phalangeal epiphyses (second phalanges of the 2, 3, 4 digits)
Retinal dystrophy	Hemeralopia (10 yo) VA: 9/10RLE	Hemeralopia (7 yo) Divergent strabismus VA: 7/10RLE RP Altered ERG	Onset (3 mo) Nystagmus Night blindness VA: 2/10RLE Maculopathy Altered ERG	Night blindness VA: 4/10RLE	Divergent strabismus (2 yo) Hemeralopia (10 yo) RP (16 yo) VA: 2.4/10 RE 1/20 LE Advanced RP VF: very reduced on right side	Congenital pendular nystagmus Strabismus Hyperopia Astigmatism Severe visual reduction	Myopia RP (4 yo)	Myopia RP (1 yo)	Nystagmus Strabismus Refraction +2 Retinal dystrophy not known Papillae edges not well differentiated	ERG: no signal left	RP slow evolution	No eye pursuit (1 mo) Bilateral divergent strabismus (5 mo) Nystagmus (6 mo) Early onset advanced RP (3 yo) ERG: severe and primary damage of cones and rods
Renal involvement	Moderate renal failure Microalbuminuria Renal US normal	Moderate renal failure Microalbuminuria Renal US normal Arterial hypertension	End stage renal failure (5 yo) Dialysis Transplant Nephronophtisis with polyria polydipsia (3 yo) Arterial hypertension Renal US (Several cortical cysts (10 yo), echogenicity, kidney's size: R76mm L60mm)	End stage renal failure (7 yo) Dialysis Transplant (7 yo) Nephronophtisis with polyria polydipsia (6 yo) Arterial hypertension (24/12) Renal US (no cysts, echogenicity, kidney's size: R55mm L59mm) Obsolescent/hyalinized glomerulus tubulo-interstitial nephritis important arterial damage	End stage renal failure Dialysis (10 yo) Transplantation (16 yo)	End stage renal failure (1.5 yo) Peritoneal dialysis (2 yo) Transplant (2.5 yo) Hydronephrosis (4 mo) Renal US (small hyperechoic kidney, right kidney with 3 chambered cysts) Arterial hypertension Renal anemia	End stage renal failure (4.7 yo) Renal US (Small kidneys (<3th centile, increased echogenicity, reduced cortico-medullar differentiation, no cysts) Proteinuria, microhematuria	End stage renal failure (23 yo) Renal US (cysts, low-normal size, increased echogenicity, loss of cortico-medullar differentiation (11 yo) Glomerular and tubular proteinuria	End stage renal failure Dialysis (3.8 yo) Transplant CKD stage V Renal US (small cysts, small kidney) Glomerular and tubular proteinuria No hematuria			Renal function normal Renal ultrasound normal
Hepatic involvement	No	No	Portal space fibrosis Hemosiderosis		No	Slightly hyperechoic liver Slightly enlarged	Yes	No	Normal liver function tests Hepatomegaly			
Pancreatic fibrosis	No	No			No	No				Yes (cystic)		
Neurological features	No	No	No			Asphyxia at birth Slight Developmental delay (head control 4 mo, sitting 8 mo, walking 24 mo, first words 18 mo, slow speech development) No ataxia, no seizures No autistic features	No	Hypotonia at birth Developmental delay	Developmental delay No cerebellar ataxia, but oculomotor apraxia Intracranial hypertension No seizures			
Cerebral anomalies	No	No				Generalized brain atrophy (CT scan 4 mo)	No		No		No	No
Dysmorphic features	Yes (high hairline, frontal bossing, ptosis, long philtrum, prominent chin)	Yes (high forehead, frontal bossing, long philtrum, prominent chin)			No	Yes (high hairline, frontal bossing, hypertelorism, broad nasal ridge, deep-set eyes, anteverted nares, long philtrum, everted lower lip, prominent chin, deep-set ears)	No	No	Yes			
Oral anomalies	No	No	No		No	No	No	No	No			
Other	Asthma					Truncal adiposity, breast tissue Splenomegaly Obstructive bronchitis Recurrent infections	Recurrent respiratory Infections	Late onset obesity	Premature birth 28 weeks of gestational age		Overweight when younger (9 yo)	Joint luxations with spontaneous subluxations

Table S2 Clinical Features of MSS-affected individuals harboring compound-heterozygous mutations in *IFT140*. CKD: Cystic Kidney Disease, CT: Computed Tomography, ERG: Electroretinogram, F: Female, L: Left, LE: Left Eye, M: Male, mo: month old, MSS: Mainzer-Saldino Syndrome, R: Right, RE: Right Eye, RLE: Right and left Eye, RP: Retinitis Pigmentosa, SD: standard deviation, US: Ultrasound, VA: Visual acuity, VF: visual field, yo: years old.

For Peer Review

Application	Gene	RefSeq	Name	Forward (5'-3')	Reverse (5'-3')
DNA c.2577+25G>A	*IFT140*	NM_014714.3	IFT140-ex20	GATGACTACTTGCCCCAGGA	CACACCTGACAGGCACACAC
DNA breakpoint TD			IFT140-intron30F_ intron26R	CCGGAGTAGCTGGGATTACA	CTGTCTAGCTGGGGTGGGTA
DNA c.1565G>A			IFT140-ex14	TGAGCGAGTGAATGAGTGGA	CGGTGTTAAAATGGGCACCT
cDNA c.2577+25G>A			IFT140-RT-PCR-ex18-19F_ex22R	ATCAAAAGTGAGGCCGTCTG	GCGTGTCCGACTTCTCGTA
cDNA TD			IFT140-RT-PCR-ex30F_ex28R	AAGCAGTGTGAGCTGCTCCT	GGGCACAAGCGTCATAAAAG
cDNA control	*PRMT9*	NM_001304458	PRMT9-RT-PCR-ex3F_ex6R	GTGGAACGCTGGCACTTTAT	CAATTAGCACTTTCACCTTTGG
qPCR	*IFT140*	NM_014714.3	IFT140-QPCR-ex30F_ex30R	CATAGAAAGTTTCCTCCTGCCAT	TCTGGTTCCTCCAGGAGCA
qPCR	*HMBS*	NM_000190	HMBS-QPCR-	CACAGCACTCCCACTGACAAC	GAGGGAGGCGATAGTAGGACC
qPCR	*HBB*	NM_000518	BGLOB-QPCR-ex1F_ex1R	ACACAACTGTGTTCACTAGC	CAACTTCATCCACGTTCACC
Duplex PCR	*IFT140*	NM_014714.3	IFT140-intron30F_ intron26R	CCCAAAGCGTCATTTCTGGTGTTC	CCCAGGGTCAGGTGATCTTCCTA
	BBS12	NM_152618	BBS12-ex2F_ex2R	AGTTCTCATTGAGGGTGACCT	GCCAGAGATGAAGCCAGCCA

Table S3 List and characteristics of primers used in this study. TD: Tandem duplication of exon 27-30

For Peer Review

Name primary antibodies	Applications	Producer	Conditions
Mouse anti-Acetylated α-Tubulin	WB, IF	Abcam #ab24610	IF: 1/200
Rabbit anti-IFT140	WB, IF	Proteintech #17460-1-AP	IF:1/100 WB: 1/1000
Mouse anti-β-Tubulin	WB	Euromedex # TUB-2A2	1/2500
Name secondary antibodies			
Donkey anti-Mouse IgG (H+L), FITC conjugate	IF	Thermo Scientific #A16012	1/500
Goat anti-Rabbit IgG (H+L) Alexa Fluor 568 conjugate	IF	Thermo Scientific #A-11011	1/500
Chicken anti-Rabbbit IgG-HRP	WB	Santa Cruz #sc-2955	1/5000
Goat anti mouse IgG-HRP	WB	Santa Cruz #sc-2060	1/5000

Table S4. List of antibodies used in this study. IF: immunofluorescence, WB Western Blot.

For Peer Review

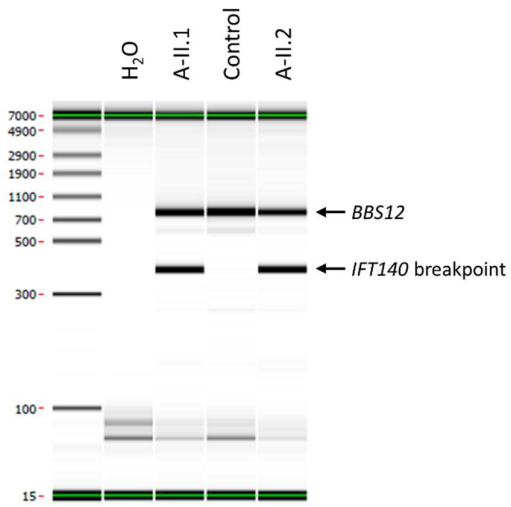


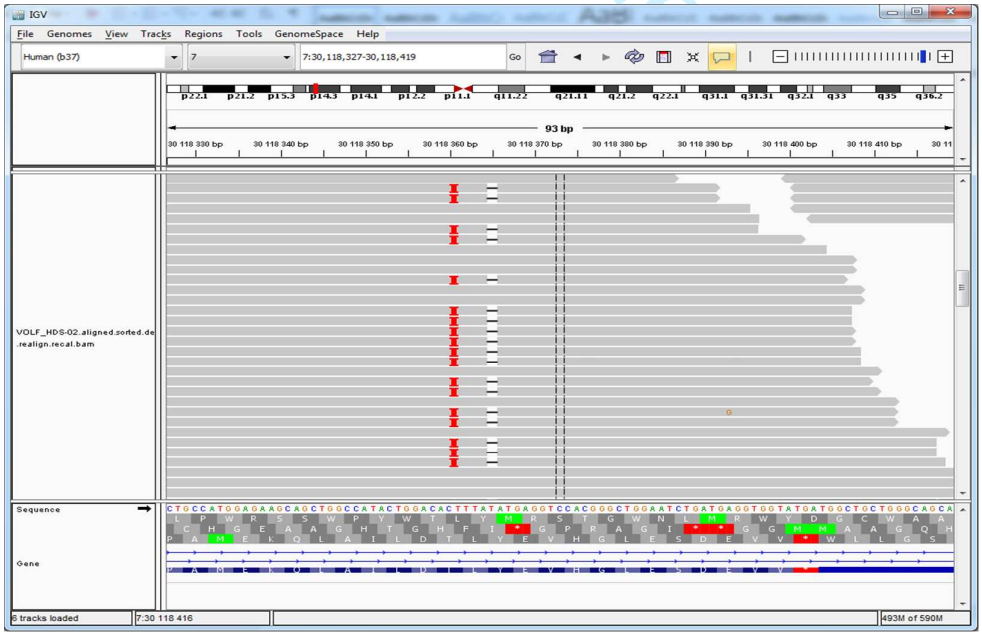
Figure S1. Duplex PCR profiles on a LabChip GX (Perkin Elmer) for a rapid detection of the *IFT140* tandem duplication (see **Table S3** for primers details). Normal samples harbor a single band (*BBS12*) whereas homozygous or heterozygous carriers for the tandem duplication show an additional band specific for the duplication breakpoint.

For Peer Review

Patient	A-II.1		A-II.2	
Type of sequence variant	**SNV**	**indel**	**SNV**	**indel**
Total number of variants	49489	2168	49418	2133
After exclusion of variants with an allele frequency >1% (extracted from dbSNP, the Exome Variant Server and the ExAC database or our internal exome database)	6528	1038	6517	1017
After exclusion of variants found at the homozygous state and more than once at the heterozygous state in 90 control exomes	1159	90	1029	72
After exclusion of 5'UTR, 3'UTR, downstream, upstream and intron locations without local splice effect prediction (from the "localSpliceEffect" field of Alamut-Batch)	517	20	458	12
After exclusion of synonymous variants without local splice effect prediction (from the "localSpliceEffect" field of Alamut-Batch)	417	20	373	12
After selection of variants consistent with recessive transmission (compound heterozygous, homozygous variants).	2 compound heterozygous in *PLEKHA8*			

Table S5 Summary of the whole exome sequencing results from family A

WES data processing and variant calling revealed respectively 51,657 and 51,551 SNV and indels in individual A-II.1 and A-II.2. Variant filtering using stringent criteria reduced the number of genetics variants to respectively 437 and 385 variants per proband. To identify variants consistent with autosomal recessive inheritance, we kept only compound shared heterozygous or homozygous variants, reducing the number of variants to 2 heterozygous variants in the *PLEKHA8* gene: NM_001197026.1:c.1517_1518insAT (p.Leu507Phefs*27) and NM_001197026.1:c.1522del (p.Tyr508Metfs*25). Both variants being always supported on the same reads (as shown in the IGV view below), they are in *cis* and are probably an artifact due to a bad reads alignment. So, no candidate gene was found by the WES analyses.



For Peer Review

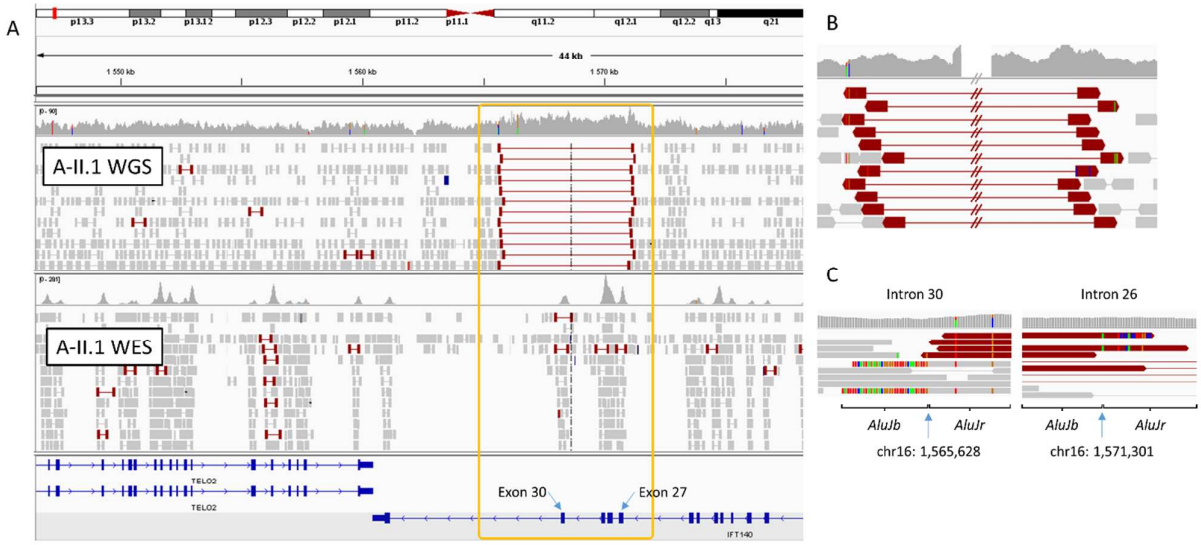


Figure S2. Next generation sequencing data from patient A-II.1 displayed around *IFT140,* in particular from exon 27 to exon 30 highlighting the tandem duplication.

(A) Comparison between whole genome sequencing (WGS) and whole exome sequencing (WES) data from patient A-II.1. The reads are displayed as pairs and sorted according to their insert size using IGV (Thorvaldsdottir, et al., 2013).

(B) Highlight of the WGS read pairs at the breakpoint junctions showing the typical drawing (read pairs pointing in opposite direction) for tandem duplications in direct orientation.

(C) Highlight of the WGS read pairs at the breakpoints including soft clipped bases. The left breakpoint in intron 30 overlaps the *AluJr* while the right breakpoint in intron 26 overlaps the *AluJb*. Analysis of the reads and the split reads that are misaligned allowed us to define the exact position of the duplication.

For Peer Review

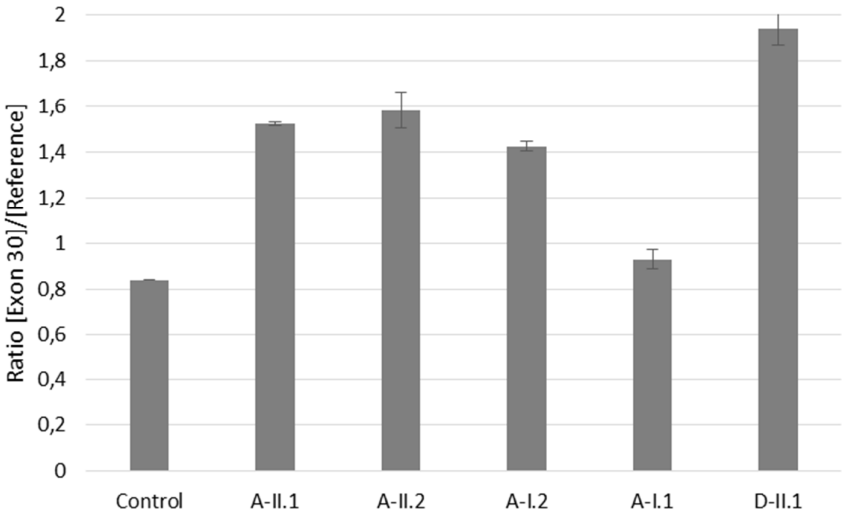


Figure S3. Homozygous duplication in *IFT140* validated by quantitative real-time PCR. Quantitative real-time PCR was performed on DNA from exon 30 in all individuals from family A and D as well as in one unrelated control. DNA quantity from exon 30 of *IFT140* was compared to 2 reference genes (*HBB* and *HMBS*) using the absolute quantification method.

For Peer Review

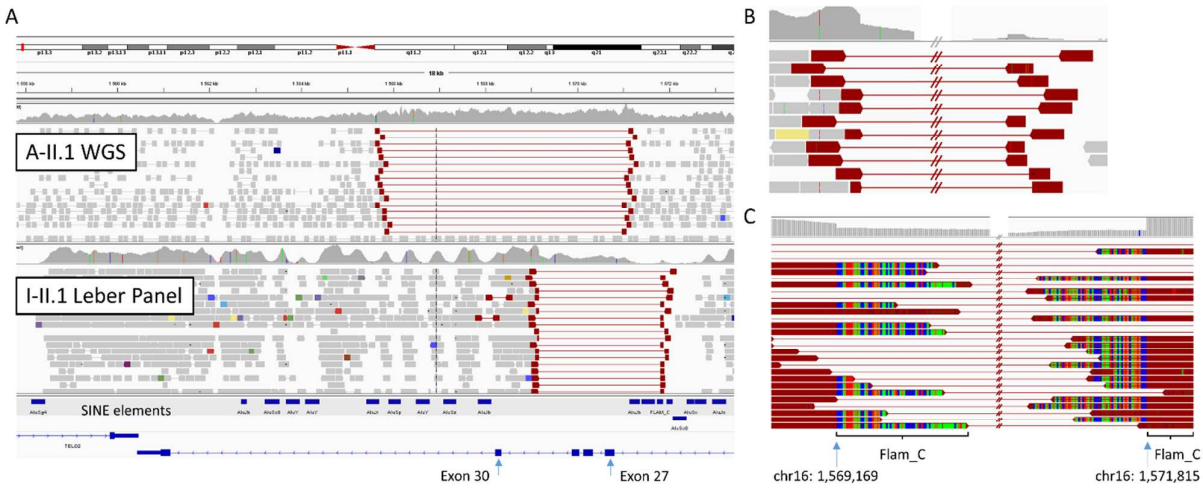


Figure S4. Next generation sequencing data for family A and I displayed around *IFT140,* in particular from exon 27 to exon 30.

(A) Comparison between whole genome sequencing (WGS) from patient A-II.1 (tandem duplication carrier) and targeted exome sequencing using the full *IFT140* gene as a target region from patient I-II.1 (deletion of exon 27 to 29). The reads are displayed as pairs and sorted according to their insert size using IGV (Thorvaldsdottir, et al., 2013). The two events do not share the same breakpoints.

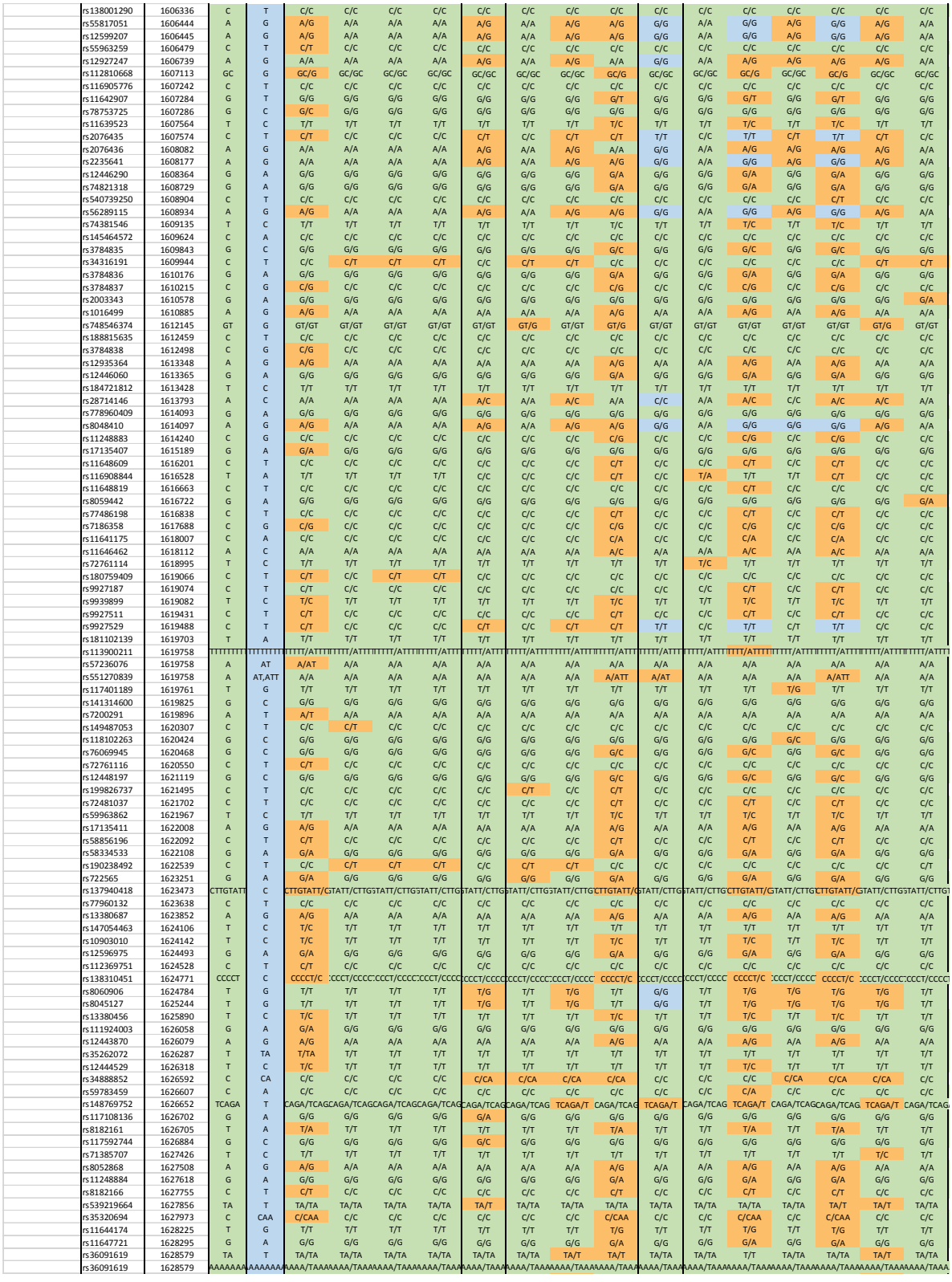
(B) Highlight of the read pairs from patient I-II.1 at the breakpoint junctions showing the typical drawing (read pairs pointing at each other) for a genomic deletion.

(C) Highlight of the read pairs from patient I-II.1 at the breakpoints including soft clipped bases. The right breakpoint overlaps a SINE element from the *Alu* family (Flam_C) but not the left breakpoint. Analysis of the split reads that are misaligned allows us to define the exact position of the deletion (c.3454-1005_4040+737delinsCCC).

For Peer Review

					WGS				TES					WGS			TES		
	rsID	Chr 16 position	ref	alt	A-I.1	A-1.2	A-II.1*	A-II.2*	C-II.1*	B-II.2*	B-I.1	B-I.2	D-II.1*	Control 1	Control 2	Control 3	Control 4	Control 5	Control 6
						Dup	Dup	Dup	Dup	Dup	Dup								
	rs149678731	1560882	T	C	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/C
	rs1053730	1560886	A	G	A/A	A/G	A/A	A/A	A/A	A/A	A/G	A/A	G/G	A/A	A/G	A/A	A/A	A/G	A/A
	rs2667688	1561446	A	G	A/A	A/G	A/A	A/A	A/A	A/A	A/G	A/A	G/G	A/A	A/G	A/A	A/A	A/G	A/A
	rs572130902	1561488	G	C	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
	rs8048290	1562026	T	C	T/T	T/C	T/T	T/T	T/T	T/T	T/C	T/T	C/C	T/T	T/C	T/T	T/T	T/C	T/T
	rs12918013	1562195	A	G	A/A	A/G	A/A	A/A	A/A	A/A	A/G	A/A	G/G	A/A	A/A	A/A	A/A	A/G	A/A
	rs10083822	1562252	C	T	C/C	C/C	C/C	C/C	C/C	C/C	C/T	C/C	T/T	C/C	C/T	C/C	C/C	C/T	C/C
	rs117143480	1562399	C	T	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/T	C/C	C/C	C/C	C/C
	rs8048980	1562460	T	C	T/T	T/C	T/T	T/T	T/T	T/T	T/C	T/T	C/C	T/T	T/C	T/T	T/T	T/C	T/T
	rs1966171	1562804	G	A	G/G	G/A	G/G	G/G	G/G	G/G	G/A	G/A	A/A	G/G	A/A	G/G	G/A	G/A	G/G
	rs58369664	1562915	C	G	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/G	C/C	C/C	C/G	C/C	C/G	C/C	C/C
	rs112450135	1562987	G	C	G/G	G/C	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
	rs7189700	1563582	G	A	G/G	G/A	G/G	G/G	G/G	G/G	G/A	G/G	A/A	G/G	G/G	G/G	G/G	G/A	G/G
	rs10558385	1563667	ACTTT	A	ACTTT/ACTT	ACTTT/ACTT	ACTTT/ACTT	ACTTT/ACTT	ACTTT/ACTT	ACTTT/ACTT	ACTTT/A	ACTTT/ACTT	A/A	ACTTT/ACTT	ACTTT/A	ACTTT/ACTT	ACTTT/ACTT	ACTTT/A	ACTTT/ACTT
	rs113508951	1563672	C	CT	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/CT	C/C	C/C	C/C	C/C	C/CT	C/C	C/C
	rs7191239	1563672	C	T	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
	rs67974359	1563705	T	C	T/T	T/C	T/T	T/T	T/T	T/T	T/C	T/C	C/C	T/T	C/C	T/T	T/C	T/C	T/T
	rs184085046	1563799	C	T	C/C	C/T	C/C	C/C	C/C	C/C	C/T	C/C	T/T	C/C	C/T	C/C	C/C	C/T	C/C
	rs73497522	1563800	A	G	A/A	A/G	A/A	A/A	A/A	A/A	A/G	A/G	G/G	A/A	G/G	A/A	A/G	A/G	A/A
	rs35773761	1563865	G	A	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/A	G/G	G/G	G/G	G/G	G/A	G/A	G/G
	rs551416387	1563876	G	A	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/A	G/G	G/G	G/G	G/G	G/A	G/G	G/G
	rs74430146	1563886	A	G	A/A	A/A	A/A	A/A	A/A	A/A	A/G	A/G	A/G	A/G	G/G	A/A	A/G	A/G	A/A
	rs62012850	1563896	C	A,G	C/C	C/G	C/C	C/C	C/C	C/C	C/G	C/A	G/G	C/C	C/C	C/C	C/A	G/G	C/C
	rs62012851	1563916	T	C	T/T	T/C	T/T	T/T	T/T	T/T	T/C	T/T	C/C	T/T	T/T	T/T	T/T	T/C	T/T
	rs62012852	1563920	C	T	C/C	C/T	C/C	C/C	C/C	C/C	C/T	C/C	T/T	C/C	C/C	C/C	C/C	C/T	C/C
	rs2745111	1564823	G	A	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/A	G/G	G/G	G/A	G/G	G/A	G/G	G/G
	rs12597753	1565116	C	A	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/A	C/C	C/C	C/A	C/C	C/A	C/C	C/C
	rs11643891	1565540	C	T	C/C	C/T	C/C	C/C	C/C	C/C	C/T	C/T	T/T	C/C	T/T	C/C	C/T	C/T	C/C
	rs79602039	1565893	A	T	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/T	A/A	A/A	A/T	A/A	A/T	A/A	A/A
	rs6600149	1566177	A	C	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
	rs6600150	1566178	A	C	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
	rs752259914	1566178	A	ACC	A/A	A/A	A/A	A/A	A/A	A/A	A/ACC	A/ACC	ACC/ACC	A/A	A/A	A/A	A/ACC	A/ACC	A/A
	rs8051374	1566284	T	C	T/T	T/C	T/T	T/T	T/T	T/T	T/C	T/C	C/C	T/T	C/C	T/T	T/C	T/C	T/T
	rs112878182	1566444	A	G	A/A	A/G	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
	rs9932727	1566586	C	T	C/C	C/T	C/C	C/C	C/C	C/C	C/T	C/C	T/T	C/C	C/T	C/C	C/C	C/T	C/C
	rs2745185	1566597	T	C	T/T	T/C	T/T	T/T	T/T	T/T	T/C	T/C	C/C	T/T	C/C	T/T	T/C	T/C	T/T
	rs71385704	1566826	G	A	G/G	G/A	G/G	G/G	G/G	G/G	G/A	G/G	A/A	G/G	G/A	G/G	G/G	G/A	G/G
	rs150027092	1567367	AAG	A	AAG/AAG	AAG/AAG	AAG/AAG	AAG/AAG	AAG/AAG	AAG/AAG	AAG/AAG	AAG/AAG	AAG/AAG	AAG/AAG	AAG/AAG	AAG/AAG	AAG/AAG	AAG/AAG	AAG/AAG
	rs147652844	1567368	AG	A	AG/AG	AG/AG	AG/AG	AG/AG	AG/AG	AG/AG	AG/AG	AG/AG	A/A	AG/AG	AG/AG	AG/AG	AG/AG	AG/A	AG/AG
	rs4787262	1567368	AG	A	AG/AG	AG/A	AG/AG	AG/AG	AG/AG	AG/AG	AG/AG	AG/AG	AG/AG	AG/AG	AG/AG	AG/AG	AG/AG	AG/AG	AG/AG
	rs76268857	1567453	GAGA	G	GAGA/GAGA	GAGA/GAGA	GAGA/GAGA	GAGA/GAGA	GAGA/GAGA	GAGA/GAGA	GAGA/GAGA	GAGA/G	GAGA/GAGA	GAGA/GAGA	GAGA/G	GAGA/GAGA	GAGA/G	GAGA/GAGA	GAGA/GAGA
IFT140 Duplication	rs2072990	1567651	A	G	A/A	A/G	A/A	A/A	A/A	A/A	A/G	A/G	G/G	A/A	G/G	A/A	A/G	A/G	A/A
	rs2235637	1567694	G	A	G/G	G/A	G/G	G/G	G/G	G/G	G/A	G/A	A/A	G/G	A/A	G/G	G/A	G/A	G/G
	rs2076434	1568402	A	G	A/A	A/G	A/A	A/A	A/A	A/A	A/G	A/A	G/G	A/A	A/G	A/A	A/A	A/G	A/A
	rs71385705	1569003	A	T	A/A	A/T	A/A	A/A	A/A	A/A	A/A	A/A	A/T	A/A	A/A	A/A	A/A	A/A	A/A
	rs117190102	1569186	G	A	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
	rs3784819	1569202	C	T	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/T	C/C	C/C	C/T	C/C	C/T	C/C	C/C
	rs3784820	1569252	G	A	G/G	G/A	G/G	G/G	G/G	G/G	G/A	G/G	A/A	G/G	G/A	G/G	G/G	G/A	G/G
	rs3784821	1569301	G	A	G/G	G/G	G/G	G/G	G/G	G/G	G/A	G/A	G/G	G/G	G/A	G/G	G/A	G/G	G/G
	rs3784822	1569308	C	T	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/T	C/C	C/C	C/T	C/C	C/T	C/C	C/C
	rs2667665	1569524	A	C	A/A	A/C	A/A	A/A	A/A	A/A	A/C	A/C	C/C	A/A	C/C	A/A	A/C	A/C	A/A
	rs199887622	1570059	G	A	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
	rs113734226	1570078	C	T	C/C	C/C	C/C	C/C	C/T	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
	rs932391	1570857	T	C	T/T	T/C	T/T	T/T	T/T	T/T	T/C	T/C	C/C	T/T	C/C	T/T	T/C	T/C	T/T
	rs932392	1571120	T	G	T/T	T/G	T/T	T/T	T/T	T/T	T/G	T/T	G/G	T/T	T/G	T/T	T/T	T/G	T/T
	rs73497551	1571811	T	C	T/T	T/C	T/T	T/T	T/T	T/T	T/C	T/C	C/C	T/T	C/C	T/T	T/C	T/C	T/T
	rs111953232	1571862	G	GA	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
	rs9328919	1572282	C	T	C/C	C/T	C/C	C/C	C/C	C/C	C/T	C/C	T/T	C/C	C/T	C/C	C/C	C/C	C/C
	rs4787271	1572314	T	C	T/T	T/C	T/T	T/T	T/T	T/T	T/C	T/T	C/C	T/T	C/C	T/T	T/T	T/C	T/T
	rs9673971	1572393	G	A	G/G	G/A	G/G	G/G	G/G	G/G	G/G	G/G	A/A	G/G	G/A	G/G	G/G	G/A	G/G
	rs8055979	1572569	T	C	T/T	T/C	T/T	T/T	T/T	T/T	T/C	T/C	C/C	T/T	C/C	T/T	T/C	T/C	T/T
	rs76393829	1572594	C	T	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/T	C/C	C/C	C/T	C/C	C/T	C/C	C/C
	rs554375875	1573219	CA	C	CA/CA	CA/CA	CA/CA	CA/CA	CA/CA	CA/C	CA/CA	CA/C	C/C	CA/CA	CA/CA	CA/CA	CA/C	CA/C	CA/CA
	rs11361072	1573219	CA	C	CA/CA	CA/C	CA/CA	CA/CA	CA/CA	CA/CA	CA/CA	CA/CA	CA/CA	CA/CA	CA/C	CA/CA	CA/CA	CA/CA	CA/CA
	rs34819117	1573485	G	A	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/A	G/G	G/G	G/A	G/G	G/A	G/G	G/G
	rs78028801	1573754	C	T	C/C	C/C	C/C	C/C	C/C	C/C	C/T	C/C	C/T	C/C	C/T	C/C	C/C	C/T	C/C
	rs2745176	1573810	A	G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
	rs2235638	1573890	G	A	G/G	G/G	G/G	G/G	G/G	G/G	G/A	G/G	G/G	G/G	G/A	G/G	G/A	G/G	G/G
	rs2076437	1574030	G	C	G/G	G/G	G/G	G/G	G/G	G/C	G/G	G/C	G/G	G/G	G/G	G/G	G/G	G/G	G/G
	rs141171038	1574170	A	G	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
	rs2235639	1574252	A	G	A/G	A/G	A/A	A/A	A/A	A/A	A/G	A/G	G/G	A/G	G/G	A/A	G/G	A/G	A/A
	rs2235640	1574863	G	A	G/G	G/A	G/G	G/G	G/G	G/G	G/G	G/A	A/A	G/G	G/A	G/G	G/G	G/A	G/G
	rs2745178	1575704	T	C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
	rs112977910	1576858	C	T	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/T	C/C	C/C	C/C
	rs2745181	1577056	A	G	A/A	A/G	A/A	A/A	A/G	A/A	A/G	A/G	G/G	A/A	G/G	A/G	A/G	A/G	A/A
	rs12599859	1577184	A	G	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/G	A/A	A/A	A/G	A/A	A/G	A/A	A/A
	rs9939655	1577963	A	G	A/A	A/G	A/A	A/A	A/G	A/A	A/G	A/G	G/G	A/A	G/G	A/G	A/G	A/G	A/A
	rs192452222	1578295	C	T	C/C	C/C	C/C	C/C	C/C	C/C	C/T	C/C	C/C	C/C	C/C	C/C	C/C	C/T	C/C
	rs140400369	1578373	C	G	C/C	C/C	C/C	C/C	C/C	C/C	C/G	C/C	C/C	C/C	C/C	C/C	C/C	C/G	C/C
	rs149176109	1578434	G	GCA	G/G	G/GCA	G/G	G/G	G/GCA	G/G	G/GCA	G/GCA	GCA/GCA	G/G	GCA/GCA	G/GCA	G/GCA	G/GCA	G/G
	rs1983600	1578514	A	C	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/C	A/A	A/A	A/C	A/A	A/C	A/A	A/A
	rs1983601	1578725	G	A	G/G	G/G	G/G	G/G	G/G	G/G	G/A	G/G	G/G	G/G	G/A	G/G	G/A	G/G	G/G
	rs3784825	1578914	A	G	A/A	A/A	A/A	A/A	A/A	A/A	A/G	A/A	A/A	A/A	A/G	A/A	A/G	A/A	A/A
	rs138773419	1578999	G	A	G/G	G/G	G/G	G/G	G/G	G/A	G/G	G/A	G/G	G/G	G/G	G/G	G/G	G/G	G/G
	rs2064291	1579011	T	C	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/C	T/T	T/T	T/C	T/T	T/C	T/T	T/T
	rs3784828	1579151	C	T	C/C	C/C	C/C	C/C	C/C	C/C	C/T	C/C	C/C	C/C	C/T	C/C	C/T	C/C	C/C
	rs148131225	1579244	T	C	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/C	T/T	T/T	T/T	T/T	T/T	T/T
	rs9934455	1579483	C	T	C/C	C/T	C/C	C/C	C/T	C/C	C/T	C/C	T/T	C/C	C/T	C/T	C/C	C/T	C/C
	rs4786275	1580025	A	G	A/G	A/A	A/A	A/A	A/A	A/A	A/A	A/G	A/A	A/A	A/G	A/A	A/G	A/A	A/A
	rs113201692	1580708	G	A	G/G	G/G	G/G	G/G	G/A	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
	rs142700555	1580752	G	T	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/T	G/G	G/G	G/G	G/G	G/G
	rs2745183	1581770	A	G	A/G	A/A	A/A	A/A	A/A	A/A	A/A	A/G	A/A	A/A	A/G	A/A	A/G	A/A	A/A
	rs67832788	1582870	C	T	C/T	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
	rs3809586	1583201	C	T	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/T	C/C	C/C	C/T	C/C	C/T	C/C	C/C
	rs118152236	1583238	C	A	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/A
	rs8048206	1583634	T	C	T/C	T/T	T/T	T/T	T/T	T/T	T/C	T/T	T/T	T/T	T/C	T/T	T/C	T/T	T/T

	rs1057612	1584446	G	C	G/C	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
	rs2076438	1584618	T	C	T/C	T/C	T/T	T/T	T/C	T/T	T/C	T/C	C/C	T/T	C/C	T/C	C/C	T/C	T/T
	rs12446975	1584628	G	A	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/A	G/G	G/G	G/A	G/G	G/A	G/G	G/G
	rs143833421	1584847	G	A	G/G	G/A	G/A	G/A	G/G	G/A	G/A	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/A
	rs2235642	1584866	T	C	T/C	T/C	T/T	T/T	T/C	T/T	T/C	T/C	C/C	T/T	C/C	T/C	C/C	T/C	T/T
	rs116903565	1584867	G	A	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/A	G/G	G/G	G/G	G/G
	rs2235643	1585115	G	A	G/G	G/A	G/G	G/G	G/A	G/G	G/A	G/G	A/A	G/G	G/A	G/A	G/A	G/A	G/G
	rs7196649	1585456	A	G	A/G	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
	rs141460785	1585650	A	G	A/G	A/A	A/G	A/G	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
	rs72761105	1585928	C	T	C/T	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
	rs2294624	1586255	G	A	G/A	G/G	G/G	G/G	G/G	G/G	G/G	G/A	G/G	G/G	G/A	G/G	G/A	G/G	G/G
	rs190265545	1586275	C	T	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/T	C/C	C/C	C/C	C/C	C/C	C/C
	rs2294625	1586596	C	T	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/T	C/C	C/C	C/T	C/C	C/T	C/C	C/C
	rs2294626	1586717	C	G	C/C	C/C	C/C	C/C	C/G	C/C	C/G	C/C	G/G	C/C	C/G	C/G	C/G	C/G	C/C
	rs2294627	1586821	G	A	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/A	G/G	G/G	G/A	G/G	G/A	G/G	G/G
	rs12923143	1587067	T	C	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/C	T/T
	rs2076439	1587277	C	T	C/T	C/C	C/C	C/C	C/T	C/C	C/T	C/T	T/T	C/C	T/T	C/T	T/T	C/T	C/C
	rs2076440	1587322	G	A	G/A	G/G	G/G	G/G	G/A	G/G	G/A	G/A	A/A	G/G	A/A	G/A	A/A	G/A	G/G
	rs2066937	1587511	G	A	G/G	G/G	G/G	G/G	G/A	G/G	G/A	G/G	A/A	G/G	G/A	G/A	G/A	G/A	G/G
	rs141085747	1587842	C	T	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/T	C/C	C/C	C/C	C/C
	rs78140717	1587938	C	T	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/T	C/C	C/C	C/T	C/C	C/T	C/C	C/C
	rs8051295	1588087	A	G	A/G	A/A	A/A	A/A	A/G	A/A	A/G	A/G	G/G	A/A	G/G	A/G	G/G	A/G	A/A
	rs28668520	1588800	G	A	G/G	G/G	G/G	G/G	G/A	G/G	G/A	G/G	A/A	G/G	G/A	G/A	G/A	G/A	G/G
	rs35521135	1588849	C	T	C/C	C/C	C/C	C/C	C/T	C/C	C/T	C/C	T/T	C/C	C/T	C/T	C/T	C/T	C/C
	rs189807580	1589671	T	C	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/C	T/T	T/T	T/T	T/T	T/T
	rs80113230	1590168	G	A	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/A	G/G	G/G	G/A	G/G	G/A	G/G	G/G
	rs116941519	1590480	G	A	G/G	G/G	G/G	G/G	G/A	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
	rs2076441	1590576	C	G	C/G	C/G	C/G	C/G	C/G	C/G	G/G	C/G	G/G	C/C	G/G	C/G	G/G	G/G	C/G
	rs6600152	1590672	G	A	G/A	G/G	G/G	G/G	G/G	G/G	G/G	G/A	G/G	G/G	G/A	G/G	G/A	G/G	G/G
	rs2076442	1591112	G	T	G/T	G/G	G/G	G/G	G/T	G/G	G/T	G/T	T/T	G/G	T/T	G/T	T/T	G/T	G/G
	rs117189215	1591284	G	A	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
	rs2050117	1591592	G	A	G/G	G/G	G/G	G/G	G/A	G/G	G/A	G/G	A/A	G/G	G/A	G/A	G/A	G/A	G/G
	rs2076443	1591962	G	A	G/G	G/G	G/G	G/G	G/A	G/G	G/A	G/G	A/A	G/G	G/A	G/A	G/A	G/A	G/G
	rs2076444	1592087	T	C	T/T	T/T	T/T	T/T	T/C	T/T	T/C	T/T	C/C	T/T	T/C	T/C	T/C	T/C	T/T
	rs2076445	1592266	A	T	A/T	A/A	A/A	A/A	A/T	A/A	A/T	A/T	T/T	A/A	T/T	A/T	T/T	A/T	A/A
	rs11649407	1592367	G	A	G/A	G/G	G/G	G/G	G/G	G/G	G/A	G/A	G/G	G/G	G/A	G/G	G/A	G/G	G/G
	rs2076446	1592524	C	T	C/C	C/C	C/C	C/C	C/T	C/C	C/T	C/C	T/T	C/C	C/T	C/T	C/T	C/T	C/C
	rs3784830	1592857	G	C	G/C	G/G	G/G	G/G	G/G	G/G	G/C	G/C	G/G	G/G	G/C	G/G	G/C	G/G	G/G
	rs2050118	1593004	G	A	G/G	G/G	G/G	G/G	G/A	G/G	G/A	G/G	A/A	G/G	G/A	G/A	G/A	G/A	G/G
	rs3784832	1593108	G	C	G/C	G/G	G/G	G/G	G/G	G/G	G/G	G/C	G/G	G/G	G/C	G/G	G/C	G/G	G/G
	rs2143285	1593171	A	G	A/G	A/A	A/A	A/A	A/G	A/A	A/G	A/G	G/G	A/A	G/G	A/G	G/G	A/G	A/A
	rs8049972	1593606	C	T	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/T	C/C	C/C	C/T	C/C	C/T	C/C	C/C
	rs2064289	1593645	C	T	C/T	C/C	C/C	C/C	C/T	C/C	C/T	C/T	T/T	C/T	T/T	C/T	T/T	C/T	C/C
	rs4558418	1593843	G	A	G/G	G/G	G/G	G/G	G/A	G/G	G/A	G/G	A/A	G/G	G/A	G/A	G/A	G/A	G/G
	rs8053605	1594158	G	A	G/A	G/G	G/G	G/G	G/G	G/G	G/A	G/A	A/A	G/G	A/A	G/A	A/A	G/A	G/G
	rs8053608	1594161	G	T	G/G	G/G	G/G	G/G	G/T	G/G	G/G	G/G	T/T	G/G	G/T	G/T	G/T	G/T	G/G
	rs569784189	1594162	GCACCTGC	G	CTGC/GCACCTGC/GCACCTGC/GCACCTGC/GCACCTGC/GCACCTGC/GCACCTGC/GCACCTGC/GCACCTGC/GCACCTGC/GCACCTGC/GCACCTGC/GCACCTGC/GCACCTGC/GCACCTGC/GCACCTGC/GCACCTGC/GCACCTGC/GCAC														
	rs150672825	1594218	T	A	T/A	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
	rs6600153	1594248	A	G	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/G	A/G	A/A	A/A	A/A
	rs35055895	1594314	A	G	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/G	A/A	A/A	A/A	A/A
	rs140189706	1594376	T	C	T/C	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
	rs192068259	1594672	T	C	T/C	T/T	T/C	T/C	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
	rs183277603	1594683	C	T	C/T	C/C	C/T	C/T	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
	rs187549062	1594688	G	C	G/C	G/G	G/C	G/C	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
	rs8056892	1594926	A	G	A/G	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
	rs12445282	1594944	C	T	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/T	C/C	C/C	C/T	C/C	C/T	C/C	C/C
	rs56342298	1595600	C	T	C/T	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
	rs2294622	1596259	C	T	C/C	C/C	C/C	C/C	C/C	C/C	C/T	C/C	C/T	C/C	C/T	C/C	C/C	C/T	C/C
	rs2294623	1597210	C	T	C/T	C/C	C/C	C/C	C/C	C/C	C/T	C/T	C/C	C/C	C/T	C/C	C/T	C/C	C/C
	rs2235644	1597814	G	A	G/A	G/G	G/G	G/G	G/G	G/G	G/G	G/A	G/G	G/G	G/A	G/G	G/A	G/G	G/G
	rs11864966	1597877	C	T	C/T	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
	rs2235645	1597924	C	T	C/C	C/C	C/C	C/C	C/T	C/C	C/T	C/C	T/T	C/C	C/T	C/T	C/T	C/T	C/C
	rs368948845	1599590	G	A	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
	rs2281227	1599913	G	A	G/G	G/G	G/G	G/G	G/A	G/G	G/A	G/G	A/A	G/G	G/A	G/A	G/A	G/A	G/G
	rs2281228	1600137	C	T	C/C	C/C	C/C	C/C	C/T	C/C	C/T	C/C	T/T	C/C	C/T	C/T	C/T	C/T	C/C
	rs189219366	1600216	C	G	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/G	C/C	C/C
	rs74605356	1600272	T	C	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/C	T/T	T/T	T/C	T/T	T/C	T/T	T/T
	rs2281229	1600738	G	A	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/A	G/G	G/G	G/A	G/G	G/A	G/G	G/G
	rs2235646	1601287	A	G	A/G	A/A	A/A	A/A	A/G	A/A	A/G	A/G	G/G	A/A	G/G	A/G	G/G	A/G	A/A
	rs2235647	1601331	C	A	C/A	C/C	C/C	C/C	C/A	C/C	C/A	C/C	A/A	C/C	C/A	C/A	C/A	C/A	C/C
	rs2667664	1601458	A	G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
	rs2281230	1601767	T	C	T/C	T/T	T/T	T/T	T/C	T/T	T/C	T/T	C/C	T/T	T/C	T/C	T/C	T/C	T/T
	rs80150449	1601816	C	T	C/T	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
	rs2281231	1601879	A	G	A/G	A/A	A/A	A/A	A/G	A/A	A/G	A/G	G/G	A/A	G/G	A/G	G/G	A/G	A/A
	rs79038790	1602035	C	T	C/C	C/C	C/C	C/C	C/C	C/T	C/C	C/T	C/C	C/C	C/C	C/C	C/C	C/C	C/C
	rs2281232	1602114	C	A	C/A	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
	rs8050274	1602260	C	T	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/T	C/C	C/C	C/T	C/C	C/T	C/C	C/C
	rs2281233	1602323	T	C	T/C	T/T	T/T	T/T	T/T	T/T	T/T	T/C	T/T	T/T	T/C	T/T	T/C	T/T	T/T
	rs8051229	1602427	A	C	A/A	A/A	A/A	A/A	A/A	A/A	A/C	A/C	A/A	A/A	A/C	A/A	A/C	A/A	A/A
	rs8049566	1602594	G	A	G/G	G/G	G/G	G/G	G/G	G/G	G/A	G/G	G/G	G/G	G/A	G/G	G/A	G/G	G/G
	rs11409894	1602649	C	CT	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
	rs559777878	1602649	C	CT	C/C	C/C	C/C	C/C	C/CT	C/C	C/CT	C/CT	C/CT	C/C	C/C	C/C	C/CT	C/CT	C/C
	rs111621426	1602738	C	T	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/T	C/C	C/C	C/C	C/C	C/C	C/C
	rs8051895	1602811	A	G	A/G	A/A	A/A	A/A	A/G	A/A	A/G	A/G	G/G	A/A	G/G	A/G	G/G	A/G	A/A
	rs8051909	1602845	A	G	A/G	A/A	A/A	A/A	A/G	A/A	A/G	A/G	G/G	A/A	G/G	A/G	G/G	A/G	A/A
	rs2281234	1602938	A	G	A/G	A/A	A/A	A/A	A/G	A/A	A/G	A/G	G/G	A/A	G/G	A/G	G/G	A/G	A/A
	rs2281235	1603535	T	C	T/C	T/T	T/T	T/T	T/C	T/T	T/C	T/C	C/C	T/T	C/C	T/C	C/C	T/C	T/T
	rs2281236	1603576	G	A	G/A	G/G	G/G	G/G	G/A	G/G	G/A	G/A	A/A	G/G	A/A	G/A	A/A	G/A	G/G
	rs79005075	1603716	C	A	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/A	C/C	C/C	C/A	C/C	C/A	C/C	C/C
	rs66798554	1604388	G	A	G/G	G/G	G/G	G/G	G/A	G/G	G/A	G/G	A/A	G/G	G/A	G/A	G/A	G/A	G/G
	rs2272976	1604701	T	C	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/C	T/T	T/T	T/C	T/T	T/C	T/T	T/T
	rs9932984	1605051	A	G	A/G	A/A	A/A	A/A	A/G	A/A	A/G	A/G	G/G	A/A	G/G	A/G	G/G	A/G	A/A
	rs12443939	1605053	C	T	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/T	C/C	C/C	C/T	C/C	C/T	C/C	C/C
	rs9922661	1605163	C	A	C/C	C/C	C/C	C/C	C/A	C/C	C/A	C/C	A/A	C/C	C/A	C/A	C/A	C/A	C/C
	rs114022362	1605234	T	C	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/C	T/T	T/T	T/T
	rs1057610	1605537	T	C	T/C	T/T	T/T	T/T	T/C	T/T	T/C	T/C	C/C	T/T	C/C	T/C	C/C	T/C	T/T
	rs17135391	1605606	C	T	C/C	C/C	C/C	C/C	C/T	C/C	C/T	C/C	T/T	C/C	C/T	C/T	C/T	C/T	C/C
	rs8048382	1605975	C	T	C/T	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
	rs8053335	1606167	A	G	A/G	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
	rs138001290	1606336	C	T	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C



	rs368082114	1628604	C	T	C/C	C/C	C/C	C/C	C/C	C/C	C/T	C/C	C/C	C/C	C/C	C/C	C/C	C/T	C/C
	rs138003160	1629061	T	C	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
	rs573551047	1629964	CA	C	CA/CA	CA/CA	CA/CA	CA/CA	CA/CA	CA/CA	CA/C	CA/CA	CA/CA	CA/CA	CA/CA	CA/CA	CA/CA	CA/CA	CA/CA
	rs11867080	1630185	T	C	T/C	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
	rs3743507	1630381	G	A	G/G	G/G	G/G	G/G	G/A	G/G	G/A	G/G	A/A	G/G	G/A	G/A	G/A	G/A	G/G
	rs56218869	1631294	A	G	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/G	A/A	A/A	A/A	A/A	A/A
	rs114623241	1631536	G	A	G/A	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
	rs140950029	1631591	A	ACT	A/ACT	A/A	A/A	A/A	A/A	A/A	A/A	A/ACT	A/A	A/A	A/ACT	A/A	A/ACT	A/A	A/A
	rs187641585	1631762	G	A	A/A	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
	rs117406049	1631809	T	G	G/G	T/T	T/T	T/T	T/T	T/T	T/T	T/G	T/T	T/T	T/T	T/T	T/G	T/T	T/T
	rs72761124	1632448	C	T	C/T	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
	rs76835012	1632901	C	T	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
	rs9934342	1632965	G	A	G/A	G/G	G/G	G/G	G/G	G/G	G/G	G/A	G/G	G/G	G/A	G/G	G/A	G/G	G/G
	rs9927202	1633100	A	G	A/G	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
	rs147086637	1633446	A	C	A/C	A/A	A/C	A/C	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
	rs8055710	1633872	G	A	G/A	G/G	G/G	G/G	G/G	G/G	G/G	G/A	G/G	G/G	G/A	G/G	G/A	G/G	G/G
	rs28406883	1633946	G	C	G/C	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
	rs34762152	1634385	C	T	C/T	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
	rs76705317	1635137	G	A	G/G	G/G	G/G	G/G	G/G	G/G	G/A	G/G	G/A	G/G	G/A	G/G	G/G	G/A	G/G
	rs55695276	1635161	A	G	A/G	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
	rs28647246	1635752	T	C	T/C	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
	rs28608068	1635872	C	T	C/T	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
	rs11640303	1635908	G	A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
	rs2236268	1636549	C	A	C/A	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
	rs71860883	1636628	GCA	G	GCA/G	GCA/GCA	GCA/GCA	GCA/GCA	GCA/G	GCA/G	GCA/G	GCA/G	GCA/GCA	GCA/GCA	G/G	GCA/GCA	GCA/GCA	GCA/G	GCA/GCA
	rs76842429	1636896	C	T	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/T	C/C	C/C	C/T	C/C	C/T	C/C	C/C
	rs113143675	1636908	A	G	A/G	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
	rs77750249	1636935	T	G	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/G	T/T	T/T	T/G	T/T	T/G	T/T	T/T
	rs574261663	1637667	A	G	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
	rs431905521	1637962	C	T	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/T	C/C	C/C	C/C	C/C	C/C	C/C
	rs35972580	1638323	A	G	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/G	A/A	A/A	A/G	A/A	A/G	A/A	A/A
	rs180758804	1638450	C	T	C/C	C/C	C/C	C/C	C/C	C/T	C/C	C/T	C/C	C/C	C/C	C/C	C/C	C/C	C/C
	rs145693119	1638528	A	ATAATAAT	A/A	A/A	A/A	A/A	AATAATAAT	A/A	AATAATAAT	A/A	A/A	A/A	A/A	A/A	A/A	AATAATAAT	A/A
	rs145693119	1638528	A	AATAATAA	A/A	A/A	A/A	A/A	A/A	A/A	A/A	ATAATAATA	A/A	A/A	A/A	A/A	A/A	A/A	A/A
	rs72761127	1638820	C	T	C/T	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
	rs9924141	1638989	T	C	T/C	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
	rs142607578	1639029	G	A	G/G	G/G	G/G	G/G	G/G	G/A	G/G	G/A	G/G	G/G	G/G	G/G	G/G	G/G	G/G
	rs12447357	1639793	C	G	C/G	C/C	C/C	C/C	C/C	C/C	C/C	C/G	C/C	C/C	C/G	C/C	C/G	C/C	C/C
	rs112431294	1639860	A	T	A/A	A/A	A/A	A/A	A/T	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
	rs72761128	1640035	T	C	T/C	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
	rs758621945	1640201	C	CA	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/CA	C/C	C/C	C/C
	rs11859318	1640301	C	G	C/G	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
	rs11641248	1641093	C	T	C/C	C/C	C/C	C/C	C/C	C/C	C/T	C/C	C/C	C/C	C/T	C/C	C/T	C/C	C/C
	rs536399679	1641983	G	A	G/G	G/G	G/G	G/G	G/G	G/G	G/A	G/G	G/G	G/G	G/G	G/G	G/G	G/A	G/G
	rs116859784	1642026	G	A	G/A	G/G	G/A	G/A	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
	rs201188361	1642177	C	T	C/C	C/C	C/C	C/C	C/T	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
	rs2273679	1642423	C	T	C/T	C/C	C/C	C/C	C/C	C/C	C/T	C/C	C/C	C/C	C/T	C/C	C/T	C/C	C/C
	rs112565100	1642660	T	C	T/C	T/T	T/T	T/T	T/T	T/T	T/C	T/T	T/T	T/C	T/T	T/C	T/T	T/T	T/T
	rs9927226	1642679	G	A	G/A	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
	rs80064098	1642906	T	C	T/T	T/T	T/T	T/T	T/T	T/T	T/C	T/T	T/T	T/C	T/T	T/C	T/T	T/T	T/T
	rs796687743	1643316	CT	C	C/T	C/T	C/T	C/T	CT/C	CT/CT	CT/CT	CT/CT	CT/CT	C/T	C/T	C/T	CT/CT	CT/C	CT/CT
	rs187540964	1643455	G	A	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
	rs12446250	1643552	C	T	C/C	C/C	C/C	C/C	C/C	C/C	C/T	C/C	C/C	C/T	C/C	C/T	C/C	C/C	C/C
	rs12449265	1643606	A	T	A/A	A/A	A/A	A/A	A/A	A/A	A/T	A/A	A/A	A/T	A/A	A/T	A/A	A/A	A/A
	rs116910642	1644032	C	T	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/T
	rs536466763	1644227	G	GT	G/G	G/G	G/G	G/G	G/GT	G/G	G/GT	G/GT	G/G	G/G	G/G	G/G	G/G	G/GT	G/GT
	rs57833692	1644227	G	GT	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/GT	G/G	G/G	G/G	G/G	G/G
	rs2755183	1644312	C	G	C/G	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
	rs2941956	1644764	A	G	A/G	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
	rs9933780	1644808	C	G	C/G	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
	rs112191146	1645058	G	A	G/A	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
	rs2755184	1645264	T	C	T/C	T/T	T/T	T/T	T/C	T/T	T/C	T/C	C/C	T/T	C/C	T/C	C/C	T/C	T/T
	rs548658471	1645604	G	GA	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/GA	G/GA	G/G	G/G	G/G	G/GA	G/G	G/GA
	rs17135416	1645745	T	C	T/C	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
	rs2755185	1645827	A	G	A/G	A/A	A/A	A/A	A/G	A/A	A/G	A/G	G/G	A/A	G/G	A/G	G/G	A/G	A/A
	rs35576736	1646018	C	T	C/C	C/C	C/C	C/C	C/T	C/C	C/T	C/C	T/T	C/C	C/T	C/T	C/T	C/T	C/C
	rs2667669	1646543	G	A	G/A	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
	rs193299834	1646787	C	T	C/C	C/C	C/C	C/C	C/T	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
	rs13334027	1646934	A	C	A/C	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A
	rs143684159	1646973	C	A	C/C	C/A	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
	rs6600156	1648126	A	C	A/C	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/C	A/C	A/A	A/A	A/A	A/A
	rs6600156	1648126	A	C	A/A	A/A	A/A	A/A	A/C	A/A	A/C	A/A	C/C	A/A	A/A	A/A	A/C	A/C	A/A
	rs2974845	1648162	T	C	T/C	T/T	T/T	T/T	T/C	T/T	T/C	T/C	C/C	T/T	C/C	T/C	C/C	T/C	T/T
	rs4786368	1648638	G	C	G/C	G/G	G/G	G/G	G/C	G/G	G/C	G/C	C/C	G/C	C/C	G/C	C/C	G/C	G/G
	rs78305748	1648685	G	C	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/C	G/G	G/G	G/C	G/G	G/C	G/G	G/G
	rs186995221	1649287	A	G	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/A	A/G	A/A	A/A	A/A	A/A
	rs201340184	1649393	T	C	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/C	T/T	T/C	T/T	T/T	T/T
	rs180744609	1649476	G	T	G/G	G/T	G/T	G/T	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
	rs769942411	1650461	AT	A	AT/AT	AT/AT	AT/AT	AT/AT	AT/AT	AT/AT	AT/AT	AT/AT	AT/A	AT/AT	AT/AT	AT/AT	AT/AT	AT/AT	AT/AT
	rs137922934	1650463	TTA	T	TTA/TTA	TTA/TTA	TTA/TTA	TTA/TTA	TTA/TTA	TTA/TTA	TTA/TTA	TTA/TT	TTA/TTA	TTA/TTA	TTA/TTA	TTA/TTA	TTA/TT	TTA/TTA	TTA/TTA
	rs137922934	1650463	TTA	T	TTA/TTA	TTA/TTA	TTA/TTA	TTA/TTA	TTA/TTA	TTA/TTA	TTA/TTA	TTA/T	TTA/T	TTA/TTA	TTA/T	TTA/TTA	TTA/T	TTA/TTA	TTA/TTA
	rs34665250	1650464	TA	T	TA/T	TA/TA	TA/TA	TA/TA	TA/T	TA/T	TA/T	TA/T	TA/T	TA/TA	TA/TA	TA/TA	TA/TA	TA/T	TA/T
	rs34665250	1650464	AAAAAAA	AAAAAAA	AAAAA/TAAAAAA/TAAAAAA/TAA

	rs138669173	1654525	C	G	C/C	C/C	C/C	C/C	C/C	C/C	C/G	C/C	C/C	C/C	C/C	C/C	C/C	C/G	C/C
	rs7197288	1654753	C	A	C/A	C/C	C/C	C/C	C/A	C/C	C/A	C/A	A/A	C/C	A/A	C/A	A/A	C/A	C/C
	rs2667679	1654946	G	A	G/A	G/G	G/G	G/G	G/A	G/G	G/A	G/G	A/A	G/G	G/A	G/A	G/A	G/A	G/G
	rs56359342	1654948	G	A	G/A	G/G	G/G	G/G	G/A	G/G	G/A	G/G	A/A	G/G	G/A	G/A	G/A	G/A	G/G
	rs117745556	1655073	T	C	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/C	T/T	T/T	T/C	T/T	T/C	T/T	T/T
	rs35220693	1655439	C	T	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/T	C/C	C/C	C/T	C/C	C/T	C/C	C/C
	rs6600157	1656321	G	A	G/A	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/G
	rs142169797	1656564	C	T	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
	rs2859310	1656765	A	G	A/G	A/A	A/A	A/A	A/G	A/A	A/G	A/G	G/G	A/A	G/G	A/G	G/G	A/G	A/A
	rs542474562	1656927	GCACA	G	CACA/GCACCACA/GCACCACA/GCACCACA/GCAC	CACA/GCAC	CACA/GCACCACA/GCACCACA/GCAC	CACA/GCACCACA/GCACCACA/GCAC	CACA/GCACCACA/GCACCACA/GCAC	CACA/GCAC	CACA/GCACCACA/GCACCACA/GCAC	CACA/GCACCACA/GCACCACA/GCAC	CACA/GCAC	CACA/GCACCACA/GCACCACA/GCAC	CACA/GCACCACA/GCACCACA/GCAC	CACA/GCACCACA/GCACCACA/GCAC	CACA/GCACCACA/GCACCACA/GCAC	CACA/GCACCACA/GCACCACA/GCAC	CACA/GCACCACA/GCAC	GCACA/G
	rs7192480	1657012	C	T	C/C	C/C	C/C	C/C	C/T	C/C	C/T	C/C	T/T	C/C	C/T	C/T	C/T	C/T	C/C
	rs34935994	1657034	CAT	C	CAT/C	CAT/CAT	CAT/CAT	CAT/CAT	CAT/CAT	CAT/CAT	CAT/CAT	CAT/CAT	CAT/CAT	CAT/CAT	CAT/CAT	CAT/CAT	CAT/CAT	CAT/CAT	CAT/CAT	CAT/CAT
	rs1894649	1657284	C	G	C/G	C/C	C/C	C/C	C/G	C/C	C/G	C/G	G/G	C/C	G/G	C/G	G/G	C/G	C/C
	rs743963	1657350	A	T	A/T	A/A	A/A	A/A	A/T	A/A	A/T	A/A	T/T	A/A	A/T	A/T	A/T	A/T	A/A
	rs3784840	1657754	C	T	C/C	C/C	C/C	C/C	C/T	C/C	C/T	C/C	T/T	C/C	C/T	C/T	C/T	C/T	C/C
	rs743964	1657853	C	T	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/T	C/C	C/C	C/T	C/C	C/T	C/C	C/C
	rs763152	1658056	T	C	T/C	T/T	T/T	T/T	T/C	T/T	T/C	T/C	C/C	T/T	C/C	T/C	C/C	T/C	T/T
	rs763153	1658086	G	C	G/G	G/G	G/G	G/G	G/G	G/G	G/G	G/C	G/G	G/G	G/C	G/G	G/C	G/G	G/G
	rs35888283	1658101	TAG	T	TAG/T	TAG/TAG	TAG/TAG	TAG/TAG	TAG/T	TAG/TAG	TAG/T	TAG/TAG	T/T	TAG/TAG	TAG/T	TAG/T	TAG/T	TAG/T	TAG/TAG
	rs9937922	1658154	C	G	C/G	C/C	C/C	C/C	C/G	C/C	C/G	C/G	G/G	C/C	G/G	C/G	G/G	C/G	C/C
	rs115729443	1658454	C	T	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/T	C/C	C/C	C/T	C/C	C/T	C/C	C/C
	rs11438070	1658632	C	CA	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
	rs28519681	1659128	T	C	T/C	T/T	T/T	T/T	T/C	T/T	T/C	T/C	C/C	T/T	C/C	T/C	C/C	T/C	T/T
	rs28633318	1659723	T	G	T/G	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T	T/T
	rs1057983	1660274	A	G	A/G	A/A	A/A	A/A	A/G	A/A	A/G	A/G	G/G	A/A	G/G	A/G	G/G	A/G	A/A
	rs150781872	1661491	C	CA	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C	C/C
	rs2859314	1661701	A	G	A/G	A/A	A/A	A/A	A/A	A/A	A/G	A/G	A/G	A/A	G/G	A/A	A/G	A/G	A/A
	rs553272330	1661814	CA	C	CA/CA	CA/CA	CA/CA	CA/CA	CA/CA	CA/C	CA/C	CA/C	CA/C	CA/CA	CA/CA	CA/CA	CA/C	CA/CA	CA/C

Table S6. Haplotype analysis for ~100kb around *IFT140* on chromosome 16.

SNP analysis at the *IFT140* locus for carrier individuals of the tandem duplication and in controls sequenced either by WGS or Target Exome Sequencing (TES, by the Leber panel which include the full *IFT140* gene). In total, 6 carrier individuals from 3 families were analysed (family A, B and C) and indicated. Affected patients are depicted with an "*". Six control samples are also shown. For each individual, SNP detected in *IFT140* were extracted and listed. Homozygous SNP are shown in green (reference allele) or in blue (alternative allele) and heterozygous SNP in orange. The duplication region is highlighted in grey.

For Peer Review

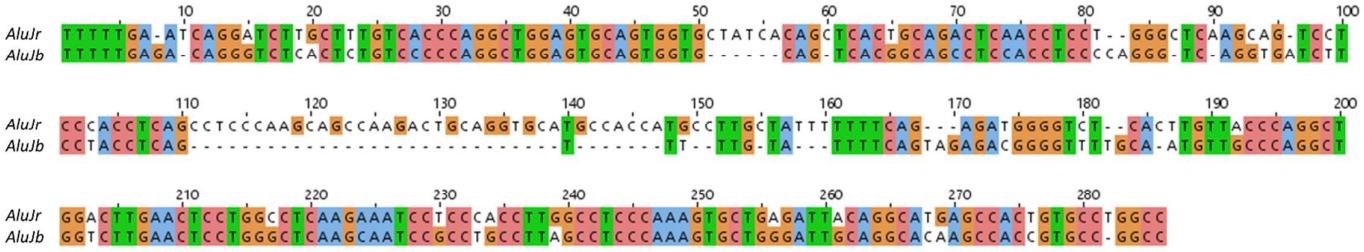


Figure S5 *AluJb* and *AluJr* alignment

AluJb and *AluJr* sequences were extracted from IGV (Thorvaldsdottir, et al., 2013) at the breakpoints of the tandem duplications. Pairwise alignment of the nucleotide sequences and percent identity has been computed using Jalview (Waterhouse, et al., 2009).

For Peer Review

Technique	Individuals	CANOES	Pindel	CNVnator
Ciliome panel	**D-II.1**	**DUP, intron26-intron30**	67 DEL, 13 INV, 35 SI	NA
Ciliome panel	**K-II.2**	2 DEL, 3DUP (**DUP, intron27-exon31**)	36 DEL, 1 INV, 6 SI, 1 TD	NA
Leber panel	**B-I.1**	**DUP, intron26-intron30**	**DUP, intron26-intron30** (675 DEL, 70 INV, 233 SI, 5 TD)	NA
Leber panel	**B-I.2**	0	(688 DEL, 62 INV, 211 SI, 2 TD)	NA
Leber panel	**B-II.1**	**DUP, intron26-intron30**	**DUP, intron26-intron30** (595 DEL, 55 INV, 193 SI, 2 TD)	NA
Leber panel	**C-II.1**	**DUP, intron26-intron30**	**DUP, intron26-intron30** (842 DEL, 120 INV, 417 SI, 5 TD)	NA
Leber panel	**I-II.1**	1 DEL	(659 DEL, 115 INV, 339 SI, 2 TD)	NA
WES	**A-II.1**	0	10 DEL, 3 INV, 10 SI, 1 TD	NA
WES	**A-II.2**	**DUP, intron27-intron30**	13 DEL, 3 INV, 9 SI	NA
WGS	**A-II.1**	**DUP, exon26-intron31**	14 DEL, 3 INV, 4 SI	0
WGS	**A-II.2**	**DUP, exon26-intron31**	10 DEL, 2 INV, 3 SI	0
WGS	**A-I.2**	**DUP, exon26-intron31**	17 DEL, 1 INV, 2 SI	0
WGS	**A-I.1**	0	53 DEL, 1 INV, 1 SI	0

Table S7. Summary of the CNV detected in *IFT140* by using 3 different tools in 13 individuals.

CNV detection has been performed on 13 individuals sequenced with 4 different approaches: the Ciliome panel, the Leber panel, the WES and the WGS. Three tools have been run with their default options: CANOES (v1.0), Pindel (v0.2.5b9) and CNVnator (v0.3.3) (Abyzov, et al., 2011; Backenroth, et al., 2014; Ye, et al., 2009). The number of CNV detected in *IFT140* are reported here, as well as the presence or not of the tandem duplication of exons 26 to 30. It is to notice that Pindel simultaneously calls small indel and structural variations, increasing thus the number of detected variants. NA, not applicable; DUP, duplication; DEL, deletion; INV, inversion; SI, short insertion; TD, tandem duplication.

Peer Review

SUPPLEMENTARY REFERENCES

Abyzov A, Urban AE, Snyder M, Gerstein M. 2011. CNVnator: an approach to discover, genotype, and characterize typical and atypical CNVs from family and population genome sequencing. Genome Res 21(6):974-84.

Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR. 2010. A method and server for predicting damaging missense mutations. Nature Methods 7(4):248-9.

Backenroth D, Homsy J, Murillo LR, Glessner J, Lin E, Brueckner M, Lifton R, Goldmuntz E, Chung WK, Shen Y. 2014. CANOES: detecting rare copy number variants from whole exome sequencing data. Nucleic Acids Res 42(12):e97.

Bayat A, Kerr B, Douzgou S, Study DDD. 2017. The evolving craniofacial phenotype of a patient with Sensenbrenner syndrome caused by IFT140 compound heterozygous mutations. Clin Dysmorphol 26(4):247-251.

Beheshtian M, Saee Rad S, Babanejad M, Mohseni M, Hashemi H, Eshghabadi A, Hajizadeh F, Akbari MR, Kahrizi K, Riazi Esfahani M and others. 2015. Impact of whole exome sequencing among Iranian patients with autosomal recessive retinitis pigmentosa. Arch Iran Med 18(11):776-85.

Bifari IN, Elkhamary SM, Bolz HJ, Khan AO. 2016. The ophthalmic phenotype of IFT140-related ciliopathy ranges from isolated to syndromic congenital retinal dystrophy. Br J Ophthalmol 100(6):829-33.

DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, Philippakis AA, del Angel G, Rivas MA, Hanna M and others. 2011. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet 43(5):491-8.

Gerber S, Alzayady KJ, Burglen L, Bremond-Gignac D, Marchesin V, Roche O, Rio M, Funalot B, Calmon R, Durr A and others. 2016. Recessive and Dominant De Novo ITPR1 Mutations Cause Gillespie Syndrome. Am J Hum Genet 98(5):971-980.

Grampa V, Delous M, Zaidan M, Odye G, Thomas S, Elkhartoufi N, Filhol E, Niel O, Silbermann F, Lebreton C and others. 2016. Novel NEK8 Mutations Cause Severe Syndromic Renal Cystic Dysplasia through YAP Dysregulation. PLoS Genet 12(3):e1005894.

Helm BM, Willer JR, Sadeghpour A, Golzio C, Crouch E, Vergano SS, Katsanis N, Davis EE. 2017. Partial uniparental isodisomy of chromosome 16 unmasks a deleterious biallelic mutation in IFT140 that causes Mainzer-Saldino syndrome. Hum Genomics 11(1):16.

Huerta-Cepas J, Szklarczyk D, Forslund K, Cook H, Heller D, Walter MC, Rattei T, Mende DR, Sunagawa S, Kuhn M and others. 2016. eggNOG 4.5: a hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. Nucleic Acids Res 44(D1):D286-93.

Hull S, Owen N, Islam F, Tracey-White D, Plagnol V, Holder GE, Michaelides M, Carss K, Raymond FL, Rozet JM and others. 2016. Nonsyndromic Retinal Dystrophy due to Bi-Allelic Mutations in the Ciliary Transport Gene IFT140. Invest Ophthalmol Vis Sci 57(3):1053-62.

Khan AO, Bolz HJ, Bergmann C. 2014. Early-onset severe retinal dystrophy as the initial presentation of IFT140-related skeletal ciliopathy. J AAPOS 18(2):203-5.

Kumar P, Henikoff S, Ng PC. 2009. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nature protocols 4(7):1073-81.

Li H, Durbin R. 2010. Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics 26(5):589-95.

Pena-Padilla C, Marshall CR, Walker S, Scherer SW, Tavares-Macias G, Razo-Jimenez G, Bobadilla-Morales L, Acosta-Fernandez E, Corona-Rivera A, Mendoza-Londono R and others. 2016. Compound heterozygous mutations in the IFT140 gene cause Opitz trigonocephaly C syndrome in a patient with typical features of a ciliopathy. Clin Genet.

For Peer Review

Perrault I, Saunier S, Hanein S, Filhol E, Bizet AA, Collins F, Salih MA, Gerber S, Delphin N, Bigot K and others. 2012. Mainzer-Saldino syndrome is a ciliopathy caused by IFT140 mutations. Am J Hum Genet 90(5):864-70.

Schmidts M, Frank V, Eisenberger T, Al Turki S, Bizet AA, Antony D, Rix S, Decker C, Bachmann N, Bald M and others. 2013. Combined NGS approaches identify mutations in the intraflagellar transport gene IFT140 in skeletal ciliopathies with early progressive kidney Disease. Hum Mutat 34(5):714-24.

Schwarz JM, Rodelsperger C, Schuelke M, Seelow D. 2010. MutationTaster evaluates disease-causing potential of sequence alterations. Nat Methods 7(8):575-6.

Thorvaldsdottir H, Robinson JT, Mesirov JP. 2013. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief Bioinform 14(2):178-92.

Waterhouse AM, Procter JB, Martin DM, Clamp M, Barton GJ. 2009. Jalview Version 2--a multiple sequence alignment editor and analysis workbench. Bioinformatics 25(9):1189-91.

Weisschuh N, Mayer AK, Strom TM, Kohl S, Glockle N, Schubach M, Andreasson S, Bernd A, Birch DG, Hamel CP and others. 2016. Mutation Detection in Patients with Retinal Dystrophies Using Targeted Next Generation Sequencing. PLoS One 11(1):e0145951.

Xu M, Yang L, Wang F, Li H, Wang X, Wang W, Ge Z, Wang K, Zhao L, Li H and others. 2015. Mutations in human IFT140 cause non-syndromic retinal degeneration. Hum Genet 134(10):1069-78.

Ye K, Schulz MH, Long Q, Apweiler R, Ning Z. 2009. Pindel: a pattern growth approach to detect break points of large deletions and medium sized insertions from paired-end short reads. Bioinformatics 25(21):2865-71.

For Peer Review