Novel IQCE variations confirm its role in postaxial polydactyly and cause ciliary defect phenotype in zebrafish

Alejandro Estrada-Cuzcano, Christelle Etard, Clarisse Delvallée, Corinne Stoetzel, Elise Schaefer, Sophie Scheidecker, Véronique Geoffroy, Aline Schneider, Fouzia Studer, Francesca Mattioli, et al.

Novel *IQCE* variations confirm its role in postaxial polydactyly and cause ciliary defect phenotype in zebrafish.

Alejandro Estrada-Cuzcano[1]*, Christelle Etard[2]*, Clarisse Delvallée[1]*, Corinne Stoetzel[1], Elise Schaefer[1,3], Sophie Scheidecker[1,4], Véronique Geoffroy[1], Aline Schneider[1], Fouzia Studer[5], Francesca Mattioli[6], Kirsley Chennen[1,7], Sabine Sigaudy[8], Damien Plassard[9], Olivier Poch[7], Amélie Piton[4,6], Uwe Strahle[2], Jean Muller[1,4]* and Hélène Dollfus[1,3,5]*

1. Laboratoire de Génétique médicale, UMR_S INSERM U1112, IGMA, Faculté de Médecine FMTS, Université de Strasbourg, Strasbourg, France.
2. Karlsruhe Institute of Technology (KIT), Institute of Toxicology and Genetics (ITG), Eggenstein-Leopoldshafen, Germany.
3. Service de Génétique Médicale, Institut de Génétique Médicale d'Alsace, Hôpitaux Universitaires de Strasbourg, Strasbourg, France.
4. Laboratoires de Diagnostic Génétique, Hôpitaux Universitaires de Strasbourg, Strasbourg, France.
5. Centre de Référence pour les affections rares en génétique ophtalmologique, CARGO, Filière SENSGENE, Hôpitaux Universitaires de Strasbourg, 67091 Strasbourg, France.
6. Institut de Génétique et de Biologie Moléculaire et Cellulaire, 67400 Illkirch-Graffenstaden, France; INSERM U1258, 67400 Illkirch-Graffenstaden, France; CNRS UMR 7104, 67400 Illkirch-Graffenstaden, France; Université de Strasbourg, 67400 Illkirch, France.
7. Complex Systems and Translational Bioinformatics, ICube UMR 7357, Université de Strasbourg, Fédération de Médecine Translationnelle, Strasbourg, France.
8. Département de Génétique Médicale, Hôpital de la Timone, Marseille, France.
9. Plateforme GenomEast, IGBMC, Illkirch, France.

*equal contributor

Corresponding Author: Jean Muller/Hélène Dollfus

Email address: jeanmuller@unistra.fr / dollfus@unistra.fr

ABSTRACT

Polydactyly is one of the most frequent inherited defects of the limbs characterized by supernumerary digits and high genetic heterogeneity. Among the many genes involved, either in isolated or syndromic forms, 8 have been implicated in postaxial polydactyly (PAP). Among those *IQCE* has been recently identified in a single consanguineous family. Using whole-exome sequencing in patients with uncharacterized ciliopathies including PAP, we identified 3 families with biallelic pathogenic variations in *IQCE*. Interestingly, the c.895_904del (p.Val301Serfs*8) was found in all families without sharing a common haplotype, suggesting a recurrent mechanism. Moreover, in 2 families, the systemic phenotype could be explained by additional pathogenic variants in known genes (*TULP1, ATP6V1B1*). RNA expression analysis on patients' fibroblasts confirms that dysfunction of IQCE leads to dysregulation of genes associated with the hedgehog-signaling pathway and zebrafish experiments demonstrate a full spectrum of phenotypes linked to defective cilia: body curvature, kidney cysts, left right asymmetry, misdirected cilia in pronephric duct and retinal defects. In conclusion, we identified 3 additional families confirming *IQCE* as a non-syndromic PAP gene. Our data emphasize, the importance of taking into account the complete set of variations of each individual as each clinical presentation could finally be explained by multiple genes.

Keywords: IQCE, cilia, polydactyly, hedgehog signaling, zebrafish, RNA-seq

INTRODUCTION

Polydactyly is one of the most frequent inherited defects of the limb and has a prevalence of 0.3-3.6/1000 in live-births (Castilla et al., 1973; Malik, Ullah, Afzal, Lal, & Haque, 2014). It is characterized by supernumerary digits, ranging from extra additive soft tissue without bone structure to an integral and fully developed digit (Umair, Ahmad, Bilal, Ahmad, & Alfadhel, 2018). According to the position of the extra digit, polydactyly has been classified into postaxial (PAP), mesoaxial and preaxial (Malik, 2014). This defect is also associated with at least 221 syndromes according to the London Dysmorphology Database (Phadke & Sankar, 2010) covering 146 genes (Umair et al., 2018) and can be a hallmark for a group of syndromes such as for ciliopathies. For this group of diseases, in addition to skeletal malformations such as PAP, the phenotypic spectrum can associate retinal degeneration, obesity, kidney dysfunction and sometimes intellectual disability as observed in the classical Bardet-Biedl syndrome (BBS; MIM# 209900) for which 22 genes have been identified. In such heterogeneous diseases, high throughput genetic testing has made possible the diagnosis for many individuals (Shaheen et al., 2016) but also revealed more complex cases. Indeed, multi loci bearing pathogenic variations have been identified in several known disease related loci in single patients with apparent syndromic presentation (Karaca et al., 2018).

To date, eight genes *FAM92A1, GLI1, GLI3*, *IQCE*, *KIAA0825*, *MIPOL1, PITX1* and *ZNF141* (Ullah et al., 2019; Umair et al., 2018) have been implicated in non-syndromic polydactyly accounting for 286 disease causing variations according to HGMD (2018.1). Among those genes, *IQCE* (IQ domain-containing protein E) encodes a ciliary protein located at the base of the primary cilia and is linked to the Hedgehog (Hh) signaling pathway (Pusapati et al., 2014). Hh signaling is one of many major pathways that controls key steps of embryonic development (Ingham, Nakano, & Seger, 2011), tissue homeostasis and regeneration (Petrova & Joyner, 2014) and related to human cancers (Barakat, Humke, & Scott, 2010). Human hand and feet development is based on

a highly conserved pentadactyl pattern (Abbasi, 2011), the conserved Hh signaling pathway regulates the precise digital shape and pattern in the limbs. IQCE together with EFCAB7 have an important regulatory role in the EvC complex (EVC-EVC2), a positive tissue regulator of Hh signaling. Recently, a homozygous splice acceptor variation in *IQCE* has been described in a single consanguineous family with PAP (Umair et al., 2017).

In our study, we report three families with biallelic pathogenic variations in *IQCE* identified by whole exome sequencing. Interestingly, these families with PAP were initially recruited as syndromic ciliopathies and two have additional pathogenic variations in other genes explaining their apparent syndromic phenotype. Functional studies based on the patient's cells or zebrafish (*Danio rerio*) assays confirm the ciliary role of IQCE.

MATERIALS AND METHODS

Subjects

The patients were recruited by the Strasbourg University Hospital Medical Genetics Department as well as by the Center for Rare Genetic Ophthalmologic Diseases (CARGO, Hôpitaux Universitaires de Strasbourg, Strasbourg) because they were suspected to have a ciliopathy. Study protocols used in each cohort have been approved by the corresponding Institutional Review Board or equivalent committees (in Strasbourg, "Comité de Protection des Personnes" EST IV, N°DC-20142222), and written informed consent was given by 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.

Whole-exome sequencing

Whole exome sequencing was performed by IntegraGen (Evry, France) for family A, B and C in 2011, 2013 and 2018 respectively. After shearing of the genomic DNA, library preparation was performed using either the Agilent Human All Exon Kits v2 (family A and B) or the Twist Human Core Exome Enrichment System (Twist Bioscience) and IntegraGen Custom additions (family C). Sequencing was performed on an Illumina HiSeq 2000/4000 to generate 75-bp paired-end reads following the manufacturer's protocols. Image analysis and base calling were performed using Illumina Real Time Analysis (1.14/2.7.7) with default parameters.

Bioinformatics analysis

The reads were mapped to the reference human genome (GRCh37/hg19) using BWA v0.7.12 (Etard, Roostalu, & Strahle, 2010). GATK UG v3.4-46 was used to call SNV and indel variations (DePristo et al., 2011). Annotation and ranking of SNV and indel were performed by VaRank 1.4.3 (Geoffroy et al., 2015) in combination with the Alamut Batch software (1.11, Interactive Biosoftware, Rouen, France). Very stringent filtering criteria were applied to filter out non-pathogenic variants (Supp. Table 1): (i) variants represented with an allele frequency of more than 1% in public variation databases including the 1000Genomes (Altshuler et al., 2015), the gnomAD database (Lek et al., 2016), the DGV (MacDonald, Ziman, Yuen, Feuk, & Scherer, 2014) or our internal exome database, (ii) variants in 5′ and 3′ UTR, downstream, upstream, intronic and synonymous locations without pathogenic prediction of local splice effect. Structural variants were predicted using CANOES (Backenroth et al., 2014) and annotated by AnnotSV 2.0 (Geoffroy, Herenger, et al., 2018). Our analysis considered all inheritance modes but given that 2 families were consanguineous we focused on compound heterozygous and homozygous variants consistent with a recessive transmission. Variations have been classified according to the ACMG

guidelines (Richards et al., 2015) and only class 3 (likely pathogenic) and class 4 (pathogenic) have been considered.

Homozygosity mapping

Based on the WES data, homozygosity mapping has been performed using PLINK 1.9 (Chang et al., 2015). We followed the parameters described by Kancheva *et al* (Kancheva et al., 2016). The results are available in Supp. Table 2.

Cell culture

Skin fibroblasts from patient A.IV-1 and age-matched healthy control fibroblasts cells were grown in DMEM supplemented with 10% fetal calf serum (FCS) and 1% PSG (penicillin–streptomycin–glutamin). To induce primary cilium formation, the cells were deprived of serum by growth for 24h in DMEM with 1% PSG but only 0.1% FCS. In order to activate the Hh pathway, the cells were incubated in a DMEM medium with 1% PSG without serum (-FCS) and with SAG (100nM). SAG (Smoothened Agonist) has been resuspended in DMSO at a stock concentration of 4mM.

RNA extraction, cDNA synthesis, q-PCR and Taqman

For RNA seq validation, cells have been cultured in normal condition DMEM 1% PSG, 10% FCS. Total RNA was prepared from patient A.IV-1 fibroblast pellets (technical triplicate) and three controls pellets (biological triplicate) using QIAshredder Kit (Qiagen Ref 79654) and RNeasy Kit (Qiagen Ref 74104) followed by a DNAse treatment with the TURBO DNA-free™. Experiments have been repeated 3 times. For the Hh genes expression analysis (qRT-PCR), cells have been cultured in DMEM 1% PSG (-FCS) + SAG (100nM). Total RNA was prepared as described above and experiments have been repeated twice with 2 different control pellets (biological duplicate)

and one patient pellet (technical triplicate). RNA integrity was assessed by gel electrophoresis and RNA concentration by Eppendorf Biophotometer Plus with the Hellma® Tray Cell. Reverse transcription of 1 µg total RNA to cDNA was performed using the BioRad iScript™ cDNA synthesis kit. Real-time quantitative polymerase chain reaction amplification was performed in a BioRad CFX96 TM Real-Time System using the iQ™ SYBR® Green Supermix and primer sets (Supp. Table 3) optimized for tested targets for SYBR Green-based real-time PCR for the real-time PCR. The normalized fold expression of the target gene was calculated using the comparative cycle threshold (C_t) method by normalizing target mRNA C_t to those for both GAPDH and HPRT reference gene using the CFX Manager Software Version 1.5 and excel calculation. Statistical analysis have been done using the "Graphpad prism" software.

Immunocytochemistry and fluorescence microscopy

A.IV-1 and control ciliated fibroblasts were seeded on 12 mm diameter cover slips. Ciliogenesis has been induced by 24h incubation in a DMEM, 1% PSG medium, without serum (-FCS). Then, cells have been fixed with PFA-PBS 4%. Subsequently, the cells were stained with the ciliary markers anti-acetylated alpha tubulin (Lys40) mouse monoclonal primary antibody (1:200, # 32-2700 Invitrogen) and anti-EVC2 rabbit polyclonal primary antibody (1:250, # ab198930, Abcam). As secondary antibody, donkey anti-mouse Alexa 594 (1:1000, Invitrogen) and donkey anti-rabbit DyLight 488 (1:1000, #611-741-127, Rockland - Tebu-bio) were used. Ciliated cells count between A.IV-1 and control was performed manually using ~100 cells per genotype (exact Fisher test was applied). Counts of EVC2 colocalization at the base of the cilia was performed manually using also ~100 cells per genotype in triplicate (Student t-test has been applied). Pictures were taken with a Zeiss axio imager Z2 and the analysis with the ZEN2012 software (Carl Zeiss Inc., Oberkochen, Germany).

Transcriptome analysis

RNA samples were extracted from two fibroblast replicates of A.IV-1 (A.IV-1.1 and A.IV-1.2), and replicates from 6 individuals affected by other ciliopathies or intellectual disabilities of known molecular origin (ARN1.1, ARN1.2, ARN2.1, ARN2.2, ARN3.1, ARN3.2, ARN4.1, ARN4.2, ARN5.1, ARN6.1, ARN6.2) using TRI reagent® (Molecular Research Center) or using the RNeasy Mini Kit Qiagen®. Both protocols included an additional step of DNAse I recombinant treatment (Sigma-Aldrich®). The integrity of the RNA was visualized on a 1% bleach agarose gel by electrophoresis (Aranda et al., 2012). Quantification and further quality analysis were performed using the Nanodrop®. Samples should have a 260/280 ratios around 2 and a 260/230 ratio above 1.7. The integrity and quality of the RNA were also evaluated by running samples on a RNA 6000 Nano Chip on the Bioanalyzer (Agilent Technologies) and samples should have a RNA integrity number (RIN) equal or above 8. Library preparation was performed at the GenomEast platform at the Institute of Genetics and Molecular and Cellular Biology (Strasbourg, France), using the TruSeq® stranded mRNA sample preparation kit (Illumina) starting from 1 µg of extracted total RNA. Libraries were then 100+75bp paired-end sequenced, with 2 samples per lane on an Illumina Hiseq4000 sequencer. The corresponding bioinformatics pipeline is described in detail in the supplementary information.

Zebrafish stocks

Fish were bred and raised at 28.5°C as described previously (Westerfield, 1993). The AB wild-type line (European Zebrafish Resource Centre (EZRC), Karlsruhe) was used for all the experiments. Zebrafish husbandry and experimental procedures were performed in accordance with German animal protection regulations (Regierungspräsidium Karlsruhe, Germany, AZ35-

9185.81/G-137/10). Fish were crossed pairwise. For experiments, fertilized eggs were raised in 1× Instant Ocean salt solution (Aquarium Systems, Inc.) supplemented with 200 μM 1-phenyl 2-thiourea (PTU) to suppress melanogenesis. For visualization of *situs inversus*, we used a transgenic line expressing *GFP* under the control of the skeletal and cardiac muscle specific promoter *unc45b* (Roostalu & Strahle, 2012).

In situ hybridization

Whole-mount *in situ* hybridization was performed as previously described (Costa, Escaleira, Rodrigues, Manasfi, & Mermelstein, 2002; Oxtoby & Jowett, 1993). *Ptc1* probe was obtained from (Concordet et al., 1996). wt1a probe was obtained from (Armant et al., 2013).

Morpholinos and microinjections

A search of the zebrafish genome (GRCz11) revealed a single orthologue of the human *iqce* gene (RefSeq: NM_001287204.1) encoding a protein sharing 58% amino acid similarity with the human IQCE protein. Injections were performed as described previously (F. Muller et al., 1999). In brief, zebrafish eggs were collected shortly after being laid. Cleaned eggs were transferred to a petri dish with a minimal amount of water. Embryos were injected (FemtoJet; Eppendorf) through the chorion into the yolk at the one-cell stage. Injection needles were pulled from borosilicate glass capillary tubes with filament (Warner Instruments) using a micropipette puller (Sutter Instrument Co). Morpholinos were designed against the start codon of *iqce* (*iqce-mo*). Morpholino controls consist of the 5 bp mismatch *iqce-mo* (*iqce-mocont*). Morpholinos (Gene Tools, LLC) were injected at the following concentrations: *iqce-Mo*: CAAGTTCTCCAGCTACCACAGACAT (0.6 μM); *iqce-mocont*; CAAcTTgTCgAGCTACCAgA cACAT (0.6 μM). All dilutions were made in distilled water. Phenol red was added to the samples before injection (0.1% final concentration). *Iqce-gfp* plasmid was injected at the final concentration of 40 ng/μl. When injected at 0.6 μM, neither the anti *iqce*

nor the *iqce-mocont* (control) morpholino elicited necrosis, an unspecific effect of morpholino injections occasionally observed with some morpholinos (Robu et al., 2007). The efficiency of *iqce-mo* to block translation was assessed by co-injection of *iqce-mo* with a plasmid containing the *iqce* cDNA in-frame with *gfp* (*iqce-gfp*). As expected the embryos were depleted of GFP expression, indicating that the morpholinos indeed target efficiently *iqce*. In contrast, the 5-mismatch *iqce-mocont* was unable to abrogate the translation of the fusion gene (Supp. Figure 3A).

PCR amplification

Total RNA was isolated from 24 to 72 hpf (hours post fertilization) embryos using Tri-reagent (Invitrogen, Carlsbad, CA). RT-PCR was carried out following standard protocols. For *iqce in situ* hybridization probe, a fragment of 580 bp was amplified by PCR with the following primers *iqcewm5*: GCTGCATCATGTAGGACGATG *iqcewm3:* CACTCTGAAGCTGACTCAAG, and cloned into pgem-t-easy vector (Promega). *Iqce* full-length sequence was amplified from zebrafish cDNA with the following primers: *iqce5':* tcgGAATTCATGTCTGTGGTAGCTGGAGAAC; *iqce3':* agactcgagggTAAATCAAGTCA TCGGCACTG and cloned into *pcs2+GFP* with EcoRI-XhoI. For the rescue experiment *iqce5'modif:* tcgGAATTCATGTCGGTTGTAGCGGGCGAACTAG was used in combination with *iqce3'*, and cloned into *pCS2+GFP* Morpholino-resistant *iqce* mRNA was synthesized with the SP6kit (Promega).

Immunohistochemistry

Fish embryos were fixed in 4% paraformaldehyde/1x PBS one hour at room temperature. Embryos were permeabilized 7 min at -20°C with cold acetone, rinsed 5 times, 5 min with 1x PBS and blocked 4 h with BDP (1% DMSO, 5% BSA, 1x PBS). Incubation with primary antibody occurred

over night at 4°C, followed by thorough washing with PTW (1x PBS, 0.1% Tween20). Embryos were incubated with secondary antibody two hours at room temperature, and washed with PTW. Antibodies dilutions: Acetylated-tubulin (Sigma-Aldrich) 1/500; engrailed (Developmental Studies Hybridoma Bank, The University of Iowa): 1/50; goat anti-rabbit alexa 488 (Invitrogen): 1/1000, goat anti-mouse alexa 488 (Invitrogen): 1/1000. For nucleus staining, we used Draq5 at 1/1000 dilution. Staining was analyzed by confocal microscopy (TCS SP2, Leica). For sectioning of the eyes, embryos were fixed in 4% paraformaldehyde/PBS, progressively dehydrated in ethanol and stepwise transferred into Epon resin. Five µm sections were cut with a Leica microtome and stained with toluidine blue.

RESULTS

Clinical features

Three families with ciliopathy like phenotypes were referred to our laboratory as they remained unsolved following standard genetic testing. They all shared postaxial polydactyly (PAP) as a common feature but each presented other symptoms as described below (Figure 1A and Table 1). The affected patient of family A (A.IV-1) was born after an unremarkable pregnancy and is the only child of related Algerian parents. Bilateral polydactyly and brachydactyly of hands and feet were noted at birth associated with a syndactyly of the 2nd and 3rd digits of the feet. At the age of six years old, he was diagnosed with a decreased visual acuity and his ophthalmologic evaluation revealed retinal degeneration. He had hypogonadism and overweight was noted during his last clinical evaluation (15 years old). Family B is composed of three children presenting with PAP of hands and feet (bilateral for B.II-1 and unilateral for B.II-2 and B.II-3). The older one and the twin sister presented overweight respectively at the age of 7 and 5 years old in a context of obese parents. The ophthalmologic evaluation, including electroretinography, was normal for all the

siblings. In family C, two affected children were reported. Individual C.IV-3 is the last child of related Moroccan parents and presented with PAP of both hands and one foot, learning disabilities and overweight. Her two sisters and parents share also the two last features. Her cousin, patient C.IV-4, born also from related parents is more severe. He had also PAP of the four limbs associated with renal anomalies, deafness and speech delay. The observed kidney disease consisted of a tubular nephropathy with nephrocalcinosis diagnosed at one year old evocative of a distal renal tubular acidosis. Deafness was detected at the age of two years old. His development was in the normal range, except for a language delay, in the context of hearing loss, which improved with speech therapy (Table 1).

The association of PAP and the other clinical features of each family member was assumed to be compatible with a syndromic ciliopathy (Wolff et al., 2004) and in particular to the Bardet-Biedl syndrome (BBS; MIM# 209900). In line with the described phenotypes, the three families have been screened without success for the known BBS genes at that time (*BBS1-BBS18*) using either Sanger or targeted exome sequencing (J. Muller et al., 2010; Redin et al., 2012).

Identification of *IQCE* pathogenic variations

Whole exome sequencing (WES) was applied to all three families, generating on average a total of 81000 single nucleotide variants (SNVs), 11400 insertions-deletions (indels) and 21 structural variants (SV) (Supp. Table 1). In family A, as the proband (A.IV-1) was supposedly born from consanguineous parents, we identified 39 sizeable homozygous regions encompassing 238 Mb (Supp. Table 2); supporting the assumption of consanguinity. In this case, autosomal recessive inheritance is more likely and we thus focused on homozygous variants, reducing the number of variants to 7 in 6 genes (*DNAAF1*, *IQCE*, *MADD*, *SHANK2*, *TULP1* and *ZDHHC2*) (Supp. Table 1). A.IV-1 is affected by polydactyly, retinal dystrophy and hypogonadism, which made us,

concentrate on 2 genes. The first one is *IQCE* (NM_152558.5), a gene that was just described as associated to PAP (Umair et al., 2017), with a homozygous 10 bp deletion in exon 12 c.895_904del (p.Val301Serfs*8) (Figure 1B) located in a conspicuous homozygous region of 2.4 Mb. This variant is present in gnomAD at a very low combined allele frequency (0.0009534) with two homozygous carriers, which is compatible with the frequency of the disease. Subsequently, we analyzed *IQCE* exons 12 in 96 matched controls of Algerian ethnicity and could not find this variant (supplementary methods). Interestingly, the second gene is *TULP1* (NM_001289395.1) a gene reported in Leber congenital amaurosis or retinitis pigmentosa (Hagstrom, North, Nishina, Berson, & Dryja, 1998; Hanein et al., 2004) for which we identified a known homozygous pathogenic (class 5) variation, c.1198C>T (p.Arg400Trp), also located in a homozygous region of 8.5 Mb. This variant does not appear in public databases (gnomAD, 1000Genomes). The variants in both genes (*IQCE* and *TULP1*) segregated with the phenotypes within the family (Figure 1B) and in combination explained most of the clinical features of this individual. In the non-consanguineous family B (Figure 1A), WES revealed only 2 class-4 variations in *IQCE* at compound heterozygous state: c.895_904del (p.Val301Serfs*8) in exon 12 and c.1350_1353delAGAG, (p.Glu451Argfs*15) in exon 16 (Figure 1B). Both variations segregated in all affected members (B.II-1, B.II-2 and B.II-3). The second variation (c.1350_1353delAGAG) was found in gnomAD with a frequency of 0.0002381 with 66 heterozygous carriers. In family C, WES analysis of the most severe case, C.IV-4 (Figure 1A), revealed a known homozygous variant in *ATP6V1B1* (NM_001692.3), a gene that is associated with distal renal tubular acidosis with deafness (Vargas-Poussou et al., 2006) (dRTA; MIM# 267300). This class-4 variant, c.175-1G>C, is located in a canonical splice site and is predicted to cause the skipping of exon 3, this variation was not found in public databases (gnomAD, 1000Genomes). Interestingly, the c.895_904del (p.Val301Serfs*8) variation in *IQCE* that has been already observed in family A and B (respectively at the homozygous and compound heterozygous state) was only found in C.IV-4 at the heterozygous state (whereas no second pathogenic allele could be identified by cDNA sequencing and quantification, data not shown).

The WES did not reveal any other variations in known PAP genes for this individual. Given that patient C.IV-3 had only PAP and given the C.IV-4 variation in *IQCE*, we sequenced the *IQCE* gene in C.IV-3 and surprisingly found the c.895_904del (p.Val301Serfs*8) at the homozygous state (Figure 1).

Given the recurrence of the c.895_904del (p.Val301Serfs*8) variant in our cohort, we investigated a possible founder effect. Nevertheless, no common haplotype could be observed in the corresponding region in family A and B (Supp. Table 4). Analysis of the gnomAD record of this variation revealed its presence in different ethnical groups at a very low allele frequency: European (221/126532 alleles), Ashkenazi (12/10134 alleles), Latino (17/34392 alleles), African (6/23966 alleles), Finish (2/25788 alleles), south Asian (1/30782 alleles). The diverse origin of individuals with this *IQCE* variant and the differences in the haplotype suggested a recurrent mutational mechanism.

In order to better understand the effect of *IQCE* variants, we investigated the expression and subcellular localization of IQCE interacting partners in patient cells. *IQCE* is ubiquitously expressed in human adult tissues with higher levels of mRNA expression in lung, brain and hippocampus (Supp. Figure 1A). Endogenous IQCE is located at the base of the cilia and forms a complex with EFCAB7, to anchor the EVC-EVC2 sub-complex at the base of the cilia (EvC zone) for propagation of Hh signaling (Pusapati et al., 2014). We investigated the impact of IQCE deficiency in the formation of the primary cilia and localization of EVC-EVC2 sub-complex. Immunohistochemical studies on ciliated fibroblasts from individual A.IV-1 showed no differences in the number of ciliated cells (Supp. Figure 1B). Nevertheless, for 77% of patient's cells, we did not observe colocalization of EVC2 at the base of the cilia (Figure 2A and B), that presumably results in mislocalization of the EVC-EVC2 sub-complex and may compromise the Hh signaling.

Transcriptome analysis and Hh signaling effect

In order to get insight into the mechanistic impact of IQCE, we explored the transcriptome variation of individual A.IV-1 by RNA sequencing. We considered the possible impact of *TULP1* variation. However, under physiological conditions, *TULP1* is expressed exclusively in the retina, where it is involved in protein trafficking which is essential for the transport of rhodopsin (North, Naggert, Yan, Noben-Trauth, & Nishina, 1997). This was confirmed as no *TULP1* expression could be observed in the patient's fibroblasts. Given the *IQCE* genotype (frameshift) for A.IV-1 no *IQCE* expression could be detected. RNA differential expression analysis resulted in 98 significantly differentially expressed genes (Figure 2C and Supp. Table 5). None of the known IQCE interactors - *EVC, EVC2, EFCAB7* and *TCC23* (Breslow et al., 2018; Pusapati et al., 2018) - showed significant variations in their expression levels neither did the downstream effectors of the Hh signaling pathway *GLI1*, *GLI2* and *GLI3*. We identified among the most significantly up and downregulated genes *SLC35D3, SLC30A8* and *GSTT1* respectively. *SLC35D3* has been shown to be implicated in metabolic syndrome and obesity (Z. Zhang et al., 2014) and *SLC30A8* is associated with a protective role against diabetes type2 (Flannick et al., 2014).

Regarding *GSTT1,* the protein catalyzes the conjugation of glutathione to a wide range of potential toxins as the first step in detoxification and is identified as risk factor for open angle glaucoma or ventricular diastolic dysfunction (Singh, Kumar, Tewari, & Agarwal, 2018; Stamenkovic et al., 2018). In order to obtain functional clues about the IQCE protein modules, we applied functional enrichment analysis (WebGestalt 2019) (Wang, Vasaikar, Shi, Greer, & Zhang, 2017); the over-representation analysis identified 10 significant clusters involving tissue development, ossification and skeletal system development among others (Figure 2D). A number of genes associated with Hh signaling and limb formation are included in these protein clusters (ALPL, ALDH1A1, FGF9, HOXC10, IGF2, IGFBP1, MECOM, MLLT3, NDP, SFRP1 and SMOC1) (Supp. Table 6). Selected genes with significant | log2 fold-change | > 0.5 were tested compared to controls by RT-qPCR and we confirmed the expression change of *ALPL, GSTT1, IGF2, IQCE, MECOM, RSPO1* and *SLC30A8* but we did not detect significant differences in *FGF9* and *SFRP1* (Figure 2E).

In order to assess whether the Hh signaling is affected, we determined the expression levels of key genes in the patient's cells both under ciliary conditions and Hh activation (-FCS + SAG treatment). qRT-PCR revealed a significant under-expression of genes involved in the EVC-zone complex including *EFCAB7, EVC, EVC2* and also *GLI2* and *GLI3* (Supp. Figure 2).

Lack of *iqce* causes ciliary defects in zebrafish embryos

Considering the ciliary localization of the human IQCE and encouraged by the previous studies, where genetic screens in zebrafish identified numerous defects of subpopulations of cilia (Zhao & Malicki, 2007), we investigated expression and function of *iqce* in zebrafish (*Danio rerio).* We first assessed the *iqce* expression in the developing embryo. *In situ* hybridisation of *iqce* mRNA showed ubiquitous expression at the 10-somites and 20 hpf stages with higher expression in the pronephric ducts at 20 hpf. At 48 hpf, higher levels of *iqce* mRNA were found in the pronephric ducts and in the mid- and hindbrain (Supp. Figure 3B).

A morpholino directed against the *iqce* start codon (*iqce-mo*) and a morpholino with a 5 mismatch control sequence (*iqce-mocont)* were used to elucidate the function of *iqce* in zebrafish embryos. Embryos injected with *iqce-mo* (77%, n= 126) but not with the *iqce-mocont* control (2%, n=221) showed ventral curvature of the body axis (Figure 3A-B) which is often found in cilia mutants (Supp. Figure 3C). Occasionally we observed hydrocephalus (Figure 3B). In order to prove that the observed phenotype is due to the specific knock-down of *iqce*, we co-injected *iqce-mo* together with *iqce* mRNA resistant to *iqce-mo*. 54% of embryos exhibited a characteristic curved phenotype compared to 77% for *iqce-mo* alone (Figure 3D, n=193). We conclude from this experiment that the *iqce-mo* effectively abrogate the *iqce* expression.

Bilaterally dilated kidney cyst were present in 37% of *iqce-mo* morphants, but neither in uninjected nor in *iqce-mocont* injected embryos (Figure 3C and Supp. Figure 3D). These cysts are localized within the glomerulus as indicated by *wt1a* staining (Supp. Figure 3E).

We next examined the cilia of the pronephric duct. Embryos at 48 hpf were stained with acetylated-tubulin antibody. In *iqce-mocont* and uninjected embryos, cilia are aligned along the duct forming a thin line (Figure 3E, G). In contrast, the cilia organization in *iqce* morphants was severely perturbed showing cilia oriented in all directions (Figure 3F). The misorientation and shortening of pronephric duct cilia can be due to dilation of the tubules and suggest that pronephric fluid flow may be affected (Zhao & Malicki, 2007). We next tested whether other cilia are also affected, the dense ring of cilia on the epithelium surrounding the olfactory placode of 48-hpf wild-type larvae was strongly reduced in morphants (Figure 3H-J). In addition, the photoreceptor outer segment connects to the cell body via a narrow constriction that forms around a ciliary axoneme, and is known as the connecting cilium. Defects of photoreceptor cilia are associated with outer segment abnormalities and degeneration (Zhao & Malicki, 2007). Therefore, we analyzed eye sections of 4 day old morphants. In non-injected embryos or control morphants all the layers are properly formed (Figure 3K). In contrast in *iqce* morphants, the lamination is disrupted and the retinal layer altered with no visible photoreceptor outer segment (Figure 3L).

The perturbation of motile cilia often results in a problem of left-right asymmetry. We thus examined the percentage of *situs inversus* in *iqce* morphants. *iqce-mo* or *iqce-mocont* were injected into embryos carrying a transgene that expresses GFP in the heart (-1.8 unc45b-TFP, (Roostalu & Strahle, 2012). This allowed us to observe the looping of the heart. 44% of morphant embryos exhibited *situs inversus* (with the ventricle on the right and atrium on the left) or heterotaxia demonstrating that cilia from the Kupffer's vesicle, the fish equivalent of the mouse node, were affected by the *iqce* knock-down. In contrast only 3% and 6% of uninjected and *iqce-mocont* injected embryos, respectively, showed reverse heart jogging (Figure 3M).

In mutants with defective cilia (*iguana*) or ciliary motility (*lrrc50*), otoliths are frequently found ectopically, untethered or fused (Stooke-Vaughan, Huang, Hammond, Schier, & Whitfield, 2012). We thus examined the number of otoliths in *iqce* morphants at 2 dpf. At 48 dpf each ear should contains 2 otoliths and it is the case for 86 % of the morphant. However, for 14 % of them one of the two ears presented either one extra otolith (3 instead of 2) or only one (1 instead of 2) (data not shown, n= 262 embryos).

Cilia are necessary for normal activation as well as repression of the hedgehog (Hh) signaling pathway (Goetz & Anderson, 2010; Huang & Schier, 2009; Wolff et al., 2004). It was shown that in the absence of the ciliary protein BBS1 and under stimulation of Hh, the ciliary localization of Smoothened (*Smo*) was further enhanced compared to wild-type cells (Nozaki, Katoh, Kobayashi, & Nakayama, 2018). We thus examined the expression of the Hh receptor Patch1 (*Ptc1*) and the downstream gene engrailed (*eng*) in the developing somites in *iqce* morphants. At 10 somites stages, transcription of *ptc1* was expanded to the myotome, whereas it is exclusively restricted to the adaxial cells in control or uninjected embryos (Figure 3N-PO and Supp. Figure 3F). Similarly, the domain of *eng* expression was significantly expanded among the fast muscle fibres by 24 hpf, indicative of an increased number of Hh-dependent muscle fast fibers (Figure 3Q-S). These results indicate that the Hh pathway is perturbed by the loss of *iqce* function.

DISCUSSION

An *IQCE* homozygous variant affecting the splicing of the gene was recently identified by whole exome sequencing in a single Pakistani family presenting with postaxial polydactyly type A limited to lower limb (Umair et al., 2017). Here, we report 3 additional families with pathogenic variations in *IQCE* referred initially as ciliopathies with postaxial polydactyly involving the upper and lower limbs. Interestingly, the 3 families shared the same c.895_904del (p.Val301Serfs*8) variation in

IQCE exon 12 in a homozygous or in a heterozygous state with another truncating variant. These two *IQCE* variations have already been observed in the largest public databases. The c.1350_1353del (p.Glu451Argfs*15) was described in Exome Variant Server (EVS) (Tennessen et al., 2012) with an average allele frequency >1.5% with 90 homozygous calls. Initially, this overestimated allele frequency prevented its identification as candidate pathogenic variation. The higher false positive rate for indel calls was a known issue solved while integrating in ExAC/gnomAD (0.024% and no homozygous calls).

The index case A.IV-1 had a suspected ciliopathy-like phenotype, due to the presence of retinitis pigmentosa (RP), hypogonadism, polydactyly and speech delay. Although the proband was assumed as a patient with one syndrome, the occurrence of pathogenic variations in both *IQCE* and *TULP1* can explain independently the polydactyly and the RP. In the family C, the phenotypes of the patients were different; C.IV-3 has only PAP and overweight compared to her cousin who has PAP, deafness and renal anomalies. Phenotypic variations in a single family is not unusual in ciliopathies even with a single gene involved (Abdelhamed et al., 2013; Gonzalez-Del Pozo et al., 2014; Katsanis et al., 2001; Ware, Aygun, & Hildebrandt, 2011; W. Zhang et al., 2018). The variations in *IQCE,* can clearly explain the isolated PAP in C.IV-3. However, in the C.IV-4 individual, the association of deafness and nephropathy could finally be explained by the identification of variations in the *ATP6V1B1* gene (dRTA; MIM# 267300), known to be responsible for renal tubular acidosis associated with deafness. Nevertheless, the identification of only a single pathogenic allele in *IQCE* and the absence of any other variation in known polydactyly genes do not provide any explanation to date for the PAP for this patient. Although, gene dosage and RNA expression have been performed for this individual revealing still 50% of the *IQCE* expression level (data not shown), one cannot rule out missed variation(s) either in *IQCE* or in another gene. Whole genome sequencing might reveal intronic, intragenic or structural variations missed by the WES (Geoffroy, Stoetzel, et al., 2018). These two examples illustrate well how the incidental co-

occurrence of 2 different conditions can mimic the clinical features of a known Mendelian disease (DeLuca, Weed, Haas, Halder, & Stone, 2015; Karaca et al., 2018). This point out also the importance for searching for multiple hits in whole exome or whole genome sequencing data to explain some apparent syndromic phenotype such as ciliopathy-like phenotype. The occurrence of the other clinical manifestations, overweight or learning disability, in these families seemed to be due to other factors than the presence of neurosensory impairment (RP or deafness). However, we should take into consideration the appreciable variations of systemic features in individuals affected by pathogenic variants in different ciliopathy genes (Carmi, Elbedour, Stone, & Sheffield, 1995; Estrada-Cuzcano et al., 2012; Waters & Beales, 2011).

Vertebrate limb patterning is controlled by two signaling centers, the zone of polarizing activity (ZPA) and the distal posterior mesenchyme and the apical ectodermal ridge (AER) at the distal limb bud. The ZPA and AER produce signaling molecules -Hh and FGFs respectively- that control the formation of anteroposterior and proximodistal limb axes in a cooperative manner (Yamamoto et al., 2012). Additionally, Hh can induce a positive feedback Igf signaling during osteoblast differentiation (mostly Igf2), regulating in this way the long bone growth (Fisher, Meyer, Garber, & Dealy, 2005; Shi, Chen, Karner, & Long, 2015). Until recently, the combination of temporal and spatial exposure to long-range Hh signaling was accepted as the main molecular mechanism underlying digit patterning (Ahn & Joyner, 2004; Harfe et al., 2004). However, current investigations in chicken limb buds show that mesenchymal cells in the posterior limb bud produce actin-based filopodia that accumulate particles of Hh, allowing the Hh particles to act at a distance without extracellular diffusion (Sanders, Llagostera, & Barna, 2013). This study suggested that proper activation of the Hh target genes is more critical for digit patterning than diffusible long-range Hh morphogen threshold. The polydactyly phenotype of other carriers of *IQCE* variations supports the involvement of IQCE in the formation of limbs, process that is related to Hh signaling. At the base of the cilia, the EvC complex (constituted by EVC-EVC2 and EFCAB7-IQCE) is not essential for Hh target gene induction. However, it can reduce the Hh-stimulated induction of GLI1

or PTCH1 (Pusapati et al., 2014). Interestingly, gene ontology analysis of misregulated transcripts revealed enrichment for skeletal system development. Notably, IGF2 – a critical protein for bone formation (Fisher et al., 2005; Shi et al., 2015)- is misregulated in absence of IQCE. Defects in *IGF2* in humans cause Silver-Russell syndrome, ectrodactyly, under masculinized genitalia, developmental delay, and placental hypoplasia (Begemann et al., 2015; Yamoto et al., 2017). Additionally, we observed the up-regulation of *IGFBP1* and *IGFBP3*, these two proteins bind IGF2 and their excess may limit IGF2 function (Chao & D'Amore, 2008).

Knockdown of *iqce* in zebrafish showed the full spectrum of phenotypes associated with defective cilia: body curvature, kidney cysts, left right asymmetry, misdirected cilia in pronephric duct (Shi, Su, Lipschutz, & Lobo, 2017) as well as retinal defects resembling very much those observed in *bbs9* morphants (Veleri et al., 2012). In addition, full *iqce* knockout mouse showed short tibias, thin retinas, metabolic and behavioural alterations as well as polydactyly. Hh expression during limb development is pivotal for defining the identity and number of digits. When its expression is activated at an ectopic site in the limb bud this define a pathogenic mechanism that leads to preaxial polydactyly (Lettice et al., 2012). The *iqce* zebrafish model exhibits modified Hh signalling. Indeed, we observed a broadening of the Hh *ptc1* receptor outside of the adaxial cell, as well as extended *Engrailed* expression. It is well known that the primary cilium acts as a Hh signal transduction machine (Goetz, Ocbina, & Anderson, 2009). Recently it was shown that in a *bbs1* KO cell line, the entry of Smo into cilia upon Hh stimulation was increased compared to the wild type cell line (Nozaki et al., 2018). This in turn activated ectopic engrailed expression as observed in *iqce* zebrafish morphants. This result was unexpected since Pusapati *et al*. provided evidence by analysing osteoblasts and the NIH3T3 cell line that the EFCAB7-IQCE module anchors the EVC-EVC2 complex in a signalling microdomain at the base of the cilia. Moreover, they showed that EFCAB7 depletion triggered the mislocalization of EVC-EVC2 within cilia and impaired activation of the transcription factor GLI2. These data suggest an impairment of Hh signalling

rather than the gain of Hh signalling activity in the *iqce* mutant background that we observed in the zebrafish. However, in zebrafish, no Evc or Evc2 orthologues have been identified but only a paralog absent in human (Pusapati et al., 2014). Transcriptome analysis of the patient's cells under normal condition did not provide evidences of GLI genes modified expression compared to controls. However, under Hh activation, the patient's cells show a reduced expression of most of the key Hh pathway genes (*GLI2, GLI3, EFCAB7, EVC, EVC2*) (Supp. Figure 2). At the protein level, we observed a significant amount of EVC2 mislocalization at the base of the cilia (Figure 2B) as well as a reduced GLI2 protein expression (data not shown).

Blindness is not present in patients with only *iqce* defect. However, it has been previously observed in other cases that KO of genes in zebrafish can give additional phenotypes not observed in patients. For example, the zebrafish cc2d2 adult mutants exhibit a scoliosis that is not observed in Joubert syndrome (JBTS; MIM# 213300) (Bachmann-Gagescu et al., 2011). The defective retina of *iqce* morphants is very similar to the *bbs9* morphant model showing a lack of lamination (Veleri et al., 2012). This observation fits with the implication of *iqce* in ciliopathies.

In conclusion, our findings confirm the role of *IQCE* in non-syndromic postaxial polydactyly, and its negative impact on the Hh signaling pathway. The differences between *iqce* zebrafish and mice models compared to humans raise the question of additional genetic factors. This study illustrates how detailed and accurate clinical data are needed to interpret exome-scale genetic results and translate those results in correct counselling and therapeutic intervention.

AVAILABILITY OF DATA AND MATERIAL

Data generated or analysed 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 using the following range of accessions numbers SCV000929968 to SCV000929972 (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.

ACKNOWLEDGMENTS

We would like to thank the patient's family for their participation and the patient's associations, Formicoeur, Bardet-Biedl France and Unadev, for their constant and strong support. We also would like to acknowledge the members of the diagnostic laboratories (molecular and cytogenetic units, in particular Manuela Antin, Anne-Sophie Leuvrey and Elsa Nourisson) at the University Hospital in Strasbourg for technical assistance. We thank Nadine Borel and the zebrafish house team. We also thank our collaborators for screening their negative cohort (KM. Bujakowska, E. Pierce, P. Beales, F. Hildebrandt). We thank the GenomEast platform (Strasbourg, France) especially Serge Vicaire for library preparation, Céline Keime for supervising the transcriptome analysis as well as Bernard Jost and Christelle Thibault. The computing resources for this work were provided by the BICS and BISTRO bioinformatics platforms in Strasbourg.

FUNDING INFORMATION

Whole exome sequencing was funded by Retina France (100 exomes Program), by Hôpitaux Universitaires de Strasbourg (API 2013-2014 HUS N° 5885). RNASeq analysis performed in patients with *IQCE* pathogenic variant and other patients were funded by Fondation Maladie Rare and Fondation Jérome Lejeune. CE and US were supported by the HGF Biointerfaces Programme. AEC is supported by Fondation JED-Belgique. CD is supported by a FRM grant (ECO20170637509).

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

REFERENCES

Abbasi, A. A. (2011). Evolution of vertebrate appendicular structures: Insight from genetic and palaeontological data. *Dev Dyn, 240*(5), 1005-1016. doi:10.1002/dvdy.22572

Abdelhamed, Z. A., Wheway, G., Szymanska, K., Natarajan, S., Toomes, C., Inglehearn, C., & Johnson, C. A. (2013). Variable expressivity of ciliopathy neurological phenotypes that encompass Meckel-Gruber syndrome and Joubert syndrome is caused by complex de-regulated ciliogenesis, Shh and Wnt signalling defects. *Hum Mol Genet, 22*(7), 1358-1372. doi:10.1093/hmg/dds546

Ahn, S., & Joyner, A. L. (2004). Dynamic changes in the response of cells to positive hedgehog signaling during mouse limb patterning. *Cell, 118*(4), 505-516. doi:10.1016/j.cell.2004.07.023

Altshuler, D. M., Durbin, R. M., Abecasis, G. R., Bentley, D. R., Chakravarti, A., Clark, A. G., . . . Consortium, G. P. (2015). A global reference for human genetic variation. *Nature, 526*(7571), 68-+. doi:10.1038/nature15393

Armant, O., Marz, M., Schmidt, R., Ferg, M., Diotel, N., Ertzer, R., . . . Rastegar, S. (2013). Genome-wide, whole mount in situ analysis of transcriptional regulators in zebrafish embryos. *Dev Biol, 380*(2), 351-362. doi:10.1016/j.ydbio.2013.05.006

Bachmann-Gagescu, R., Phelps, I. G., Stearns, G., Link, B. A., Brockerhoff, S. E., Moens, C. B., & Doherty, D. (2011). The ciliopathy gene cc2d2a controls zebrafish photoreceptor outer segment development through a role in Rab8-dependent vesicle trafficking. *Hum Mol Genet, 20*(20), 4041-4055. doi:10.1093/hmg/ddr332

Backenroth, D., Homsy, J., Murillo, L. R., Glessner, J., Lin, E., Brueckner, M., . . . Shen, Y. F. (2014). CANOES: detecting rare copy number variants from whole exome sequencing data. *Nucleic Acids Research, 42*(12). doi:ARTN e97 10.1093/nar/gku345

Barakat, M. T., Humke, E. W., & Scott, M. P. (2010). Learning from Jekyll to control Hyde: Hedgehog signaling in development and cancer. *Trends Mol Med, 16*(8), 337-348. doi:10.1016/j.molmed.2010.05.003

Begemann, M., Zirn, B., Santen, G., Wirthgen, E., Soellner, L., Buttel, H. M., . . . Eggermann, T. (2015). Paternally Inherited IGF2 Mutation and Growth Restriction. *N Engl J Med, 373*(4), 349-356. doi:10.1056/NEJMoa1415227

Breslow, D. K., Hoogendoorn, S., Kopp, A. R., Morgens, D. W., Vu, B. K., Kennedy, M. C., . . . Nachury, M. V. (2018). A CRISPR-based screen for Hedgehog signaling provides insights into ciliary function and ciliopathies. *Nat Genet*. doi:10.1038/s41588-018-0054-7

Carmi, R., Elbedour, K., Stone, E. M., & Sheffield, V. C. (1995). Phenotypic differences among patients with Bardet-Biedl syndrome linked to three different chromosome loci. *Am J Med Genet, 59*(2), 199-203. doi:10.1002/ajmg.1320590216

Castilla, E., Paz, J., Mutchinick, O., Munoz, E., Giorgiutti, E., & Gelman, Z. (1973). Polydactyly: a genetic study in South America. *Am J Hum Genet, 25*(4), 405-412. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/4716659

Chang, C. C., Chow, C. C., Tellier, L. C., Vattikuti, S., Purcell, S. M., & Lee, J. J. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience, 4*, 7. doi:10.1186/s13742-015-0047-8

Chao, W., & D'Amore, P. A. (2008). IGF2: epigenetic regulation and role in development and disease. *Cytokine Growth Factor Rev, 19*(2), 111-120. doi:10.1016/j.cytogfr.2008.01.005

Concordet, J. P., Lewis, K. E., Moore, J. W., Goodrich, L. V., Johnson, R. L., Scott, M. P., & Ingham, P. W. (1996). Spatial regulation of a zebrafish patched homologue reflects the roles of sonic hedgehog and protein kinase A in neural tube and somite patterning. *Development, 122*(9), 2835-2846. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8787757

Costa, M. L., Escaleira, R. C., Rodrigues, V. B., Manasfi, M., & Mermelstein, C. S. (2002). Some distinctive features of zebrafish myogenesis based on unexpected distributions of the muscle cytoskeletal proteins actin, myosin, desmin, alpha-actinin, troponin and titin. *Mechanisms of Development, 116*(1-2), 95-104. doi:Pii S0925-4773(02)00149-1

Doi 10.1016/S0925-4773(02)00149-1

DeLuca, A. P., Weed, M. C., Haas, C. M., Halder, J. A., & Stone, E. M. (2015). Apparent Usher Syndrome Caused by the Combination of BBS1-Associated Retinitis Pigmentosa and SLC26A4-Associated Deafness. *JAMA Ophthalmol, 133*(8), 967-968. doi:10.1001/jamaophthalmol.2015.1463

DePristo, M. A., Banks, E., Poplin, R., Garimella, K. V., Maguire, J. R., Hartl, C., . . . Daly, M. J. (2011). A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet, 43*(5), 491-+. doi:10.1038/ng.806

Estrada-Cuzcano, A., Koenekoop, R. K., Senechal, A., De Baere, E. B., de Ravel, T., Banfi, S., . . . Klevering, B. J. (2012). BBS1 mutations in a wide spectrum of phenotypes ranging from nonsyndromic retinitis pigmentosa to Bardet-Biedl syndrome. *Arch Ophthalmol, 130*(11), 1425-1432. doi:10.1001/archophthalmol.2012.2434

Etard, C., Roostalu, U., & Strahle, U. (2010). Lack of Apobec2-related proteins causes a dystrophic muscle phenotype in zebrafish embryos. *J Cell Biol, 189*(3), 527-539. doi:10.1083/jcb.200912125

Fisher, M. C., Meyer, C., Garber, G., & Dealy, C. N. (2005). Role of IGFBP2, IGF-I and IGF-II in regulating long bone growth. *Bone, 37*(6), 741-750. doi:10.1016/j.bone.2005.07.024

Flannick, J., Thorleifsson, G., Beer, N. L., Jacobs, S. B., Grarup, N., Burtt, N. P., . . . Altshuler, D. (2014). Loss-of-function mutations in SLC30A8 protect against type 2 diabetes. *Nat Genet, 46*(4), 357-363. doi:10.1038/ng.2915

Geoffroy, V., Herenger, Y., Kress, A., Stoetzel, C., Piton, A., Dollfus, H., & Muller, J. (2018). AnnotSV: An integrated tool for Structural Variations annotation. *Bioinformatics*. doi:10.1093/bioinformatics/bty304

Geoffroy, V., Pizot, C., Redin, C., Piton, A., Vasli, N., Stoetzel, C., . . . Muller, J. (2015). VaRank: a simple and powerful tool for ranking genetic variants. *PeerJ, 3*, e796. doi:10.7717/peerj.796

Geoffroy, V., Stoetzel, C., Scheidecker, S., Schaefer, E., Perrault, I., Bar, S., . . . Muller, J. (2018). Whole-genome sequencing in patients with ciliopathies uncovers a novel recurrent tandem duplication in IFT140. *Hum Mutat, 39*(7), 983-992. doi:10.1002/humu.23539

Goetz, S. C., & Anderson, K. V. (2010). The primary cilium: a signalling centre during vertebrate development. *Nature Reviews Genetics, 11*(5), 331-344. doi:10.1038/nrg2774

Goetz, S. C., Ocbina, P. J. R., & Anderson, K. V. (2009). The Primary Cilium as a Hedgehog Signal Transduction Machine. *Primary Cilia, 94*, 199-+. doi:10.1016/S0091-679x(08)94010-3

Gonzalez-Del Pozo, M., Mendez-Vidal, C., Santoyo-Lopez, J., Vela-Boza, A., Bravo-Gil, N., Rueda, A., . . . Antinolo, G. (2014). Deciphering intrafamilial phenotypic variability by exome sequencing in a Bardet-Biedl family. *Mol Genet Genomic Med, 2*(2), 124-133. doi:10.1002/mgg3.50

Hagstrom, S. A., North, M. A., Nishina, P. L., Berson, E. L., & Dryja, T. P. (1998). Recessive mutations in the gene encoding the tubby-like protein TULP1 in patients with retinitis pigmentosa. *Nat Genet, 18*(2), 174-176. doi:10.1038/ng0298-174

Hanein, S., Perrault, I., Gerber, S., Tanguy, G., Barbet, F., Ducroq, D., . . . Kaplan, J. (2004). Leber congenital amaurosis: comprehensive survey of the genetic heterogeneity, refinement of the clinical definition, and genotype-phenotype correlations as a strategy for molecular diagnosis. *Hum Mutat, 23*(4), 306-317. doi:10.1002/humu.20010

Harfe, B. D., Scherz, P. J., Nissim, S., Tian, H., McMahon, A. P., & Tabin, C. J. (2004). Evidence for an expansion-based temporal Shh gradient in specifying vertebrate digit identities. *Cell, 118*(4), 517-528. doi:10.1016/j.cell.2004.07.024

Huang, P., & Schier, A. F. (2009). Dampened Hedgehog signaling but normal Wnt signaling in zebrafish without cilia. *Development, 136*(18), 3089-3098. doi:10.1242/dev.041343

Ingham, P. W., Nakano, Y., & Seger, C. (2011). Mechanisms and functions of Hedgehog signalling across the metazoa. *Nature Reviews Genetics, 12*(6), 393-406. doi:10.1038/nrg2984

Kancheva, D., Atkinson, D., De Rijk, P., Zimon, M., Chamova, T., Mitev, V., . . . Jordanova, A. (2016). Novel mutations in genes causing hereditary spastic paraplegia and Charcot-Marie-Tooth neuropathy identified by an optimized protocol for homozygosity mapping based on whole-exome sequencing. *Genet Med, 18*(6), 600-607. doi:10.1038/gim.2015.139

Karaca, E., Posey, J. E., Coban Akdemir, Z., Pehlivan, D., Harel, T., Jhangiani, S. N., . . . Lupski, J. R. (2018). Phenotypic expansion illuminates multilocus pathogenic variation. *Genet Med*. doi:10.1038/gim.2018.33

Katsanis, N., Ansley, S. J., Badano, J. L., Eichers, E. R., Lewis, R. A., Hoskins, B. E., . . . Lupski, J. R. (2001). Triallelic inheritance in Bardet-Biedl syndrome, a Mendelian recessive disorder. *Science, 293*(5538), 2256-2259. doi:10.1126/science.1063525

Lek, M., Karczewski, K. J., Minikel, E. V., Samocha, K. E., Banks, E., Fennell, T., . . . Consortium, E. A. (2016). Analysis of protein-coding genetic variation in 60,706 humans. *Nature, 536*(7616), 285-+. doi:10.1038/nature19057

Lettice, L. A., Williamson, I., Wiltshire, J. H., Peluso, S., Devenney, P. S., Hill, A. E., . . . Hill, R. E. (2012). Opposing functions of the ETS factor family define Shh spatial expression in limb buds and underlie polydactyly. *Dev Cell, 22*(2), 459-467. doi:10.1016/j.devcel.2011.12.010

MacDonald, J. R., Ziman, R., Yuen, R. K. C., Feuk, L., & Scherer, S. W. (2014). The Database of Genomic Variants: a curated collection of structural variation in the human genome. *Nucleic Acids Research, 42*(D1), D986-D992. doi:10.1093/nar/gkt958

Malik, S. (2014). Polydactyly: phenotypes, genetics and classification. *Clin Genet, 85*(3), 203-212. doi:10.1111/cge.12276

Malik, S., Ullah, S., Afzal, M., Lal, K., & Haque, S. (2014). Clinical and descriptive genetic study of polydactyly: a Pakistani experience of 313 cases. *Clin Genet, 85*(5), 482-486. doi:10.1111/cge.12217

Metsalu, T., & Vilo, J. (2015). ClustVis: a web tool for visualizing clustering of multivariate data using Principal Component Analysis and heatmap. *Nucleic Acids Research, 43*(W1), W566-570. doi:10.1093/nar/gkv468

Muller, F., Blader, P., Rastegar, S., Fischer, N., Knochel, W., & Strahle, U. (1999). Characterization of zebrafish smad1, smad2 and smad5: the amino-terminus of Smad1 and Smad5 is required for specific function in the embryo. *Mechanisms of Development, 88*(1), 73-88. doi:Doi 10.1016/S0925-4773(99)00173-2

Muller, J., Stoetzel, C., Vincent, M. C., Leitch, C. C., Laurier, V., Danse, J. M., . . . Dollfus, H. (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), 583-593. doi:10.1007/s00439-010-0804-9

North, M. A., Naggert, J. K., Yan, Y., Noben-Trauth, K., & Nishina, P. M. (1997). Molecular characterization of TUB, TULP1, and TULP2, members of the novel tubby gene family and their possible relation to ocular diseases. *Proc Natl Acad Sci U S A, 94*(7), 3128-3133. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9096357

Nozaki, S., Katoh, Y., Kobayashi, T., & Nakayama, K. (2018). BBS1 is involved in retrograde trafficking of ciliary GPCRs in the context of the BBSome complex. *PLoS One, 13*(3), e0195005. doi:10.1371/journal.pone.0195005

Oxtoby, E., & Jowett, T. (1993). Cloning of the Zebrafish Krox-20 Gene (Krx-20) and Its Expression during Hindbrain Development. *Nucleic Acids Research, 21*(5), 1087-1095. doi:DOI 10.1093/nar/21.5.1087

Petrova, R., & Joyner, A. L. (2014). Roles for Hedgehog signaling in adult organ homeostasis and repair. *Development, 141*(18), 3445-3457. doi:10.1242/dev.083691

Phadke, S. R., & Sankar, V. H. (2010). Polydactyly and genes. *Indian J Pediatr, 77*(3), 277-281. doi:10.1007/s12098-010-0033-1

Pusapati, G. V., Hughes, C. E., Dorn, K. V., Zhang, D., Sugianto, P., Aravind, L., & Rohatgi, R. (2014). EFCAB7 and IQCE regulate hedgehog signaling by tethering the EVC-EVC2 complex to the base of primary cilia. *Dev Cell, 28*(5), 483-496. doi:10.1016/j.devcel.2014.01.021

Pusapati, G. V., Kong, J. H., Patel, B. B., Krishnan, A., Sagner, A., Kinnebrew, M., . . . Rohatgi, R. (2018). CRISPR Screens Uncover Genes that Regulate Target Cell Sensitivity to the Morphogen Sonic Hedgehog. *Dev Cell, 44*(2), 271. doi:10.1016/j.devcel.2018.01.002

Redin, C., Le Gras, S., Mhamdi, O., Geoffroy, V., Stoetzel, C., Vincent, M. C., . . . Muller, J. (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), 502-512. doi:10.1136/jmedgenet-2012-100875

Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., . . . Committee, A. L. Q. A. (2015). Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med, 17*(5), 405-424. doi:10.1038/gim.2015.30

Robu, M. E., Larson, J. D., Nasevicius, A., Beiraghi, S., Brenner, C., Farber, S. A., & Ekker, S. C. (2007). p53 activation by knockdown technologies. *PLoS Genet, 3*(5), e78. doi:10.1371/journal.pgen.0030078

Roostalu, U., & Strahle, U. (2012). In Vivo Imaging of Molecular Interactions at Damaged Sarcolemma. *Dev Cell, 22*(3), 515-529. doi:10.1016/j.devcel.2011.12.008

Sanders, T. A., Llagostera, E., & Barna, M. (2013). Specialized filopodia direct long-range transport of SHH during vertebrate tissue patterning. *Nature, 497*(7451), 628-632. doi:10.1038/nature12157

Shaheen, R., Szymanska, K., Basu, B., Patel, N., Ewida, N., Faqeih, E., . . . Alkuraya, F. S. (2016). Characterizing the morbid genome of ciliopathies. *Genome Biol, 17*(1), 242. doi:10.1186/s13059-016-1099-5

Shi, Y., Chen, J., Karner, C. M., & Long, F. (2015). Hedgehog signaling activates a positive feedback mechanism involving insulin-like growth factors to induce osteoblast differentiation. *Proc Natl Acad Sci U S A, 112*(15), 4678-4683. doi:10.1073/pnas.1502301112

Shi, Y., Su, Y., Lipschutz, J. H., & Lobo, G. P. (2017). Zebrafish as models to study ciliopathies of the eye and kidney. *Clin Nephrol Res, 1*(1), 6-9. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/29553143

Singh, M. M., Kumar, R., Tewari, S., & Agarwal, S. (2018). Association of GSTT1/GSTM1 and ApoE variants with left ventricular diastolic dysfunction in thalassaemia major patients. *Hematology*, 1-6. doi:10.1080/10245332.2018.1502397

Stamenkovic, M., Lukic, V., Suvakov, S., Simic, T., Sencanic, I., Pljesa-Ercegovac, M., . . . Djukic, T. (2018). GSTM1-null and GSTT1-active genotypes as risk determinants of primary open angle glaucoma among smokers. *Int J Ophthalmol, 11*(9), 1514-1520. doi:10.18240/ijo.2018.09.14

Stooke-Vaughan, G. A., Huang, P., Hammond, K. L., Schier, A. F., & Whitfield, T. T. (2012). The role of hair cells, cilia and ciliary motility in otolith formation in the zebrafish otic vesicle. *Development, 139*(10), 1777-1787. doi:10.1242/dev.079947

Tennessen, J. A., Bigham, A. W., O'Connor, T. D., Fu, W., Kenny, E. E., Gravel, S., . . . Project, N. E. S. (2012). Evolution and functional impact of rare coding variation from deep sequencing of human exomes. *Science, 337*(6090), 64-69. doi:10.1126/science.1219240

Ullah, I., Kakar, N., Schrauwen, I., Hussain, S., Chakchouk, I., Liaqat, K., . . . Leal, S. M. (2019). Variants in KIAA0825 underlie autosomal recessive postaxial polydactyly. *Hum Genet*. doi:10.1007/s00439-019-02000-0

Umair, M., Ahmad, F., Bilal, M., Ahmad, W., & Alfadhel, M. (2018). Clinical Genetics of Polydactyly: An Updated Review. *Front Genet, 9*, 447. doi:10.3389/fgene.2018.00447

Umair, M., Shah, K., Alhaddad, B., Haack, T. B., Graf, E., Strom, T. M., . . . Ahmad, W. (2017). Exome sequencing revealed a splice site variant in the IQCE gene underlying post-axial polydactyly type A restricted to lower limb. *Eur J Hum Genet, 25*(8), 960-965. doi:10.1038/ejhg.2017.83

Vargas-Poussou, R., Houillier, P., Le Pottier, N., Strompf, L., Loirat, C., Baudouin, V., . . . Blanchard, A. (2006). Genetic investigation of autosomal recessive distal renal tubular acidosis: evidence for early sensorineural hearing loss associated with mutations in the ATP6V0A4 gene. *J Am Soc Nephrol, 17*(5), 1437-1443. doi:10.1681/ASN.2005121305

Veleri, S., Bishop, K., Nogare, D. E. D., English, M. A., Foskett, T. J., Chitnis, A., . . . Swaroop, A. (2012). Knockdown of Bardet-Biedl Syndrome Gene BBS9/PTHB1 Leads to Cilia Defects. *PLoS One, 7*(3). doi:ARTN e34389 10.1371/journal.pone.0034389

Wang, J., Vasaikar, S., Shi, Z., Greer, M., & Zhang, B. (2017). WebGestalt 2017: a more comprehensive, powerful, flexible and interactive gene set enrichment analysis toolkit. *Nucleic Acids Research, 45*(W1), W130-W137. doi:10.1093/nar/gkx356

Ware, S. M., Aygun, M. G., & Hildebrandt, F. (2011). Spectrum of clinical diseases caused by disorders of primary cilia. *Proc Am Thorac Soc, 8*(5), 444-450. doi:10.1513/pats.201103-025SD

Waters, A. M., & Beales, P. L. (2011). Ciliopathies: an expanding disease spectrum. *Pediatr Nephrol, 26*(7), 1039-1056. doi:10.1007/s00467-010-1731-7

Westerfield, M. (1993). The zebrafish book: a guide for the laboratory use of zebrafish (Brachydanio reriro). 300.

Wolff, C., Roy, S., Lewis, K. E., Schauerte, H., Joerg-Rauch, G., Kirn, A., . . . Ingham, P. W. (2004). iguana encodes a novel zinc-finger protein with coiled-coil domains essential for Hedgehog signal transduction in the zebrafish embryo. *Genes & Development, 18*(13), 1565-1576. doi:10.1101/gad.296004

Yamamoto, M., Matsuzaki, T., Takahashi, R., Adachi, E., Maeda, Y., Yamaguchi, S., . . . Noda, M. (2012). The transformation suppressor gene Reck is required for postaxial patterning in mouse forelimbs. *Biol Open, 1*(5), 458-466. doi:10.1242/bio.2012638

Yamoto, K., Saitsu, H., Nakagawa, N., Nakajima, H., Hasegawa, T., Fujisawa, Y., . . . Ogata, T. (2017). De novo IGF2 mutation on the paternal allele in a patient with Silver-Russell syndrome and ectrodactyly. *Hum Mutat, 38*(8), 953-958. doi:10.1002/humu.23253

Zhang, W., Taylor, S. P., Ennis, H. A., Forlenza, K. N., Duran, I., Li, B., . . . Cohn, D. H. (2018). Expanding the genetic architecture and phenotypic spectrum in the skeletal ciliopathies. *Hum Mutat, 39*(1), 152-166. doi:10.1002/humu.23362

Zhang, Z., Hao, C. J., Li, C. G., Zang, D. J., Zhao, J., Li, X. N., . . . Li, W. (2014). Mutation of SLC35D3 causes metabolic syndrome by impairing dopamine signaling in striatal D1 neurons. *PLoS Genet, 10*(2), e1004124. doi:10.1371/journal.pgen.1004124

Zhao, C. T., & Malicki, J. (2007). Genetic defects of pronephric cilia in zebrafish. *Mechanisms of Development, 124*(7-8), 605-616. doi:10.1016/j.mod.2007.04.004

	Family A		Family B		Family C	
	A.IV-1	B.II-1	B.II-2	B.II-3	C.IV-3	C.IV-4
IQCE	p.[Val301Serfs*8];[Val301Serfs*8]	p.[Val301Serfs*8];[Glu451fs*15]	p.[Val301Serfs*8];[Glu451fs*15]	p.[Val301Serfs*8];[Glu451fs*15]	p.[Val301Serfs*8];[Val301Serfs*8]	p.[Val301Serfs*8];[=]
Other gene	*TULP1* p.[Arg400Trp];[Arg400Trp]					*ATP6V1B1* c.[175-1G>C];[175-1G>C]
Clinical features						
Hands and feet abnormalities						
Postaxial polydactyly	Y	Y	Y	Y	Y	Y
Hands	Y	Y	Y*	Y*	Y	Y
Feet	Y	Y	Y*	Y*	Y	Y
Brachydactyly	Y	Y	Y	Y	N	N
Syndactyly	Y	N	N	Y	N	N
Neurosensory disease						
Retinal dystrophy	Y	N	N	N	N	N
Deafness	N	N	N	N	N	Y
Pondostatural evaluation						
Overweight	Y	Y	Y	N	Y	N
Development psychomotor						
Intellectual abnormalities	N	N	N	N	N	N
Learning disabilities	N	N	N	N	Y	N
Speech delay	Y	N	N	N	Y	Y
Ataxia	N	N	N	N	N	N
Kidney function						
Renal abnormality	N	N	N	N	N	Tubulopathy - Nephrocalcinosis
Genital abnormalities						
Hypogonadism	Micropenis	N			N	N

Table 1. Clinical characteristic of individuals with variants in *IQCE.* Y=yes, N= not present, *= unilateral

Figure 1. Clinical description and molecular genetic characterization of *IQCE* variations. **(A)** Graphic representation of three autosomal recessive families with complicated polydactyly in whom *IQCE, TULP1* and *ATP6V1B1* variations were identified. Sanger sequencing traces confirming the presence of the pathogenic variant and segregation in available family members is demonstrated. The "WT" symbol represents wild type and "M" symbol represents pathogenic variant. Blackened symbols indicate affected individuals. **(B)** Genomic structure (NC_000007.14) and schematic representation of the protein layout (NP_689771.3) of IQCE and location of currently known disease-associated variant (in bold, novel pathogenic variants identified in this study). Topology of IQCE represents one hydrophobic segment (HS), coil-coil conserved domain regions, three IQ calmodulin binding motifs (IQ) and one acid region (AR).

Figure 2. (A). Immunostaining of A.IV-1 and control ciliated fibroblasts, antibodies against human EVC2 (in green) and the ciliary marker anti acetylated alpha tubulin (in red), revealed that in absence of endogenous IQCE, EVC2 fails to localize at the base of the cilia. **(B).** Count of mislocalized EVC2 at the base of the cilia in fibroblasts under ciliated conditions (–FCS) between A.IV-1 and control (3 times ~100 cells per condition). Statistical significance was determined using the Student's t-test, NS, non-significant, p=0.0004 (***). **(C).** Heat map generated using ClustVis (Metsalu & Vilo, 2015) from the transcriptome sequencing data reflecting gene counts of the significantly differentially expressed genes across patient and control samples. **(D).** Volcano plot of overrepresentation enrichment analysis of significantly misregulated transcripts in A.IV-1 fibroblast compared to wild type. These results were obtained using WebGestalt 2019, (Supp. Table 6). **(E).** Real-time quantitative polymerase chain reaction analysis from 3 technical replicates of A.IV-1 and 3 controls fibroblasts showed relative expression levels (normalized to GAPDH) of 9 significantly differentially expressed genes, data includes ±SD. NS= no significative, *= p-value < 0.05, **= p-value < 0.01, ***= p-value < 0.001 (alpha risk 5%). Scale bar 5 µm.

Figure 3. *iqce* morphants present ciliopathic phenotypes. **(A-B).** *iqce* morphants exhibit a curved body axis A: uninjected B: *iqce-mo* morphants, star indicates hydrocephalus. **(C).** Cysts in the paired pronephric glomeruli in *iqce* morphants (arrows). **(D).** Rescue of *iqce-mo* (6 µM) injected embryos with different concentrations of *iqce* mRNA. 10 ng/µl (n=121), 20 ng/µl (n=163), 30 ng/µl (n=193). The rescue was evaluated by the loss of curved body, data includes ± SD only when

needed (WT and the best condition (i.e. 30 ng/µl)). **(E-G).** Cilia in the pronephric duct of *iqce* morphants are misdirected, shortened, and malformed (F) whereas in uninjected (E) or control morpholinos (G) they are elongated. bar: 11.6 µm. **(H-J).** Cilia in nasal placode are reduced in number and size in morphants (I, n=5) compared to uninjected (H, n=5) and morpholino control (J, n=5). bar: 11.6 µm. Scheme: Square shows the area of observation for pronephric cilia, and circle for nasal placode cilia. **(K-L).** 5 µM sections through eyes of uninjected or *iqce-mo* injected embryos stained with toluidine blue (4 dpf). The uninjected embryos show normal lamination whereas morphants lack a properly formed retina. PRs: photoreceptors; ONL: outer nuclear layer; INL: inner nuclear layer; GCL: ganglion cell layer. Star indicates the outer segments of photoreceptors. Scale bar: 0.2 cm. **(M).** Chart indicating the percentage of normal and *situs inversus* or heterotaxia in uninjected (n=70), *iqce-mo* (n=250) or *iqce-mocont* (n=160) injected embryos, data includes ± SD. **(N-P)**. *patched1* (*ptc1*) *in situ* hybridization revealed expanded expression in *iqce* morphant (O, n=81) but not in uninjected (N, n=88) or control morpholino (P, n=30) injected embryos. **(Q-S)**. *In situ* immunohistochemistry showing expanded engrailed expression in *iqce* morphants (R, n=10) compared to uninjected (Q, n=10) and *iqce-mocont* (S, n=10) injected embryos. Scale bars: 47.2 µM. A-J: 48 hpf. K-M: 72 hpf. N-P: 10 somites. Q-S: 24 hpf.

Novel *IQCE* variations confirm its role in postaxial polydactyly and cause ciliary defect phenotype in zebrafish.

Alejandro Estrada-Cuzcano[1*], Christelle Etard[2*], Clarisse Delvallée[1*], Corinne Stoetzel[1], Elise Schaefer[1,3], Sophie Scheidecker[1,4], Véronique Geoffroy[1], Aline Schneider[1], Fouzia Studer[5], Francesca Mattioli[6], Kirsley Chennen[1,7], Sabine Sigaudy[8], Damien Plassard[9], Olivier Poch[7], Amélie Piton[4,6], Uwe Strahle[2], Jean Muller[1,4*] and Hélène Dollfus[1,3,5*]

1. Laboratoire de Génétique médicale, UMR_S INSERM U1112, IGMA, Faculté de Médecine FMTS, Université de Strasbourg, Strasbourg, France.
2. Karlsruhe Institute of Technology (KIT), Institute of Toxicology and Genetics (ITG), Eggenstein-Leopoldshafen, Germany.
3. Service de Génétique Médicale, Institut de Génétique Médicale d'Alsace, Hôpitaux Universitaires de Strasbourg, Strasbourg, France.
4. Laboratoires de Diagnostic Génétique, Hôpitaux Universitaires de Strasbourg, Strasbourg, France.
5. Centre de Référence pour les affections rares en génétique ophtalmologique, CARGO, Filière SENSGENE, Hôpitaux Universitaires de Strasbourg, 67091 Strasbourg, France.
6. Institut de Génétique et de Biologie Moléculaire et Cellulaire, 67400 Illkirch-Graffenstaden, France; INSERM U1258, 67400 Illkirch-Graffenstaden, France; CNRS UMR 7104, 67400 Illkirch-Graffenstaden, France; Université de Strasbourg, 67400 Illkirch, France.
7. Complex Systems and Translational Bioinformatics, ICube UMR 7357, Université de Strasbourg, Fédération de Médecine Translationnelle, Strasbourg, France.
8. Département de Génétique Médicale, Hôpital de la Timone, Marseille, France.
9. Plateforme GenomEast, IGBMC, Illkirch, France.

*equal contributor

Corresponding Author: Jean Muller/Hélène Dollfus

Email address: jeanmuller@unistra.fr / dollfus@unistra.fr

ABSTRACT

Polydactyly is one of the most frequent inherited defects of the limbs characterized by supernumerary digits and high genetic heterogeneity. Among the many genes involved, either in isolated or syndromic forms, 8 have been implicated in postaxial polydactyly (PAP). Among those *IQCE* has been recently identified in a single consanguineous family. Using whole-exome sequencing in patients with uncharacterized ciliopathies including PAP, we identified 3 families with biallelic pathogenic variations in *IQCE*. Interestingly, the c.895_904del (p.Val301Serfs*8) was found in all families without sharing a common haplotype, suggesting a recurrent mechanism. Moreover, in 2 families, the systemic phenotype could be explained by additional pathogenic variants in known genes (*TULP1, ATP6V1B1*). RNA expression analysis on patients' fibroblasts confirms that dysfunction of IQCE leads to dysregulation of genes associated with the hedgehog-signaling pathway and zebrafish experiments demonstrate a full spectrum of phenotypes linked to defective cilia: body curvature, kidney cysts, left right asymmetry, misdirected cilia in pronephric duct and retinal defects. In conclusion, we identified 3 additional families confirming *IQCE* as a non-syndromic PAP gene. Our data emphasize, the importance of taking into account the complete set of variations of each individual as each clinical presentation could finally be explained by multiple genes.

Keywords: IQCE, cilia, polydactyly, hedgehog signaling, zebrafish, RNA-seq

INTRODUCTION

Polydactyly is one of the most frequent inherited defects of the limb and has a prevalence of 0.3-3.6/1000 in live-births (Castilla et al., 1973; Malik, Ullah, Afzal, Lal, & Haque, 2014). It is characterized by supernumerary digits, ranging from extra additive soft tissue without bone structure to an integral and fully developed digit (Umair, Ahmad, Bilal, Ahmad, & Alfadhel, 2018). According to the position of the extra digit, polydactyly has been classified into postaxial (PAP), mesoaxial and preaxial (Malik, 2014). This defect is also associated with at least 221 syndromes according to the London Dysmorphology Database (Phadke & Sankar, 2010) covering 146 genes (Umair et al., 2018) and can be a hallmark for a group of syndromes such as for ciliopathies. For this group of diseases, in addition to skeletal malformations such as PAP, the phenotypic spectrum can associate retinal degeneration, obesity, kidney dysfunction and sometimes intellectual disability as observed in the classical Bardet-Biedl syndrome (BBS; MIM# 209900) for which 22 genes have been identified. In such heterogeneous diseases, high throughput genetic testing has made possible the diagnosis for many individuals (Shaheen et al., 2016) but also revealed more complex cases. Indeed, multi loci bearing pathogenic variations have been identified in several known disease related loci in single patients with apparent syndromic presentation (Karaca et al., 2018).

To date, eight genes *FAM92A1, GLI1, GLI3*, *IQCE*, *KIAA0825*, *MIPOL1, PITX1* and *ZNF141* (Ullah et al., 2019; Umair et al., 2018) have been implicated in non-syndromic polydactyly accounting for 286 disease causing variations according to HGMD (2018.1). Among those genes, *IQCE* (IQ domain-containing protein E) encodes a ciliary protein located at the base of the primary cilia and is linked to the Hedgehog (Hh) signaling pathway (Pusapati et al., 2014). Hh signaling is one of many major pathways that controls key steps of embryonic development (Ingham, Nakano, & Seger, 2011), tissue homeostasis and regeneration (Petrova & Joyner, 2014) and related to human cancers (Barakat, Humke, & Scott, 2010). Human hand and feet development is based on

a highly conserved pentadactyl pattern (Abbasi, 2011), the conserved Hh signaling pathway regulates the precise digital shape and pattern in the limbs. IQCE together with EFCAB7 have an important regulatory role in the EvC complex (EVC-EVC2), a positive tissue regulator of Hh signaling. Recently, a homozygous splice acceptor variation in *IQCE* has been described in a single consanguineous family with PAP (Umair et al., 2017).

In our study, we report three families with biallelic pathogenic variations in *IQCE* identified by whole exome sequencing. Interestingly, these families with PAP were initially recruited as syndromic ciliopathies and two have additional pathogenic variations in other genes explaining their apparent syndromic phenotype. Functional studies based on the patient's cells or zebrafish (*Danio rerio*) assays confirm the ciliary role of IQCE.

MATERIALS AND METHODS

Subjects

The patients were recruited by the Strasbourg University Hospital Medical Genetics Department as well as by the Center for Rare Genetic Ophthalmologic Diseases (CARGO, Hôpitaux Universitaires de Strasbourg, Strasbourg) because they were suspected to have a ciliopathy. Study protocols used in each cohort have been approved by the corresponding Institutional Review Board or equivalent committees (in Strasbourg, "Comité de Protection des Personnes" EST IV, N°DC-20142222), and written informed consent was given by 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.

Whole-exome sequencing

Whole exome sequencing was performed by IntegraGen (Evry, France) for family A, B and C in 2011, 2013 and 2018 respectively. After shearing of the genomic DNA, library preparation was performed using either the Agilent Human All Exon Kits v2 (family A and B) or the Twist Human Core Exome Enrichment System (Twist Bioscience) and IntegraGen Custom additions (family C). Sequencing was performed on an Illumina HiSeq 2000/4000 to generate 75-bp paired-end reads following the manufacturer's protocols. Image analysis and base calling were performed using Illumina Real Time Analysis (1.14/2.7.7) with default parameters.

Bioinformatics analysis

The reads were mapped to the reference human genome (GRCh37/hg19) using BWA v0.7.12 (Etard, Roostalu, & Strahle, 2010). GATK UG v3.4-46 was used to call SNV and indel variations (DePristo et al., 2011). Annotation and ranking of SNV and indel were performed by VaRank 1.4.3 (Geoffroy et al., 2015) in combination with the Alamut Batch software (1.11, Interactive Biosoftware, Rouen, France). Very stringent filtering criteria were applied to filter out non-pathogenic variants (Supp. Table 1): (i) variants represented with an allele frequency of more than 1% in public variation databases including the 1000Genomes (Altshuler et al., 2015), the gnomAD database (Lek et al., 2016), the DGV (MacDonald, Ziman, Yuen, Feuk, & Scherer, 2014) or our internal exome database, (ii) variants in 5′ and 3′ UTR, downstream, upstream, intronic and synonymous locations without pathogenic prediction of local splice effect. Structural variants were predicted using CANOES (Backenroth et al., 2014) and annotated by AnnotSV 2.0 (Geoffroy, Herenger, et al., 2018). Our analysis considered all inheritance modes but given that 2 families were consanguineous we focused on compound heterozygous and homozygous variants consistent with a recessive transmission. Variations have been classified according to the ACMG

guidelines (Richards et al., 2015) and only class 3 (likely pathogenic) and class 4 (pathogenic) have been considered.

Homozygosity mapping

Based on the WES data, homozygosity mapping has been performed using PLINK 1.9 (Chang et al., 2015). We followed the parameters described by Kancheva *et al* (Kancheva et al., 2016). The results are available in Supp. Table 2.

Cell culture

Skin fibroblasts from patient A.IV-1 and age-matched healthy control fibroblasts cells were grown in DMEM supplemented with 10% fetal calf serum (FCS) and 1% PSG (penicillin–streptomycin–glutamin). To induce primary cilium formation, the cells were deprived of serum by growth for 24h in DMEM with 1% PSG but only 0.1% FCS. In order to activate the Hh pathway, the cells were incubated in a DMEM medium with 1% PSG without serum (-FCS) and with SAG (100nM). SAG (Smoothened Agonist) has been resuspended in DMSO at a stock concentration of 4mM.

RNA extraction, cDNA synthesis, q-PCR and Taqman

For RNA seq validation, cells have been cultured in normal condition DMEM 1% PSG, 10% FCS. Total RNA was prepared from patient A.IV-1 fibroblast pellets (technical triplicate) and three controls pellets (biological triplicate) using QIAshredder Kit (Qiagen Ref 79654) and RNeasy Kit (Qiagen Ref 74104) followed by a DNAse treatment with the TURBO DNA-free™. Experiments have been repeated 3 times. For the Hh genes expression analysis (qRT-PCR), cells have been cultured in DMEM 1% PSG (-FCS) + SAG (100nM). Total RNA was prepared as described above and experiments have been repeated twice with 2 different control pellets (biological duplicate)

and one patient pellet (technical triplicate). RNA integrity was assessed by gel electrophoresis and RNA concentration by Eppendorf Biophotometer Plus with the Hellma® Tray Cell. Reverse transcription of 1 µg total RNA to cDNA was performed using the BioRad iScript™ cDNA synthesis kit. Real-time quantitative polymerase chain reaction amplification was performed in a BioRad CFX96 TM Real-Time System using the iQ™ SYBR® Green Supermix and primer sets (Supp. Table 3) optimized for tested targets for SYBR Green-based real-time PCR for the real-time PCR. The normalized fold expression of the target gene was calculated using the comparative cycle threshold (C_t) method by normalizing target mRNA C_t to those for both GAPDH and HPRT reference gene using the CFX Manager Software Version 1.5 and excel calculation. Statistical analysis have been done using the "Graphpad prism" software.

Immunocytochemistry and fluorescence microscopy

A.IV-1 and control ciliated fibroblasts were seeded on 12 mm diameter cover slips. Ciliogenesis has been induced by 24h incubation in a DMEM, 1% PSG medium, without serum (-FCS). Then, cells have been fixed with PFA-PBS 4%. Subsequently, the cells were stained with the ciliary markers anti-acetylated alpha tubulin (Lys40) mouse monoclonal primary antibody (1:200, # 32-2700 Invitrogen) and anti-EVC2 rabbit polyclonal primary antibody (1:250, # ab198930, Abcam).

As secondary antibody, donkey anti-mouse Alexa 594 (1:1000, Invitrogen) and donkey anti-rabbit DyLight 488 (1:1000, #611-741-127, Rockland - Tebu-bio) were used. Ciliated cells count between A.IV-1 and control was performed manually using ~100 cells per genotype (exact Fisher test was applied). Counts of EVC2 colocalization at the base of the cilia was performed manually using also ~100 cells per genotype in triplicate (Student t-test has been applied). Pictures were taken with a Zeiss axio imager Z2 and the analysis with the ZEN2012 software (Carl Zeiss Inc., Oberkochen, Germany).

Transcriptome analysis

RNA samples were extracted from two fibroblast replicates of A.IV-1 (A.IV-1.1 and A.IV-1.2), and replicates from 6 individuals affected by other ciliopathies or intellectual disabilities of known molecular origin (ARN1.1, ARN1.2, ARN2.1, ARN2.2, ARN3.1, ARN3.2, ARN4.1, ARN4.2, ARN5.1, ARN6.1, ARN6.2) using TRI reagent® (Molecular Research Center) or using the RNeasy Mini Kit Qiagen®. Both protocols included an additional step of DNAse I recombinant treatment (Sigma-Aldrich®). The integrity of the RNA was visualized on a 1% bleach agarose gel by electrophoresis (Aranda et al., 2012). Quantification and further quality analysis were performed using the Nanodrop®. Samples should have a 260/280 ratios around 2 and a 260/230 ratio above 1.7. The integrity and quality of the RNA were also evaluated by running samples on a RNA 6000 Nano Chip on the Bioanalyzer (Agilent Technologies) and samples should have a RNA integrity number (RIN) equal or above 8. Library preparation was performed at the GenomEast platform at the Institute of Genetics and Molecular and Cellular Biology (Strasbourg, France), using the TruSeq® stranded mRNA sample preparation kit (Illumina) starting from 1 µg of extracted total RNA. Libraries were then 100+75bp paired-end sequenced, with 2 samples per lane on an Illumina Hiseq4000 sequencer. The corresponding bioinformatics pipeline is described in detail in the supplementary information.

Zebrafish stocks

Fish were bred and raised at 28.5°C as described previously (Westerfield, 1993). The AB wild-type line (European Zebrafish Resource Centre (EZRC), Karlsruhe) was used for all the experiments. Zebrafish husbandry and experimental procedures were performed in accordance with German animal protection regulations (Regierungspräsidium Karlsruhe, Germany, AZ35-

9185.81/G-137/10). Fish were crossed pairwise. For experiments, fertilized eggs were raised in 1× Instant Ocean salt solution (Aquarium Systems, Inc.) supplemented with 200 μM 1-phenyl 2-thiourea (PTU) to suppress melanogenesis. For visualization of *situs inversus*, we used a transgenic line expressing *GFP* under the control of the skeletal and cardiac muscle specific promoter *unc45b* (Roostalu & Strahle, 2012).

In situ hybridization

Whole-mount *in situ* hybridization was performed as previously described (Costa, Escaleira, Rodrigues, Manasfi, & Mermelstein, 2002; Oxtoby & Jowett, 1993). *Ptc1* probe was obtained from (Concordet et al., 1996). wt1a probe was obtained from (Armant et al., 2013).

Morpholinos and microinjections

A search of the zebrafish genome (GRCz11) revealed a single orthologue of the human *iqce* gene (RefSeq: NM_001287204.1) encoding a protein sharing 58% amino acid similarity with the human IQCE protein. Injections were performed as described previously (F. Muller et al., 1999). In brief, zebrafish eggs were collected shortly after being laid. Cleaned eggs were transferred to a petri dish with a minimal amount of water. Embryos were injected (FemtoJet; Eppendorf) through the chorion into the yolk at the one-cell stage. Injection needles were pulled from borosilicate glass capillary tubes with filament (Warner Instruments) using a micropipette puller (Sutter Instrument Co). Morpholinos were designed against the start codon of *iqce* (*iqce-mo*). Morpholino controls consist of the 5 bp mismatch *iqce-mo* (*iqce-mocont*). Morpholinos (Gene Tools, LLC) were injected at the following concentrations: *iqce-Mo*: CAAGTTCTCCAGCTACCACAGACAT (0.6 μM); *iqce-mocont*; CAAcTTgTCgAGCTACCAgA cACAT (0.6 μM). All dilutions were made in distilled water. Phenol red was added to the samples before injection (0.1% final concentration). *Iqce-gfp* plasmid was injected at the final concentration of 40 ng/μl. When injected at 0.6 μM, neither the anti *iqce*

nor the *iqce-mocont* (control) morpholino elicited necrosis, an unspecific effect of morpholino injections occasionally observed with some morpholinos (Robu et al., 2007). The efficiency of *iqce-mo* to block translation was assessed by co-injection of *iqce-mo* with a plasmid containing the *iqce* cDNA in-frame with *gfp* (*iqce-gfp*). As expected the embryos were depleted of GFP expression, indicating that the morpholinos indeed target efficiently *iqce*. In contrast, the 5-mismatch *iqce-mocont* was unable to abrogate the translation of the fusion gene (Supp. Figure 3A).

PCR amplification

Total RNA was isolated from 24 to 72 hpf (hours post fertilization) embryos using Tri-reagent (Invitrogen, Carlsbad, CA). RT-PCR was carried out following standard protocols. For *iqce in situ* hybridization probe, a fragment of 580 bp was amplified by PCR with the following primers *iqcewm5*: GCTGCATCATGTAGGACGATG *iqcewm3:* CACTCTGAAGCTGACTCAAG, and cloned into pgem-t-easy vector (Promega). *Iqce* full-length sequence was amplified from zebrafish cDNA with the following primers: *iqce5':* tcgGAATTCATGTCTGTGGTAGCTGGAGAAC; *iqce3':* agactcgagggTAAATCAAGTCA TCGGCACTG and cloned into *pcs2+GFP* with EcoRI-XhoI. For the rescue experiment *iqce5'modif:* tcgGAATTCATGTCGGTTGTAGCGGGCGAACTAG was used in combination with *iqce3'*, and cloned into *pCS2+GFP* Morpholino-resistant *iqce* mRNA was synthesized with the SP6kit (Promega).

Immunohistochemistry

Fish embryos were fixed in 4% paraformaldehyde/1x PBS one hour at room temperature. Embryos were permeabilized 7 min at -20°C with cold acetone, rinsed 5 times, 5 min with 1x PBS and blocked 4 h with BDP (1% DMSO, 5% BSA, 1x PBS). Incubation with primary antibody occurred

over night at 4°C, followed by thorough washing with PTW (1x PBS, 0.1% Tween20). Embryos were incubated with secondary antibody two hours at room temperature, and washed with PTW. Antibodies dilutions: Acetylated-tubulin (Sigma-Aldrich) 1/500; engrailed (Developmental Studies Hybridoma Bank, The University of Iowa): 1/50; goat anti-rabbit alexa 488 (Invitrogen): 1/1000, goat anti-mouse alexa 488 (Invitrogen): 1/1000. For nucleus staining, we used Draq5 at 1/1000 dilution. Staining was analyzed by confocal microscopy (TCS SP2, Leica). For sectioning of the eyes, embryos were fixed in 4% paraformaldehyde/PBS, progressively dehydrated in ethanol and stepwise transferred into Epon resin. Five µm sections were cut with a Leica microtome and stained with toluidine blue.

RESULTS

Clinical features

Three families with ciliopathy like phenotypes were referred to our laboratory as they remained unsolved following standard genetic testing. They all shared postaxial polydactyly (PAP) as a common feature but each presented other symptoms as described below (Figure 1A and Table 1). The affected patient of family A (A.IV-1) was born after an unremarkable pregnancy and is the only child of related Algerian parents. Bilateral polydactyly and brachydactyly of hands and feet were noted at birth associated with a syndactyly of the 2nd and 3rd digits of the feet. At the age of six years old, he was diagnosed with a decreased visual acuity and his ophthalmologic evaluation revealed retinal degeneration. He had hypogonadism and overweight was noted during his last clinical evaluation (15 years old). Family B is composed of three children presenting with PAP of hands and feet (bilateral for B.II-1 and unilateral for B.II-2 and B.II-3). The older one and the twin sister presented overweight respectively at the age of 7 and 5 years old in a context of obese parents. The ophthalmologic evaluation, including electroretinography, was normal for all the

siblings. In family C, two affected children were reported. Individual C.IV-3 is the last child of related Moroccan parents and presented with PAP of both hands and one foot, learning disabilities and overweight. Her two sisters and parents share also the two last features. Her cousin, patient C.IV-4, born also from related parents is more severe. He had also PAP of the four limbs associated with renal anomalies, deafness and speech delay. The observed kidney disease consisted of a tubular nephropathy with nephrocalcinosis diagnosed at one year old evocative of a distal renal tubular acidosis. Deafness was detected at the age of two years old. His development was in the normal range, except for a language delay, in the context of hearing loss, which improved with speech therapy (Table 1).

The association of PAP and the other clinical features of each family member was assumed to be compatible with a syndromic ciliopathy (Wolff et al., 2004) and in particular to the Bardet-Biedl syndrome (BBS; MIM# 209900). In line with the described phenotypes, the three families have been screened without success for the known BBS genes at that time (*BBS1-BBS18*) using either Sanger or targeted exome sequencing (J. Muller et al., 2010; Redin et al., 2012).

Identification of *IQCE* pathogenic variations

Whole exome sequencing (WES) was applied to all three families, generating on average a total of 81000 single nucleotide variants (SNVs), 11400 insertions-deletions (indels) and 21 structural variants (SV) (Supp. Table 1). In family A, as the proband (A.IV-1) was supposedly born from consanguineous parents, we identified 39 sizeable homozygous regions encompassing 238 Mb (Supp. Table 2); supporting the assumption of consanguinity. In this case, autosomal recessive inheritance is more likely and we thus focused on homozygous variants, reducing the number of variants to 7 in 6 genes (*DNAAF1*, *IQCE*, *MADD*, *SHANK2*, *TULP1* and *ZDHHC2*) (Supp. Table 1). A.IV-1 is affected by polydactyly, retinal dystrophy and hypogonadism, which made us,

concentrate on 2 genes. The first one is *IQCE* (NM_152558.5), a gene that was just described as associated to PAP (Umair et al., 2017), with a homozygous 10 bp deletion in exon 12 c.895_904del (p.Val301Serfs*8) (Figure 1B) located in a conspicuous homozygous region of 2.4 Mb. This variant is present in gnomAD at a very low combined allele frequency (0.0009534) with two homozygous carriers, which is compatible with the frequency of the disease. Subsequently, we analyzed *IQCE* exons 12 in 96 matched controls of Algerian ethnicity and could not find this variant (supplementary methods). Interestingly, the second gene is *TULP1* (NM_001289395.1) a gene reported in Leber congenital amaurosis or retinitis pigmentosa (Hagstrom, North, Nishina, Berson, & Dryja, 1998; Hanein et al., 2004) for which we identified a known homozygous pathogenic (class 5) variation, c.1198C>T (p.Arg400Trp), also located in a homozygous region of 8.5 Mb. This variant does not appear in public databases (gnomAD, 1000Genomes). The variants in both genes (*IQCE* and *TULP1*) segregated with the phenotypes within the family (Figure 1B) and in combination explained most of the clinical features of this individual. In the non-consanguineous family B (Figure 1A), WES revealed only 2 class-4 variations in *IQCE* at compound heterozygous state: c.895_904del (p.Val301Serfs*8) in exon 12 and c.1350_1353delAGAG, (p.Glu451Argfs*15) in exon 16 (Figure 1B). Both variations segregated in all affected members (B.II-1, B.II-2 and B.II-3). The second variation (c.1350_1353delAGAG) was found in gnomAD with a frequency of 0.0002381 with 66 heterozygous carriers. In family C, WES analysis of the most severe case, C.IV-4 (Figure 1A), revealed a known homozygous variant in *ATP6V1B1* (NM_001692.3), a gene that is associated with distal renal tubular acidosis with deafness (Vargas-Poussou et al., 2006) (dRTA; MIM# 267300). This class-4 variant, c.175-1G>C, is located in a canonical splice site and is predicted to cause the skipping of exon 3, this variation was not found in public databases (gnomAD, 1000Genomes). Interestingly, the c.895_904del (p.Val301Serfs*8) variation in *IQCE* that has been already observed in family A and B (respectively at the homozygous and compound heterozygous state) was only found in C.IV-4 at the heterozygous state (whereas no second pathogenic allele could be identified by cDNA sequencing and quantification, data not shown).

The WES did not reveal any other variations in known PAP genes for this individual. Given that patient C.IV-3 had only PAP and given the C.IV-4 variation in *IQCE*, we sequenced the *IQCE* gene in C.IV-3 and surprisingly found the c.895_904del (p.Val301Serfs*8) at the homozygous state (Figure 1).

Given the recurrence of the c.895_904del (p.Val301Serfs*8) variant in our cohort, we investigated a possible founder effect. Nevertheless, no common haplotype could be observed in the corresponding region in family A and B (Supp. Table 4). Analysis of the gnomAD record of this variation revealed its presence in different ethnical groups at a very low allele frequency: European (221/126532 alleles), Ashkenazi (12/10134 alleles), Latino (17/34392 alleles), African (6/23966 alleles), Finish (2/25788 alleles), south Asian (1/30782 alleles). The diverse origin of individuals with this *IQCE* variant and the differences in the haplotype suggested a recurrent mutational mechanism.

In order to better understand the effect of *IQCE* variants, we investigated the expression and subcellular localization of IQCE interacting partners in patient cells. *IQCE* is ubiquitously expressed in human adult tissues with higher levels of mRNA expression in lung, brain and hippocampus (Supp. Figure 1A). Endogenous IQCE is located at the base of the cilia and forms a complex with EFCAB7, to anchor the EVC-EVC2 sub-complex at the base of the cilia (EvC zone) for propagation of Hh signaling (Pusapati et al., 2014). We investigated the impact of IQCE deficiency in the formation of the primary cilia and localization of EVC-EVC2 sub-complex. Immunohistochemical studies on ciliated fibroblasts from individual A.IV-1 showed no differences in the number of ciliated cells (Supp. Figure 1B). Nevertheless, for 77% of patient's cells, we did not observe colocalization of EVC2 at the base of the cilia (Figure 2A and B), that presumably results in mislocalization of the EVC-EVC2 sub-complex and may compromise the Hh signaling.

Transcriptome analysis and Hh signaling effect

In order to get insight into the mechanistic impact of IQCE, we explored the transcriptome variation of individual A.IV-1 by RNA sequencing. We considered the possible impact of *TULP1* variation. However, under physiological conditions, *TULP1* is expressed exclusively in the retina, where it is involved in protein trafficking which is essential for the transport of rhodopsin (North, Naggert, Yan, Noben-Trauth, & Nishina, 1997). This was confirmed as no *TULP1* expression could be observed in the patient's fibroblasts. Given the *IQCE* genotype (frameshift) for A.IV-1 no *IQCE* expression could be detected. RNA differential expression analysis resulted in 98 significantly differentially expressed genes (Figure 2C and Supp. Table 5). None of the known IQCE interactors - *EVC, EVC2, EFCAB7* and *TCC23* (Breslow et al., 2018; Pusapati et al., 2018) - showed significant variations in their expression levels neither did the downstream effectors of the Hh signaling pathway *GLI1*, *GLI2* and *GLI3*. We identified among the most significantly up and downregulated genes *SLC35D3, SLC30A8* and *GSTT1* respectively. *SLC35D3* has been shown to be implicated in metabolic syndrome and obesity (Z. Zhang et al., 2014) and *SLC30A8* is associated with a protective role against diabetes type2 (Flannick et al., 2014).

Regarding *GSTT1,* the protein catalyzes the conjugation of glutathione to a wide range of potential toxins as the first step in detoxification and is identified as risk factor for open angle glaucoma or ventricular diastolic dysfunction (Singh, Kumar, Tewari, & Agarwal, 2018; Stamenkovic et al., 2018). In order to obtain functional clues about the IQCE protein modules, we applied functional enrichment analysis (WebGestalt 2019) (Wang, Vasaikar, Shi, Greer, & Zhang, 2017); the over-representation analysis identified 10 significant clusters involving tissue development, ossification and skeletal system development among others (Figure 2D). A number of genes associated with Hh signaling and limb formation are included in these protein clusters (ALPL, ALDH1A1, FGF9, HOXC10, IGF2, IGFBP1, MECOM, MLLT3, NDP, SFRP1 and SMOC1) (Supp. Table 6). Selected genes with significant | log2 fold-change | > 0.5 were tested compared to controls by RT-qPCR and we confirmed the expression change of *ALPL, GSTT1, IGF2, IQCE, MECOM, RSPO1* and *SLC30A8* but we did not detect significant differences in *FGF9* and *SFRP1* (Figure 2E).

In order to assess whether the Hh signaling is affected, we determined the expression levels of key genes in the patient's cells both under ciliary conditions and Hh activation (-FCS + SAG treatment). qRT-PCR revealed a significant under-expression of genes involved in the EVC-zone complex including *EFCAB7, EVC, EVC2* and also *GLI2* and *GLI3* (Supp. Figure 2).

Lack of *iqce* causes ciliary defects in zebrafish embryos

Considering the ciliary localization of the human IQCE and encouraged by the previous studies, where genetic screens in zebrafish identified numerous defects of subpopulations of cilia (Zhao & Malicki, 2007), we investigated expression and function of *iqce* in zebrafish (*Danio rerio).* We first assessed the *iqce* expression in the developing embryo. *In situ* hybridisation of *iqce* mRNA showed ubiquitous expression at the 10-somites and 20 hpf stages with higher expression in the pronephric ducts at 20 hpf. At 48 hpf, higher levels of *iqce* mRNA were found in the pronephric ducts and in the mid- and hindbrain (Supp. Figure 3B).

A morpholino directed against the *iqce* start codon (*iqce-mo*) and a morpholino with a 5 mismatch control sequence (*iqce-mocont)* were used to elucidate the function of *iqce* in zebrafish embryos. Embryos injected with *iqce-mo* (77%, n= 126) but not with the *iqce-mocont* control (2%, n=221) showed ventral curvature of the body axis (Figure 3A-B) which is often found in cilia mutants (Supp. Figure 3C). Occasionally we observed hydrocephalus (Figure 3B). In order to prove that the observed phenotype is due to the specific knock-down of *iqce*, we co-injected *iqce-mo* together with *iqce* mRNA resistant to *iqce-mo*. 54% of embryos exhibited a characteristic curved phenotype compared to 77% for *iqce-mo* alone (Figure 3D, n=193). We conclude from this experiment that the *iqce-mo* effectively abrogate the *iqce* expression.

Bilaterally dilated kidney cyst were present in 37% of *iqce-mo* morphants, but neither in uninjected nor in *iqce-mocont* injected embryos (Figure 3C and Supp. Figure 3D). These cysts are localized within the glomerulus as indicated by *wt1a* staining (Supp. Figure 3E).

We next examined the cilia of the pronephric duct. Embryos at 48 hpf were stained with acetylated-tubulin antibody. In *iqce-mocont* and uninjected embryos, cilia are aligned along the duct forming a thin line (Figure 3E, G). In contrast, the cilia organization in *iqce* morphants was severely perturbed showing cilia oriented in all directions (Figure 3F). The misorientation and shortening of pronephric duct cilia can be due to dilation of the tubules and suggest that pronephric fluid flow may be affected (Zhao & Malicki, 2007). We next tested whether other cilia are also affected, the dense ring of cilia on the epithelium surrounding the olfactory placode of 48-hpf wild-type larvae was strongly reduced in morphants (Figure 3H-J). In addition, the photoreceptor outer segment connects to the cell body via a narrow constriction that forms around a ciliary axoneme, and is known as the connecting cilium. Defects of photoreceptor cilia are associated with outer segment abnormalities and degeneration (Zhao & Malicki, 2007). Therefore, we analyzed eye sections of 4 day old morphants. In non-injected embryos or control morphants all the layers are properly formed (Figure 3K). In contrast in *iqce* morphants, the lamination is disrupted and the retinal layer altered with no visible photoreceptor outer segment (Figure 3L).

The perturbation of motile cilia often results in a problem of left-right asymmetry. We thus examined the percentage of *situs inversus* in *iqce* morphants. *iqce-mo* or *iqce-mocont* were injected into embryos carrying a transgene that expresses GFP in the heart (-1.8 unc45b-TFP, (Roostalu & Strahle, 2012). This allowed us to observe the looping of the heart. 44% of morphant embryos exhibited *situs inversus* (with the ventricle on the right and atrium on the left) or heterotaxia demonstrating that cilia from the Kupffer's vesicle, the fish equivalent of the mouse node, were affected by the *iqce* knock-down. In contrast only 3% and 6% of uninjected and *iqce-mocont* injected embryos, respectively, showed reverse heart jogging (Figure 3M).

In mutants with defective cilia (*iguana*) or ciliary motility (*lrrc50*), otoliths are frequently found ectopically, untethered or fused (Stooke-Vaughan, Huang, Hammond, Schier, & Whitfield, 2012). We thus examined the number of otoliths in *iqce* morphants at 2 dpf. At 48 dpf each ear should contains 2 otoliths and it is the case for 86 % of the morphant. However, for 14 % of them one of the two ears presented either one extra otolith (3 instead of 2) or only one (1 instead of 2) (data not shown, n= 262 embryos).

Cilia are necessary for normal activation as well as repression of the hedgehog (Hh) signaling pathway (Goetz & Anderson, 2010; Huang & Schier, 2009; Wolff et al., 2004). It was shown that in the absence of the ciliary protein BBS1 and under stimulation of Hh, the ciliary localization of Smoothened (*Smo*) was further enhanced compared to wild-type cells (Nozaki, Katoh, Kobayashi, & Nakayama, 2018). We thus examined the expression of the Hh receptor Patch1 (*Ptc1*) and the downstream gene engrailed (*eng*) in the developing somites in *iqce* morphants. At 10 somites stages, transcription of *ptc1* was expanded to the myotome, whereas it is exclusively restricted to the adaxial cells in control or uninjected embryos (Figure 3N-PO and Supp. Figure 3F). Similarly, the domain of *eng* expression was significantly expanded among the fast muscle fibres by 24 hpf, indicative of an increased number of Hh-dependent muscle fast fibers (Figure 3Q-S). These results indicate that the Hh pathway is perturbed by the loss of *iqce* function.

DISCUSSION

An *IQCE* homozygous variant affecting the splicing of the gene was recently identified by whole exome sequencing in a single Pakistani family presenting with postaxial polydactyly type A limited to lower limb (Umair et al., 2017). Here, we report 3 additional families with pathogenic variations in *IQCE* referred initially as ciliopathies with postaxial polydactyly involving the upper and lower limbs. Interestingly, the 3 families shared the same c.895_904del (p.Val301Serfs*8) variation in

IQCE exon 12 in a homozygous or in a heterozygous state with another truncating variant. These two *IQCE* variations have already been observed in the largest public databases. The c.1350_1353del (p.Glu451Argfs*15) was described in Exome Variant Server (EVS) (Tennessen et al., 2012) with an average allele frequency >1.5% with 90 homozygous calls. Initially, this overestimated allele frequency prevented its identification as candidate pathogenic variation. The higher false positive rate for indel calls was a known issue solved while integrating in ExAC/gnomAD (0.024% and no homozygous calls).

The index case A.IV-1 had a suspected ciliopathy-like phenotype, due to the presence of retinitis pigmentosa (RP), hypogonadism, polydactyly and speech delay. Although the proband was assumed as a patient with one syndrome, the occurrence of pathogenic variations in both *IQCE* and *TULP1* can explain independently the polydactyly and the RP. In the family C, the phenotypes of the patients were different; C.IV-3 has only PAP and overweight compared to her cousin who has PAP, deafness and renal anomalies. Phenotypic variations in a single family is not unusual in ciliopathies even with a single gene involved (Abdelhamed et al., 2013; Gonzalez-Del Pozo et al., 2014; Katsanis et al., 2001; Ware, Aygun, & Hildebrandt, 2011; W. Zhang et al., 2018). The variations in *IQCE,* can clearly explain the isolated PAP in C.IV-3. However, in the C.IV-4 individual, the association of deafness and nephropathy could finally be explained by the identification of variations in the *ATP6V1B1* gene (dRTA; MIM# 267300), known to be responsible for renal tubular acidosis associated with deafness. Nevertheless, the identification of only a single pathogenic allele in *IQCE* and the absence of any other variation in known polydactyly genes do not provide any explanation to date for the PAP for this patient. Although, gene dosage and RNA expression have been performed for this individual revealing still 50% of the *IQCE* expression level (data not shown), one cannot rule out missed variation(s) either in *IQCE* or in another gene. Whole genome sequencing might reveal intronic, intragenic or structural variations missed by the WES (Geoffroy, Stoetzel, et al., 2018). These two examples illustrate well how the incidental co-

occurrence of 2 different conditions can mimic the clinical features of a known Mendelian disease (DeLuca, Weed, Haas, Halder, & Stone, 2015; Karaca et al., 2018). This point out also the importance for searching for multiple hits in whole exome or whole genome sequencing data to explain some apparent syndromic phenotype such as ciliopathy-like phenotype. The occurrence of the other clinical manifestations, overweight or learning disability, in these families seemed to be due to other factors than the presence of neurosensory impairment (RP or deafness). However, we should take into consideration the appreciable variations of systemic features in individuals affected by pathogenic variants in different ciliopathy genes (Carmi, Elbedour, Stone, & Sheffield, 1995; Estrada-Cuzcano et al., 2012; Waters & Beales, 2011).

Vertebrate limb patterning is controlled by two signaling centers, the zone of polarizing activity (ZPA) and the distal posterior mesenchyme and the apical ectodermal ridge (AER) at the distal limb bud. The ZPA and AER produce signaling molecules -Hh and FGFs respectively- that control the formation of anteroposterior and proximodistal limb axes in a cooperative manner (Yamamoto et al., 2012). Additionally, Hh can induce a positive feedback Igf signaling during osteoblast differentiation (mostly Igf2), regulating in this way the long bone growth (Fisher, Meyer, Garber, & Dealy, 2005; Shi, Chen, Karner, & Long, 2015). Until recently, the combination of temporal and spatial exposure to long-range Hh signaling was accepted as the main molecular mechanism underlying digit patterning (Ahn & Joyner, 2004; Harfe et al., 2004). However, current investigations in chicken limb buds show that mesenchymal cells in the posterior limb bud produce actin-based filopodia that accumulate particles of Hh, allowing the Hh particles to act at a distance without extracellular diffusion (Sanders, Llagostera, & Barna, 2013). This study suggested that proper activation of the Hh target genes is more critical for digit patterning than diffusible long-range Hh morphogen threshold. The polydactyly phenotype of other carriers of *IQCE* variations supports the involvement of IQCE in the formation of limbs, process that is related to Hh signaling. At the base of the cilia, the EvC complex (constituted by EVC-EVC2 and EFCAB7-IQCE) is not essential for Hh target gene induction. However, it can reduce the Hh-stimulated induction of GLI1

or PTCH1 (Pusapati et al., 2014). Interestingly, gene ontology analysis of misregulated transcripts revealed enrichment for skeletal system development. Notably, IGF2 – a critical protein for bone formation (Fisher et al., 2005; Shi et al., 2015)- is misregulated in absence of IQCE. Defects in *IGF2* in humans cause Silver-Russell syndrome, ectrodactyly, under masculinized genitalia, developmental delay, and placental hypoplasia (Begemann et al., 2015; Yamoto et al., 2017). Additionally, we observed the up-regulation of *IGFBP1* and *IGFBP3*, these two proteins bind IGF2 and their excess may limit IGF2 function (Chao & D'Amore, 2008).

Knockdown of *iqce* in zebrafish showed the full spectrum of phenotypes associated with defective cilia: body curvature, kidney cysts, left right asymmetry, misdirected cilia in pronephric duct (Shi, Su, Lipschutz, & Lobo, 2017) as well as retinal defects resembling very much those observed in *bbs9* morphants (Veleri et al., 2012). In addition, full *iqce* knockout mouse showed short tibias, thin retinas, metabolic and behavioural alterations as well as polydactyly. Hh expression during limb development is pivotal for defining the identity and number of digits. When its expression is activated at an ectopic site in the limb bud this define a pathogenic mechanism that leads to preaxial polydactyly (Lettice et al., 2012). The *iqce* zebrafish model exhibits modified Hh signalling. Indeed, we observed a broadening of the Hh *ptc1* receptor outside of the adaxial cell, as well as extended *Engrailed* expression. It is well known that the primary cilium acts as a Hh signal transduction machine (Goetz, Ocbina, & Anderson, 2009). Recently it was shown that in a *bbs1* KO cell line, the entry of Smo into cilia upon Hh stimulation was increased compared to the wild type cell line (Nozaki et al., 2018). This in turn activated ectopic engrailed expression as observed in *iqce* zebrafish morphants. This result was unexpected since Pusapati *et al*. provided evidence by analysing osteoblasts and the NIH3T3 cell line that the EFCAB7-IQCE module anchors the EVC-EVC2 complex in a signalling microdomain at the base of the cilia. Moreover, they showed that EFCAB7 depletion triggered the mislocalization of EVC-EVC2 within cilia and impaired activation of the transcription factor GLI2. These data suggest an impairment of Hh signalling

rather than the gain of Hh signalling activity in the *iqce* mutant background that we observed in the zebrafish. However, in zebrafish, no Evc or Evc2 orthologues have been identified but only a paralog absent in human (Pusapati et al., 2014). Transcriptome analysis of the patient's cells under normal condition did not provide evidences of GLI genes modified expression compared to controls. However, under Hh activation, the patient's cells show a reduced expression of most of the key Hh pathway genes (*GLI2, GLI3, EFCAB7, EVC, EVC2*) (Supp. Figure 2). At the protein level, we observed a significant amount of EVC2 mislocalization at the base of the cilia (Figure 2B) as well as a reduced GLI2 protein expression (data not shown).

Blindness is not present in patients with only *iqce* defect. However, it has been previously observed in other cases that KO of genes in zebrafish can give additional phenotypes not observed in patients. For example, the zebrafish cc2d2 adult mutants exhibit a scoliosis that is not observed in Joubert syndrome (JBTS; MIM# 213300) (Bachmann-Gagescu et al., 2011). The defective retina of *iqce* morphants is very similar to the *bbs9* morphant model showing a lack of lamination (Veleri et al., 2012). This observation fits with the implication of *iqce* in ciliopathies.

In conclusion, our findings confirm the role of *IQCE* in non-syndromic postaxial polydactyly, and its negative impact on the Hh signaling pathway. The differences between *iqce* zebrafish and mice models compared to humans raise the question of additional genetic factors. This study illustrates how detailed and accurate clinical data are needed to interpret exome-scale genetic results and translate those results in correct counselling and therapeutic intervention.

AVAILABILITY OF DATA AND MATERIAL

Data generated or analysed 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 using the following range of accessions numbers SCV000929968 to SCV000929972 (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.

ACKNOWLEDGMENTS

We would like to thank the patient's family for their participation and the patient's associations, Formicoeur, Bardet-Biedl France and Unadev, for their constant and strong support. We also would like to acknowledge the members of the diagnostic laboratories (molecular and cytogenetic units, in particular Manuela Antin, Anne-Sophie Leuvrey and Elsa Nourisson) at the University Hospital in Strasbourg for technical assistance. We thank Nadine Borel and the zebrafish house team. We also thank our collaborators for screening their negative cohort (KM. Bujakowska, E. Pierce, P. Beales, F. Hildebrandt). We thank the GenomEast platform (Strasbourg, France) especially Serge Vicaire for library preparation, Céline Keime for supervising the transcriptome analysis as well as Bernard Jost and Christelle Thibault. The computing resources for this work were provided by the BICS and BISTRO bioinformatics platforms in Strasbourg.

FUNDING INFORMATION

Whole exome sequencing was funded by Retina France (100 exomes Program), by Hôpitaux Universitaires de Strasbourg (API 2013-2014 HUS N° 5885). RNASeq analysis performed in patients with *IQCE* pathogenic variant and other patients were funded by Fondation Maladie Rare and Fondation Jérome Lejeune. CE and US were supported by the HGF Biointerfaces Programme. AEC is supported by Fondation JED-Belgique. CD is supported by a FRM grant (ECO20170637509).

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

REFERENCES

Abbasi, A. A. (2011). Evolution of vertebrate appendicular structures: Insight from genetic and palaeontological data. *Dev Dyn, 240*(5), 1005-1016. doi:10.1002/dvdy.22572

Abdelhamed, Z. A., Wheway, G., Szymanska, K., Natarajan, S., Toomes, C., Inglehearn, C., & Johnson, C. A. (2013). Variable expressivity of ciliopathy neurological phenotypes that encompass Meckel-Gruber syndrome and Joubert syndrome is caused by complex de-regulated ciliogenesis, Shh and Wnt signalling defects. *Hum Mol Genet, 22*(7), 1358-1372. doi:10.1093/hmg/dds546

Ahn, S., & Joyner, A. L. (2004). Dynamic changes in the response of cells to positive hedgehog signaling during mouse limb patterning. *Cell, 118*(4), 505-516. doi:10.1016/j.cell.2004.07.023

Altshuler, D. M., Durbin, R. M., Abecasis, G. R., Bentley, D. R., Chakravarti, A., Clark, A. G., . . . Consortium, G. P. (2015). A global reference for human genetic variation. *Nature, 526*(7571), 68-+. doi:10.1038/nature15393

Armant, O., Marz, M., Schmidt, R., Ferg, M., Diotel, N., Ertzer, R., . . . Rastegar, S. (2013). Genome-wide, whole mount in situ analysis of transcriptional regulators in zebrafish embryos. *Dev Biol, 380*(2), 351-362. doi:10.1016/j.ydbio.2013.05.006

Bachmann-Gagescu, R., Phelps, I. G., Stearns, G., Link, B. A., Brockerhoff, S. E., Moens, C. B., & Doherty, D. (2011). The ciliopathy gene cc2d2a controls zebrafish photoreceptor outer segment development through a role in Rab8-dependent vesicle trafficking. *Hum Mol Genet, 20*(20), 4041-4055. doi:10.1093/hmg/ddr332

Backenroth, D., Homsy, J., Murillo, L. R., Glessner, J., Lin, E., Brueckner, M., . . . Shen, Y. F. (2014). CANOES: detecting rare copy number variants from whole exome sequencing data. *Nucleic Acids Research, 42*(12). doi:ARTN e97 10.1093/nar/gku345

Barakat, M. T., Humke, E. W., & Scott, M. P. (2010). Learning from Jekyll to control Hyde: Hedgehog signaling in development and cancer. *Trends Mol Med, 16*(8), 337-348. doi:10.1016/j.molmed.2010.05.003

Begemann, M., Zirn, B., Santen, G., Wirthgen, E., Soellner, L., Buttel, H. M., . . . Eggermann, T. (2015). Paternally Inherited IGF2 Mutation and Growth Restriction. *N Engl J Med, 373*(4), 349-356. doi:10.1056/NEJMoa1415227

Breslow, D. K., Hoogendoorn, S., Kopp, A. R., Morgens, D. W., Vu, B. K., Kennedy, M. C., . . . Nachury, M. V. (2018). A CRISPR-based screen for Hedgehog signaling provides insights into ciliary function and ciliopathies. *Nat Genet*. doi:10.1038/s41588-018-0054-7

Carmi, R., Elbedour, K., Stone, E. M., & Sheffield, V. C. (1995). Phenotypic differences among patients with Bardet-Biedl syndrome linked to three different chromosome loci. *Am J Med Genet, 59*(2), 199-203. doi:10.1002/ajmg.1320590216

Castilla, E., Paz, J., Mutchinick, O., Munoz, E., Giorgiutti, E., & Gelman, Z. (1973). Polydactyly: a genetic study in South America. *Am J Hum Genet, 25*(4), 405-412. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/4716659

Chang, C. C., Chow, C. C., Tellier, L. C., Vattikuti, S., Purcell, S. M., & Lee, J. J. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience, 4*, 7. doi:10.1186/s13742-015-0047-8

Chao, W., & D'Amore, P. A. (2008). IGF2: epigenetic regulation and role in development and disease. *Cytokine Growth Factor Rev, 19*(2), 111-120. doi:10.1016/j.cytogfr.2008.01.005

Concordet, J. P., Lewis, K. E., Moore, J. W., Goodrich, L. V., Johnson, R. L., Scott, M. P., & Ingham, P. W. (1996). Spatial regulation of a zebrafish patched homologue reflects the roles of sonic hedgehog and protein kinase A in neural tube and somite patterning. *Development, 122*(9), 2835-2846. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8787757

Costa, M. L., Escaleira, R. C., Rodrigues, V. B., Manasfi, M., & Mermelstein, C. S. (2002). Some distinctive features of zebrafish myogenesis based on unexpected distributions of the muscle cytoskeletal proteins actin, myosin, desmin, alpha-actinin, troponin and titin. *Mechanisms of Development, 116*(1-2), 95-104. doi:Pii S0925-4773(02)00149-1

Doi 10.1016/S0925-4773(02)00149-1

DeLuca, A. P., Weed, M. C., Haas, C. M., Halder, J. A., & Stone, E. M. (2015). Apparent Usher Syndrome Caused by the Combination of BBS1-Associated Retinitis Pigmentosa and SLC26A4-Associated Deafness. *JAMA Ophthalmol, 133*(8), 967-968. doi:10.1001/jamaophthalmol.2015.1463

DePristo, M. A., Banks, E., Poplin, R., Garimella, K. V., Maguire, J. R., Hartl, C., . . . Daly, M. J. (2011). A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet, 43*(5), 491-+. doi:10.1038/ng.806

Estrada-Cuzcano, A., Koenekoop, R. K., Senechal, A., De Baere, E. B., de Ravel, T., Banfi, S., . . . Klevering, B. J. (2012). BBS1 mutations in a wide spectrum of phenotypes ranging from nonsyndromic retinitis pigmentosa to Bardet-Biedl syndrome. *Arch Ophthalmol, 130*(11), 1425-1432. doi:10.1001/archophthalmol.2012.2434

Etard, C., Roostalu, U., & Strahle, U. (2010). Lack of Apobec2-related proteins causes a dystrophic muscle phenotype in zebrafish embryos. *J Cell Biol, 189*(3), 527-539. doi:10.1083/jcb.200912125

Fisher, M. C., Meyer, C., Garber, G., & Dealy, C. N. (2005). Role of IGFBP2, IGF-I and IGF-II in regulating long bone growth. *Bone, 37*(6), 741-750. doi:10.1016/j.bone.2005.07.024

Flannick, J., Thorleifsson, G., Beer, N. L., Jacobs, S. B., Grarup, N., Burtt, N. P., . . . Altshuler, D. (2014). Loss-of-function mutations in SLC30A8 protect against type 2 diabetes. *Nat Genet, 46*(4), 357-363. doi:10.1038/ng.2915

Geoffroy, V., Herenger, Y., Kress, A., Stoetzel, C., Piton, A., Dollfus, H., & Muller, J. (2018). AnnotSV: An integrated tool for Structural Variations annotation. *Bioinformatics*. doi:10.1093/bioinformatics/bty304

Geoffroy, V., Pizot, C., Redin, C., Piton, A., Vasli, N., Stoetzel, C., . . . Muller, J. (2015). VaRank: a simple and powerful tool for ranking genetic variants. *PeerJ, 3*, e796. doi:10.7717/peerj.796

Geoffroy, V., Stoetzel, C., Scheidecker, S., Schaefer, E., Perrault, I., Bar, S., . . . Muller, J. (2018). Whole-genome sequencing in patients with ciliopathies uncovers a novel recurrent tandem duplication in IFT140. *Hum Mutat, 39*(7), 983-992. doi:10.1002/humu.23539

Goetz, S. C., & Anderson, K. V. (2010). The primary cilium: a signalling centre during vertebrate development. *Nature Reviews Genetics, 11*(5), 331-344. doi:10.1038/nrg2774

Goetz, S. C., Ocbina, P. J. R., & Anderson, K. V. (2009). The Primary Cilium as a Hedgehog Signal Transduction Machine. *Primary Cilia, 94*, 199-+. doi:10.1016/S0091-679x(08)94010-3

Gonzalez-Del Pozo, M., Mendez-Vidal, C., Santoyo-Lopez, J., Vela-Boza, A., Bravo-Gil, N., Rueda, A., . . . Antinolo, G. (2014). Deciphering intrafamilial phenotypic variability by exome sequencing in a Bardet-Biedl family. *Mol Genet Genomic Med, 2*(2), 124-133. doi:10.1002/mgg3.50

Hagstrom, S. A., North, M. A., Nishina, P. L., Berson, E. L., & Dryja, T. P. (1998). Recessive mutations in the gene encoding the tubby-like protein TULP1 in patients with retinitis pigmentosa. *Nat Genet, 18*(2), 174-176. doi:10.1038/ng0298-174

Hanein, S., Perrault, I., Gerber, S., Tanguy, G., Barbet, F., Ducroq, D., . . . Kaplan, J. (2004). Leber congenital amaurosis: comprehensive survey of the genetic heterogeneity, refinement of the clinical definition, and genotype-phenotype correlations as a strategy for molecular diagnosis. *Hum Mutat, 23*(4), 306-317. doi:10.1002/humu.20010

Harfe, B. D., Scherz, P. J., Nissim, S., Tian, H., McMahon, A. P., & Tabin, C. J. (2004). Evidence for an expansion-based temporal Shh gradient in specifying vertebrate digit identities. *Cell, 118*(4), 517-528. doi:10.1016/j.cell.2004.07.024

Huang, P., & Schier, A. F. (2009). Dampened Hedgehog signaling but normal Wnt signaling in zebrafish without cilia. *Development, 136*(18), 3089-3098. doi:10.1242/dev.041343

Ingham, P. W., Nakano, Y., & Seger, C. (2011). Mechanisms and functions of Hedgehog signalling across the metazoa. *Nature Reviews Genetics, 12*(6), 393-406. doi:10.1038/nrg2984

Kancheva, D., Atkinson, D., De Rijk, P., Zimon, M., Chamova, T., Mitev, V., . . . Jordanova, A. (2016). Novel mutations in genes causing hereditary spastic paraplegia and Charcot-Marie-Tooth neuropathy identified by an optimized protocol for homozygosity mapping based on whole-exome sequencing. *Genet Med, 18*(6), 600-607. doi:10.1038/gim.2015.139

Karaca, E., Posey, J. E., Coban Akdemir, Z., Pehlivan, D., Harel, T., Jhangiani, S. N., . . . Lupski, J. R. (2018). Phenotypic expansion illuminates multilocus pathogenic variation. *Genet Med*. doi:10.1038/gim.2018.33

Katsanis, N., Ansley, S. J., Badano, J. L., Eichers, E. R., Lewis, R. A., Hoskins, B. E., . . . Lupski, J. R. (2001). Triallelic inheritance in Bardet-Biedl syndrome, a Mendelian recessive disorder. *Science, 293*(5538), 2256-2259. doi:10.1126/science.1063525

Lek, M., Karczewski, K. J., Minikel, E. V., Samocha, K. E., Banks, E., Fennell, T., . . . Consortium, E. A. (2016). Analysis of protein-coding genetic variation in 60,706 humans. *Nature, 536*(7616), 285-+. doi:10.1038/nature19057

Lettice, L. A., Williamson, I., Wiltshire, J. H., Peluso, S., Devenney, P. S., Hill, A. E., . . . Hill, R. E. (2012). Opposing functions of the ETS factor family define Shh spatial expression in limb buds and underlie polydactyly. *Dev Cell, 22*(2), 459-467. doi:10.1016/j.devcel.2011.12.010

MacDonald, J. R., Ziman, R., Yuen, R. K. C., Feuk, L., & Scherer, S. W. (2014). The Database of Genomic Variants: a curated collection of structural variation in the human genome. *Nucleic Acids Research, 42*(D1), D986-D992. doi:10.1093/nar/gkt958

Malik, S. (2014). Polydactyly: phenotypes, genetics and classification. *Clin Genet, 85*(3), 203-212. doi:10.1111/cge.12276

Malik, S., Ullah, S., Afzal, M., Lal, K., & Haque, S. (2014). Clinical and descriptive genetic study of polydactyly: a Pakistani experience of 313 cases. *Clin Genet, 85*(5), 482-486. doi:10.1111/cge.12217

Metsalu, T., & Vilo, J. (2015). ClustVis: a web tool for visualizing clustering of multivariate data using Principal Component Analysis and heatmap. *Nucleic Acids Research, 43*(W1), W566-570. doi:10.1093/nar/gkv468

Muller, F., Blader, P., Rastegar, S., Fischer, N., Knochel, W., & Strahle, U. (1999). Characterization of zebrafish smad1, smad2 and smad5: the amino-terminus of Smad1 and Smad5 is required for specific function in the embryo. *Mechanisms of Development, 88*(1), 73-88. doi:Doi 10.1016/S0925-4773(99)00173-2

Muller, J., Stoetzel, C., Vincent, M. C., Leitch, C. C., Laurier, V., Danse, J. M., . . . Dollfus, H. (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), 583-593. doi:10.1007/s00439-010-0804-9

North, M. A., Naggert, J. K., Yan, Y., Noben-Trauth, K., & Nishina, P. M. (1997). Molecular characterization of TUB, TULP1, and TULP2, members of the novel tubby gene family and their possible relation to ocular diseases. *Proc Natl Acad Sci U S A, 94*(7), 3128-3133. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9096357

Nozaki, S., Katoh, Y., Kobayashi, T., & Nakayama, K. (2018). BBS1 is involved in retrograde trafficking of ciliary GPCRs in the context of the BBSome complex. *PLoS One, 13*(3), e0195005. doi:10.1371/journal.pone.0195005

Oxtoby, E., & Jowett, T. (1993). Cloning of the Zebrafish Krox-20 Gene (Krx-20) and Its Expression during Hindbrain Development. *Nucleic Acids Research, 21*(5), 1087-1095. doi:DOI 10.1093/nar/21.5.1087

Petrova, R., & Joyner, A. L. (2014). Roles for Hedgehog signaling in adult organ homeostasis and repair. *Development, 141*(18), 3445-3457. doi:10.1242/dev.083691

Phadke, S. R., & Sankar, V. H. (2010). Polydactyly and genes. *Indian J Pediatr, 77*(3), 277-281. doi:10.1007/s12098-010-0033-1

Pusapati, G. V., Hughes, C. E., Dorn, K. V., Zhang, D., Sugianto, P., Aravind, L., & Rohatgi, R. (2014). EFCAB7 and IQCE regulate hedgehog signaling by tethering the EVC-EVC2 complex to the base of primary cilia. *Dev Cell, 28*(5), 483-496. doi:10.1016/j.devcel.2014.01.021

Pusapati, G. V., Kong, J. H., Patel, B. B., Krishnan, A., Sagner, A., Kinnebrew, M., . . . Rohatgi, R. (2018). CRISPR Screens Uncover Genes that Regulate Target Cell Sensitivity to the Morphogen Sonic Hedgehog. *Dev Cell, 44*(2), 271. doi:10.1016/j.devcel.2018.01.002

Redin, C., Le Gras, S., Mhamdi, O., Geoffroy, V., Stoetzel, C., Vincent, M. C., . . . Muller, J. (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), 502-512. doi:10.1136/jmedgenet-2012-100875

Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., . . . Committee, A. L. Q. A. (2015). Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med, 17*(5), 405-424. doi:10.1038/gim.2015.30

Robu, M. E., Larson, J. D., Nasevicius, A., Beiraghi, S., Brenner, C., Farber, S. A., & Ekker, S. C. (2007). p53 activation by knockdown technologies. *PLoS Genet, 3*(5), e78. doi:10.1371/journal.pgen.0030078

Roostalu, U., & Strahle, U. (2012). In Vivo Imaging of Molecular Interactions at Damaged Sarcolemma. *Dev Cell, 22*(3), 515-529. doi:10.1016/j.devcel.2011.12.008

Sanders, T. A., Llagostera, E., & Barna, M. (2013). Specialized filopodia direct long-range transport of SHH during vertebrate tissue patterning. *Nature, 497*(7451), 628-632. doi:10.1038/nature12157

Shaheen, R., Szymanska, K., Basu, B., Patel, N., Ewida, N., Faqeih, E., . . . Alkuraya, F. S. (2016). Characterizing the morbid genome of ciliopathies. *Genome Biol, 17*(1), 242. doi:10.1186/s13059-016-1099-5

Shi, Y., Chen, J., Karner, C. M., & Long, F. (2015). Hedgehog signaling activates a positive feedback mechanism involving insulin-like growth factors to induce osteoblast differentiation. *Proc Natl Acad Sci U S A, 112*(15), 4678-4683. doi:10.1073/pnas.1502301112

Shi, Y., Su, Y., Lipschutz, J. H., & Lobo, G. P. (2017). Zebrafish as models to study ciliopathies of the eye and kidney. *Clin Nephrol Res, 1*(1), 6-9. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/29553143

Singh, M. M., Kumar, R., Tewari, S., & Agarwal, S. (2018). Association of GSTT1/GSTM1 and ApoE variants with left ventricular diastolic dysfunction in thalassaemia major patients. *Hematology*, 1-6. doi:10.1080/10245332.2018.1502397

Stamenkovic, M., Lukic, V., Suvakov, S., Simic, T., Sencanic, I., Pljesa-Ercegovac, M., . . . Djukic, T. (2018). GSTM1-null and GSTT1-active genotypes as risk determinants of primary open angle glaucoma among smokers. *Int J Ophthalmol, 11*(9), 1514-1520. doi:10.18240/ijo.2018.09.14

Stooke-Vaughan, G. A., Huang, P., Hammond, K. L., Schier, A. F., & Whitfield, T. T. (2012). The role of hair cells, cilia and ciliary motility in otolith formation in the zebrafish otic vesicle. *Development, 139*(10), 1777-1787. doi:10.1242/dev.079947

Tennessen, J. A., Bigham, A. W., O'Connor, T. D., Fu, W., Kenny, E. E., Gravel, S., . . . Project, N. E. S. (2012). Evolution and functional impact of rare coding variation from deep sequencing of human exomes. *Science, 337*(6090), 64-69. doi:10.1126/science.1219240

Ullah, I., Kakar, N., Schrauwen, I., Hussain, S., Chakchouk, I., Liaqat, K., . . . Leal, S. M. (2019). Variants in KIAA0825 underlie autosomal recessive postaxial polydactyly. *Hum Genet*. doi:10.1007/s00439-019-02000-0

Umair, M., Ahmad, F., Bilal, M., Ahmad, W., & Alfadhel, M. (2018). Clinical Genetics of Polydactyly: An Updated Review. *Front Genet, 9*, 447. doi:10.3389/fgene.2018.00447

Umair, M., Shah, K., Alhaddad, B., Haack, T. B., Graf, E., Strom, T. M., . . . Ahmad, W. (2017). Exome sequencing revealed a splice site variant in the IQCE gene underlying post-axial polydactyly type A restricted to lower limb. *Eur J Hum Genet, 25*(8), 960-965. doi:10.1038/ejhg.2017.83

Vargas-Poussou, R., Houillier, P., Le Pottier, N., Strompf, L., Loirat, C., Baudouin, V., . . . Blanchard, A. (2006). Genetic investigation of autosomal recessive distal renal tubular acidosis: evidence for early sensorineural hearing loss associated with mutations in the ATP6V0A4 gene. *J Am Soc Nephrol, 17*(5), 1437-1443. doi:10.1681/ASN.2005121305

Veleri, S., Bishop, K., Nogare, D. E. D., English, M. A., Foskett, T. J., Chitnis, A., . . . Swaroop, A. (2012). Knockdown of Bardet-Biedl Syndrome Gene BBS9/PTHB1 Leads to Cilia Defects. *PLoS One, 7*(3). doi:ARTN e34389 10.1371/journal.pone.0034389

Wang, J., Vasaikar, S., Shi, Z., Greer, M., & Zhang, B. (2017). WebGestalt 2017: a more comprehensive, powerful, flexible and interactive gene set enrichment analysis toolkit. *Nucleic Acids Research, 45*(W1), W130-W137. doi:10.1093/nar/gkx356

Ware, S. M., Aygun, M. G., & Hildebrandt, F. (2011). Spectrum of clinical diseases caused by disorders of primary cilia. *Proc Am Thorac Soc, 8*(5), 444-450. doi:10.1513/pats.201103-025SD

Waters, A. M., & Beales, P. L. (2011). Ciliopathies: an expanding disease spectrum. *Pediatr Nephrol, 26*(7), 1039-1056. doi:10.1007/s00467-010-1731-7

Westerfield, M. (1993). The zebrafish book: a guide for the laboratory use of zebrafish (Brachydanio reriro). 300.

Wolff, C., Roy, S., Lewis, K. E., Schauerte, H., Joerg-Rauch, G., Kirn, A., . . . Ingham, P. W. (2004). iguana encodes a novel zinc-finger protein with coiled-coil domains essential for Hedgehog signal transduction in the zebrafish embryo. *Genes & Development, 18*(13), 1565-1576. doi:10.1101/gad.296004

Yamamoto, M., Matsuzaki, T., Takahashi, R., Adachi, E., Maeda, Y., Yamaguchi, S., . . . Noda, M. (2012). The transformation suppressor gene Reck is required for postaxial patterning in mouse forelimbs. *Biol Open, 1*(5), 458-466. doi:10.1242/bio.2012638

Yamoto, K., Saitsu, H., Nakagawa, N., Nakajima, H., Hasegawa, T., Fujisawa, Y., . . . Ogata, T. (2017). De novo IGF2 mutation on the paternal allele in a patient with Silver-Russell syndrome and ectrodactyly. *Hum Mutat, 38*(8), 953-958. doi:10.1002/humu.23253

Zhang, W., Taylor, S. P., Ennis, H. A., Forlenza, K. N., Duran, I., Li, B., . . . Cohn, D. H. (2018). Expanding the genetic architecture and phenotypic spectrum in the skeletal ciliopathies. *Hum Mutat, 39*(1), 152-166. doi:10.1002/humu.23362

Zhang, Z., Hao, C. J., Li, C. G., Zang, D. J., Zhao, J., Li, X. N., . . . Li, W. (2014). Mutation of SLC35D3 causes metabolic syndrome by impairing dopamine signaling in striatal D1 neurons. *PLoS Genet, 10*(2), e1004124. doi:10.1371/journal.pgen.1004124

Zhao, C. T., & Malicki, J. (2007). Genetic defects of pronephric cilia in zebrafish. *Mechanisms of Development, 124*(7-8), 605-616. doi:10.1016/j.mod.2007.04.004

	Family A	Family B			Family C	
	A.IV-1	B.II-1	B.II-2	B.II-3	C.IV-3	C.IV-4
IQCE	p.[Val301Serfs*8];[Val301Serfs*8]	p.[Val301Serfs*8];[Glu451fs*15]	p.[Val301Serfs*8];[Glu451fs*15]	p.[Val301Serfs*8];[Glu451fs*15]	p.[Val301Serfs*8];[Val301Serfs*8]	p.[Val301Serfs*8];[=]
Other gene	*TULP1* p.[Arg400Trp];[Arg400Trp]					*ATP6V1B1* c.[175-1G>C];[175-1G>C]
Clinical features						
Hands and feet abnormalities						
Postaxial polydactyly	Y	Y	Y	Y	Y	Y
Hands	Y	Y	Y*	Y*	Y	Y
Feet	Y	Y	Y*	Y*	Y	Y
Brachydactyly	Y	Y	Y	Y	N	N
Syndactyly	Y	N	N	Y	N	N
Neurosensory disease						
Retinal dystrophy	Y	N	N	N	N	N
Deafness	N	N	N	N	N	Y
Pondostatural evaluation						
Overweight	Y	Y	Y	N	Y	N
Development psychomotor						
Intellectual abnormalities	N	N	N	N	N	N
Learning disabilities	N	N	N	N	Y	N
Speech delay	Y	N	N	N	Y	Y
Ataxia	N	N	N	N	N	N
Kidney function						
Renal abnormality	N	N	N	N	N	Tubulopathy - Nephrocalcinosis
Genital abnormalities						
Hypogonadism	Micropenis	N			N	N

Table 1. Clinical characteristic of individuals with variants in *IQCE.* Y=yes, N= not present, *= unilateral

Figure 1. Clinical description and molecular genetic characterization of *IQCE* variations. **(A)** Graphic representation of three autosomal recessive families with complicated polydactyly in whom *IQCE, TULP1* and *ATP6V1B1* variations were identified. Sanger sequencing traces confirming the presence of the pathogenic variant and segregation in available family members is demonstrated. The "WT" symbol represents wild type and "M" symbol represents pathogenic variant. Blackened symbols indicate affected individuals. **(B)** Genomic structure (NC_000007.14) and schematic representation of the protein layout (NP_689771.3) of IQCE and location of currently known disease-associated variant (in bold, novel pathogenic variants identified in this study). Topology of IQCE represents one hydrophobic segment (HS), coil-coil conserved domain regions, three IQ calmodulin binding motifs (IQ) and one acid region (AR).

Figure 2. (A). Immunostaining of A.IV-1 and control ciliated fibroblasts, antibodies against human EVC2 (in green) and the ciliary marker anti acetylated alpha tubulin (in red), revealed that in absence of endogenous IQCE, EVC2 fails to localize at the base of the cilia. **(B).** Count of mislocalized EVC2 at the base of the cilia in fibroblasts under ciliated conditions (–FCS) between A.IV-1 and control (3 times ~100 cells per condition). Statistical significance was determined using the Student's t-test, NS, non-significant, p=0.0004 (***). **(C).** Heat map generated using ClustVis (Metsalu & Vilo, 2015) from the transcriptome sequencing data reflecting gene counts of the significantly differentially expressed genes across patient and control samples. **(D).** Volcano plot of overrepresentation enrichment analysis of significantly misregulated transcripts in A.IV-1 fibroblast compared to wild type. These results were obtained using WebGestalt 2019, (Supp. Table 6). **(E).** Real-time quantitative polymerase chain reaction analysis from 3 technical replicates of A.IV-1 and 3 controls fibroblasts showed relative expression levels (normalized to GAPDH) of 9 significantly differentially expressed genes, data includes ±SD. NS= no significative, *= p-value < 0.05, **= p-value < 0.01, ***= p-value < 0.001 (alpha risk 5%). Scale bar 5 µm.

Figure 3. *iqce* morphants present ciliopathic phenotypes. **(A-B).** *iqce* morphants exhibit a curved body axis A: uninjected B: *iqce-mo* morphants, star indicates hydrocephalus. **(C).** Cysts in the paired pronephric glomeruli in *iqce* morphants (arrows). **(D).** Rescue of *iqce-mo* (6 µM) injected embryos with different concentrations of *iqce* mRNA. 10 ng/µl (n=121), 20 ng/µl (n=163), 30 ng/µl (n=193). The rescue was evaluated by the loss of curved body, data includes ± SD only when

needed (WT and the best condition (i.e. 30 ng/µl)). **(E-G).** Cilia in the pronephric duct of *iqce* morphants are misdirected, shortened, and malformed (F) whereas in uninjected (E) or control morpholinos (G) they are elongated. bar: 11.6 µm. **(H-J).** Cilia in nasal placode are reduced in number and size in morphants (I, n=5) compared to uninjected (H, n=5) and morpholino control (J, n=5). bar: 11.6 µm. Scheme: Square shows the area of observation for pronephric cilia, and circle for nasal placode cilia. **(K-L).** 5 µM sections through eyes of uninjected or *iqce-mo* injected embryos stained with toluidine blue (4 dpf). The uninjected embryos show normal lamination whereas morphants lack a properly formed retina. PRs: photoreceptors; ONL: outer nuclear layer; INL: inner nuclear layer; GCL: ganglion cell layer. Star indicates the outer segments of photoreceptors. Scale bar: 0.2 cm. **(M).** Chart indicating the percentage of normal and *situs inversus* or heterotaxia in uninjected (n=70), *iqce-mo* (n=250) or *iqce-mocont* (n=160) injected embryos, data includes ± SD. **(N-P)**. *patched1* (*ptc1*) *in situ* hybridization revealed expanded expression in *iqce* morphant (O, n=81) but not in uninjected (N, n=88) or control morpholino (P, n=30) injected embryos. **(Q-S)**. *In situ* immunohistochemistry showing expanded engrailed expression in *iqce* morphants (R, n=10) compared to uninjected (Q, n=10) and *iqce-mocont* (S, n=10) injected embryos. Scale bars: 47.2 µM. A-J: 48 hpf. K-M: 72 hpf. N-P: 10 somites. Q-S: 24 hpf.

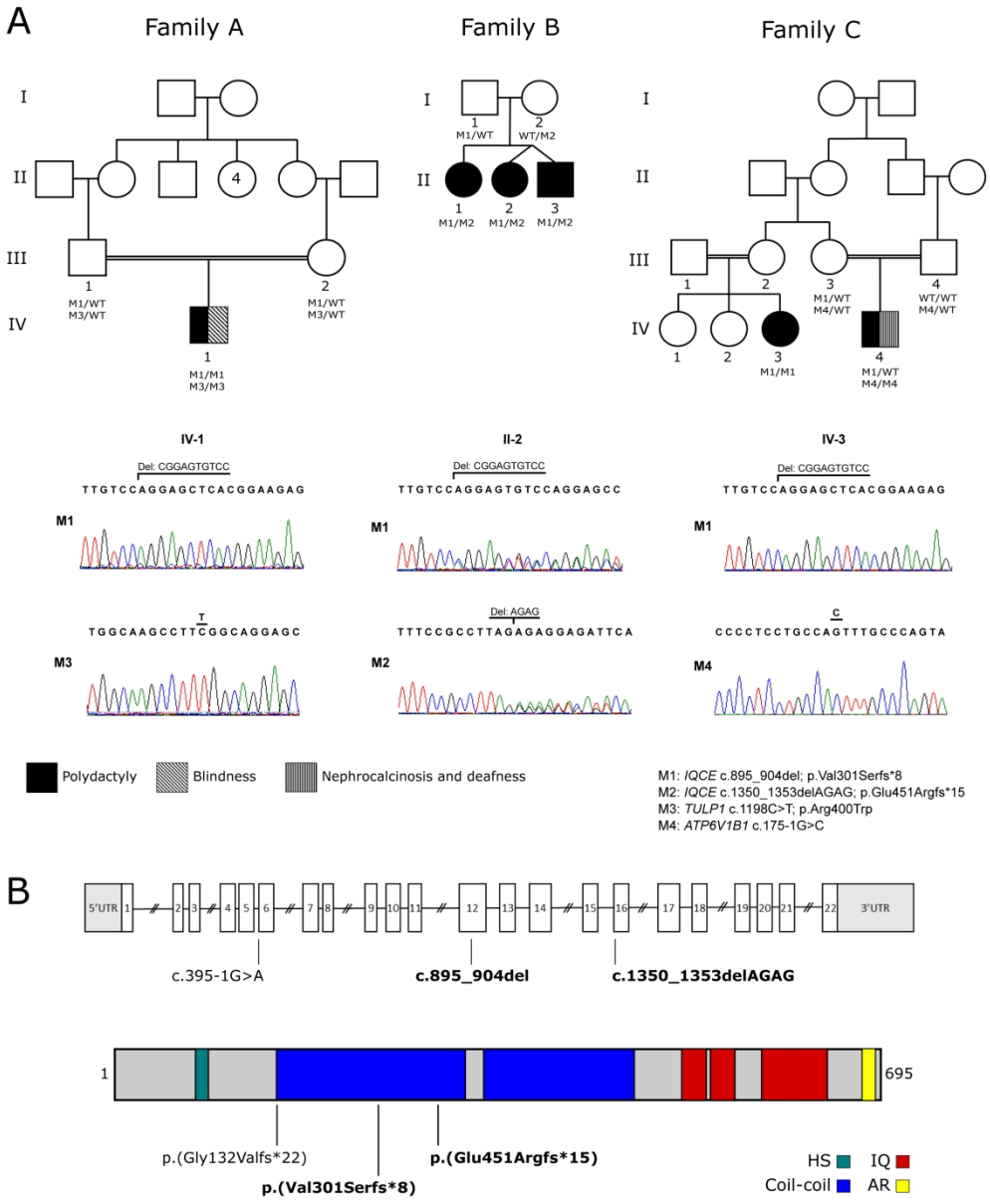


Figure 1. Clinical description and molecular genetic characterization of IQCE variations. (A) Graphic representation of three autosomal recessive families with complicated polydactyly in whom IQCE, TULP1 and ATP6V1B1 variations were identified. Sanger sequencing traces confirming the presence of the pathogenic variant and segregation in available family members is demonstrated. The "WT" symbol represents wild type and "M" symbol represents pathogenic variant. Blackened symbols indicate affected individuals. (B) Genomic structure (NC_000007.14) and schematic representation of the protein layout (NP_689771.3) of IQCE and location of currently known disease-associated variant (in bold, pathogenic variants identified in this study). Topology of IQCE represents one hydrophobic segment (HS), coil-coil conserved domain regions, three IQ calmodulin binding motifs (IQ) and one acid region (AR).

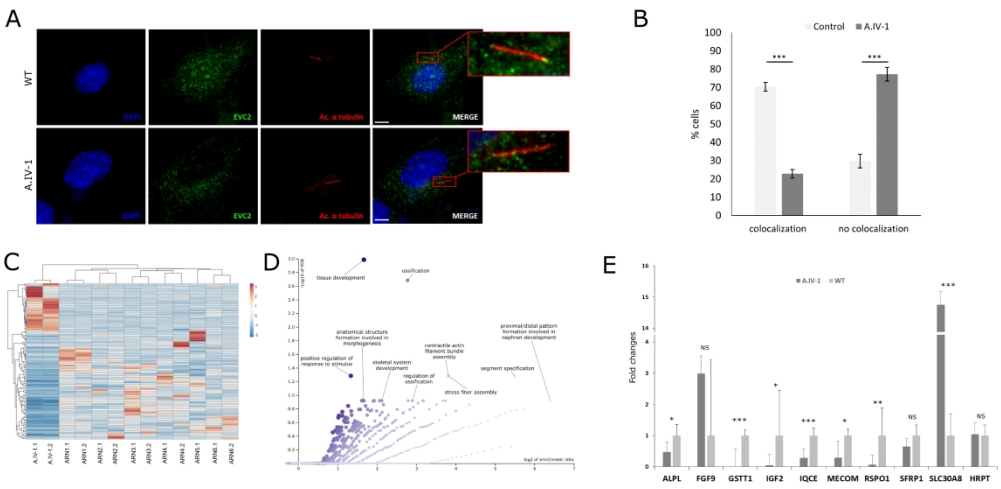


Figure 2. (A). Immunostaining of A.IV-1 and control ciliated fibroblasts, antibodies against human EVC2 (in green) and the ciliary marker anti acetylated alpha tubulin (in red), revealed that in absence of endogenous IQCE, EVC2 fails to localize at the base of the cilia. (B). Count of mislocalized EVC2 at the base of the cilia in fibroblasts under ciliated conditions (–FCS) between A.IV-1 and control (3 times ~100 cells per condition). Statistical significance was determined using the Student's t-test, NS, non-significant, p=0.0004 (***). (C). Heat map generated using ClustVis (Metsalu & Vilo, 2015) from the transcriptome sequencing data reflecting gene counts of the significantly differentially expressed genes across patient and control samples. (D). Volcano plot of overrepresentation enrichment analysis of significantly misregulated transcripts in A.IV-1 fibroblast compared to wild type. These results were obtained using WebGestalt 2019, (Supp. Table 6). (E). Real-time quantitative polymerase chain reaction analysis from 3 technical replicates of A.IV-1 and 3 controls fibroblasts showed relative expression levels (normalized to GAPDH) of 9 significantly differentially expressed genes, data includes ±SD. NS= no significative, *= p-value < 0.05, **= p-value < 0.01, ***= p-value < 0.001 (alpha risk 5%). Scale bar 5 µm.

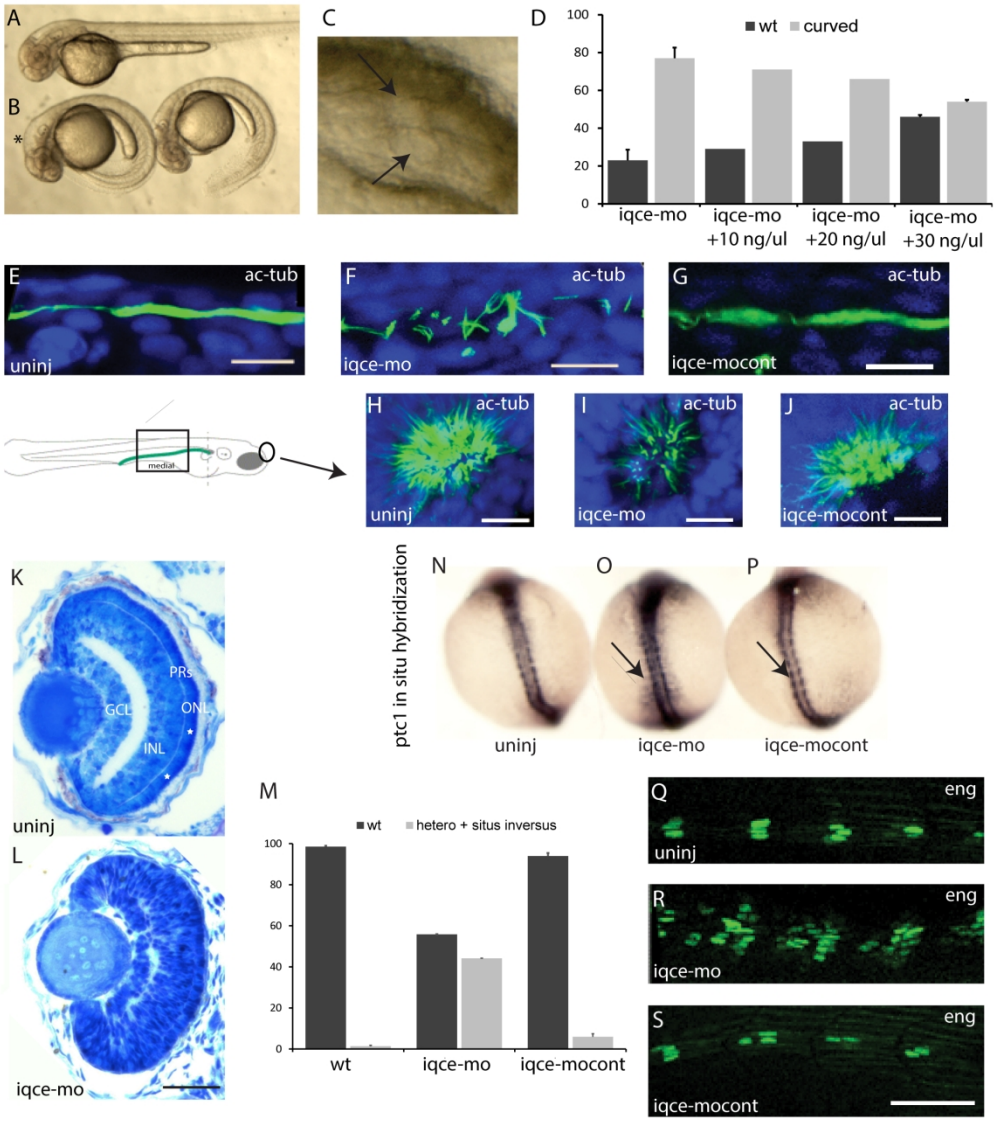


Figure 3. iqce morphants present ciliopathic phenotypes. (A-B). iqce morphants exhibit a curved body axis A: uninjected B: iqce-mo morphants, star indicates hydrocephalus. (C). Cysts in the paired pronephric glomeruli in iqce morphants (arrows). (D). Rescue of iqce-mo (6 µM) injected embryos with different concentrations of iqce mRNA. 10 ng/µl (n=121), 20 ng/µl (n=163), 30 ng/µl (n=193). The rescue was evaluated by the loss of curved body, data includes ± SD only when needed (WT and the best condition (i.e. 30 ng/µl)). (E-G). Cilia in the pronephric duct of iqce morphants are misdirected, shortened, and malformed (F) whereas in uninjected (E) or control morpholinos (G) they are elongated. bar: 11.6 µm. (H-J). Cilia in nasal placode are reduced in number and size in morphants (I, n=5) compared to uninjected (H, n=5) and morpholino control (J, n=5). bar: 11.6 µm. Scheme: Square shows the area of observation for pronephric cilia, and circle for nasal placode cilia. (K-L). 5 µM sections through eyes of uninjected or iqce-mo injected embryos stained with toluidine blue (4 dpf). The uninjected embryos show normal lamination whereas morphants lack a properly formed retina. PRs: photoreceptors; ONL: outer nuclear layer; INL: inner nuclear layer; GCL: ganglion cell layer. Star indicates the outer segments of photoreceptors. Scale bar: 0.2 cm. (M). Chart indicating the percentage of normal and situs inversus or heterotaxia in uninjected (n=70), iqce-mo (n=250) or iqce-mocont (n=160) injected embryos, data includes ± SD. (N-P). patched1 (ptc1) in situ hybridization revealed expanded expression in iqce morphant (O, n=81) but not in uninjected (N, n=88) or

control morpholino (P, n=30) injected embryos. (Q-S). In situ immunohistochemistry showing expanded engrailed expression in iqce morphants (R, n=10) compared to uninjected (Q, n=10) and iqce-mocont (S, n=10) injected embryos. Scale bars: 47.2 μM. A-J: 48 hpf. K-M: 72 hpf. N-P: 10 somites. Q-S: 24 hpf.

	Family A		Family B		Family C	
	A.IV-1	B.II-1	B.II-2	B.II-3	C.IV-3	C.IV-4
IQCE	p.[Val301Serfs*8];[Val301Serfs*8]	p.[Val301Serfs*8];[Glu451fs*15]	p.[Val301Serfs*8];[Glu451fs*15]	p.[Val301Serfs*8];[Glu451fs*15]	p.[Val301Serfs*8];[Val301Serfs*8]	p.[Val301Serfs*8];[=]
Other gene	*TULP1* p.[Arg400Trp];[Arg400Trp]					*ATP6V1B1* c.[175-1G>C];[175-1G>C]
Clinical features						
Hands and feet abnormalities						
Postaxial polydactyly	Y	Y	Y	Y	Y	Y
Hands	Y	Y	Y*	Y*	Y	Y
Feet	Y	Y	Y*	Y*	Y	Y
Brachydactyly	Y	Y	Y	Y	N	N
Syndactyly	Y	N	N	Y	N	N
Neurosensory disease						
Retinal dystrophy	Y	N	N	N	N	N
Deafness	N	N	N	N	N	Y
Pondostatural evaluation						
Overweight	Y	Y	Y	N	Y	N
Development psychomotor						
Intellectual abnormalities	N	N	N	N	N	N
Learning disabilities	N	N	N	N	Y	N
Speech delay	Y	N	N	N	Y	Y
Ataxia	N	N	N	N	N	N
Kidney function						
Renal abnormality	N	N	N	N	N	Tubulopathy - Nephrocalcinosis
Genital abnormalities						
Hypogonadism	Micropenis	N			N	N

Table 1. Clinical characteristic of individuals with variants in *IQCE.* Y=yes, N= not present, *= unilateral

	Family A		Family B		Family C	
	A.IV-1	B.II-1	B.II-2	B.II-3	C.IV-3	C.IV-4
IQCE	p.[Val301Serfs*8];[Val301Serfs*8]	p.[Val301Serfs*8];[Glu451fs*15]	p.[Val301Serfs*8];[Glu451fs*15]	p.[Val301Serfs*8];[Glu451fs*15]	p.[Val301Serfs*8];[Val301Serfs*8]	p.[Val301Serfs*8];[=]
Other gene	*TULP1* p.[Arg400Trp];[Arg400Trp]					*ATP6V1B1* c.[175-1G>C];[175-1G>C]
Clinical features						
Hands and feet abnormalities						
Postaxial polydactyly	Y	Y	Y	Y	Y	Y
Hands	Y	Y	Y*	Y*	Y	Y
Feet	Y	Y	Y*	Y*	Y	Y
Brachydactyly	Y	Y	Y	Y	N	N
Syndactyly	Y	N	N	Y	N	N
Neurosensory disease						
Retinal dystrophy	Y	N	N	N	N	N
Deafness	N	N	N	N	N	Y
Pondostatural evaluation						
Overweight	Y	Y	Y	N	Y	N
Development psychomotor						
Intellectual abnormalities	N	N	N	N	N	N
Learning disabilities	N	N	N	N	Y	N
Speech delay	Y	N	N	N	Y	Y
Ataxia	N	N	N	N	N	N
Kidney function						
Renal abnormality	N	N	N	N	N	Tubulopathy - Nephrocalcinosis
Genital abnormalities						
Hypogonadism	Micropenis	N			N	N

Table 1. Clinical characteristic of individuals with variants in *IQCE.* Y=yes, N= not present, *= unilateral

SUPPLEMENTARY INFORMATION

Sanger Sequencing

Sanger sequencing was performed by way of PCR amplification with 50ng of genomic DNA template. The primers were designed with Primer 3 (http://frodo.wi.mit.edu/primer3). Bidirectional sequencing of the purified PCR products was performed by Eurofins Genomics Sequencing Facilities (https://www.eurofinsgenomics.eu/en/custom-dna-sequencing).

Control screening

IQCE exon 12 was tested in 93 control DNA of matched origin (Algerian) and 93 multi ethnic. High resolution melting (HRM) assay was performed in a 20 µl reaction volume containing 50 ng of template DNA, 15 µl Precision Melt Supermix (Bio-Rad) and 10 µmol of each primer of the primer set. PCR was performed using a CFX96 (Bio-Rad) thermal cycler and consisted of predenaturation (98°C for 2 min) followed by 35 cycles of amplification involving denaturation (98°C for 5 s), annealing (60°C for 10 s) and a melting curve involving denaturation (95°C for 30 s), annealing (70°C for 30 s) and a melt curve (72–95°C with an increment of 0.2°C per 10 s). Data generated were analyzed using Precision Melt Analysis software (Bio-Rad). In HRM analysis, differences in Tm and normalized curve shape were used together to discriminate even the most difficult–to–detect sequence variation.

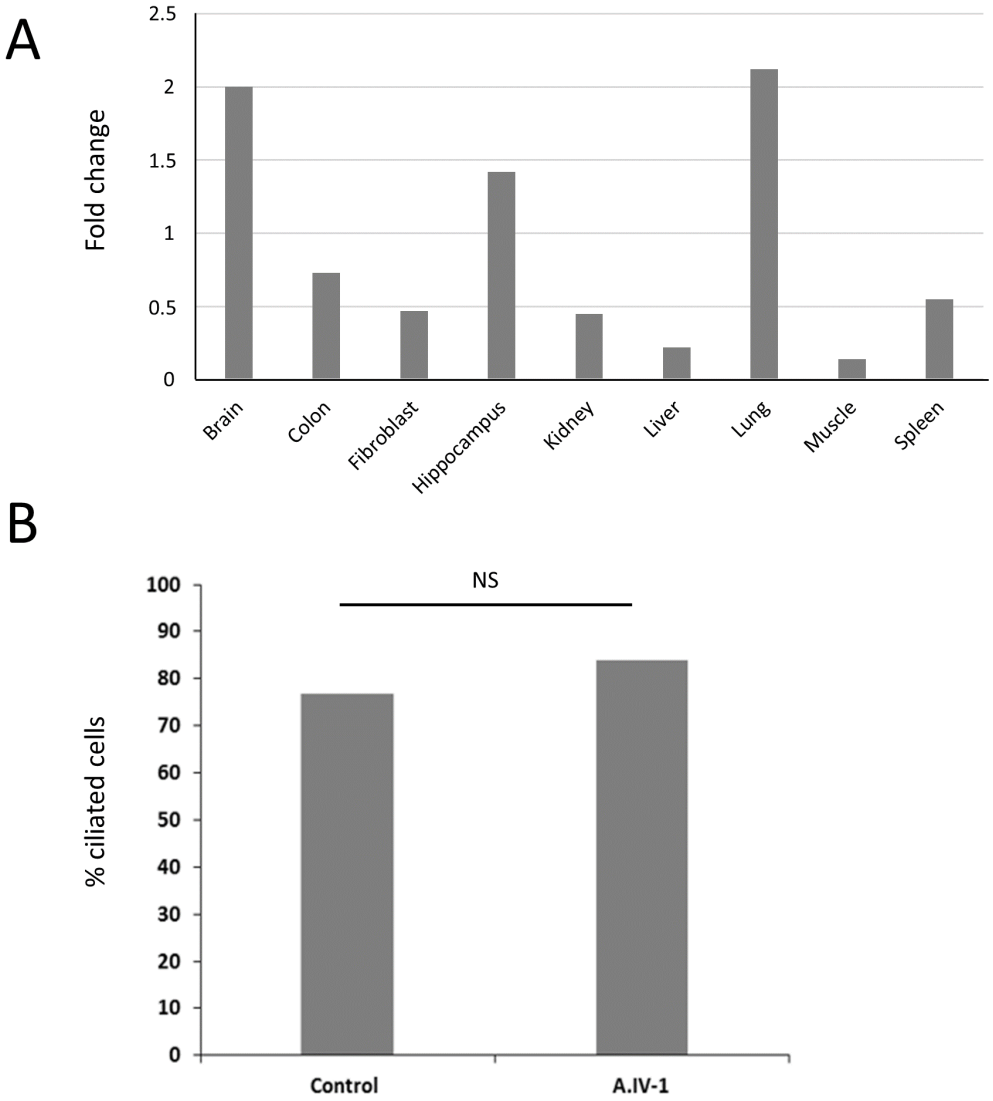
Bioinformatic pipeline RNA seq

Image analysis and base calling were performed using RTA v.2.7.3 and blc2fastq v.2.17.1.14 (Illumina). Sequence reads were mapped onto hg19/GRCh37 assembly of human genome using Tophat 2.0.14 (Kim et al., 2013) and bowtie version 2-2.1.0 (Langmead, Trapnell, Pop, & Salzberg, 2009), A summary of the mapping results is available in Supplementary Table 7. Gene expression was quantified from uniquely aligned reads using HTSeq-0.6.1 with gene annotations from Ensembl release 75 and intersection non-empty mode.

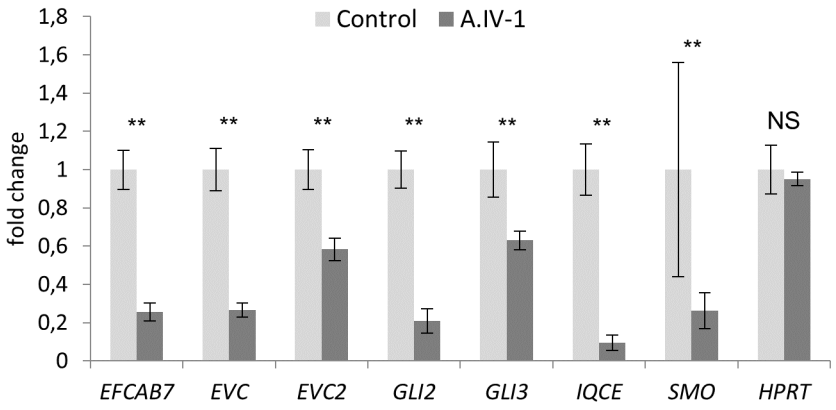
Differential expression RNA-seq

In order to identify variation in mRNA expression due to the absence of IQCE in the patient cells, comparisons of read counts were performed using R 3.2.5 with the statistical method proposed by Anders and Huber implemented in the DESeq2 v.1.10.1 Bioconductor package

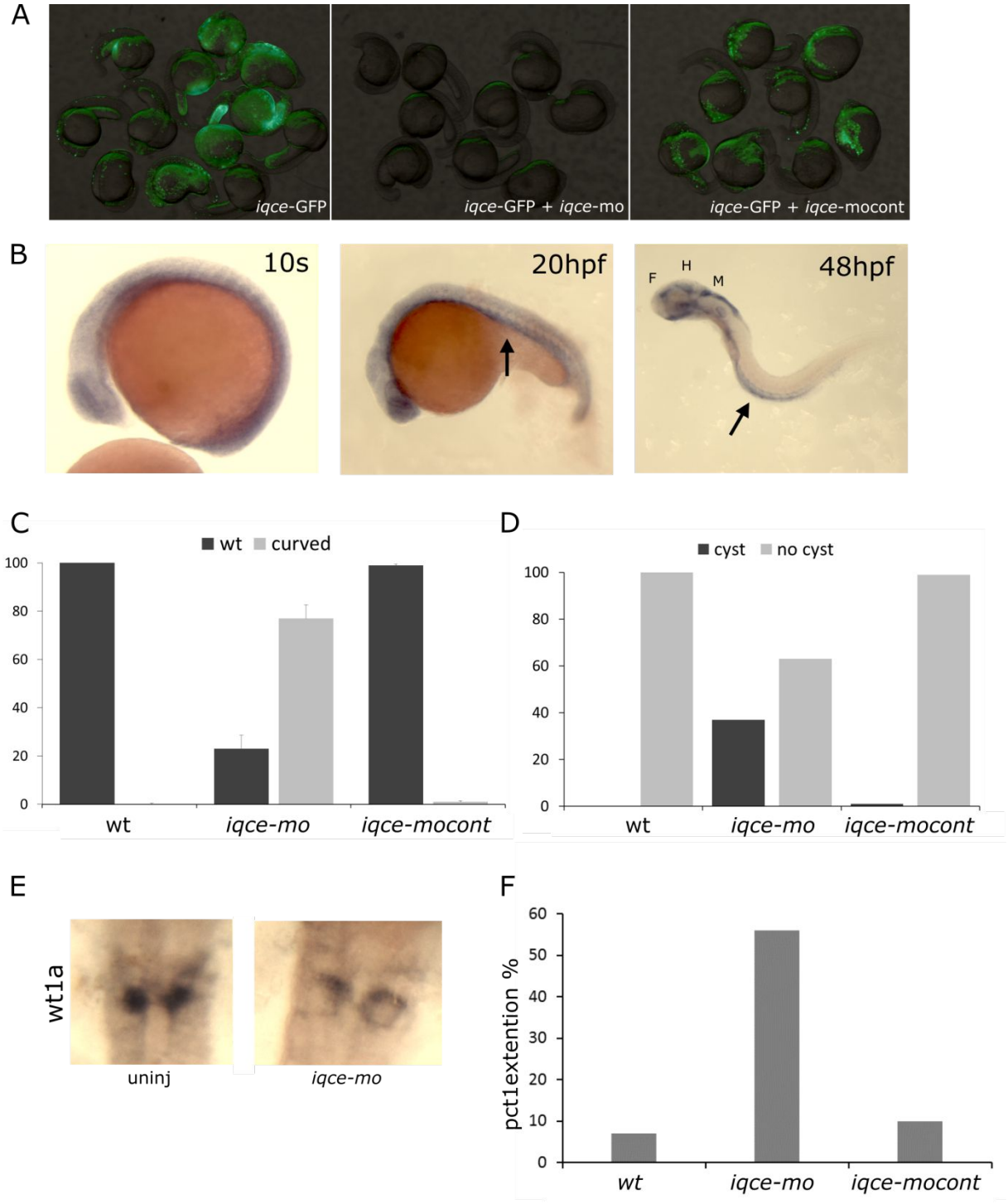
(Love, Huber, & Anders, 2014). The generated data were explored and visualized to assess and check data quality and to eventually remove bad quality data. Especially, PCA was built after stabilizing variance using the regularized log transformation method or RSeqQC was used to do some quality controls such as gene body coverage. Only genes with | log2 fold-change | > 0.5 were considered, as we expected to identify a large amount of change between transcriptomic results from one patient compared to other individuals (indirect consequences of the pathogenic mutation or unrelated events due to variability between humans). To avoid as much as possible false positive variants, we filtered out all genes with a Benjamini and Hochberg adjusted p-value above 0.05.



Supplementary Figure S1. **A.** Highest mRNA expression of *IQCE* in adult human brain/hippocampus and lung. **B.** Count of ciliated fibroblasts between A.IV-1 and control (~100 cells per condition). Percentages are shown in histogram, N.S.: non-significant, Fisher exact test 0.224, $P < 0.05$).



Supplementary Figure S2. Expression levels of EVC-zone and GLI genes upon ciliogenesis and Hh activation. Patient's (A.IV-1) and control cells upon treatment (–FCS + SAG 100nM) were compared (n=4 controls, n=6 patient). Statistical significance has been performed using the non-parametric Mann-Whitney test, p=0.0095 (**), NS : non-significant.



Supplementary Figure S3. Efficacy of *iqce* morpholinos. A. Morpholinos directed against the start codon of *iqce* blocks the expression of a construct containing *iqce* cDNA in frame with gfp (central panel). In contrast, 5 mismatch morpholinos did not reduce expression of the chimeric protein (left panel) compared to *iqce-gfp* alone (right panel). 24 hpf. **B.** *in situ* hybridization showing *iqce* expression pattern in 10 somites, 20 and 48 hpf embryos. Arrow indicates the pronephric duct. F: forebrain ventricule; H: hindbrain ventricule; M: midbrain ventricule. **C.** Chart showing the percentage of curved embryos in uninjected (n=200), *iqce-*

mo (n=120); and *iqce-mocont* (n=221) injected embryos. 48 hpf, data includes ± SD and statistical significance was determined using the Student's t-test, all comparisons were significant with a p-value<0.05. **D.** Chart showing the percentage of cysts in uninjected (n=101), *iqce-mo* (n=70) and *iqce-mocont* (n=90) injected embryos. Statistical significance was determined using the Student's t-test, all comparisons were significant with a p-value<0.05. **E.** wt1a *in situ* hybridization showing expression within glomeruli in non-injected embryos and at the periphery of the cyst in morphants (48 hpf). **F.** Chart showing the percentage of embryos with *ptc1* extension in 10 somites embryos (n=191).

Supplementary Table S1. Summary of the whole exome sequencing results. SNV: Single nucleotide variation, indel: gain or loss of up to 50 nucleotides at a single locus, SV: Structural Variation. Annotations are gathered using Alamut Batch v1.11, especially for the variation databases including gnomAD (v2.0.2, Oct. 2017), 1000 Genomes Project phase3 release (version 20150813 v5b) and the following predictions including phastCons (UCSC, 44 vertebrates). Effect on the splice has been evaluated using the MaxEntScan (Yeo & Burge, 2004), NNSPLICE 0.9 (Reese, Eeckman, Kulp, & Haussler, 1997) and Splice Site Finder (Shapiro & Senapathy, 1987) by calculating score change between the wild type and the mutated sequences expressed as a percent differences. Missenses have been evaluated using default parameters from PolyPhen-2 (2.2.2) (Adzhubei et al., 2010) and SIFT 4.0.3 (Kumar, Henikoff, & Ng, 2009). Default cut-offs used have been described in VaRank (Geoffroy et al., 2015) for both type of predictions. Exclusion of SV with a DGV (Gold standard from 20160515) frequency > 1% is done only with studies of more than 1000 individuals.

	A.IV-1			**B.II-2**			**B.II-3**			**C.IV-4**		
	SNV	**indel**	**SV**	**SNV**	**indel**	**SV**	**SNV**	**Indel**	**SV**	**SNV**	**Indel**	**SV**
Total number of variants	83129	11664	31	81972	11319	17	82116	11195	16	78906	11727	31
After exclusion of variants with an allele frequency >1% (gnomAD, 1000G, internal exome database, DGV)	9771	1125	16	7069	993	6	6856	943	9	5323	982	17
After exclusion of SNV/indel in 5'UTR, 3'UTR, downstream, upstream, intron and synonymous locations without local splice effect prediction	788	76	/	553	59	/	560	65	/	753	113	/
After exclusion of missense with no deleterious effect according to SIFT, PPH2 or PhastCons	669	76	/	458	59	/	465	65	/	641	113	/
After exclusion of variants not in a ciliary genes list (Nevers et al., 2017) and not in non-syndromic polydactyly known genes	52	6	1	31	2	2	40	4	2	78	15	1
After selection of variants consistent with recessive transmission (compound heterozygous, homozygous variants)	7 hom.: *DNAAF1*, ***IQCE***, *MADD, SHANK2,* ***TULP1***, *ZDHHC2* 3 comp. het.: *ARMC3, DNAH11, MGA*			0 homozygous variant 1 comp. het. In ***IQCE***						3 hom.: *ZMAT1, AGTPBP1*, *KDM4C* 11 comp. het: *CTDSPL2, DDX1, DNAH7, DYNLL1, EP400, IFT20, LRRCC1, LRRC46, MRPL10, PRKDC, SMARCA5, TMEM97, TPTE*		

Supplementary Table S2. Homozygous regions identified in individual A.IV-1.

Chr	SNP1	SNP2	POS1	POS2	MB	#SNP	DENSITY	
1	rs4147830	rs11431894	94544276	100680594	6.14	33	185.949	
1	rs4915053	rs3761934	108113526	114189067	6.08	136	44.673	
1	1_148741720_T_C	rs11428742	148741720	150550020	1.81	22	82.196	
1	rs4636	rs10927011	234744413	243579112	8.83	125	70.678	
2	rs7577088	rs3764906	39233500	42588195	3.35	20	167.735	
2	rs522893	2_188250301_T_A	169690729	188250301	18.56	224	82.855	
2	rs2304704	rs12477076	190430177	191848312	1.42	23	61.658	
3	rs901854	rs113574694	4856234	10089550	5.23	58	90.230	
3	rs9814557	rs2246945	135720540	137843476	2.12	21	101.092	
4	rs13146318	rs1051447	41615690	49063872	7.45	68	109.532	
4	rs4864727	rs12621	54011424	57897570	3.89	61	63.707	
4	rs10434219	rs344122	62598689	77637502	15.04	175	85.936	
6	rs2076506	6_32485853_C_T	22570064	32485853	9.92	646	15.350	
6	rs1059547	rs12189627	32557529	41059458	8.50	309	27.514	*TULP1*
7	rs4916935	rs798479	197267	2691201	2.49	41	60.828	*IQCE*
7	rs7794563	rs80024169	5347295	14775824	9.43	78	120.879	
7	rs4718101	rs6460315	64152353	66289040	2.14	38	56.229	
8	rs3989699	rs13271468	7218621	12236202	5.02	132	38.012	
8	8_12273335_G_A	rs66480340	12273335	24211885	11.94	202	59.102	
8	rs10086200	rs59331088	61193588	74005131	12.81	72	177.938	
8	rs11777189	rs12545587	99057150	136560952	37.50	204	183.842	
9	rs4977555	rs11515	19550308	21968199	2.42	30	80.596	
9	rs7863560	rs3217100	119495697	123734288	4.24	30	141.286	
11	11_32452000_GG_-	rs2076623	32452000	33564123	1.11	25	44.485	
11	rs12274095	rs7101792	69063767	71249386	2.19	28	78.058	
11	rs3740912	rs4937391	125891269	128786294	2.90	32	90.470	
12	rs11054683	rs1600	12247616	14656768	2.41	31	77.715	
12	rs3088008	rs3210837	49724955	51685831	1.96	55	35.652	
12	rs74651927	rs2242497	52695754	56992630	4.30	185	23.226	
12	rs1663564	rs7488309	105546172	118520170	12.97	152	85.355	
12	rs1794956	rs11059841	124848137	129190021	4.34	37	117.348	
14	rs34800262	rs2249922	51370852	52906081	1.54	34	45.154	
14	rs2297113	rs863091	58036494	59112475	1.08	23	46.782	
14	rs35533709	rs2057482	60433392	62213848	1.78	30	59.349	
15	rs4923929	rs555001	42439376	43545728	1.11	37	29.901	
15	rs11351249	rs2623989	72105929	73545732	1.44	23	62.600	
15	rs752270	rs4777755	87099537	93510603	6.41	171	37.492	
16	rs6564764	rs931713	80581631	85023855	4.44	121	36.713	
17	rs8071836	rs67749219	47238028	49711042	2.47	40	61.825	

Supplementary Table S3. List and characteristics of primers used in this study

Application	Gene	Exon	Forward (5'-3')	Reverse (5'-3')	Size (bp)
DNA	*IQCE*	12	CTCTGAGCTCACAACCAACAG	GTTGCCCAAGGGTTCACA	399
		16	TGGTTTCTGGTGTTGTCAGG	TTTTCAGTGGCTGGTCAGAA	382
	ATP6V1B1	3	AGCCGTGGGAAGTAGGTGTT	ATGTCCCATTGACCCCACT	385
	TULP1	12/13	GGGATGTAGGATCCCCTCA	CATGCCAGGAATGATGACG	481
cDNA	*IQCE*	3-4/5	ACATCGCCAAAGTCACCTTATC	AGGGCCTGGGTCAGACTT	158
	GSTT1	2-3/3	ACCTTGACGGAGAGTGTGG	GCAGCTTCTCCGCAGAGT	144
	SLC30A8	9-10/10	CCATCTTACTCATGGAAGGTGTG	GCTGTAGCAACATGAGCTGAGA	154
	RSPO1	4/5	TGGAGAGGAACGACATCCG	CCTTACACTTGGTGCAGAAGTTA	156
	FGF9	2/3	GGCCTGGTCAGCATTCGAG	GTATCGCCTTCCAGTGTCCAC	180
	SFRP1	1/2	GCTTCTACTGGCCCGAGATG	TGGCCTCAGATTTCAACTCGT	180
	MECOM	4/5	AAACTCGAAAGCGAGAATGATCT	TGGTGGCGAATTAAATTGGACTT	185
	ALPL	6-7/7	AACATCAGGGACATTGACGTG	GTATCTCGGTTTGAAGCTCTTCC	159
	IGF2	2-3/3-4	TTCTACTTCAGCAGGCCCG	GAAGTTGTCCGGAAGCACG	168
	GAPDH	4-5/6	GGAGCGAGATCCCTCCAAAAT	GGCTGTTGTCATACTTCTCATGG	197
	HPRT1	1-2/2-3	CCTGGCGTCGTGATTAGTGAT	AGACGTTCAGTCCTGTCCATAA	131
	EFCAB7	6/6-7	GTCAGCAACCAGGAAGTTCA	CATGTGTTGCCAGTCCTTTATT	115
	EVC2	6-7/8	CTCGTCACGGAACAGAACAC	CAGGTCAGCACAAGGGAGAG	135
	EVC	2/4	TCGAAGGACAAGGAAGCTGT	TTTCATGCAGAGACGGGTTG	148
	GLI2	2/3	AAGCAAGAAGCCAAAAGTGG	TGGTACCTTCCTTCCTGGTG	188
	GLI3	2/3	GGCATTTTTGGTCGAAGAGA	GGACATTCTGTGGCTGCATA	238
	SMO	2/3	TGCCCAAGTGTGAGAATGAC	TACCAGCTCTTGGGGTTGTC	229

Supplementary Table S4. Comparison of haplotypes of individuals A.IV-1 and B.II-2.

ID	Gene	Chr	Position	Allele_1_A.IV-1	Allele_2_A.IV-1	Allele_1_II-2_Family2	Allele_1_II-2_Family2
7_2255974_C_T	*MAD1L1*	7	2255974	C	C	C	T
7_2256017_C_A	*MAD1L1*	7	2256017	C	C	C	A
7_2257612_G_A	*MAD1L1*	7	2257612	A	A	A	G
7_2259134_C_T	*MAD1L1*	7	2259134	C	C	C	T
7_2260715_G_T	*MAD1L1*	7	2260715	T	T	T	T
7_2269552_G_T	*MAD1L1*	7	2269552	T	T	T	G
7_2278663_T_A	*FTSJ2*	7	2278663	T	T	A	A
7_2279482_A_C	*FTSJ2*	7	2279482	A	A	C	C
7_2279851_T_C	*FTSJ2*	7	2279851	T	T	C	C
7_2281754_C_T	*FTSJ2*	7	2281754	C	C	C	T
7_2290328_C_T	*SNX8*	7	2290328	C	C	C	T
7_2290336_T_G	*SNX8*	7	2290336	T	T	T	G
7_2290372_C_T	*SNX8*	7	2290372	C	C	T	T
7_2290522_C_T	*SNX8*	7	2290522	T	T	C	C
7_2292881_-_TG	*SNX8*	7	2292881	-	-	TG	TG
7_2296493_A_G	*SNX8*	7	2296493	G	G	G	G
7_2297006_A_G	*SNX8*	7	2297006	A	A	G	G
7_2303000_G_A	*SNX8*	7	2303000	A	A	G	G
7_2303109_G_A	*SNX8*	7	2303109	G	G	G	A
7_2303986_T_C	*SNX8*	7	2303986	C	C	C	C
7_2304191_A_C	*SNX8*	7	2304191	A	A	C	C
7_2309159_-_A	*SNX8*	7	2309159	-	-	-	A
7_2309174_G_A	*SNX8*	7	2309174	G	G	G	A
7_2309185_G_A	*SNX8*	7	2309185	G	G	G	A
7_2309209_C_T	*SNX8*	7	2309209	C	C	C	T
7_2314652_G_A	*SNX8*	7	2314652	G	G	G	A
7_2315008_G_C	*SNX8*	7	2315008	G	G	C	C
7_2318254_ACTC_-	*SNX8*	7	2318254	-	-	ACTC	ACTC
7_2318268_C_T	*SNX8*	7	2318268	T	T	C	C
7_2318463_C_T	*SNX8*	7	2318463	C	T	C	C
7_2318472_C_A	*SNX8*	7	2318472	C	A	C	C
7_2349470_T_C	*SNX8*	7	2349470	T	T	T	C
7_2394991_C_T	*EIF3B*	7	2394991	C	C	C	T
7_2400306_T_C	*EIF3B*	7	2400306	C	C	C	T
7_2405885_C_A	*EIF3B*	7	2405885	A	A	A	C
7_2409035_C_T	*EIF3B*	7	2409035	C	C	C	T
7_2414142_A_G	*EIF3B*	7	2414142	G	G	G	A
7_2414365_A_G	*EIF3B*	7	2414365	A	A	A	G
7_2416822_A_G	*EIF3B*	7	2416822	A	A	A	G
7_2418645_A_G	*EIF3B*	7	2418645	A	A	A	G
7_2433729_C_A	NA	7	2433729	C	C	C	A
7_2434522_G_A	NA	7	2434522	G	G	A	A
7_2434596_TT_-	NA	7	2434596	TT	TT	TT	-
7_2434597_T_-	NA	7	2434597	T	T	T	-

7_2472429_C_A	*CHST12*	7	2472429	C	C	C	A
7_2472455_A_T	*CHST12*	7	2472455	A	A	A	T
7_2489311_C_G	NA	7	2489311	C	C	C	G
7_2489336_GTGTGTGCACCT_-	NA	7	2489336	GTGTGTGCACCT	GTGTGTGCACCT	GTGTGTGCACCT	-
7_2489382_A_G	NA	7	2489382	A	A	A	G
7_2515382_-_C	NA	7	2515382	-	-	C	C
7_2517473_C_T	NA	7	2517473	C	C	C	T
7_2552986_A_C	*LFNG*	7	2552986	A	A	A	C
7_2553030_T_C	*LFNG*	7	2553030	T	T	T	C
7_2557266_T_C	*LFNG*	7	2557266	T	T	C	C
7_2557377_T_C	*LFNG*	7	2557377	T	T	T	C
7_2565268_C_A	*LFNG/MIR4648*	7	2565268	C	C	A	A
7_2566433_G_A	*LFNG/MIR4648*	7	2566433	G	G	G	A
7_2577691_C_G	*BRAT1*	7	2577691	C	C	C	G
7_2577781_T_C	*BRAT1*	7	2577781	T	T	T	C
7_2578237_T_C	*BRAT1*	7	2578237	T	T	T	C
7_2578238_C_T	*BRAT1*	7	2578238	C	C	C	T
7_2578434_G_T	*BRAT1*	7	2578434	G	G	G	T
7_2578455_G_A	*BRAT1*	7	2578455	G	G	G	A
7_2580914_C_T	*BRAT1*	7	2580914	C	C	C	T
7_2581994_T_C	*BRAT1*	7	2581994	T	T	T	C
7_2582154_-_CCCCAGCCTCCCGGGTGT	*BRAT1*	7	2582154	CCCCAGCCTCCCGGGTGT	CCCCAGCCTCCCGGGTGT	-	-
7_2582233_A_G	*BRAT1*	7	2582233	A	A	A	G
7_2582381_T_A	*BRAT1*	7	2582381	T	T	T	A
7_2582584_G_A	*BRAT1*	7	2582584	G	G	G	A
7_2583165_G_A	*BRAT1*	7	2583165	G	G	G	A
7_2583328_C_T	*BRAT1*	7	2583328	C	C	C	T
7_2584523_T_C	*BRAT1*	7	2584523	T	T	T	C
7_2587122_T_C	*BRAT1*	7	2587122	T	T	T	C
7_2598757_T_C	*IQCE*	7	2598757	T	T	T	C
7_2604768_GTGT_-	*IQCE*	7	2604768	GTGT	GTGT	GTGT	-
7_2604769_T_*	*IQCE*	7	2604769	T	T	T	*
7_2604816_C_T	*IQCE*	7	2604816	C	C	C	T
7_2611878_T_C	*IQCE*	7	2611878	T	T	T	C
7_2612225_G_A	*IQCE*	7	2612225	A	A	G	G
7_2613042_G_C	*IQCE*	7	2613042	G	G	G	C
7_2618070_-_A	*IQCE*	7	2618070	-	-	-	A
7_2619341_GT_-	*IQCE*	7	2619341	GT	GT	GT	-
7_2619363_G_C	*IQCE*	7	2619363	G	G	G	C
7_2623351_G_A	*IQCE*	7	2623351	G	G	G	A
7_2623801_G_T	*IQCE*	7	2623801	G	G	G	T
7_2625907_TGTCCCGGAG_-	*IQCE*	7	2625907	-	-	-	TGTCCCGGAG
7_2626039_CAGGGAATGG_-	*IQCE*	7	2626039	CAGGGAATGG	CAGGGAATGG	CAGGGAATGG	-
7_2627310_T_A	*IQCE*	7	2627310	T	T	T	A
7_2627579_AG_-	*IQCE*	7	2627579	AG	AG	-	-
7_2632477_A_G	*IQCE*	7	2632477	A	A	A	G

7_2634518_AGAG_-	*IQCE*	7	2634518	AGAG	AGAG	AGAG	-
7_2636689_C_T	*IQCE*	7	2636689	C	C	C	T
7_2638347_C_G	*IQCE*	7	2638347	C	C	C	G
7_2640985_G_C	*IQCE*	7	2640985	G	G	G	C
7_2647048_C_G	*IQCE*	7	2647048	C	C	C	G
7_2691201_C_A	*TTYH3*	7	2691201	A	A	C	C
7_2699687_-_G	*TTYH3*	7	2699687	-	-	G	G
7_2701809_T_C	*TTYH3*	7	2701809	T	C	T	T
7_2742097_T_C	*AMZ1*	7	2742097	C	C	T	T
7_2749565_G_A	*AMZ1*	7	2749565	G	G	G	A
7_2752152_G_A	*AMZ1*	7	2752152	G	G	G	A
7_2752487_G_A	*AMZ1*	7	2752487	A	A	G	G
7_2773035_-_AT	*GNA12*	7	2773035	-	-	AT	AT
7_2802173_T_C	*GNA12*	7	2802173	T	T	T	C
7_2802522_T_C	*GNA12*	7	2802522	T	T	T	C
7_2834869_C_T	*GNA12*	7	2834869	C	C	T	T
7_2854010_G_A	*GNA12*	7	2854010	G	G	A	A
7_2946461_T_C	*CARD11*	7	2946461	C	C	C	T
7_2952912_G_A	*CARD11*	7	2952912	G	G	G	A
7_2957005_T_C	*CARD11*	7	2957005	C	C	C	T
7_2958298_C_T	*CARD11*	7	2958298	C	C	T	T
7_2962241_G_A	*CARD11*	7	2962241	A	A	G	G
7_2962753_G_A	*CARD11*	7	2962753	G	G	G	A
7_2966445_T_G	*CARD11*	7	2966445	T	T	T	G
7_2966466_GT_-	*CARD11*	7	2966466	GT	GT	GT	-
7_2968195_G_A	*CARD11*	7	2968195	G	G	G	A
7_2968361_G_A	*CARD11*	7	2968361	G	G	G	A
7_2968486_G_C	*CARD11*	7	2968486	G	G	G	C
7_2985364_AC_-	*CARD11*	7	2985364	A	C	-	-

Supplementary Table S5. Differentially expressed transcripts in A.IV-1 fibroblasts.

Gene name	Ensembl gene id	Description	log2(FC)	P-value	Adjusted p-value
ABCA8	ENSG00000141338	ATP-binding cassette, sub-family A (ABC1), member 8	-4.205829134	3.52E-05	0.019362186
AC011294.3	ENSG00000233539	Uncharacterized protein	3.130226421	2.55E-07	0.000427239
ACCS	ENSG00000110455	1-aminocyclopropane-1-carboxylate synthase homolog (Arabidopsis)(non-functional)	-0.953037549	1.28E-07	0.00025011
ADAMTSL4	ENSG00000143382	ADAMTS-like 4	-3.153983754	5.00E-07	0.000732904
ALDH1A1	ENSG00000165092	aldehyde dehydrogenase 1 family, member A1	-6.206001924	1.99E-08	5.40E-05
ALPL	ENSG00000162551	alkaline phosphatase, liver/bone/kidney	-3.039649649	0.000113858	0.038536558
APBA2	ENSG00000034053	amyloid beta (A4) precursor protein-binding, family A, member 2	-2.732442847	6.79E-05	0.028790694
ARHGAP28	ENSG00000088756	Rho GTPase activating protein 28	2.800539497	1.58E-05	0.010684692
C10orf54	ENSG00000107738	chromosome 10 open reading frame 54	-1.64124697	9.82E-05	0.035643913
C1QTNF9B-AS1	ENSG00000205861	C1QTNF9B antisense RNA 1	-2.047191404	3.88E-05	0.019780947
C2CD4A	ENSG00000198535	C2 calcium-dependent domain containing 4A	5.96982663	8.95E-07	0.001085914
C4A	ENSG00000244731	complement component 4A (Rodgers blood group)	-6.329532092	5.21E-11	3.66E-07
CCDC180	ENSG00000197816	coiled-coil domain containing 180	2.351169896	3.81E-05	0.019780947
CD200	ENSG00000091972	CD200 molecule	-4.053753384	0.000135641	0.044622222
CHI3L1	ENSG00000133048	chitinase 3-like 1 (cartilage glycoprotein-39)	4.852100597	6.43E-07	0.000837885
CHN1	ENSG00000128656	chimerin 1	-1.617295158	2.25E-09	1.32E-05
CHRDL2	ENSG00000054938	chordin-like 2	-4.241249176	0.000101018	0.03591767
CNTN1	ENSG00000018236	contactin 1	-4.135254047	8.76E-05	0.032879001
COL11A1	ENSG00000060718	collagen, type XI, alpha 1	-4.558619592	6.30E-05	0.027357715
CSF2RA	ENSG00000198223	colony stimulating factor 2 receptor, alpha, low-affinity (granulocyte-macrophage)	3.072895701	6.90E-05	0.028906369
CYP21A1P	ENSG00000204338	cytochrome P450, family 21, subfamily A, polypeptide 1 pseudogene	-4.61429374	6.95E-06	0.005439088
DUSP2	ENSG00000158050	dual specificity phosphatase 2	3.527490256	1.70E-05	0.01128586
ECM1	ENSG00000143369	extracellular matrix protein 1	-1.261675626	5.81E-07	0.000786019
EMCN	ENSG00000164035	endomucin	3.361027658	4.95E-05	0.0229427
EYA4	ENSG00000112319	eyes absent homolog 4 (Drosophila)	-4.200503943	4.26E-05	0.020550766
FAM21B	ENSG00000152726	family with sequence similarity 21, member B	-4.995162913	7.93E-08	0.000164108
FGF9	ENSG00000102678	fibroblast growth factor 9	4.555259473	5.95E-08	0.000133394
FOSB	ENSG00000125740	FBJ murine osteosarcoma viral oncogene homolog B	3.628551087	8.80E-05	0.032879001
FYB	ENSG00000082074	FYN binding protein	-4.90648728	3.03E-05	0.016914087
GRPR	ENSG00000126010	gastrin-releasing peptide receptor	-4.422707449	0.000144048	0.046095412
GSTT1	ENSG00000184674	glutathione S-transferase theta 1	-9.843482395	6.94E-60	2.44E-55
HIVEP3	ENSG00000127124	human immunodeficiency virus type I enhancer binding protein 3	-1.599082154	6.64E-05	0.028486515
HOXC10	ENSG00000180818	homeobox C10	-4.942455716	2.36E-05	0.014065434
HOXC-AS3	ENSG00000251151	HOXC cluster antisense RNA 3	-5.14737488	6.01E-05	0.026775571
IGF2	ENSG00000167244	insulin-like growth factor 2 (somatomedin A)	-4.063105964	3.83E-05	0.019780947
IGFBP1	ENSG00000146678	insulin-like growth factor binding protein 1	2.926653654	1.58E-05	0.010684692

IGFBP3	ENSG00000146674	insulin-like growth factor binding protein 3	1.932377253	2.71E-06	0.002574523
IL17RB	ENSG00000056736	interleukin 17 receptor B	3.133016005	8.91E-05	0.032879001
IL20RA	ENSG00000016402	interleukin 20 receptor, alpha	3.104556086	3.12E-06	0.002749475
IQCE	ENSG00000106012	IQ motif containing E	-2.198435274	1.75E-43	3.08E-39
IRF6	ENSG00000117595	interferon regulatory factor 6	-3.931349479	7.74E-05	0.030946745
IRX1	ENSG00000170549	iroquois homeobox 1	1.496134127	2.70E-05	0.015596484
IRX2	ENSG00000170561	iroquois homeobox 2	1.161748374	4.51E-09	1.77E-05
JAZF1-AS1	ENSG00000234336	JAZF1 antisense RNA 1	-2.132735824	0.000107563	0.037119743
KLKB1	ENSG00000164344	kallikrein B, plasma (Fletcher factor) 1	2.478417669	8.84E-05	0.032879001
LAMC1	ENSG00000135862	laminin, gamma 1 (formerly LAMB2)	-1.201150815	2.84E-07	0.000453983
LHX9	ENSG00000143355	LIM homeobox 9	4.298162126	1.45E-06	0.001650048
LINC00537	ENSG00000232815	long intergenic non-protein coding RNA 537	-3.494641854	3.83E-05	0.019780947
LINC00578	ENSG00000228221	long intergenic non-protein coding RNA 578	-4.541164267	7.46E-09	2.39E-05
LINC01048	ENSG00000230390	long intergenic non-protein coding RNA 1048	2.214309895	3.99E-05	0.019780947
LINC01117	ENSG00000224577	long intergenic non-protein coding RNA 1117	-2.17227763	2.17E-05	0.01316806
LRRC16A	ENSG00000079691	leucine rich repeat containing 16A	-1.784218256	2.43E-05	0.014245821
MAP6	ENSG00000171533	microtubule-associated protein 6	-1.79656788	1.30E-06	0.001530605
MAP7D2	ENSG00000184368	MAP7 domain containing 2	2.359306267	2.66E-06	0.002574523
MBP	ENSG00000197971	myelin basic protein	-3.402601353	7.14E-05	0.029220727
MCCC1	ENSG00000078070	methylcrotonoyl-CoA carboxylase 1 (alpha)	-0.554719912	0.000128238	0.042584679
MECOM	ENSG00000085276	MDS1 and EVI1 complex locus	-4.511721791	3.14E-09	1.58E-05
METTL9	ENSG00000197006	methyltransferase like 9	0.906519962	3.86E-08	9.70E-05
MFGE8	ENSG00000140545	milk fat globule-EGF factor 8 protein	-2.694104724	7.44E-07	0.000934855
MLLT3	ENSG00000171843	myeloid/lymphoid or mixed-lineage leukemia (trithorax homolog, Drosophila); translocated to, 3	0.969013225	2.10E-05	0.01316806
MRVI1-AS1	ENSG00000177112	MRVI1 antisense RNA 1	-3.884388307	1.04E-05	0.007509783
MYO16	ENSG00000041515	myosin XVI	-1.962481605	0.000117031	0.039233207
NDP	ENSG00000124479	Norrie disease (pseudoglioma)	3.545635871	1.83E-05	0.011727901
NR5A2	ENSG00000116833	nuclear receptor subfamily 5, group A, member 2	-4.138425423	8.61E-06	0.006447287
NTRK3	ENSG00000140538	neurotrophic tyrosine kinase, receptor, type 3	2.777060262	2.95E-05	0.016740854
PKIA	ENSG00000171033	protein kinase (cAMP-dependent, catalytic) inhibitor alpha	-2.779337899	4.26E-05	0.020550766
PNLDC1	ENSG00000146453	poly(A)-specific ribonuclease (PARN)-like domain containing 1	2.95490495	2.27E-07	0.000399745
PNMA6C	ENSG00000235961	paraneoplastic Ma antigen family member 6C	-4.033899706	0.000151809	0.047958078
POMZP3	ENSG00000146707	POM121 and ZP3 fusion	-1.717492119	3.12E-06	0.002749475
POSTN	ENSG00000133110	periostin, osteoblast specific factor	2.83484857	2.14E-05	0.01316806
PRKAR1B	ENSG00000188191	protein kinase, cAMP-dependent, regulatory, type I, beta	-1.732349926	1.78E-05	0.011579369
PRLR	ENSG00000113494	prolactin receptor	3.300099077	0.000140582	0.045398939
PTK2B	ENSG00000120899	protein tyrosine kinase 2 beta	-1.47679232	6.10E-05	0.026827045
RGCC	ENSG00000102760	regulator of cell cycle	-3.93550099	7.06E-05	0.029220727
RHBDD3	ENSG00000100263	rhomboid domain containing 3	-0.996433377	1.05E-05	0.007509783
RSPO1	ENSG00000169218	R-spondin 1	-4.313022549	5.98E-05	0.026775571
S1PR1	ENSG00000170989	sphingosine-1-phosphate receptor 1	-3.093512898	5.52E-06	0.00452094

SCN1A	ENSG00000144285	sodium channel, voltage-gated, type I, alpha subunit	-4.906276861	3.80E-06	0.003184547
SCN9A	ENSG00000169432	sodium channel, voltage-gated, type IX, alpha subunit	-3.142530704	1.83E-06	0.001898572
SDK1	ENSG00000146555	sidekick cell adhesion molecule 1	3.188897352	4.55E-05	0.021372366
SFRP1	ENSG00000104332	secreted frizzled-related protein 1	-4.10661345	5.21E-07	0.000732904
SHE	ENSG00000169291	Src homology 2 domain containing E	3.455259022	1.63E-06	0.001733572
SLA	ENSG00000155926	Src-like-adaptor	-2.701260288	3.99E-05	0.019780947
SLC16A11	ENSG00000174326	solute carrier family 16, member 11	-2.677392243	4.37E-05	0.020765473
SLC30A8	ENSG00000164756	solute carrier family 30 (zinc transporter), member 8	6.104148946	1.35E-07	0.000250798
SLC35D3	ENSG00000182747	solute carrier family 35, member D3	9.179494017	2.46E-31	2.89E-27
SLC4A3	ENSG00000114923	solute carrier family 4 (anion exchanger), member 3	-2.366325383	7.28E-05	0.029469479
SMOC1	ENSG00000198732	SPARC related modular calcium binding 1	-5.03813069	0.000104067	0.036268803
SPON2	ENSG00000159674	spondin 2, extracellular matrix protein	-2.639592845	0.000112611	0.038484373
SRD5A2	ENSG00000049319	steroid-5-alpha-reductase, alpha polypeptide 2 (3-oxo-5 alpha-steroid delta 4-dehydrogenase alpha 2)	4.448074003	8.97E-05	0.032879001
STAG3L2	ENSG00000160828	stromal antigen 3-like 2 (pseudogene)	-0.881204831	9.94E-05	0.035716372
TCF21	ENSG00000118526	transcription factor 21	-6.04609974	1.52E-06	0.001676294
ULK4	ENSG00000168038	unc-51 like kinase 4	1.147447551	6.46E-09	2.28E-05
WASF3	ENSG00000132970	WAS protein family, member 3	0.972046987	2.81E-06	0.002603617
XKR9	ENSG00000221947	XK, Kell blood group complex subunit-related family, member 9	-1.989973823	0.000159916	0.049814665
ZBTB16	ENSG00000109906	zinc finger and BTB domain containing 16	-4.844890791	7.83E-05	0.030946745
ZFPM2	ENSG00000169946	zinc finger protein, FOG family member 2	-3.60011069	5.25E-14	4.62E-10
ZNF804A	ENSG00000170396	zinc finger protein 804A	-5.483825868	4.34E-09	1.77E-05

Supplementary Table S6. GO enrichment analysis. Show the distribution of GO terms exhibiting statistical significant differences of mis-regulated transcripts in A.IV-1 fibroblast using WebGestalt 2019 (Wang, Vasaikar, Shi, Greer, & Zhang, 2017).

GO terms	Description	Gene set size	Overlap	Expected Value	Enrichment Ratio	pValue	FDR	overlap_Id	user_Id
GO:0009888	tissue development	1613	24	7.625171279802685	3.1474702822178497	1.528766712111107e-7	0.0012670418509976855	249;1116;1301;1844;1893;1901;2185;2254;2494;3484;3664;4300;4916;6422;6943;7704;10631;23414;25884;28984;54507;59269;79192;153572	TCF21;ZBTB16;COL11A1;CHRDL2;NR5A2;SFRP1;RGCC;IRF6;ZFPM2;ADAMTSL4;S1PR1;ALPL;HIVEP3;PTK2B;ECM1;MLLT3;IRX2;IRX1;NTRK3;POSTN;IGFBP1;DUSP2;FGF9;CHI3L1
GO:0001503	ossification	339	11	1.6025623458481775	6.864007524261469	4.940246167706164e-7	0.0020472380118974343	249;1301;1893;1901;2185;2254;3486;6422;7704;25884;64093	SMOC1;ZBTB16;COL11A1;CHRDL2;SFRP1;S1PR1;ALPL;PTK2B;ECM1;IGFBP3;FGF9
GO:0030038	contractile actin filament bundle assembly	76	5	0.35927651411345574	13.916857360793287	2.9031950272262463e-5	0.060154200964127824	1901;2185;6422;28984;55604	SFRP1;RGCC;S1PR1;CARMIL1;PTK2B
GO:0043149	stress fiber assembly	76	5	0.35927651411345574	13.916857360793287	2.9031950272262463e-5	0.060154200964127824	1901;2185;6422;28984;55604	SFRP1;RGCC;S1PR1;CARMIL1;PTK2B
GO:0048584	positive regulation of response to stimulus	1768	21	8.35790627569197	2.512590989573087	4.138463705938289e-5	0.06101540527461857	1116;1123;1893;1901;2070;2185;2254;2533;3486;3818;4300;4916;5618;6422;6503;10417;28984;53832;55540;126669;284654	FYB;RSPO1;EYA4;SFRP1;RGCC;S1PR1;SLA;SPON2;CHN1;PTK2B;ECM1;MLLT3;IGFBP3;KLKB1;NTRK3;IL20RA;IL17RB;PRLR;SHE;FGF9;CHI3L1
GO:0007379	segment specification	15	3	0.07090983831186626	42.3072463768116	4.417138412737831e-5	0.06101540527461857	4300;79192;153572	MLLT3;IRX2;IRX1
GO:0030278	regulation of ossification	169	6	0.7989175116470265	7.510162078723952	1.4457514683041683e-4	0.12546232977013955	1893;1901;2185;6422;7704;64093	SMOC1;ZBTB16;SFRP1;S1PR1;PTK2B;ECM1
GO:0048646	anatomical structure formation involved in morphogenesis	876	13	4.14113455741299	3.1392363179141025	2.015277408380367e-4	0.12546232977013955	1116;1301;1844;1893;1901;2185;2254;6422;6943;28984;79192;153572;221935	TCF21;COL11A1;SFRP1;RGCC;S1PR1;PTK2B;ECM1;IRX2;IRX1;SDK1;DUSP2;FGF9;CHI3L1
GO:0001501	skeletal system development	439	9	2.075294601260619	4.33673368327226	2.1232290176698143e-4	0.12546232977013955	249;1116;1301;2254;3226;6422;7704;10631;25884	HOXC10;ZBTB16;COL11A1;CHRDL2;SFRP1;ALPL;POSTN;FGF9;CHI3L1
GO:0072047	proximal/distal pattern formation involved in nephron development	5	2	0.023636612770622088	84.61449275362318	2.182366244216638e-4	0.12546232977013955	79192;153572	IRX2;IRX1

Supplementary Table S7. Summary of transcriptome mapping results.

Sample ID	#total reads	#aligned reads	% aligned reads	#uniquely mapped	% uniquely aligned reads
A.IV-1.1	306,265,974	303,190,132	99.00	293,189,972	96.70
A.IV-1.2	344,782,588	341,370,324	99.01	330,182,914	96.72
ARN1.1	338,129,498	335,557,854	99.24	325,618,112	97.04
ARN1.2	301,339,526	298,096,006	98.92	289,521,446	97.12
ARN2.1	311,917,536	307,641,094	98.63	296,837,798	96.49
ARN2.2	336,751,354	333,728,602	99.10	322,903,810	96.76
ARN3.1	403,718,870	400,433,942	99.19	383,740,470	95.83
ARN3.2	299,879,030	296,763,108	98.96	286,203,188	96.44
ARN4.1	318,588,494	315,510,058	99.03	304,438,002	96.49
ARN4.2	320,107,070	316,729,460	98.94	304,559,776	96.16
ARN5.1	363,584,984	360,496,138	99.15	349,081,194	96.83
ARN6.1	324,984,406	321,703,444	98.99	311,631,678	96.87
ARN6.2	364,763,088	361,575,516	99.13	349,853,468	96.76

SUPPLEMENTARY REFERENCES

Adzhubei, I. A., Schmidt, S., Peshkin, L., Ramensky, V. E., Gerasimova, A., Bork, P., . . . Sunyaev, S. R. (2010). A method and server for predicting damaging missense mutations. *Nature Methods, 7*(4), 248-249. doi:10.1038/nmeth0410-248

Geoffroy, V., Pizot, C., Redin, C., Piton, A., Vasli, N., Stoetzel, C., . . . Muller, J. (2015). VaRank: a simple and powerful tool for ranking genetic variants. *PeerJ, 3*, e796. doi:10.7717/peerj.796

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., & Salzberg, S. L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol, 14*(4), R36. doi:10.1186/gb-2013-14-4-r36

Kumar, P., Henikoff, S., & Ng, P. C. (2009). Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nature protocols, 4*(7), 1073-1081. doi:10.1038/nprot.2009.86

Langmead, B., Trapnell, C., Pop, M., & Salzberg, S. L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol, 10*(3), R25. doi:10.1186/gb-2009-10-3-r25

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol, 15*(12), 550. doi:10.1186/s13059-014-0550-8

Nevers, Y., Prasad, M. K., Poidevin, L., Chennen, K., Allot, A., Kress, A., . . . Lecompte, O. (2017). Insights into Ciliary Genes and Evolution from Multi-Level Phylogenetic Profiling. *Mol Biol Evol, 34*(8), 2016-2034. doi:10.1093/molbev/msx146

Reese, M. G., Eeckman, F. H., Kulp, D., & Haussler, D. (1997). Improved splice site detection in Genie. *J Comput Biol, 4*(3), 311-323.

Shapiro, M. B., & Senapathy, P. (1987). RNA splice junctions of different classes of eukaryotes: sequence statistics and functional implications in gene expression. *Nucleic Acids Res, 15*(17), 7155-7174.

Wang, J., Vasaikar, S., Shi, Z., Greer, M., & Zhang, B. (2017). WebGestalt 2017: a more comprehensive, powerful, flexible and interactive gene set enrichment analysis toolkit. *Nucleic Acids Research, 45*(W1), W130-W137. doi:10.1093/nar/gkx356

Yeo, G., & Burge, C. B. (2004). Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. *J Comput Biol, 11*(2-3), 377-394. doi:10.1089/1066527041410418

SUPPLEMENTARY INFORMATION

Sanger Sequencing

Sanger sequencing was performed by way of PCR amplification with 50ng of genomic DNA template. The primers were designed with Primer 3 (http://frodo.wi.mit.edu/primer3). Bidirectional sequencing of the purified PCR products was performed by Eurofins Genomics Sequencing Facilities (https://www.eurofinsgenomics.eu/en/custom-dna-sequencing).

Control screening

IQCE exon 12 was tested in 93 control DNA of matched origin (Algerian) and 93 multi ethnic. High resolution melting (HRM) assay was performed in a 20 µl reaction volume containing 50 ng of template DNA, 15 µl Precision Melt Supermix (Bio-Rad) and 10 µmol of each primer of the primer set. PCR was performed using a CFX96 (Bio-Rad) thermal cycler and consisted of predenaturation (98°C for 2 min) followed by 35 cycles of amplification involving denaturation (98°C for 5 s), annealing (60°C for 10 s) and a melting curve involving denaturation (95°C for 30 s), annealing (70°C for 30 s) and a melt curve (72–95°C with an increment of 0.2°C per 10 s). Data generated were analyzed using Precision Melt Analysis software (Bio-Rad). In HRM analysis, differences in Tm and normalized curve shape were used together to discriminate even the most difficult–to–detect sequence variation.

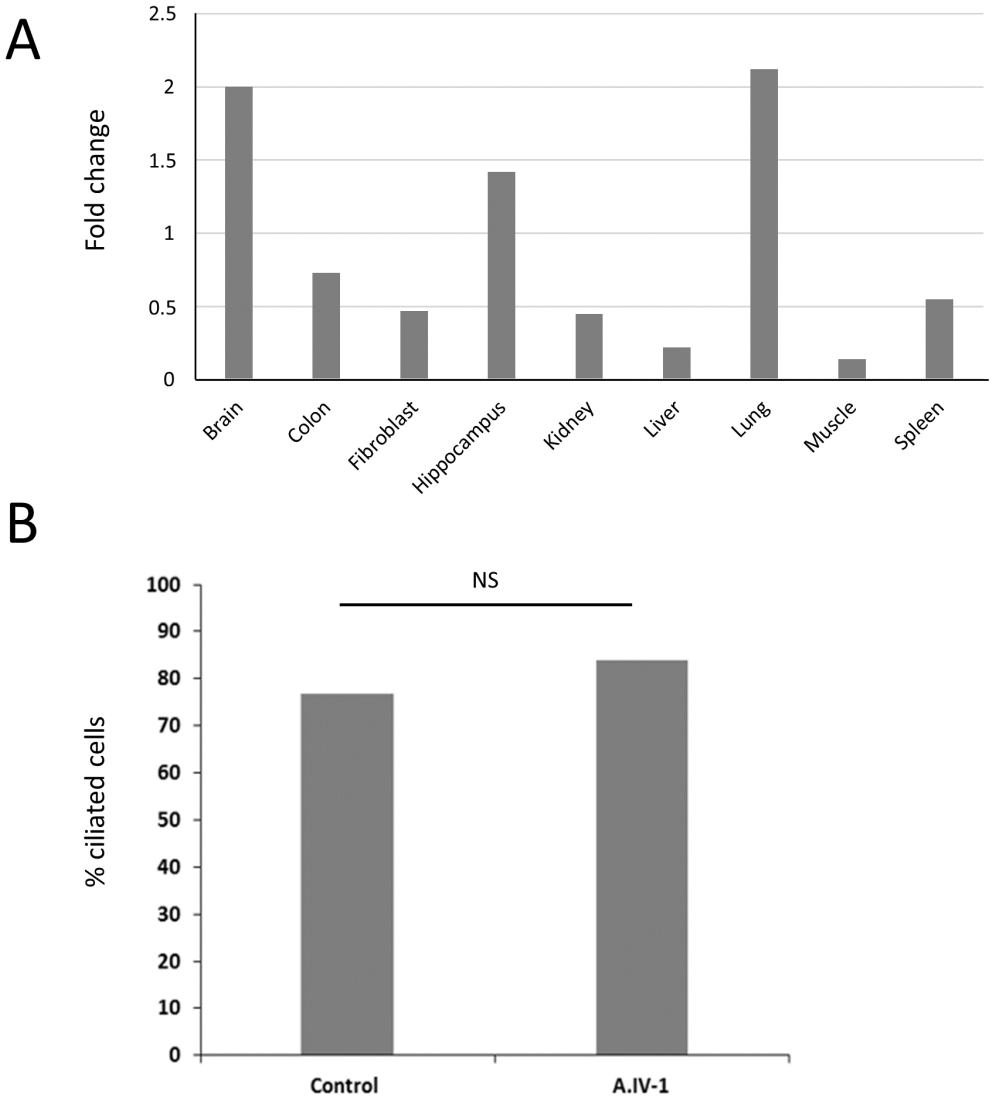
Bioinformatic pipeline RNA seq

Image analysis and base calling were performed using RTA v.2.7.3 and blc2fastq v.2.17.1.14 (Illumina). Sequence reads were mapped onto hg19/GRCh37 assembly of human genome using Tophat 2.0.14 (Kim et al., 2013) and bowtie version 2-2.1.0 (Langmead, Trapnell, Pop, & Salzberg, 2009), A summary of the mapping results is available in Supplementary Table 7. Gene expression was quantified from uniquely aligned reads using HTSeq-0.6.1 with gene annotations from Ensembl release 75 and intersection non-empty mode.

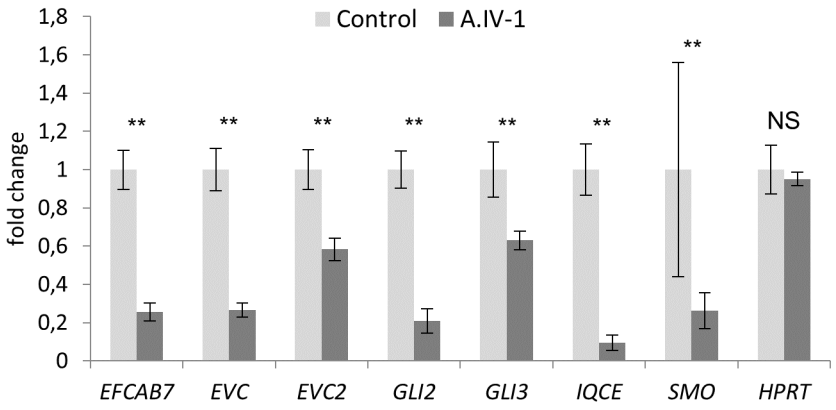
Differential expression RNA-seq

In order to identify variation in mRNA expression due to the absence of IQCE in the patient cells, comparisons of read counts were performed using R 3.2.5 with the statistical method proposed by Anders and Huber implemented in the DESeq2 v.1.10.1 Bioconductor package

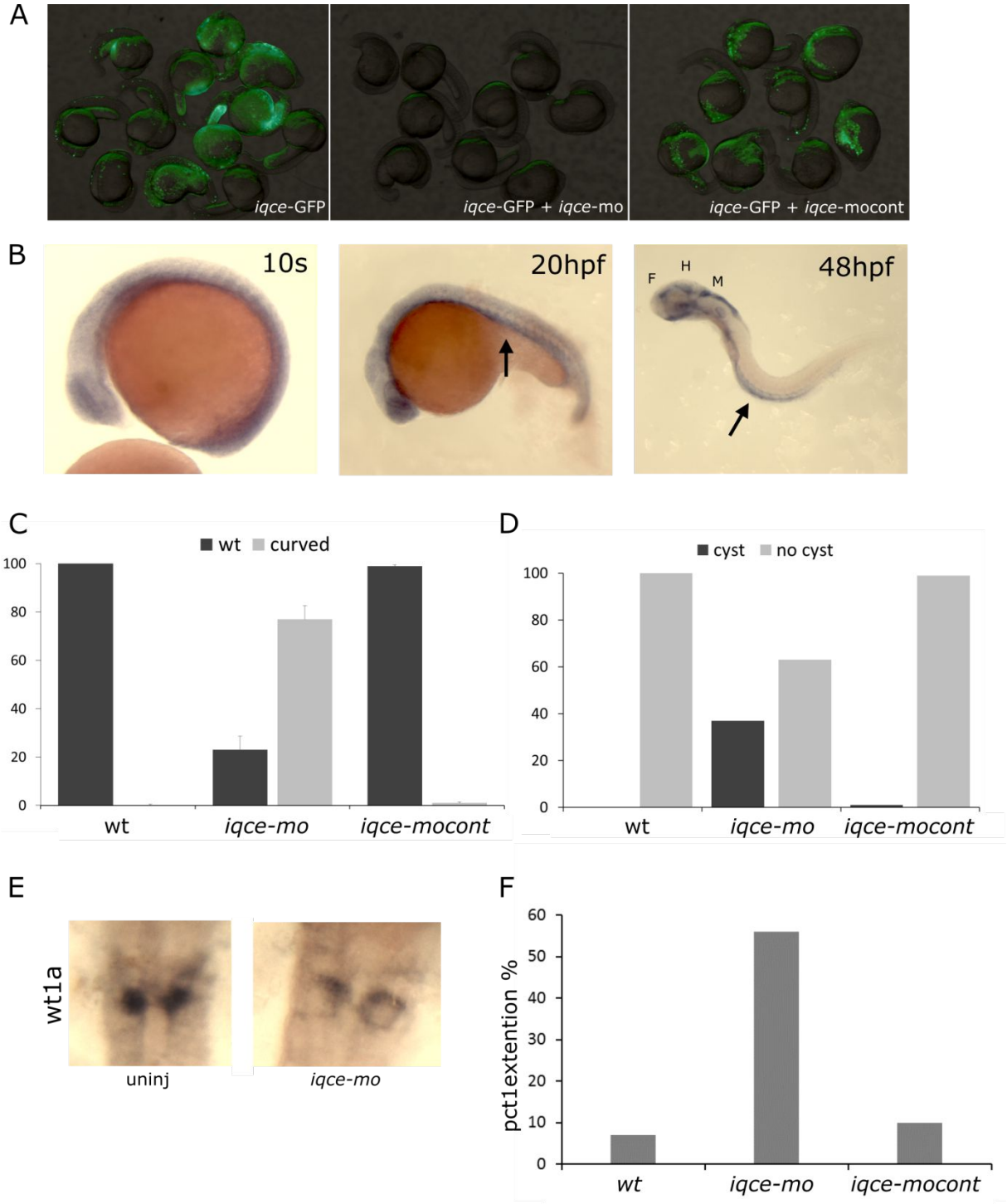
(Love, Huber, & Anders, 2014). The generated data were explored and visualized to assess and check data quality and to eventually remove bad quality data. Especially, PCA was built after stabilizing variance using the regularized log transformation method or RSeqQC was used to do some quality controls such as gene body coverage. Only genes with | log2 fold-change | > 0.5 were considered, as we expected to identify a large amount of change between transcriptomic results from one patient compared to other individuals (indirect consequences of the pathogenic mutation or unrelated events due to variability between humans). To avoid as much as possible false positive variants, we filtered out all genes with a Benjamini and Hochberg adjusted p-value above 0.05.



Supplementary Figure S1. **A.** Highest mRNA expression of *IQCE* in adult human brain/hippocampus and lung. **B.** Count of ciliated fibroblasts between A.IV-1 and control (~100 cells per condition). Percentages are shown in histogram, N.S.: non-significant, Fisher exact test 0.224, $P < 0.05$).



Supplementary Figure S2. Expression levels of EVC-zone and GLI genes upon ciliogenesis and Hh activation. Patient's (A.IV-1) and control cells upon treatment (–FCS + SAG 100nM) were compared (n=4 controls, n=6 patient). Statistical significance has been performed using the non-parametric Mann-Whitney test, p=0.0095 (**), NS : non-significant.



Supplementary Figure S3. Efficacy of *iqce* morpholinos. A. Morpholinos directed against the start codon of *iqce* blocks the expression of a construct containing *iqce* cDNA in frame with gfp (central panel). In contrast, 5 mismatch morpholinos did not reduce expression of the chimeric protein (left panel) compared to *iqce-gfp* alone (right panel). 24 hpf. **B.** *in situ* hybridization showing *iqce* expression pattern in 10 somites, 20 and 48 hpf embryos. Arrow indicates the pronephric duct. F: forebrain ventricule; H: hindbrain ventricule; M: midbrain ventricule. **C.** Chart showing the percentage of curved embryos in uninjected (n=200), *iqce-*

mo (n=120); and *iqce-mocont* (n=221) injected embryos. 48 hpf, data includes ± SD and statistical significance was determined using the Student's t-test, all comparisons were significant with a p-value<0.05. **D.** Chart showing the percentage of cysts in uninjected (n=101), *iqce-mo* (n=70) and *iqce-mocont* (n=90) injected embryos. Statistical significance was determined using the Student's t-test, all comparisons were significant with a p-value<0.05. **E.** wt1a *in situ* hybridization showing expression within glomeruli in non-injected embryos and at the periphery of the cyst in morphants (48 hpf). **F.** Chart showing the percentage of embryos with *ptc1* extension in 10 somites embryos (n=191).

Supplementary Table S1. Summary of the whole exome sequencing results. SNV: Single nucleotide variation, indel: gain or loss of up to 50 nucleotides at a single locus, SV: Structural Variation. Annotations are gathered using Alamut Batch v1.11, especially for the variation databases including gnomAD (v2.0.2, Oct. 2017), 1000 Genomes Project phase3 release (version 20150813 v5b) and the following predictions including phastCons (UCSC, 44 vertebrates). Effect on the splice has been evaluated using the MaxEntScan (Yeo & Burge, 2004), NNSPLICE 0.9 (Reese, Eeckman, Kulp, & Haussler, 1997) and Splice Site Finder (Shapiro & Senapathy, 1987) by calculating score change between the wild type and the mutated sequences expressed as a percent differences. Missenses have been evaluated using default parameters from PolyPhen-2 (2.2.2) (Adzhubei et al., 2010) and SIFT 4.0.3 (Kumar, Henikoff, & Ng, 2009). Default cut-offs used have been described in VaRank (Geoffroy et al., 2015) for both type of predictions. Exclusion of SV with a DGV (Gold standard from 20160515) frequency > 1% is done only with studies of more than 1000 individuals.

	A.IV-1			B.II-2			B.II-3			C.IV-4		
	SNV	**indel**	**SV**	**SNV**	**indel**	**SV**	**SNV**	**Indel**	**SV**	**SNV**	**Indel**	**SV**
Total number of variants	83129	11664	31	81972	11319	17	82116	11195	16	78906	11727	31
After exclusion of variants with an allele frequency >1% (gnomAD, 1000G, internal exome database, DGV)	9771	1125	16	7069	993	6	6856	943	9	5323	982	17
After exclusion of SNV/indel in 5'UTR, 3'UTR, downstream, upstream, intron and synonymous locations without local splice effect prediction	788	76	/	553	59	/	560	65	/	753	113	/
After exclusion of missense with no deleterious effect according to SIFT, PPH2 or PhastCons	669	76	/	458	59	/	465	65	/	641	113	/
After exclusion of variants not in a ciliary genes list (Nevers et al., 2017) and not in non-syndromic polydactyly known genes	52	6	1	31	2	2	40	4	2	78	15	1
After selection of variants consistent with recessive transmission (compound heterozygous, homozygous variants)	7 hom.: *DNAAF1*, ***IQCE***, *MADD, SHANK2,* ***TULP1***, *ZDHHC2* 3 comp. het.: *ARMC3, DNAH11, MGA*			0 homozygous variant 1 comp. het. In ***IQCE***						3 hom.: *ZMAT1, AGTPBP1*, *KDM4C* 11 comp. het: *CTDSPL2, DDX1, DNAH7, DYNLL1, EP400, IFT20, LRRCC1, LRRC46, MRPL10, PRKDC, SMARCA5, TMEM97, TPTE*		

Supplementary Table S2. Homozygous regions identified in individual A.IV-1.

Chr	SNP1	SNP2	POS1	POS2	MB	#SNP	DENSITY	
1	rs4147830	rs11431894	94544276	100680594	6.14	33	185.949	
1	rs4915053	rs3761934	108113526	114189067	6.08	136	44.673	
1	1_148741720_T_C	rs11428742	148741720	150550020	1.81	22	82.196	
1	rs4636	rs10927011	234744413	243579112	8.83	125	70.678	
2	rs7577088	rs3764906	39233500	42588195	3.35	20	167.735	
2	rs522893	2_188250301_T_A	169690729	188250301	18.56	224	82.855	
2	rs2304704	rs12477076	190430177	191848312	1.42	23	61.658	
3	rs901854	rs113574694	4856234	10089550	5.23	58	90.230	
3	rs9814557	rs2246945	135720540	137843476	2.12	21	101.092	
4	rs13146318	rs1051447	41615690	49063872	7.45	68	109.532	
4	rs4864727	rs12621	54011424	57897570	3.89	61	63.707	
4	rs10434219	rs344122	62598689	77637502	15.04	175	85.936	
6	rs2076506	6_32485853_C_T	22570064	32485853	9.92	646	15.350	
6	rs1059547	rs12189627	32557529	41059458	8.50	309	27.514	*TULP1*
7	rs4916935	rs798479	197267	2691201	2.49	41	60.828	*IQCE*
7	rs7794563	rs80024169	5347295	14775824	9.43	78	120.879	
7	rs4718101	rs6460315	64152353	66289040	2.14	38	56.229	
8	rs3989699	rs13271468	7218621	12236202	5.02	132	38.012	
8	8_12273335_G_A	rs66480340	12273335	24211885	11.94	202	59.102	
8	rs10086200	rs59331088	61193588	74005131	12.81	72	177.938	
8	rs11777189	rs12545587	99057150	136560952	37.50	204	183.842	
9	rs4977555	rs11515	19550308	21968199	2.42	30	80.596	
9	rs7863560	rs3217100	119495697	123734288	4.24	30	141.286	
11	11_32452000_GG_-	rs2076623	32452000	33564123	1.11	25	44.485	
11	rs12274095	rs7101792	69063767	71249386	2.19	28	78.058	
11	rs3740912	rs4937391	125891269	128786294	2.90	32	90.470	
12	rs11054683	rs1600	12247616	14656768	2.41	31	77.715	
12	rs3088008	rs3210837	49724955	51685831	1.96	55	35.652	
12	rs74651927	rs2242497	52695754	56992630	4.30	185	23.226	
12	rs1663564	rs7488309	105546172	118520170	12.97	152	85.355	
12	rs1794956	rs11059841	124848137	129190021	4.34	37	117.348	
14	rs34800262	rs2249922	51370852	52906081	1.54	34	45.154	
14	rs2297113	rs863091	58036494	59112475	1.08	23	46.782	
14	rs35533709	rs2057482	60433392	62213848	1.78	30	59.349	
15	rs4923929	rs555001	42439376	43545728	1.11	37	29.901	
15	rs11351249	rs2623989	72105929	73545732	1.44	23	62.600	
15	rs752270	rs4777755	87099537	93510603	6.41	171	37.492	
16	rs6564764	rs931713	80581631	85023855	4.44	121	36.713	
17	rs8071836	rs67749219	47238028	49711042	2.47	40	61.825	

Supplementary Table S3. List and characteristics of primers used in this study

Application	Gene	Exon	Forward (5'-3')	Reverse (5'-3')	Size (bp)
DNA	*IQCE*	12	CTCTGAGCTCACAACCAACAG	GTTGCCCAAGGGTTCACA	399
		16	TGGTTTCTGGTGTTGTCAGG	TTTTCAGTGGCTGGTCAGAA	382
	ATP6V1B1	3	AGCCGTGGGAAGTAGGTGTT	ATGTCCCATTGACCCCACT	385
	TULP1	12/13	GGGATGTAGGATCCCCTCA	CATGCCAGGAATGATGACG	481
cDNA	*IQCE*	3-4/5	ACATCGCCAAAGTCACCTTATC	AGGGCCTGGGTCAGACTT	158
	GSTT1	2-3/3	ACCTTGACGGAGAGTGTGG	GCAGCTTCTCCGCAGAGT	144
	SLC30A8	9-10/10	CCATCTTACTCATGGAAGGTGTG	GCTGTAGCAACATGAGCTGAGA	154
	RSPO1	4/5	TGGAGAGGAACGACATCCG	CCTTACACTTGGTGCAGAAGTTA	156
	FGF9	2/3	GGCCTGGTCAGCATTCGAG	GTATCGCCTTCCAGTGTCCAC	180
	SFRP1	1/2	GCTTCTACTGGCCCGAGATG	TGGCCTCAGATTTCAACTCGT	180
	MECOM	4/5	AAACTCGAAAGCGAGAATGATCT	TGGTGGCGAATTAAATTGGACTT	185
	ALPL	6-7/7	AACATCAGGGACATTGACGTG	GTATCTCGGTTTGAAGCTCTTCC	159
	IGF2	2-3/3-4	TTCTACTTCAGCAGGCCCG	GAAGTTGTCCGGAAGCACG	168
	GAPDH	4-5/6	GGAGCGAGATCCCTCCAAAAT	GGCTGTTGTCATACTTCTCATGG	197
	HPRT1	1-2/2-3	CCTGGCGTCGTGATTAGTGAT	AGACGTTCAGTCCTGTCCATAA	131
	EFCAB7	6/6-7	GTCAGCAACCAGGAAGTTCA	CATGTGTTGCCAGTCCTTTATT	115
	EVC2	6-7/8	CTCGTCACGGAACAGAACAC	CAGGTCAGCACAAGGGAGAG	135
	EVC	2/4	TCGAAGGACAAGGAAGCTGT	TTTCATGCAGAGACGGGTTG	148
	GLI2	2/3	AAGCAAGAAGCCAAAAGTGG	TGGTACCTTCCTTCCTGGTG	188
	GLI3	2/3	GGCATTTTTGGTCGAAGAGA	GGACATTCTGTGGCTGCATA	238
	SMO	2/3	TGCCCAAGTGTGAGAATGAC	TACCAGCTCTTGGGGTTGTC	229

Supplementary Table S4. Comparison of haplotypes of individuals A.IV-1 and B.II-2.

ID	Gene	Chr	Position	Allele_1_A.IV-1	Allele_2_A.IV-1	Allele_1_II-2_Family2	Allele_1_II-2_Family2
7_2255974_C_T	*MAD1L1*	7	2255974	C	C	C	T
7_2256017_C_A	*MAD1L1*	7	2256017	C	C	C	A
7_2257612_G_A	*MAD1L1*	7	2257612	A	A	A	G
7_2259134_C_T	*MAD1L1*	7	2259134	C	C	C	T
7_2260715_G_T	*MAD1L1*	7	2260715	T	T	T	T
7_2269552_G_T	*MAD1L1*	7	2269552	T	T	T	G
7_2278663_T_A	*FTSJ2*	7	2278663	T	T	A	A
7_2279482_A_C	*FTSJ2*	7	2279482	A	A	C	C
7_2279851_T_C	*FTSJ2*	7	2279851	T	T	C	C
7_2281754_C_T	*FTSJ2*	7	2281754	C	C	C	T
7_2290328_C_T	*SNX8*	7	2290328	C	C	C	T
7_2290336_T_G	*SNX8*	7	2290336	T	T	T	G
7_2290372_C_T	*SNX8*	7	2290372	C	C	T	T
7_2290522_C_T	*SNX8*	7	2290522	T	T	C	C
7_2292881_-_TG	*SNX8*	7	2292881	-	-	TG	TG
7_2296493_A_G	*SNX8*	7	2296493	G	G	G	G
7_2297006_A_G	*SNX8*	7	2297006	A	A	G	G
7_2303000_G_A	*SNX8*	7	2303000	A	A	G	G
7_2303109_G_A	*SNX8*	7	2303109	G	G	G	A
7_2303986_T_C	*SNX8*	7	2303986	C	C	C	C
7_2304191_A_C	*SNX8*	7	2304191	A	A	C	C
7_2309159_-_A	*SNX8*	7	2309159	-	-	-	A
7_2309174_G_A	*SNX8*	7	2309174	G	G	G	A
7_2309185_G_A	*SNX8*	7	2309185	G	G	G	A
7_2309209_C_T	*SNX8*	7	2309209	C	C	C	T
7_2314652_G_A	*SNX8*	7	2314652	G	G	G	A
7_2315008_G_C	*SNX8*	7	2315008	G	G	C	C
7_2318254_ACTC_-	*SNX8*	7	2318254	-	-	ACTC	ACTC
7_2318268_C_T	*SNX8*	7	2318268	T	T	C	C
7_2318463_C_T	*SNX8*	7	2318463	C	T	C	C
7_2318472_C_A	*SNX8*	7	2318472	C	A	C	C
7_2349470_T_C	*SNX8*	7	2349470	T	T	T	C
7_2394991_C_T	*EIF3B*	7	2394991	C	C	C	T
7_2400306_T_C	*EIF3B*	7	2400306	C	C	C	T
7_2405885_C_A	*EIF3B*	7	2405885	A	A	A	C
7_2409035_C_T	*EIF3B*	7	2409035	C	C	C	T
7_2414142_A_G	*EIF3B*	7	2414142	G	G	G	A
7_2414365_A_G	*EIF3B*	7	2414365	A	A	A	G
7_2416822_A_G	*EIF3B*	7	2416822	A	A	A	G
7_2418645_A_G	*EIF3B*	7	2418645	A	A	A	G
7_2433729_C_A	NA	7	2433729	C	C	C	A
7_2434522_G_A	NA	7	2434522	G	G	A	A
7_2434596_TT_-	NA	7	2434596	TT	TT	TT	-
7_2434597_T_-	NA	7	2434597	T	T	T	-

7_2472429_C_A	*CHST12*	7	2472429	C	C	C	A
7_2472455_A_T	*CHST12*	7	2472455	A	A	A	T
7_2489311_C_G	NA	7	2489311	C	C	C	G
7_2489336_GTGTGTGCACCT_-	NA	7	2489336	GTGTGTGCACCT	GTGTGTGCACCT	GTGTGTGCACCT	-
7_2489382_A_G	NA	7	2489382	A	A	A	G
7_2515382_-_C	NA	7	2515382	-	-	C	C
7_2517473_C_T	NA	7	2517473	C	C	C	T
7_2552986_A_C	*LFNG*	7	2552986	A	A	A	C
7_2553030_T_C	*LFNG*	7	2553030	T	T	T	C
7_2557266_T_C	*LFNG*	7	2557266	T	T	C	C
7_2557377_T_C	*LFNG*	7	2557377	T	T	T	C
7_2565268_C_A	*LFNG/MIR4648*	7	2565268	C	C	A	A
7_2566433_G_A	*LFNG/MIR4648*	7	2566433	G	G	G	A
7_2577691_C_G	*BRAT1*	7	2577691	C	C	C	G
7_2577781_T_C	*BRAT1*	7	2577781	T	T	T	C
7_2578237_T_C	*BRAT1*	7	2578237	T	T	T	C
7_2578238_C_T	*BRAT1*	7	2578238	C	C	C	T
7_2578434_G_T	*BRAT1*	7	2578434	G	G	G	T
7_2578455_G_A	*BRAT1*	7	2578455	G	G	G	A
7_2580914_C_T	*BRAT1*	7	2580914	C	C	C	T
7_2581994_T_C	*BRAT1*	7	2581994	T	T	T	C
7_2582154_-_CCCCAGCCTCCCGGGTGT	*BRAT1*	7	2582154	CCCCAGCCTCCCGGGTGT	CCCCAGCCTCCCGGGTGT	-	-
7_2582233_A_G	*BRAT1*	7	2582233	A	A	A	G
7_2582381_T_A	*BRAT1*	7	2582381	T	T	T	A
7_2582584_G_A	*BRAT1*	7	2582584	G	G	G	A
7_2583165_G_A	*BRAT1*	7	2583165	G	G	G	A
7_2583328_C_T	*BRAT1*	7	2583328	C	C	C	T
7_2584523_T_C	*BRAT1*	7	2584523	T	T	T	C
7_2587122_T_C	*BRAT1*	7	2587122	T	T	T	C
7_2598757_T_C	*IQCE*	7	2598757	T	T	T	C
7_2604768_GTGT_-	*IQCE*	7	2604768	GTGT	GTGT	GTGT	-
7_2604769_T_*	*IQCE*	7	2604769	T	T	T	*
7_2604816_C_T	*IQCE*	7	2604816	C	C	C	T
7_2611878_T_C	*IQCE*	7	2611878	T	T	T	C
7_2612225_G_A	*IQCE*	7	2612225	A	A	G	G
7_2613042_G_C	*IQCE*	7	2613042	G	G	G	C
7_2618070_-_A	*IQCE*	7	2618070	-	-	-	A
7_2619341_GT_-	*IQCE*	7	2619341	GT	GT	GT	-
7_2619363_G_C	*IQCE*	7	2619363	G	G	G	C
7_2623351_G_A	*IQCE*	7	2623351	G	G	G	A
7_2623801_G_T	*IQCE*	7	2623801	G	G	G	T
7_2625907_TGTCCCGGAG_-	*IQCE*	7	2625907	-	-	-	TGTCCCGGAG
7_2626039_CAGGGAATGG_-	*IQCE*	7	2626039	CAGGGAATGG	CAGGGAATGG	CAGGGAATGG	-
7_2627310_T_A	*IQCE*	7	2627310	T	T	T	A
7_2627579_AG_-	*IQCE*	7	2627579	AG	AG	-	-
7_2632477_A_G	*IQCE*	7	2632477	A	A	A	G

7_2634518_AGAG_-	*IQCE*	7	2634518	AGAG	AGAG	AGAG	-
7_2636689_C_T	*IQCE*	7	2636689	C	C	C	T
7_2638347_C_G	*IQCE*	7	2638347	C	C	C	G
7_2640985_G_C	*IQCE*	7	2640985	G	G	G	C
7_2647048_C_G	*IQCE*	7	2647048	C	C	C	G
7_2691201_C_A	*TTYH3*	7	2691201	A	A	C	C
7_2699687_-_G	*TTYH3*	7	2699687	-	-	G	G
7_2701809_T_C	*TTYH3*	7	2701809	T	C	T	T
7_2742097_T_C	*AMZ1*	7	2742097	C	C	T	T
7_2749565_G_A	*AMZ1*	7	2749565	G	G	G	A
7_2752152_G_A	*AMZ1*	7	2752152	G	G	G	A
7_2752487_G_A	*AMZ1*	7	2752487	A	A	G	G
7_2773035_-_AT	*GNA12*	7	2773035	-	-	AT	AT
7_2802173_T_C	*GNA12*	7	2802173	T	T	T	C
7_2802522_T_C	*GNA12*	7	2802522	T	T	T	C
7_2834869_C_T	*GNA12*	7	2834869	C	C	T	T
7_2854010_G_A	*GNA12*	7	2854010	G	G	A	A
7_2946461_T_C	*CARD11*	7	2946461	C	C	C	T
7_2952912_G_A	*CARD11*	7	2952912	G	G	G	A
7_2957005_T_C	*CARD11*	7	2957005	C	C	C	T
7_2958298_C_T	*CARD11*	7	2958298	C	C	T	T
7_2962241_G_A	*CARD11*	7	2962241	A	A	G	G
7_2962753_G_A	*CARD11*	7	2962753	G	G	G	A
7_2966445_T_G	*CARD11*	7	2966445	T	T	T	G
7_2966466_GT_-	*CARD11*	7	2966466	GT	GT	GT	-
7_2968195_G_A	*CARD11*	7	2968195	G	G	G	A
7_2968361_G_A	*CARD11*	7	2968361	G	G	G	A
7_2968486_G_C	*CARD11*	7	2968486	G	G	G	C
7_2985364_AC_-	*CARD11*	7	2985364	A	C	-	-

Supplementary Table S5. Differentially expressed transcripts in A.IV-1 fibroblasts.

Gene name	Ensembl gene id	Description	log2(FC)	P-value	Adjusted p-value
ABCA8	ENSG00000141338	ATP-binding cassette, sub-family A (ABC1), member 8	-4.205829134	3.52E-05	0.019362186
AC011294.3	ENSG00000233539	Uncharacterized protein	3.130226421	2.55E-07	0.000427239
ACCS	ENSG00000110455	1-aminocyclopropane-1-carboxylate synthase homolog (Arabidopsis)(non-functional)	-0.953037549	1.28E-07	0.00025011
ADAMTSL4	ENSG00000143382	ADAMTS-like 4	-3.153983754	5.00E-07	0.000732904
ALDH1A1	ENSG00000165092	aldehyde dehydrogenase 1 family, member A1	-6.206001924	1.99E-08	5.40E-05
ALPL	ENSG00000162551	alkaline phosphatase, liver/bone/kidney	-3.039649649	0.000113858	0.038536558
APBA2	ENSG00000034053	amyloid beta (A4) precursor protein-binding, family A, member 2	-2.732442847	6.79E-05	0.028790694
ARHGAP28	ENSG00000088756	Rho GTPase activating protein 28	2.800539497	1.58E-05	0.010684692
C10orf54	ENSG00000107738	chromosome 10 open reading frame 54	-1.64124697	9.82E-05	0.035643913
C1QTNF9B-AS1	ENSG00000205861	C1QTNF9B antisense RNA 1	-2.047191404	3.88E-05	0.019780947
C2CD4A	ENSG00000198535	C2 calcium-dependent domain containing 4A	5.96982663	8.95E-07	0.001085914
C4A	ENSG00000244731	complement component 4A (Rodgers blood group)	-6.329532092	5.21E-11	3.66E-07
CCDC180	ENSG00000197816	coiled-coil domain containing 180	2.351169896	3.81E-05	0.019780947
CD200	ENSG00000091972	CD200 molecule	-4.053753384	0.000135641	0.044622222
CHI3L1	ENSG00000133048	chitinase 3-like 1 (cartilage glycoprotein-39)	4.852100597	6.43E-07	0.000837885
CHN1	ENSG00000128656	chimerin 1	-1.617295158	2.25E-09	1.32E-05
CHRDL2	ENSG00000054938	chordin-like 2	-4.241249176	0.000101018	0.03591767
CNTN1	ENSG00000018236	contactin 1	-4.135254047	8.76E-05	0.032879001
COL11A1	ENSG00000060718	collagen, type XI, alpha 1	-4.558619592	6.30E-05	0.027357715
CSF2RA	ENSG00000198223	colony stimulating factor 2 receptor, alpha, low-affinity (granulocyte-macrophage)	3.072895701	6.90E-05	0.028906369
CYP21A1P	ENSG00000204338	cytochrome P450, family 21, subfamily A, polypeptide 1 pseudogene	-4.61429374	6.95E-06	0.005439088
DUSP2	ENSG00000158050	dual specificity phosphatase 2	3.527490256	1.70E-05	0.01128586
ECM1	ENSG00000143369	extracellular matrix protein 1	-1.261675626	5.81E-07	0.000786019
EMCN	ENSG00000164035	endomucin	3.361027658	4.95E-05	0.0229427
EYA4	ENSG00000112319	eyes absent homolog 4 (Drosophila)	-4.200503943	4.26E-05	0.020550766
FAM21B	ENSG00000152726	family with sequence similarity 21, member B	-4.995162913	7.93E-08	0.000164108
FGF9	ENSG00000102678	fibroblast growth factor 9	4.555259473	5.95E-08	0.000133394
FOSB	ENSG00000125740	FBJ murine osteosarcoma viral oncogene homolog B	3.628551087	8.80E-05	0.032879001
FYB	ENSG00000082074	FYN binding protein	-4.90648728	3.03E-05	0.016914087
GRPR	ENSG00000126010	gastrin-releasing peptide receptor	-4.422707449	0.000144048	0.046095412
GSTT1	ENSG00000184674	glutathione S-transferase theta 1	-9.843482395	6.94E-60	2.44E-55
HIVEP3	ENSG00000127124	human immunodeficiency virus type I enhancer binding protein 3	-1.599082154	6.64E-05	0.028486515
HOXC10	ENSG00000180818	homeobox C10	-4.942455716	2.36E-05	0.014065434
HOXC-AS3	ENSG00000251151	HOXC cluster antisense RNA 3	-5.14737488	6.01E-05	0.026775571
IGF2	ENSG00000167244	insulin-like growth factor 2 (somatomedin A)	-4.063105964	3.83E-05	0.019780947
IGFBP1	ENSG00000146678	insulin-like growth factor binding protein 1	2.926653654	1.58E-05	0.010684692

IGFBP3	ENSG00000146674	insulin-like growth factor binding protein 3	1.932377253	2.71E-06	0.002574523
IL17RB	ENSG00000056736	interleukin 17 receptor B	3.133016005	8.91E-05	0.032879001
IL20RA	ENSG00000016402	interleukin 20 receptor, alpha	3.104556086	3.12E-06	0.002749475
IQCE	ENSG00000106012	IQ motif containing E	-2.198435274	1.75E-43	3.08E-39
IRF6	ENSG00000117595	interferon regulatory factor 6	-3.931349479	7.74E-05	0.030946745
IRX1	ENSG00000170549	iroquois homeobox 1	1.496134127	2.70E-05	0.015596484
IRX2	ENSG00000170561	iroquois homeobox 2	1.161748374	4.51E-09	1.77E-05
JAZF1-AS1	ENSG00000234336	JAZF1 antisense RNA 1	-2.132735824	0.000107563	0.037119743
KLKB1	ENSG00000164344	kallikrein B, plasma (Fletcher factor) 1	2.478417669	8.84E-05	0.032879001
LAMC1	ENSG00000135862	laminin, gamma 1 (formerly LAMB2)	-1.201150815	2.84E-07	0.000453983
LHX9	ENSG00000143355	LIM homeobox 9	4.298162126	1.45E-06	0.001650048
LINC00537	ENSG00000232815	long intergenic non-protein coding RNA 537	-3.494641854	3.83E-05	0.019780947
LINC00578	ENSG00000228221	long intergenic non-protein coding RNA 578	-4.541164267	7.46E-09	2.39E-05
LINC01048	ENSG00000230390	long intergenic non-protein coding RNA 1048	2.214309895	3.99E-05	0.019780947
LINC01117	ENSG00000224577	long intergenic non-protein coding RNA 1117	-2.17227763	2.17E-05	0.01316806
LRRC16A	ENSG00000079691	leucine rich repeat containing 16A	-1.784218256	2.43E-05	0.014245821
MAP6	ENSG00000171533	microtubule-associated protein 6	-1.79656788	1.30E-06	0.001530605
MAP7D2	ENSG00000184368	MAP7 domain containing 2	2.359306267	2.66E-06	0.002574523
MBP	ENSG00000197971	myelin basic protein	-3.402601353	7.14E-05	0.029220727
MCCC1	ENSG00000078070	methylcrotonoyl-CoA carboxylase 1 (alpha)	-0.554719912	0.000128238	0.042584679
MECOM	ENSG00000085276	MDS1 and EVI1 complex locus	-4.511721791	3.14E-09	1.58E-05
METTL9	ENSG00000197006	methyltransferase like 9	0.906519962	3.86E-08	9.70E-05
MFGE8	ENSG00000140545	milk fat globule-EGF factor 8 protein	-2.694104724	7.44E-07	0.000934855
MLLT3	ENSG00000171843	myeloid/lymphoid or mixed-lineage leukemia (trithorax homolog, Drosophila); translocated to, 3	0.969013225	2.10E-05	0.01316806
MRVI1-AS1	ENSG00000177112	MRVI1 antisense RNA 1	-3.884388307	1.04E-05	0.007509783
MYO16	ENSG00000041515	myosin XVI	-1.962481605	0.000117031	0.039233207
NDP	ENSG00000124479	Norrie disease (pseudoglioma)	3.545635871	1.83E-05	0.011727901
NR5A2	ENSG00000116833	nuclear receptor subfamily 5, group A, member 2	-4.138425423	8.61E-06	0.006447287
NTRK3	ENSG00000140538	neurotrophic tyrosine kinase, receptor, type 3	2.777060262	2.95E-05	0.016740854
PKIA	ENSG00000171033	protein kinase (cAMP-dependent, catalytic) inhibitor alpha	-2.779337899	4.26E-05	0.020550766
PNLDC1	ENSG00000146453	poly(A)-specific ribonuclease (PARN)-like domain containing 1	2.95490495	2.27E-07	0.000399745
PNMA6C	ENSG00000235961	paraneoplastic Ma antigen family member 6C	-4.033899706	0.000151809	0.047958078
POMZP3	ENSG00000146707	POM121 and ZP3 fusion	-1.717492119	3.12E-06	0.002749475
POSTN	ENSG00000133110	periostin, osteoblast specific factor	2.83484857	2.14E-05	0.01316806
PRKAR1B	ENSG00000188191	protein kinase, cAMP-dependent, regulatory, type I, beta	-1.732349926	1.78E-05	0.011579369
PRLR	ENSG00000113494	prolactin receptor	3.300099077	0.000140582	0.045398939
PTK2B	ENSG00000120899	protein tyrosine kinase 2 beta	-1.47679232	6.10E-05	0.026827045
RGCC	ENSG00000102760	regulator of cell cycle	-3.93550099	7.06E-05	0.029220727
RHBDD3	ENSG00000100263	rhomboid domain containing 3	-0.996433377	1.05E-05	0.007509783
RSPO1	ENSG00000169218	R-spondin 1	-4.313022549	5.98E-05	0.026775571
S1PR1	ENSG00000170989	sphingosine-1-phosphate receptor 1	-3.093512898	5.52E-06	0.00452094

SCN1A	ENSG00000144285	sodium channel, voltage-gated, type I, alpha subunit	-4.906276861	3.80E-06	0.003184547
SCN9A	ENSG00000169432	sodium channel, voltage-gated, type IX, alpha subunit	-3.142530704	1.83E-06	0.001898572
SDK1	ENSG00000146555	sidekick cell adhesion molecule 1	3.188897352	4.55E-05	0.021372366
SFRP1	ENSG00000104332	secreted frizzled-related protein 1	-4.10661345	5.21E-07	0.000732904
SHE	ENSG00000169291	Src homology 2 domain containing E	3.455259022	1.63E-06	0.001733572
SLA	ENSG00000155926	Src-like-adaptor	-2.701260288	3.99E-05	0.019780947
SLC16A11	ENSG00000174326	solute carrier family 16, member 11	-2.677392243	4.37E-05	0.020765473
SLC30A8	ENSG00000164756	solute carrier family 30 (zinc transporter), member 8	6.104148946	1.35E-07	0.000250798
SLC35D3	ENSG00000182747	solute carrier family 35, member D3	9.179494017	2.46E-31	2.89E-27
SLC4A3	ENSG00000114923	solute carrier family 4 (anion exchanger), member 3	-2.366325383	7.28E-05	0.029469479
SMOC1	ENSG00000198732	SPARC related modular calcium binding 1	-5.03813069	0.000104067	0.036268803
SPON2	ENSG00000159674	spondin 2, extracellular matrix protein	-2.639592845	0.000112611	0.038484373
SRD5A2	ENSG00000049319	steroid-5-alpha-reductase, alpha polypeptide 2 (3-oxo-5 alpha-steroid delta 4-dehydrogenase alpha 2)	4.448074003	8.97E-05	0.032879001
STAG3L2	ENSG00000160828	stromal antigen 3-like 2 (pseudogene)	-0.881204831	9.94E-05	0.035716372
TCF21	ENSG00000118526	transcription factor 21	-6.04609974	1.52E-06	0.001676294
ULK4	ENSG00000168038	unc-51 like kinase 4	1.147447551	6.46E-09	2.28E-05
WASF3	ENSG00000132970	WAS protein family, member 3	0.972046987	2.81E-06	0.002603617
XKR9	ENSG00000221947	XK, Kell blood group complex subunit-related family, member 9	-1.989973823	0.000159916	0.049814665
ZBTB16	ENSG00000109906	zinc finger and BTB domain containing 16	-4.844890791	7.83E-05	0.030946745
ZFPM2	ENSG00000169946	zinc finger protein, FOG family member 2	-3.60011069	5.25E-14	4.62E-10
ZNF804A	ENSG00000170396	zinc finger protein 804A	-5.483825868	4.34E-09	1.77E-05

Supplementary Table S6. GO enrichment analysis. Show the distribution of GO terms exhibiting statistical significant differences of mis-regulated transcripts in A.IV-1 fibroblast using WebGestalt 2019 (Wang, Vasaikar, Shi, Greer, & Zhang, 2017).

GO terms	Description	Gene set size	Overlap	Expected Value	Enrichment Ratio	pValue	FDR	overlap_Id	user_Id
GO:0009888	tissue development	1613	24	7.625171279802685	3.1474702822178497	1.528766712111107e-7	0.0012670418509976855	249;1116;1301;1844;1893;1901;2185;2254;2494;3484;3664;4300;4916;6422;6943;7704;10631;23414;25884;28984;54507;59269;79192;153572	TCF21;ZBTB16;COL11A1;CHRDL2;NR5A2;SFRP1;RGCC;IRF6;ZFPM2;ADAMTSL4;S1PR1;ALPL;HIVEP3;PTK2B;ECM1;MLLT3;IRX2;IRX1;NTRK3;POSTN;IGFBP1;DUSP2;FGF9;CHI3L1
GO:0001503	ossification	339	11	1.6025623458481775	6.864007524261469	4.940246167706164e-7	0.0020472380118974343	249;1301;1893;1901;2185;2254;3486;6422;7704;25884;64093	SMOC1;ZBTB16;COL11A1;CHRDL2;SFRP1;S1PR1;ALPL;PTK2B;ECM1;IGFBP3;FGF9
GO:0030038	contractile actin filament bundle assembly	76	5	0.35927651411345574	13.916857360793287	2.9031950272262463e-5	0.060154200964127824	1901;2185;6422;28984;55604	SFRP1;RGCC;S1PR1;CARMIL1;PTK2B
GO:0043149	stress fiber assembly	76	5	0.35927651411345574	13.916857360793287	2.9031950272262463e-5	0.060154200964127824	1901;2185;6422;28984;55604	SFRP1;RGCC;S1PR1;CARMIL1;PTK2B
GO:0048584	positive regulation of response to stimulus	1768	21	8.35790627569197	2.512590989573087	4.138463705938289e-5	0.06101540527461857	1116;1123;1893;1901;2070;2185;2254;2533;3486;3818;4300;4916;5618;6422;6503;10417;28984;53832;55540;126669;284654	FYB;RSPO1;EYA4;SFRP1;RGCC;S1PR1;SLA;SPON2;CHN1;PTK2B;ECM1;MLLT3;IGFBP3;KLKB1;NTRK3;IL20RA;IL17RB;PRLR;SHE;FGF9;CHI3L1
GO:0007379	segment specification	15	3	0.07090983831186626	42.3072463768116	4.417138412737831e-5	0.06101540527461857	4300;79192;153572	MLLT3;IRX2;IRX1
GO:0030278	regulation of ossification	169	6	0.7989175116470265	7.510162078723952	1.4457514683041683e-4	0.12546232977013955	1893;1901;2185;6422;7704;64093	SMOC1;ZBTB16;SFRP1;S1PR1;PTK2B;ECM1
GO:0048646	anatomical structure formation involved in morphogenesis	876	13	4.14113455741299	3.1392363179141025	2.015277408380367e-4	0.12546232977013955	1116;1301;1844;1893;1901;2185;2254;6422;6943;28984;79192;153572;221935	TCF21;COL11A1;SFRP1;RGCC;S1PR1;PTK2B;ECM1;IRX2;IRX1;SDK1;DUSP2;FGF9;CHI3L1
GO:0001501	skeletal system development	439	9	2.075294601260619	4.33673368327226	2.1232290176698143e-4	0.12546232977013955	249;1116;1301;2254;3226;6422;7704;10631;25884	HOXC10;ZBTB16;COL11A1;CHRDL2;SFRP1;ALPL;POSTN;FGF9;CHI3L1
GO:0072047	proximal/distal pattern formation involved in nephron development	5	2	0.02363612770622088	84.61449275362318	2.182366244216638e-4	0.12546232977013955	79192;153572	IRX2;IRX1

Supplementary Table S7. Summary of transcriptome mapping results.

Sample ID	#total reads	#aligned reads	% aligned reads	#uniquely mapped	% uniquely aligned reads
A.IV-1.1	306,265,974	303,190,132	99.00	293,189,972	96.70
A.IV-1.2	344,782,588	341,370,324	99.01	330,182,914	96.72
ARN1.1	338,129,498	335,557,854	99.24	325,618,112	97.04
ARN1.2	301,339,526	298,096,006	98.92	289,521,446	97.12
ARN2.1	311,917,536	307,641,094	98.63	296,837,798	96.49
ARN2.2	336,751,354	333,728,602	99.10	322,903,810	96.76
ARN3.1	403,718,870	400,433,942	99.19	383,740,470	95.83
ARN3.2	299,879,030	296,763,108	98.96	286,203,188	96.44
ARN4.1	318,588,494	315,510,058	99.03	304,438,002	96.49
ARN4.2	320,107,070	316,729,460	98.94	304,559,776	96.16
ARN5.1	363,584,984	360,496,138	99.15	349,081,194	96.83
ARN6.1	324,984,406	321,703,444	98.99	311,631,678	96.87
ARN6.2	364,763,088	361,575,516	99.13	349,853,468	96.76

SUPPLEMENTARY REFERENCES

Adzhubei, I. A., Schmidt, S., Peshkin, L., Ramensky, V. E., Gerasimova, A., Bork, P., . . . Sunyaev, S. R. (2010). A method and server for predicting damaging missense mutations. *Nature Methods, 7*(4), 248-249. doi:10.1038/nmeth0410-248

Geoffroy, V., Pizot, C., Redin, C., Piton, A., Vasli, N., Stoetzel, C., . . . Muller, J. (2015). VaRank: a simple and powerful tool for ranking genetic variants. *PeerJ, 3*, e796. doi:10.7717/peerj.796

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., & Salzberg, S. L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol, 14*(4), R36. doi:10.1186/gb-2013-14-4-r36

Kumar, P., Henikoff, S., & Ng, P. C. (2009). Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nature protocols, 4*(7), 1073-1081. doi:10.1038/nprot.2009.86

Langmead, B., Trapnell, C., Pop, M., & Salzberg, S. L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol, 10*(3), R25. doi:10.1186/gb-2009-10-3-r25

Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol, 15*(12), 550. doi:10.1186/s13059-014-0550-8

Nevers, Y., Prasad, M. K., Poidevin, L., Chennen, K., Allot, A., Kress, A., . . . Lecompte, O. (2017). Insights into Ciliary Genes and Evolution from Multi-Level Phylogenetic Profiling. *Mol Biol Evol, 34*(8), 2016-2034. doi:10.1093/molbev/msx146

Reese, M. G., Eeckman, F. H., Kulp, D., & Haussler, D. (1997). Improved splice site detection in Genie. *J Comput Biol, 4*(3), 311-323.

Shapiro, M. B., & Senapathy, P. (1987). RNA splice junctions of different classes of eukaryotes: sequence statistics and functional implications in gene expression. *Nucleic Acids Res, 15*(17), 7155-7174.

Wang, J., Vasaikar, S., Shi, Z., Greer, M., & Zhang, B. (2017). WebGestalt 2017: a more comprehensive, powerful, flexible and interactive gene set enrichment analysis toolkit. *Nucleic Acids Research, 45*(W1), W130-W137. doi:10.1093/nar/gkx356

Yeo, G., & Burge, C. B. (2004). Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. *J Comput Biol, 11*(2-3), 377-394. doi:10.1089/1066527041410418

SUPPLEMENTARY INFORMATION

Sanger Sequencing

Sanger sequencing was performed by way of PCR amplification with 50ng of genomic DNA template. The primers were designed with Primer 3 (http://frodo.wi.mit.edu/primer3). Bidirectional sequencing of the purified PCR products was performed by Eurofins Genomics Sequencing Facilities (https://www.eurofinsgenomics.eu/en/custom-dna-sequencing).

Control screening

IQCE exon 12 was tested in 93 control DNA of matched origin (Algerian) and 93 multi ethnic. High resolution melting (HRM) assay was performed in a 20 µl reaction volume containing 50 ng of template DNA, 15 µl Precision Melt Supermix (Bio-Rad) and 10 µmol of each primer of the primer set. PCR was performed using a CFX96 (Bio-Rad) thermal cycler and consisted of predenaturation (98°C for 2 min) followed by 35 cycles of amplification involving denaturation (98°C for 5 s), annealing (60°C for 10 s) and a melting curve involving denaturation (95°C for 30 s), annealing (70°C for 30 s) and a melt curve (72–95°C with an increment of 0.2°C per 10 s). Data generated were analyzed using Precision Melt Analysis software (Bio-Rad). In HRM analysis, differences in Tm and normalized curve shape were used together to discriminate even the most difficult–to–detect sequence variation.

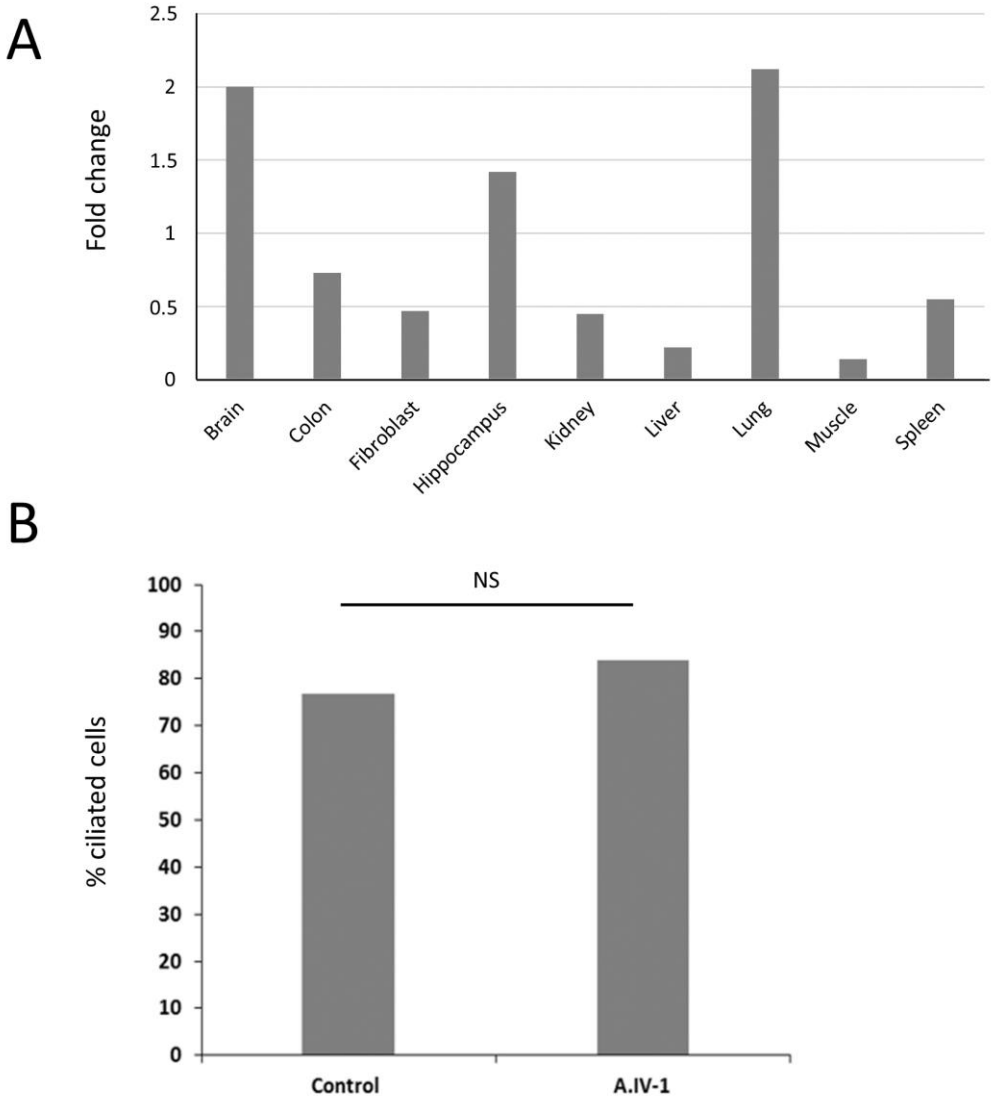
Bioinformatic pipeline RNA seq

Image analysis and base calling were performed using RTA v.2.7.3 and blc2fastq v.2.17.1.14 (Illumina). Sequence reads were mapped onto hg19/GRCh37 assembly of human genome using Tophat 2.0.14 (Kim et al., 2013) and bowtie version 2-2.1.0 (Langmead, Trapnell, Pop, & Salzberg, 2009), A summary of the mapping results is available in Supplementary Table 6. Gene expression was quantified from uniquely aligned reads using HTSeq-0.6.1 with gene annotations from Ensembl release 75 and intersection non-empty mode.

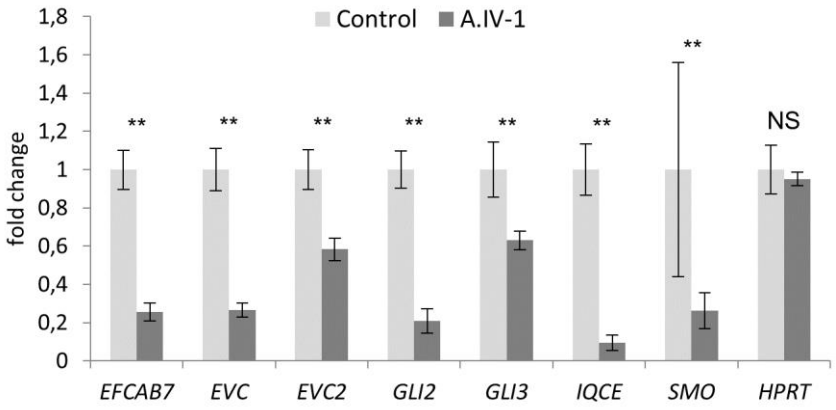
Differential expression RNA-seq

In order to identify variation in mRNA expression due to the absence of IQCE in the patient cells, comparisons of read counts were performed using R 3.2.5 with the statistical method proposed by Anders and Huber implemented in the DESeq2 v.1.10.1 Bioconductor package

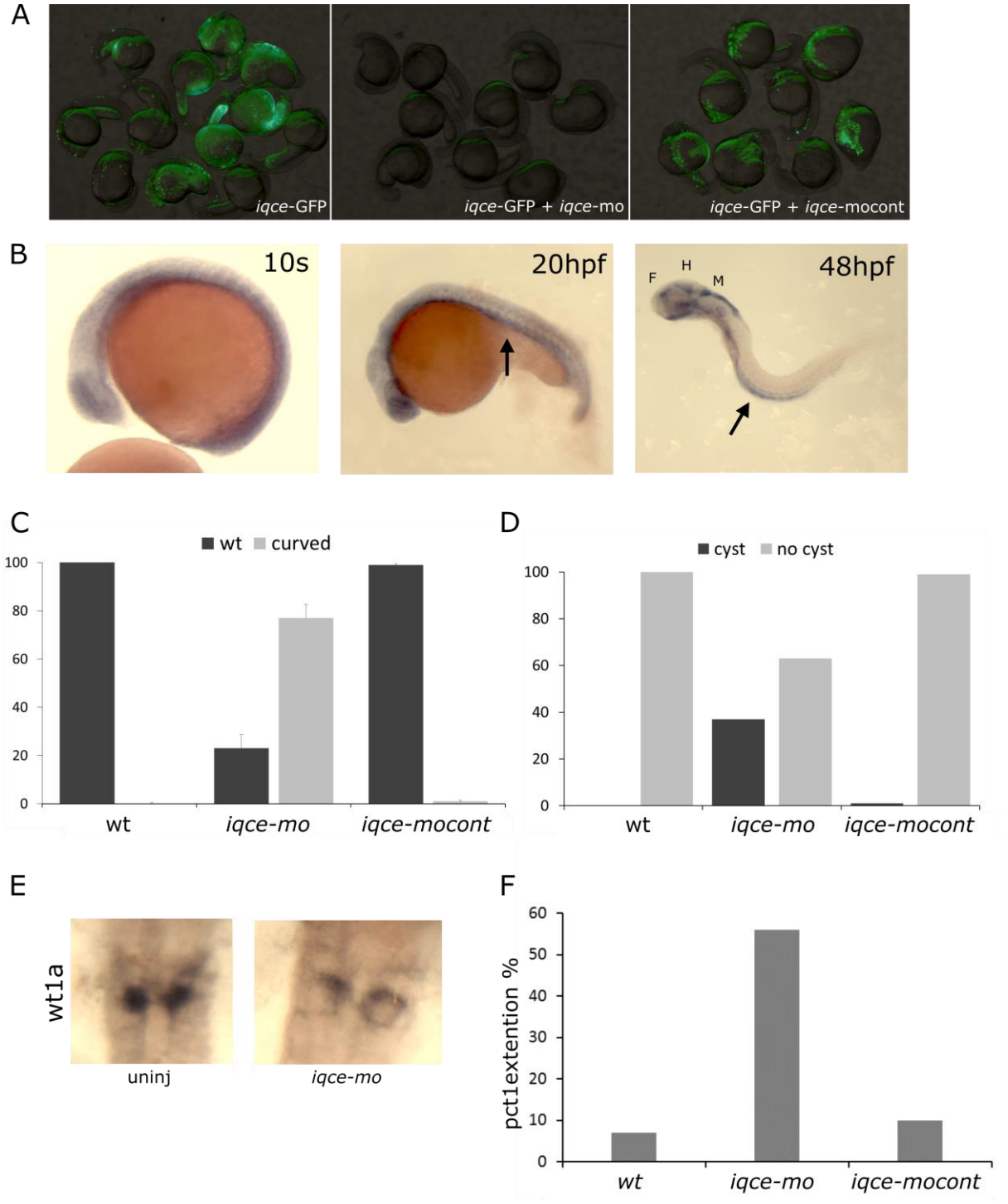
(Love, et al., 2014). The generated data were explored and visualized to assess and check data quality and to eventually remove bad quality data. Especially, PCA was built after stabilizing variance using the regularized log transformation method or RSeqQC was used to do some quality controls such as gene body coverage. Only genes with | log2 fold-change | > 0,5 were considered, as we expected to identify a large amount of change between transcriptomic results from one patient compared to other individuals (indirect consequences of the pathogenic mutation or unrelated events due to variability between humans). To avoid as much as possible false positive variants, we filtered out all genes with a Benjamini and Hochberg adjusted p-value above 0.05.



Supplementary Figure S1. **A.** Highest mRNA expression of *IQCE* in adult human brain/hippocampus and lung. **B.** Count of ciliated fibroblasts between A.IV-1 and control (~100 cells per condition). Percentages are shown in histogram, N.S.: non-significant, Fisher exact test 0.224, $P < 0.05$).



Supplementary Figure S2. Expression levels of EVC-zone and GLI genes upon ciliogenesis and Hh activation. Patient's (A.IV-1) and control cells upon treatment (–FCS + SAG 100nM) were compared (n=4 controls, n=6 patient). Statistical significance has been performed using the non-parametric Mann-Whitney test, p=0.0095 (**), NS : non-significant.



Supplementary Figure S3. Efficacy of *iqce* morpholinos. A. Morpholinos directed against the start codon of *iqce* blocks the expression of a construct containing *iqce* cDNA in frame with gfp (central panel). In contrast, 5 mismatch morpholinos did not reduce expression of the chimeric protein (left panel) compared to *iqce-gfp* alone (right panel). 24 hpf. **B.** *in situ* hybridization showing *iqce* expression pattern in 10 somites, 20 and 48 hpf embryos. Arrow indicates the pronephric duct. F: forebrain ventricule; H: hindbrain ventricule; M: midbrain ventricule. **C.** Chart showing the percentage of curved embryos in uninjected (n=200), *iqce-*

mo (n=120); and *iqce-mocont* (n=221) injected embryos. 48 hpf, data includes ± SD and statistical significance was determined using the Student's t-test, all comparisons were significant with a p-value<0.05. **D.** Chart showing the percentage of cysts in uninjected (n=101), *iqce-mo* (n=70) and *iqce-mocont* (n=90) injected embryos. Statistical significance was determined using the Student's t-test, all comparisons were significant with a p-value<0.05. **E.** wt1a *in situ* hybridization showing expression within glomeruli in non-injected embryos and at the periphery of the cyst in morphants (48 hpf). **F.** Chart showing the percentage of embryos with *ptc1* extension in 10 somites embryos (n=191).

Supplementary Table S1. Summary of the whole exome sequencing results. SNV: Single nucleotide variation, indel: gain or loss of up to 50 nucleotides at a single locus, SV: Structural Variation. Exclusion of SV with a DGV frequency > 1% is done only with studies of more than 1000 individuals

	A.IV-1			B.II-2			B.II-3			C.IV-4		
	SNV	**indel**	**SV**	**SNV**	**indel**	**SV**	**SNV**	**Indel**	**SV**	**SNV**	**Indel**	**SV**
Total number of variants	83129	11664	31	81972	11319	17	82116	11195	16	78906	11727	31
After exclusion of variants with an allele frequency >1% (gnomAD, 1000G, internal exome database, DGV)	9771	1125	16	7069	993	6	6856	943	9	5323	982	17
After exclusion of SNV/indel in 5'UTR, 3'UTR, downstream, upstream, intron and synonymous locations without local splice effect prediction	788	76	/	553	59	/	560	65	/	753	113	/
After exclusion of missense with no deleterious effect according to SIFT, PPH2 or PhastCons	669	76	/	458	59	/	465	65	/	641	113	/
After exclusion of variants not in a ciliary genes list (Nevers et al., 2017) and not in non-syndromic polydactyly known genes	52	6	1	31	2	2	40	4	2	78	15	1
After selection of variants consistent with recessive transmission (compound heterozygous, homozygous variants)	7 hom.: *DNAAF1*, ***IQCE***, *MADD, SHANK2,* ***TULP1***, *ZDHHC2* 3 comp. het.: *ARMC3, DNAH11, MGA*			0 homozygous variant 1 comp. het. In ***IQCE***						3 hom.: *ZMAT1, AGTPBP1*, *KDM4C* 11 comp. het: *CTDSPL2, DDX1, DNAH7, DYNLL1, EP400, IFT20, LRRCC1, LRRC46, MRPL10, PRKDC, SMARCA5, TMEM97, TPTE*		

Supplementary Table S2. Homozygous regions identified in individual A.IV-1.

Chr	SNP1	SNP2	POS1	POS2	MB	#SNP	DENSITY	
1	rs4147830	rs11431894	94544276	100680594	6.14	33	185.949	
1	rs4915053	rs3761934	108113526	114189067	6.08	136	44.673	
1	1_148741720_T_C	rs11428742	148741720	150550020	1.81	22	82.196	
1	rs4636	rs10927011	234744413	243579112	8.83	125	70.678	
2	rs7577088	rs3764906	39233500	42588195	3.35	20	167.735	
2	rs522893	2_188250301_T_A	169690729	188250301	18.56	224	82.855	
2	rs2304704	rs12477076	190430177	191848312	1.42	23	61.658	
3	rs901854	rs113574694	4856234	10089550	5.23	58	90.230	
3	rs9814557	rs2246945	135720540	137843476	2.12	21	101.092	
4	rs13146318	rs1051447	41615690	49063872	7.45	68	109.532	
4	rs4864727	rs12621	54011424	57897570	3.89	61	63.707	
4	rs10434219	rs344122	62598689	77637502	15.04	175	85.936	
6	rs2076506	6_32485853_C_T	22570064	32485853	9.92	646	15.350	
6	rs1059547	rs12189627	32557529	41059458	8.50	309	27.514	*TULP1*
7	rs4916935	rs798479	197267	2691201	2.49	41	60.828	*IQCE*
7	rs7794563	rs80024169	5347295	14775824	9.43	78	120.879	
7	rs4718101	rs6460315	64152353	66289040	2.14	38	56.229	
8	rs3989699	rs13271468	7218621	12236202	5.02	132	38.012	
8	8_12273335_G_A	rs66480340	12273335	24211885	11.94	202	59.102	
8	rs10086200	rs59331088	61193588	74005131	12.81	72	177.938	
8	rs11777189	rs12545587	99057150	136560952	37.50	204	183.842	
9	rs4977555	rs11515	19550308	21968199	2.42	30	80.596	
9	rs7863560	rs3217100	119495697	123734288	4.24	30	141.286	
11	11_32452000_GG_-	rs2076623	32452000	33564123	1.11	25	44.485	
11	rs12274095	rs7101792	69063767	71249386	2.19	28	78.058	
11	rs3740912	rs4937391	125891269	128786294	2.90	32	90.470	
12	rs11054683	rs1600	12247616	14656768	2.41	31	77.715	
12	rs3088008	rs3210837	49724955	51685831	1.96	55	35.652	
12	rs74651927	rs2242497	52695754	56992630	4.30	185	23.226	
12	rs1663564	rs7488309	105546172	118520170	12.97	152	85.355	
12	rs1794956	rs11059841	124848137	129190021	4.34	37	117.348	
14	rs34800262	rs2249922	51370852	52906081	1.54	34	45.154	
14	rs2297113	rs863091	58036494	59112475	1.08	23	46.782	
14	rs35533709	rs2057482	60433392	62213848	1.78	30	59.349	
15	rs4923929	rs555001	42439376	43545728	1.11	37	29.901	
15	rs11351249	rs2623989	72105929	73545732	1.44	23	62.600	
15	rs752270	rs4777755	87099537	93510603	6.41	171	37.492	
16	rs6564764	rs931713	80581631	85023855	4.44	121	36.713	
17	rs8071836	rs67749219	47238028	49711042	2.47	40	61.825	

Supplementary Table S3. List and characteristics of primers used in this study

Application	Gene	Exon	Forward (5'-3')	Reverse (5'-3')	Size (bp)
DNA	*IQCE*	12	CTCTGAGCTCACAACCAACAG	GTTGCCCAAGGGTTCACA	399
		16	TGGTTTCTGGTGTTGTCAGG	TTTTCAGTGGCTGGTCAGAA	382
	ATP6V1B1	3	AGCCGTGGGAAGTAGGTGTT	ATGTCCCATTGACCCCACT	385
	TULP1	12/13	GGGATGTAGGATCCCCTCA	CATGCCAGGAATGATGACG	481
cDNA	*IQCE*	3-4/5	ACATCGCCAAAGTCACCTTATC	AGGGCCTGGGTCAGACTT	158
	GSTT1	2-3/3	ACCTTGACGGAGAGTGTGG	GCAGCTTCTCCGCAGAGT	144
	SLC30A8	9-10/10	CCATCTTACTCATGGAAGGTGTG	GCTGTAGCAACATGAGCTGAGA	154
	RSPO1	4/5	TGGAGAGGAACGACATCCG	CCTTACACTTGGTGCAGAAGTTA	156
	FGF9	2/3	GGCCTGGTCAGCATTCGAG	GTATCGCCTTCCAGTGTCCAC	180
	SFRP1	1/2	GCTTCTACTGGCCCGAGATG	TGGCCTCAGATTTCAACTCGT	180
	MECOM	4/5	AAACTCGAAAGCGAGAATGATCT	TGGTGGCGAATTAAATTGGACTT	185
	ALPL	6-7/7	AACATCAGGGACATTGACGTG	GTATCTCGGTTTGAAGCTCTTCC	159
	IGF2	2-3/3-4	TTCTACTTCAGCAGGCCCG	GAAGTTGTCCGGAAGCACG	168
	GAPDH	4-5/6	GGAGCGAGATCCCTCCAAAAT	GGCTGTTGTCATACTTCTCATGG	197
	HPRT1	1-2/2-3	CCTGGCGTCGTGATTAGTGAT	AGACGTTCAGTCCTGTCCATAA	131
	EFCAB7	6/6-7	GTCAGCAACCAGGAAGTTCA	CATGTGTTGCCAGTCCTTTATT	115
	EVC2	6-7/8	CTCGTCACGGAACAGAACAC	CAGGTCAGCACAAGGGAGAG	135
	EVC	2/4	TCGAAGGACAAGGAAGCTGT	TTTCATGCAGAGACGGGTTG	148
	GLI2	2/3	AAGCAAGAAGCCAAAAGTGG	TGGTACCTTCCTTCCTGGTG	188
	GLI3	2/3	GGCATTTTTGGTCGAAGAGA	GGACATTCTGTGGCTGCATA	238
	SMO	2/3	TGCCCAAGTGTGAGAATGAC	TACCAGCTCTTGGGGTTGTC	229

Supplementary Table S4. Comparison of haplotypes of individuals A.IV-1 and B.II-2.

ID	Gene	Chr	Position	Allele_1_A.IV-1	Allele_2_A.IV-1	Allele_1_II-2_Family2	Allele_1_II-2_Family2
7_2255974_C_T	*MAD1L1*	7	2255974	C	C	C	T
7_2256017_C_A	*MAD1L1*	7	2256017	C	C	C	A
7_2257612_G_A	*MAD1L1*	7	2257612	A	A	A	G
7_2259134_C_T	*MAD1L1*	7	2259134	C	C	C	T
7_2260715_G_T	*MAD1L1*	7	2260715	T	T	T	T
7_2269552_G_T	*MAD1L1*	7	2269552	T	T	T	G
7_2278663_T_A	*FTSJ2*	7	2278663	T	T	A	A
7_2279482_A_C	*FTSJ2*	7	2279482	A	A	C	C
7_2279851_T_C	*FTSJ2*	7	2279851	T	T	C	C
7_2281754_C_T	*FTSJ2*	7	2281754	C	C	C	T
7_2290328_C_T	*SNX8*	7	2290328	C	C	C	T
7_2290336_T_G	*SNX8*	7	2290336	T	T	T	G
7_2290372_C_T	*SNX8*	7	2290372	C	C	T	T
7_2290522_C_T	*SNX8*	7	2290522	T	T	C	C
7_2292881_-_TG	*SNX8*	7	2292881	-	-	TG	TG
7_2296493_A_G	*SNX8*	7	2296493	G	G	G	G
7_2297006_A_G	*SNX8*	7	2297006	A	A	G	G
7_2303000_G_A	*SNX8*	7	2303000	A	A	G	G
7_2303109_G_A	*SNX8*	7	2303109	G	G	G	A
7_2303986_T_C	*SNX8*	7	2303986	C	C	C	C
7_2304191_A_C	*SNX8*	7	2304191	A	A	C	C
7_2309159_-_A	*SNX8*	7	2309159	-	-	-	A
7_2309174_G_A	*SNX8*	7	2309174	G	G	G	A
7_2309185_G_A	*SNX8*	7	2309185	G	G	G	A
7_2309209_C_T	*SNX8*	7	2309209	C	C	C	T
7_2314652_G_A	*SNX8*	7	2314652	G	G	G	A
7_2315008_G_C	*SNX8*	7	2315008	G	G	C	C
7_2318254_ACTC_-	*SNX8*	7	2318254	-	-	ACTC	ACTC
7_2318268_C_T	*SNX8*	7	2318268	T	T	C	C
7_2318463_C_T	*SNX8*	7	2318463	C	T	C	C
7_2318472_C_A	*SNX8*	7	2318472	C	A	C	C
7_2349470_T_C	*SNX8*	7	2349470	T	T	T	C
7_2394991_C_T	*EIF3B*	7	2394991	C	C	C	T
7_2400306_T_C	*EIF3B*	7	2400306	C	C	C	T
7_2405885_C_A	*EIF3B*	7	2405885	A	A	A	C
7_2409035_C_T	*EIF3B*	7	2409035	C	C	C	T
7_2414142_A_G	*EIF3B*	7	2414142	G	G	G	A
7_2414365_A_G	*EIF3B*	7	2414365	A	A	A	G
7_2416822_A_G	*EIF3B*	7	2416822	A	A	A	G
7_2418645_A_G	*EIF3B*	7	2418645	A	A	A	G
7_2433729_C_A	NA	7	2433729	C	C	C	A
7_2434522_G_A	NA	7	2434522	G	G	A	A
7_2434596_TT_-	NA	7	2434596	TT	TT	TT	-
7_2434597_T_-	NA	7	2434597	T	T	T	-

7_2472429_C_A	*CHST12*	7	2472429	C	C	C	A
7_2472455_A_T	*CHST12*	7	2472455	A	A	A	T
7_2489311_C_G	NA	7	2489311	C	C	C	G
7_2489336_GTGTGTGCACCT_-	NA	7	2489336	GTGTGTGCACCT	GTGTGTGCACCT	GTGTGTGCACCT	-
7_2489382_A_G	NA	7	2489382	A	A	A	G
7_2515382_-_C	NA	7	2515382	-	-	C	C
7_2517473_C_T	NA	7	2517473	C	C	C	T
7_2552986_A_C	*LFNG*	7	2552986	A	A	A	C
7_2553030_T_C	*LFNG*	7	2553030	T	T	T	C
7_2557266_T_C	*LFNG*	7	2557266	T	T	C	C
7_2557377_T_C	*LFNG*	7	2557377	T	T	T	C
7_2565268_C_A	*LFNG/MIR4648*	7	2565268	C	C	A	A
7_2566433_G_A	*LFNG/MIR4648*	7	2566433	G	G	G	A
7_2577691_C_G	*BRAT1*	7	2577691	C	C	C	G
7_2577781_T_C	*BRAT1*	7	2577781	T	T	T	C
7_2578237_T_C	*BRAT1*	7	2578237	T	T	T	C
7_2578238_C_T	*BRAT1*	7	2578238	C	C	C	T
7_2578434_G_T	*BRAT1*	7	2578434	G	G	G	T
7_2578455_G_A	*BRAT1*	7	2578455	G	G	G	A
7_2580914_C_T	*BRAT1*	7	2580914	C	C	C	T
7_2581994_T_C	*BRAT1*	7	2581994	T	T	T	C
7_2582154_-_CCCCAGCCTCCCGGGTGT	*BRAT1*	7	2582154	CCCCAGCCTCCCGGGTGT	CCCCAGCCTCCCGGGTGT	-	-
7_2582233_A_G	*BRAT1*	7	2582233	A	A	A	G
7_2582381_T_A	*BRAT1*	7	2582381	T	T	T	A
7_2582584_G_A	*BRAT1*	7	2582584	G	G	G	A
7_2583165_G_A	*BRAT1*	7	2583165	G	G	G	A
7_2583328_C_T	*BRAT1*	7	2583328	C	C	C	T
7_2584523_T_C	*BRAT1*	7	2584523	T	T	T	C
7_2587122_T_C	*BRAT1*	7	2587122	T	T	T	C
7_2598757_T_C	*IQCE*	7	2598757	T	T	T	C
7_2604768_GTGT_-	*IQCE*	7	2604768	GTGT	GTGT	GTGT	-
7_2604769_T_*	*IQCE*	7	2604769	T	T	T	*
7_2604816_C_T	*IQCE*	7	2604816	C	C	C	T
7_2611878_T_C	*IQCE*	7	2611878	T	T	T	C
7_2612225_G_A	*IQCE*	7	2612225	A	A	G	G
7_2613042_G_C	*IQCE*	7	2613042	G	G	G	C
7_2618070_-_A	*IQCE*	7	2618070	-	-	-	A
7_2619341_GT_-	*IQCE*	7	2619341	GT	GT	GT	-
7_2619363_G_C	*IQCE*	7	2619363	G	G	G	C
7_2623351_G_A	*IQCE*	7	2623351	G	G	G	A
7_2623801_G_T	*IQCE*	7	2623801	G	G	G	T
7_2625907_TGTCCCGGAG_-	*IQCE*	7	2625907	-	-	-	TGTCCCGGAG
7_2626039_CAGGGAATGG_-	*IQCE*	7	2626039	CAGGGAATGG	CAGGGAATGG	CAGGGAATGG	-
7_2627310_T_A	*IQCE*	7	2627310	T	T	T	A
7_2627579_AG_-	*IQCE*	7	2627579	AG	AG	-	-
7_2632477_A_G	*IQCE*	7	2632477	A	A	A	G

7_2634518_AGAG_-	*IQCE*	7	2634518	AGAG	AGAG	AGAG	-
7_2636689_C_T	*IQCE*	7	2636689	C	C	C	T
7_2638347_C_G	*IQCE*	7	2638347	C	C	C	G
7_2640985_G_C	*IQCE*	7	2640985	G	G	G	C
7_2647048_C_G	*IQCE*	7	2647048	C	C	C	G
7_2691201_C_A	*TTYH3*	7	2691201	A	A	C	C
7_2699687_-_G	*TTYH3*	7	2699687	-	-	G	G
7_2701809_T_C	*TTYH3*	7	2701809	T	C	T	T
7_2742097_T_C	*AMZ1*	7	2742097	C	C	T	T
7_2749565_G_A	*AMZ1*	7	2749565	G	G	G	A
7_2752152_G_A	*AMZ1*	7	2752152	G	G	G	A
7_2752487_G_A	*AMZ1*	7	2752487	A	A	G	G
7_2773035_-_AT	*GNA12*	7	2773035	-	-	AT	AT
7_2802173_T_C	*GNA12*	7	2802173	T	T	T	C
7_2802522_T_C	*GNA12*	7	2802522	T	T	T	C
7_2834869_C_T	*GNA12*	7	2834869	C	C	T	T
7_2854010_G_A	*GNA12*	7	2854010	G	G	A	A
7_2946461_T_C	*CARD11*	7	2946461	C	C	C	T
7_2952912_G_A	*CARD11*	7	2952912	G	G	G	A
7_2957005_T_C	*CARD11*	7	2957005	C	C	C	T
7_2958298_C_T	*CARD11*	7	2958298	C	C	T	T
7_2962241_G_A	*CARD11*	7	2962241	A	A	G	G
7_2962753_G_A	*CARD11*	7	2962753	G	G	G	A
7_2966445_T_G	*CARD11*	7	2966445	T	T	T	G
7_2966466_GT_-	*CARD11*	7	2966466	GT	GT	GT	-
7_2968195_G_A	*CARD11*	7	2968195	G	G	G	A
7_2968361_G_A	*CARD11*	7	2968361	G	G	G	A
7_2968486_G_C	*CARD11*	7	2968486	G	G	G	C
7_2985364_AC_-	*CARD11*	7	2985364	A	C	-	-

Supplementary Table S5. Differentially expressed transcripts in A.IV-1 fibroblasts.

Gene name	Ensembl gene id	Description	log2(FC)	P-value	Adjusted p-value
ABCA8	ENSG00000141338	ATP-binding cassette, sub-family A (ABC1), member 8	-4.205829134	3.52E-05	0.019362186
AC011294.3	ENSG00000233539	Uncharacterized protein	3.130226421	2.55E-07	0.000427239
ACCS	ENSG00000110455	1-aminocyclopropane-1-carboxylate synthase homolog (Arabidopsis)(non-functional)	-0.953037549	1.28E-07	0.00025011
ADAMTSL4	ENSG00000143382	ADAMTS-like 4	-3.153983754	5.00E-07	0.000732904
ALDH1A1	ENSG00000165092	aldehyde dehydrogenase 1 family, member A1	-6.206001924	1.99E-08	5.40E-05
ALPL	ENSG00000162551	alkaline phosphatase, liver/bone/kidney	-3.039649649	0.000113858	0.038536558
APBA2	ENSG00000034053	amyloid beta (A4) precursor protein-binding, family A, member 2	-2.732442847	6.79E-05	0.028790694
ARHGAP28	ENSG00000088756	Rho GTPase activating protein 28	2.800539497	1.58E-05	0.010684692
C10orf54	ENSG00000107738	chromosome 10 open reading frame 54	-1.64124697	9.82E-05	0.035643913
C1QTNF9B-AS1	ENSG00000205861	C1QTNF9B antisense RNA 1	-2.047191404	3.88E-05	0.019780947
C2CD4A	ENSG00000198535	C2 calcium-dependent domain containing 4A	5.96982663	8.95E-07	0.001085914
C4A	ENSG00000244731	complement component 4A (Rodgers blood group)	-6.329532092	5.21E-11	3.66E-07
CCDC180	ENSG00000197816	coiled-coil domain containing 180	2.351169896	3.81E-05	0.019780947
CD200	ENSG00000091972	CD200 molecule	-4.053753384	0.000135641	0.044622222
CHI3L1	ENSG00000133048	chitinase 3-like 1 (cartilage glycoprotein-39)	4.852100597	6.43E-07	0.000837885
CHN1	ENSG00000128656	chimerin 1	-1.617295158	2.25E-09	1.32E-05
CHRDL2	ENSG00000054938	chordin-like 2	-4.241249176	0.000101018	0.03591767
CNTN1	ENSG00000018236	contactin 1	-4.135254047	8.76E-05	0.032879001
COL11A1	ENSG00000060718	collagen, type XI, alpha 1	-4.558619592	6.30E-05	0.027357715
CSF2RA	ENSG00000198223	colony stimulating factor 2 receptor, alpha, low-affinity (granulocyte-macrophage)	3.072895701	6.90E-05	0.028906369
CYP21A1P	ENSG00000204338	cytochrome P450, family 21, subfamily A, polypeptide 1 pseudogene	-4.61429374	6.95E-06	0.005439088
DUSP2	ENSG00000158050	dual specificity phosphatase 2	3.527490256	1.70E-05	0.01128586
ECM1	ENSG00000143369	extracellular matrix protein 1	-1.261675626	5.81E-07	0.000786019
EMCN	ENSG00000164035	endomucin	3.361027658	4.95E-05	0.0229427
EYA4	ENSG00000112319	eyes absent homolog 4 (Drosophila)	-4.200503943	4.26E-05	0.020550766
FAM21B	ENSG00000152726	family with sequence similarity 21, member B	-4.995162913	7.93E-08	0.000164108
FGF9	ENSG00000102678	fibroblast growth factor 9	4.555259473	5.95E-08	0.000133394
FOSB	ENSG00000125740	FBJ murine osteosarcoma viral oncogene homolog B	3.628551087	8.80E-05	0.032879001
FYB	ENSG00000082074	FYN binding protein	-4.90648728	3.03E-05	0.016914087
GRPR	ENSG00000126010	gastrin-releasing peptide receptor	-4.422707449	0.000144048	0.046095412
GSTT1	ENSG00000184674	glutathione S-transferase theta 1	-9.843482395	6.94E-60	2.44E-55
HIVEP3	ENSG00000127124	human immunodeficiency virus type I enhancer binding protein 3	-1.599082154	6.64E-05	0.028486515
HOXC10	ENSG00000180818	homeobox C10	-4.942455716	2.36E-05	0.014065434
HOXC-AS3	ENSG00000251151	HOXC cluster antisense RNA 3	-5.14737488	6.01E-05	0.026775571
IGF2	ENSG00000167244	insulin-like growth factor 2 (somatomedin A)	-4.063105964	3.83E-05	0.019780947
IGFBP1	ENSG00000146678	insulin-like growth factor binding protein 1	2.926653654	1.58E-05	0.010684692

IGFBP3	ENSG00000146674	insulin-like growth factor binding protein 3	1.932377253	2.71E-06	0.002574523
IL17RB	ENSG00000056736	interleukin 17 receptor B	3.133016005	8.91E-05	0.032879001
IL20RA	ENSG00000016402	interleukin 20 receptor, alpha	3.104556086	3.12E-06	0.002749475
IQCE	ENSG00000106012	IQ motif containing E	-2.198435274	1.75E-43	3.08E-39
IRF6	ENSG00000117595	interferon regulatory factor 6	-3.931349479	7.74E-05	0.030946745
IRX1	ENSG00000170549	iroquois homeobox 1	1.496134127	2.70E-05	0.015596484
IRX2	ENSG00000170561	iroquois homeobox 2	1.161748374	4.51E-09	1.77E-05
JAZF1-AS1	ENSG00000234336	JAZF1 antisense RNA 1	-2.132735824	0.000107563	0.037119743
KLKB1	ENSG00000164344	kallikrein B, plasma (Fletcher factor) 1	2.478417669	8.84E-05	0.032879001
LAMC1	ENSG00000135862	laminin, gamma 1 (formerly LAMB2)	-1.201150815	2.84E-07	0.000453983
LHX9	ENSG00000143355	LIM homeobox 9	4.298162126	1.45E-06	0.001650048
LINC00537	ENSG00000232815	long intergenic non-protein coding RNA 537	-3.494641854	3.83E-05	0.019780947
LINC00578	ENSG00000228221	long intergenic non-protein coding RNA 578	-4.541164267	7.46E-09	2.39E-05
LINC01048	ENSG00000230390	long intergenic non-protein coding RNA 1048	2.214309895	3.99E-05	0.019780947
LINC01117	ENSG00000224577	long intergenic non-protein coding RNA 1117	-2.17227763	2.17E-05	0.01316806
LRRC16A	ENSG00000079691	leucine rich repeat containing 16A	-1.784218256	2.43E-05	0.014245821
MAP6	ENSG00000171533	microtubule-associated protein 6	-1.79656788	1.30E-06	0.001530605
MAP7D2	ENSG00000184368	MAP7 domain containing 2	2.359306267	2.66E-06	0.002574523
MBP	ENSG00000197971	myelin basic protein	-3.402601353	7.14E-05	0.029220727
MCCC1	ENSG00000078070	methylcrotonoyl-CoA carboxylase 1 (alpha)	-0.554719912	0.000128238	0.042584679
MECOM	ENSG00000085276	MDS1 and EVI1 complex locus	-4.511721791	3.14E-09	1.58E-05
METTL9	ENSG00000197006	methyltransferase like 9	0.906519962	3.86E-08	9.70E-05
MFGE8	ENSG00000140545	milk fat globule-EGF factor 8 protein	-2.694104724	7.44E-07	0.000934855
MLLT3	ENSG00000171843	myeloid/lymphoid or mixed-lineage leukemia (trithorax homolog, Drosophila); translocated to, 3	0.969013225	2.10E-05	0.01316806
MRVI1-AS1	ENSG00000177112	MRVI1 antisense RNA 1	-3.884388307	1.04E-05	0.007509783
MYO16	ENSG00000041515	myosin XVI	-1.962481605	0.000117031	0.039233207
NDP	ENSG00000124479	Norrie disease (pseudoglioma)	3.545635871	1.83E-05	0.011727901
NR5A2	ENSG00000116833	nuclear receptor subfamily 5, group A, member 2	-4.138425423	8.61E-06	0.006447287
NTRK3	ENSG00000140538	neurotrophic tyrosine kinase, receptor, type 3	2.777060262	2.95E-05	0.016740854
PKIA	ENSG00000171033	protein kinase (cAMP-dependent, catalytic) inhibitor alpha	-2.779337899	4.26E-05	0.020550766
PNLDC1	ENSG00000146453	poly(A)-specific ribonuclease (PARN)-like domain containing 1	2.95490495	2.27E-07	0.000399745
PNMA6C	ENSG00000235961	paraneoplastic Ma antigen family member 6C	-4.033899706	0.000151809	0.047958078
POMZP3	ENSG00000146707	POM121 and ZP3 fusion	-1.717492119	3.12E-06	0.002749475
POSTN	ENSG00000133110	periostin, osteoblast specific factor	2.83484857	2.14E-05	0.01316806
PRKAR1B	ENSG00000188191	protein kinase, cAMP-dependent, regulatory, type I, beta	-1.732349926	1.78E-05	0.011579369
PRLR	ENSG00000113494	prolactin receptor	3.300099077	0.000140582	0.045398939
PTK2B	ENSG00000120899	protein tyrosine kinase 2 beta	-1.47679232	6.10E-05	0.026827045
RGCC	ENSG00000102760	regulator of cell cycle	-3.93550099	7.06E-05	0.029220727
RHBDD3	ENSG00000100263	rhomboid domain containing 3	-0.996433377	1.05E-05	0.007509783
RSPO1	ENSG00000169218	R-spondin 1	-4.313022549	5.98E-05	0.026775571
S1PR1	ENSG00000170989	sphingosine-1-phosphate receptor 1	-3.093512898	5.52E-06	0.00452094

SCN1A	ENSG00000144285	sodium channel, voltage-gated, type I, alpha subunit	-4.906276861	3.80E-06	0.003184547
SCN9A	ENSG00000169432	sodium channel, voltage-gated, type IX, alpha subunit	-3.142530704	1.83E-06	0.001898572
SDK1	ENSG00000146555	sidekick cell adhesion molecule 1	3.188897352	4.55E-05	0.021372366
SFRP1	ENSG00000104332	secreted frizzled-related protein 1	-4.10661345	5.21E-07	0.000732904
SHE	ENSG00000169291	Src homology 2 domain containing E	3.455259022	1.63E-06	0.001733572
SLA	ENSG00000155926	Src-like-adaptor	-2.701260288	3.99E-05	0.019780947
SLC16A11	ENSG00000174326	solute carrier family 16, member 11	-2.677392243	4.37E-05	0.020765473
SLC30A8	ENSG00000164756	solute carrier family 30 (zinc transporter), member 8	6.104148946	1.35E-07	0.000250798
SLC35D3	ENSG00000182747	solute carrier family 35, member D3	9.179494017	2.46E-31	2.89E-27
SLC4A3	ENSG00000114923	solute carrier family 4 (anion exchanger), member 3	-2.366325383	7.28E-05	0.029469479
SMOC1	ENSG00000198732	SPARC related modular calcium binding 1	-5.03813069	0.000104067	0.036268803
SPON2	ENSG00000159674	spondin 2, extracellular matrix protein	-2.639592845	0.000112611	0.038484373
SRD5A2	ENSG00000049319	steroid-5-alpha-reductase, alpha polypeptide 2 (3-oxo-5 alpha-steroid delta 4-dehydrogenase alpha 2)	4.448074003	8.97E-05	0.032879001
STAG3L2	ENSG00000160828	stromal antigen 3-like 2 (pseudogene)	-0.881204831	9.94E-05	0.035716372
TCF21	ENSG00000118526	transcription factor 21	-6.04609974	1.52E-06	0.001676294
ULK4	ENSG00000168038	unc-51 like kinase 4	1.147447551	6.46E-09	2.28E-05
WASF3	ENSG00000132970	WAS protein family, member 3	0.972046987	2.81E-06	0.002603617
XKR9	ENSG00000221947	XK, Kell blood group complex subunit-related family, member 9	-1.989973823	0.000159916	0.049814665
ZBTB16	ENSG00000109906	zinc finger and BTB domain containing 16	-4.844890791	7.83E-05	0.030946745
ZFPM2	ENSG00000169946	zinc finger protein, FOG family member 2	-3.60011069	5.25E-14	4.62E-10
ZNF804A	ENSG00000170396	zinc finger protein 804A	-5.483825868	4.34E-09	1.77E-05

Supplementary Table S6. GO enrichment analysis. Show the distribution of GO terms exhibiting statistical significant differences of mis-regulated transcripts in A.IV-1 fibroblast using WebGestalt 2019 (Wang, Vasaikar, Shi, Greer, & Zhang, 2017).

GO terms	Description	Gene set size	Overlap	Expected Value	Enrichment Ratio	pValue	FDR	overlap_Id	user_Id
GO:0009888	tissue development	1613	24	7.625171279802685	3.1474702822178497	1.528766712111107e-7	0.0012670418509976855	249;1116;1301;1844;1893;1901;2185;2254;2494;3484;3664;4300;4916;6422;6943;7704;10631;23414;25884;28984;54507;59269;79192;153572	TCF21;ZBTB16;COL11A1;CHRDL2;NR5A2;SFRP1;RGCC;IRF6;ZFPM2;ADAMTSL4;S1PR1;ALPL;HIVEP3;PTK2B;ECM1;MLLT3;IRX2;IRX1;NTRK3;POSTN;IGFBP1;DUSP2;FGF9;CHI3L1
GO:0001503	ossification	339	11	1.6025623458481775	6.864007524261469	4.940246167706164e-7	0.0020472380118974343	249;1301;1893;1901;2185;2254;3486;6422;7704;25884;64093	SMOC1;ZBTB16;COL11A1;CHRDL2;SFRP1;S1PR1;ALPL;PTK2B;ECM1;IGFBP3;FGF9
GO:0030038	contractile actin filament bundle assembly	76	5	0.35927651411345574	13.916857360793287	2.9031950272262463e-5	0.060154200964127824	1901;2185;6422;28984;55604	SFRP1;RGCC;S1PR1;CARMIL1;PTK2B
GO:0043149	stress fiber assembly	76	5	0.35927651411345574	13.916857360793287	2.9031950272262463e-5	0.060154200964127824	1901;2185;6422;28984;55604	SFRP1;RGCC;S1PR1;CARMIL1;PTK2B
GO:0048584	positive regulation of response to stimulus	1768	21	8.35790627569197	2.512590989573087	4.138463705938289e-5	0.06101540527461857	1116;1123;1893;1901;2070;2185;2254;2533;3486;3818;4300;4916;5618;6422;6503;10417;28984;53832;55540;126669;284654	FYB;RSPO1;EYA4;SFRP1;RGCC;S1PR1;SLA;SPON2;CHN1;PTK2B;ECM1;MLLT3;IGFBP3;KLKB1;NTRK3;IL20RA;IL17RB;PRLR;SHE;FGF9;CHI3L1
GO:0007379	segment specification	15	3	0.07090983831186626	42.3072463768116	4.417138412737831e-5	0.06101540527461857	4300;79192;153572	MLLT3;IRX2;IRX1
GO:0030278	regulation of ossification	169	6	0.7989175116470265	7.510162078723952	1.4457514683041683e-4	0.12546232977013955	1893;1901;2185;6422;7704;64093	SMOC1;ZBTB16;SFRP1;S1PR1;PTK2B;ECM1
GO:0048646	anatomical structure formation involved in morphogenesis	876	13	4.14113455741299	3.1392363179141025	2.015277408380367e-4	0.12546232977013955	1116;1301;1844;1893;1901;2185;2254;6422;6943;28984;79192;153572;221935	TCF21;COL11A1;SFRP1;RGCC;S1PR1;PTK2B;ECM1;IRX2;IRX1;SDK1;DUSP2;FGF9;CHI3L1
GO:0001501	skeletal system development	439	9	2.075294601260619	4.33673368327226	2.1232290176698143e-4	0.12546232977013955	249;1116;1301;2254;3226;6422;7704;10631;25884	HOXC10;ZBTB16;COL11A1;CHRDL2;SFRP1;ALPL;POSTN;FGF9;CHI3L1
GO:0072047	proximal/distal pattern formation involved in nephron development	5	2	0.023636612770622088	84.61449275362318	2.182366244216638e-4	0.12546232977013955	79192;153572	IRX2;IRX1

Supplementary Table S7. Summary of transcriptome mapping results.

Sample ID	#total reads	#aligned reads	% aligned reads	#uniquely mapped	% uniquely aligned reads
A.IV-1.1	306,265,974	303,190,132	99.00	293,189,972	96.70
A.IV-1.2	344,782,588	341,370,324	99.01	330,182,914	96.72
ARN1.1	338,129,498	335,557,854	99.24	325,618,112	97.04
ARN1.2	301,339,526	298,096,006	98.92	289,521,446	97.12
ARN2.1	311,917,536	307,641,094	98.63	296,837,798	96.49
ARN2.2	336,751,354	333,728,602	99.10	322,903,810	96.76
ARN3.1	403,718,870	400,433,942	99.19	383,740,470	95.83
ARN3.2	299,879,030	296,763,108	98.96	286,203,188	96.44
ARN4.1	318,588,494	315,510,058	99.03	304,438,002	96.49
ARN4.2	320,107,070	316,729,460	98.94	304,559,776	96.16
ARN5.1	363,584,984	360,496,138	99.15	349,081,194	96.83
ARN6.1	324,984,406	321,703,444	98.99	311,631,678	96.87
ARN6.2	364,763,088	361,575,516	99.13	349,853,468	96.76

SUPPLEMENTARY REFERENCES

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., & Salzberg, S. L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol, 14*(4), R36. doi:10.1186/gb-2013-14-4-r36

Langmead, B., Trapnell, C., Pop, M., & Salzberg, S. L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol, 10*(3), R25. doi:10.1186/gb-2009-10-3-r25

Nevers, Y., Prasad, M. K., Poidevin, L., Chennen, K., Allot, A., Kress, A., . . . Lecompte, O. (2017). Insights into Ciliary Genes and Evolution from Multi-Level Phylogenetic Profiling. *Mol Biol Evol, 34*(8), 2016-2034. doi:10.1093/molbev/msx146

Wang, J., Vasaikar, S., Shi, Z., Greer, M., & Zhang, B. (2017). WebGestalt 2017: a more comprehensive, powerful, flexible and interactive gene set enrichment analysis toolkit. *Nucleic Acids Research, 45*(W1), W130-W137. doi:10.1093/nar/gkx356

Input	Status
NM_152558.4:c.895_904del	OK
NM_003322.5:c.1198C>T	OK
NM_152558.4:c.895_904del	OK
NM_152558.4:c.1350_1353del	OK
NM_001692.3:c.175-1G>C	OK