UCLA
UCLA Previously Published Works

Title

INTS13 variants causing a recessive developmental ciliopathy disrupt assembly of the Integrator complex.

Permalink

Journal

Authors

Mascibroda, Lauren
Shboul, Mohammad
Elrod, Nathan
et al.

Publication Date

DOI

Peer reviewed

INTS13 variants causing a recessive developmental ciliopathy disrupt assembly of the Integrator complex

Lauren G. Mascibroda [1,21], Mohammad Shboul [2,21], Nathan D. Elrod [1], Laurence Colleaux[3], Hanan Hamamy[4], Kai-Lieh Huang [1,5,6], Natoya Peart [1], Moirangthem Kiran Singh [1], Hane Lee[7,20], Barry Merriman[7], Jeanne N. Jodoin[8], Poojitha Sitaram [8], Laura A. Lee[9], Raja Fathalla[10], Baeth Al-Rawashdeh [11], Osama Ababneh[11], Mohammad El-Khateeb[10], Nathalie Escande-Beillard [12,13], Stanley F. Nelson[7], Yixuan Wu[14], Liang Tong [14], Linda J. Kenney [1], Sudipto Roy[13,15,16], William K. Russell [1], Jeanne Amiel[17], Bruno Reversade [12,13,15,18,19] ✉ & Eric J. Wagner [1,5,6] ✉

Oral-facial-digital (OFD) syndromes are a heterogeneous group of congenital disorders characterized by malformations of the face and oral cavity, and digit anomalies. Mutations within 12 cilia-related genes have been identified that cause several types of OFD, suggesting that OFDs constitute a subgroup of developmental ciliopathies. Through homozygosity mapping and exome sequencing of two families with variable OFD type 2, we identified distinct germline variants in *INTS13*, a subunit of the Integrator complex. This multi-protein complex associates with RNA Polymerase II and cleaves nascent RNA to modulate gene expression. We determined that INTS13 utilizes its C-terminus to bind the Integrator cleavage module, which is disrupted by the identified germline variants p.S652L and p.K668Nfs*9. Depletion of *INTS13* disrupts ciliogenesis in human cultured cells and causes dysregulation of a broad collection of ciliary genes. Accordingly, its knockdown in *Xenopus* embryos leads to motile cilia anomalies. Altogether, we show that mutations in *INTS13* cause an autosomal recessive ciliopathy, which reveals key interactions between components of the Integrator complex.

Ciliopathies encompass a large and expanding group of human genetic disorders caused by mutations in genes either encoding ciliary proteins or that participate indirectly in the formation and function of cilia[1-4]. There have been several hundred diseases identified as ciliopathies that range from renal cystic disease[5] and isolated blindness[6] to multiple organ disorders such as a Bardet-Biedl syndrome (BBS)[7] or Meckel Gruber syndrome (MKS)[8]. The clinical manifestations associated with ciliopathies are broad and highly variable and include: retinal degeneration, specific nervous system malformations, hepatic, kidney and pancreatic cysts, heterotaxia, polydactyly, encephalocele,

hydrocephalus, hearing loss, anosmia, intellectual disability, developmental delay, skeletal malformation, craniofacial anomalies, respiratory function impairment, infertility and obesity[4,9-13].

Among ciliopathies, oral-facial-digital syndromes (OFDs) are a group of clinically and genetically heterogeneous developmental disorders characterized by defects in the development of the face and oral cavity along with digit anomalies. Symptoms associated with OFDs can also be found in other organs, which has led to the classification of at least 14 different types of OFD[14,15]. To date, 12 genes have been identified to be mutated in OFD syndromes and the encoded proteins

have been implicated in primary cilia biogenesis and function. The phenotypic spectrum of OFD syndromes can overlap with other known ciliopathies, and mutations in OFD genes can also cause other types of ciliopathies[16–22]. Of the 12 known genes involved in OFDs, the majority encode structural components of the primary cilium, such as *OFD1*[23] and *TMEM231*[24]. However, not all OFD-causing gene mutations such as those in the RNA helicase DDX59[20] localize to the primary cilium, and are thus thought to contribute indirectly to disease pathogenesis. Despite significant progress, seven OFD syndromes are still without a clear genetic aetiology and at least three unclassified forms need further clinical and molecular characterization.

The Integrator complex (INT), which is unique to metazoans, is composed of at least 15 subunits, and it associates with RNA polymerase II (RNAPII) where it modulates transcription[25–29]. Critical among its protein components is Integrator subunit 11 (INTS11)[30], which is a member of the metallo-β-lactamase/β-CASP family of RNA endonucleases[31] and is a paralog of the cleavage and polyadenylation specificity factor of 73 kDa (CPSF73), which cleaves pre-mRNA[32]. INTS11 has been found to associate with INTS4 and INTS9 to form what is called the Integrator Cleavage Module (ICM), which is thought to be the key enzymatic factor within the complex[33]. The ICM, through the activity of INTS11, has been shown to be essential to cleave the 3′ ends of a variety of noncoding RNA and nascent mRNA of protein-coding genes with paused RNAPII. This catalytic activity thus serves as a key functional component of Integrator-mediated transcriptional repression and for the biogenesis of noncoding RNA[34–38].

The potential connection between Integrator function and ciliogenesis was uncovered initially in *Drosophila*, where depletion of nearly any Integrator subunit disrupts proper centriole localization, an early event in ciliogenesis[39,40]. Further evidence was obtained using human retinal pigment epithelial (RPE) cells, where depletion of multiple Integrator subunits can lead to inhibition of primary cilium biogenesis[41]. More recently, an unbiased CRISPR screen for factors required for ciliogenesis also identified several Integrator subunits as essential to form primary cilia[42]. Collectively, these data imply that the overall function of the Integrator complex, and not any specific subunit, is required for proper cilia formation and function. Given its role in modulating RNAPII behavior at diverse gene types, the precise function for Integrator in ciliogenesis is most likely at the level of gene expression. However, the ciliary genes subject to Integrator regulation are currently unknown, and no ciliopathies have been found associated with Integrator subunit mutations in humans.

Here, we used a combination of homozygosity mapping and deep sequencing of patients with orphan diseases to identify two distinct recessive germline variants in the *INTS13* (MIM615079) gene. These private mutations were associated with a variable form of OFD syndrome type 2 (OFD2), also known as Mohr Syndrome (MIM252100)[43]. Using multiple molecular and biochemical approaches, we demonstrate that the C-terminus of INTS13 interacts with the ICM and that mutations observed in these patients specifically disrupt this interaction. We also find that human cells depleted of INTS13 not only have reduced primary cilia, but also exhibit broad transcriptional dysregulation that impacts numerous known genes required for proper ciliogenesis. These observations were validated in vivo, where morpholino-mediated depletion of *ints13* in *Xenopus laevis* leads to profound underdevelopment, with a specific reduction in the formation and beating of multiciliated cells. Taken together, these results reveal a key role for Integrator in ciliogenesis and identify a vital interaction between two Integrator subcomplexes.

Results
Characterization of 4 patients with recessive loss-of-function *INTS13* variants
In Family 1, two affected sisters (individuals II.4 and II.5 in Fig. 1a) born to a consanguineous Jordanian couple presented with a congenital

disorder consistent with OFD2-like syndrome (MIM252100). The two sisters displayed similar phenotypes encompassing craniofacial dysmorphisms, including: a bilateral cleft lip (CL), microcephaly, hypertelorism, epicanthal fold, broad nasal root with depressed nasal tip, large and low set ears, long philtrum and thick septum. In addition to bilateral CL, their oral cavity had omega-shaped epiglottis, underdeveloped epiglottic folds and hypertrophied false vocal cords. Dental abnormalities were also observed, such as central median upper incisor in II.4 and single central upper incisor in II.5. Supernumerary teeth were observed in II.4. Both sisters have clinodactyly of the 4th and 5th fingers and toes, and pre-axial polydactyly was observed in one hand of II.4. Single transverse palmar creases on both hands and rough hair were also observed in both sisters. Audiologic examination of the two sibs revealed the presence of tympanic membrane retraction and calcification (myringosclerosis). Ophthalmic examination revealed dilated retinal vessels and a crowded optic disk. Patients were reported to have chronic urinary tract infection and recurrent respiratory tract infections with chronic cough. Supravalvular pulmonary stenosis and partial anomalous pulmonary venous return were reported in patient II.5 soon after birth. Taken together, both patients share the cardinal features of an Oral-Facial-Digital Syndrome and exhibit a wide range of clinical manifestations. The clinical findings of these patients are illustrated in Supplementary Fig. 1a–m, and a summary is provided in Supplementary Data 1. Inter- and intra-familial variations have been reported in OFD syndromes and were clearly observed between the two affected siblings. While we contend that these phenotypes overlap with OFD2, it is also possible that some of the unique features seen in these sisters delineate a new type of OFD with its own genetic aetiology.

To identify the genetic origin of this developmental disorder, the parents, both affected siblings, and healthy siblings were screened on a 610k Illumina SNP array. A subsequent homozygosity identical-by-descent (IBD) mapping revealed five homozygous candidate regions that were unique to the proposita, with a total size of 45.76 Mb on chromosomes 6, 8, 12, 13, and 16 (Fig. 1b). We next employed targeted genomic loci capture followed by massive parallel sequencing to screen 395 candidate genes in these loci[44]. Out of 894 mismatches, one homozygous missense mutation (c.6585 A > T) in *SACS* (MIM604490) and a frameshift (c.2004delA) mutation in *INTS13* (MIM615079) were found in the two affected sisters. The mutation in *SACS* was ruled out as the likely cause due to the gene's known role in causing spastic ataxia (MIM 604490), which did not align with the patients' phenotypes[45,46]. The single base pair (bp) deletion in exon 16 of *INTS13* was never observed in GnomAD and is predicted to be damaging according to SIFT or MutationTaster. It alters the reading frame, leading to a premature termination codon (PTC) at p.K668Nfs*9 and the deletion of the remaining 31 amino acids of the INTS13 protein (Fig. 1c). This private allele segregated with the disease in all family members available (Supplementary Fig. 1n).

The second family had two affected siblings born to French parents not known to be endogamous. The affected 27-year-old woman (II.2) and her affected 20-year-old brother (II.4) (Family 2, Fig. 1a) presented with some overlapping phenotypes, but also distinct features from those reported in the Jordanian patients. Both affected sibs were born at term with low birth parameters for weight, length and occipitofrontal circumference (OFC). II.4 presented with hypospadias. Both remained short and became overweight from the age of 3 years, while OFC was in the low normal range at this age. II.2 walked unaided before 18 months and II.4 walked at 22 months. For both, language was delayed and remained limited to a few spoken words, although they could understand simple conversational topics. II.2 was treated for early-onset puberty for 18 months starting at 8 years. Both presented with short extremities with brachydactyly of fingers and toes, hyperlordosis when standing, flessum of the knees and pes valgus. Facial features include upslanting palpebral fissures, macrostomia, gum
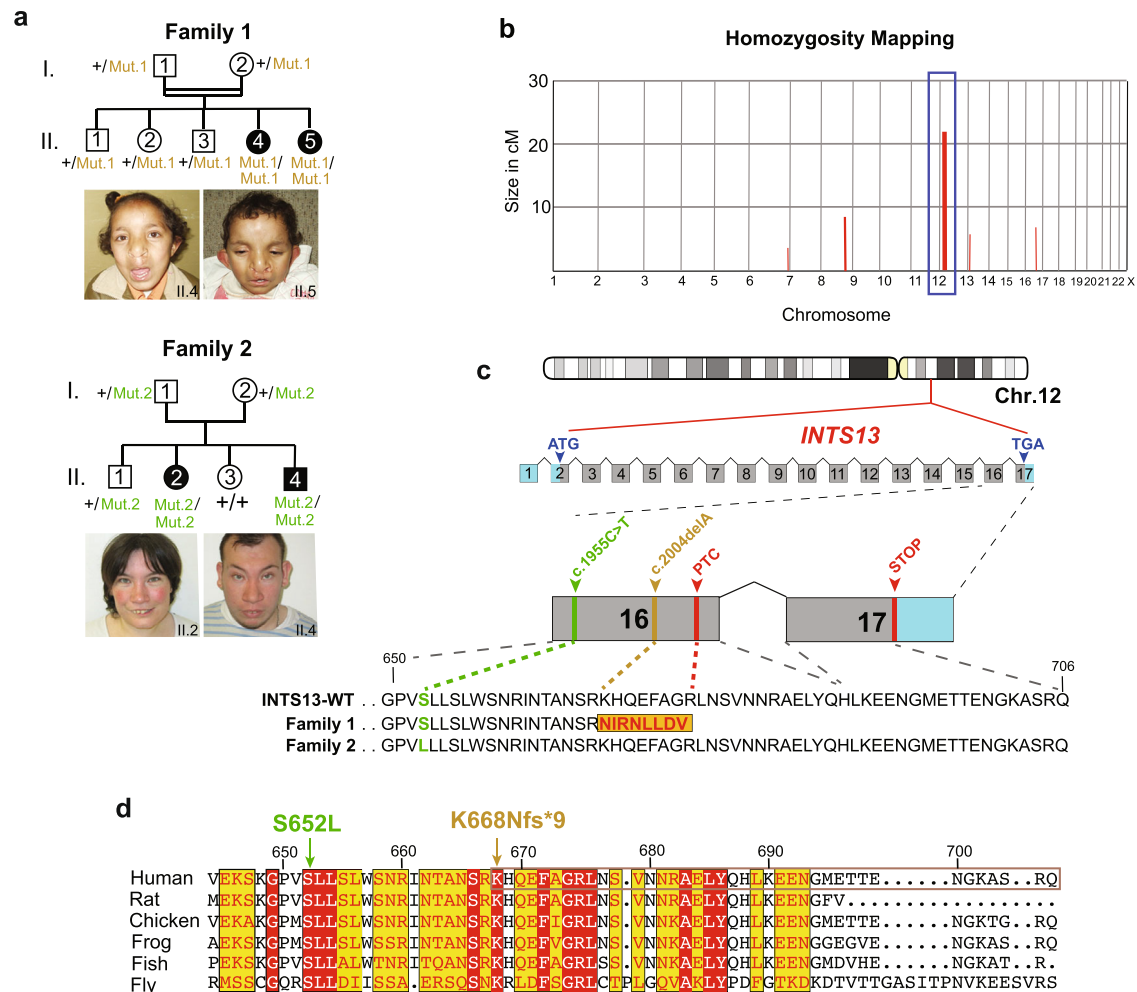
Fig. 1 | Identification of two families segregating recessive INTS13 variants.
a Pedigree of Family 1 (top): affected individuals (II.4 and II.5) were born to first-cousin parents (I.1 and I.2) with three unaffected children (II.1-II.3). Double lines indicate a consanguineous marriage. Pedigree of Family 2 (bottom): affected individuals (II.2 and II.4) were born to parents with two unaffected children (II.1 and II.3). **b** Homozygosity mapping of Family 1 delineated five candidate loci totaling 46 cM on chromosomes 6, 8, 12, 13, and 16. **c** Schematic representation of human *INTS13* located on chromosome 12 (ch12p13.2-p11.22) consisting of 17 exons. Locus capture followed by massive parallel sequencing identified a homozygous

frameshift mutation as a disease-causing mutation in Family 1. The c.2004delA (p.K668Nfs*9) mutation is caused by a single base pair deletion in exon 16 (marked in brown), resulting in a frameshift and premature termination codon (PTC) which alters nine amino acids (marked in red) and deletes the last 31. In Family 2, the c.1955C > T (p.S652L) mutation is caused by a single base pair substitution in exon 16 (marked in green). **d** Amino acid alignment of the C-terminus of the INTS13 protein shows that both mutations occur at highly conserved residues. Red and yellow shading indicates regions of conservation. Source data are provided as a Source Data file.

hypertrophy and a bulbous nasal tip. They wear glasses for correction of hyperopia and astigmatism. At 20 years, regression with less social interactions, vocal stereotypies and tip toe walking were noted in II.4, while his brain CT-scan showed some cerebellar atrophy and calcification of the cerebellum tent. II.2 has remained stable with no regression observed thus far. The clinical findings of these patients are illustrated in Supplementary Figure 1o–w, and a summary is provided in Supplementary Data 1.

Agilent 60k Array CGH showed no disequilibrium in either sibling. Exome capture was performed followed by sequencing of both affected siblings and their healthy parents. Each sample had an overall mean depth of coverage greater than 90X and more than 97% of the exome covered at least 15X. Filtering for germline compound heterozygous variants, X-linked or shared heterozygous de novo variants did not provide any candidate; however, a single shared homozygous variant in *INTS13* (NM_018164.2: c.1955C > T) emerged as the only candidate mutation (Fig. 1c). This mutation converts Serine 652 to a Leucine and is located 16 amino acids upstream of where the frameshift mutation was observed in Family 1.

INTS13 consists of 17 exons (Fig. 1c). Its encoded protein is 706 amino-acid residues long and was initially termed Mat89Bb (Maternal transcript 89Bb) due to its maternal expression in *Drosophila* embryos[47,48]. Mat89Bb was later renamed Asunder/ASUN given its mutant phenotype of altered association of the nuclei with asters and disrupted spermatogenesis[49]. Importantly, ASUN was later shown to be a component of the Integrator complex, although little is known of its distinct function in that context[50]. There is an overall high degree of sequence similarity within the C-termini of INTS13 proteins found in multiple organisms, and both mutated residues are highly conserved between vertebrate and invertebrate species (Fig. 1d), suggesting that alteration of these amino acids may lead to a pathogenic function of the INTS13 protein.

INTS13 mutations affect endogenous INTS13 protein levels in patient cells
To investigate the impact of *INTS13* mutations on its expression, we isolated primary cutaneous fibroblasts from affected patients in both families as well as fibroblasts from healthy individuals. We first
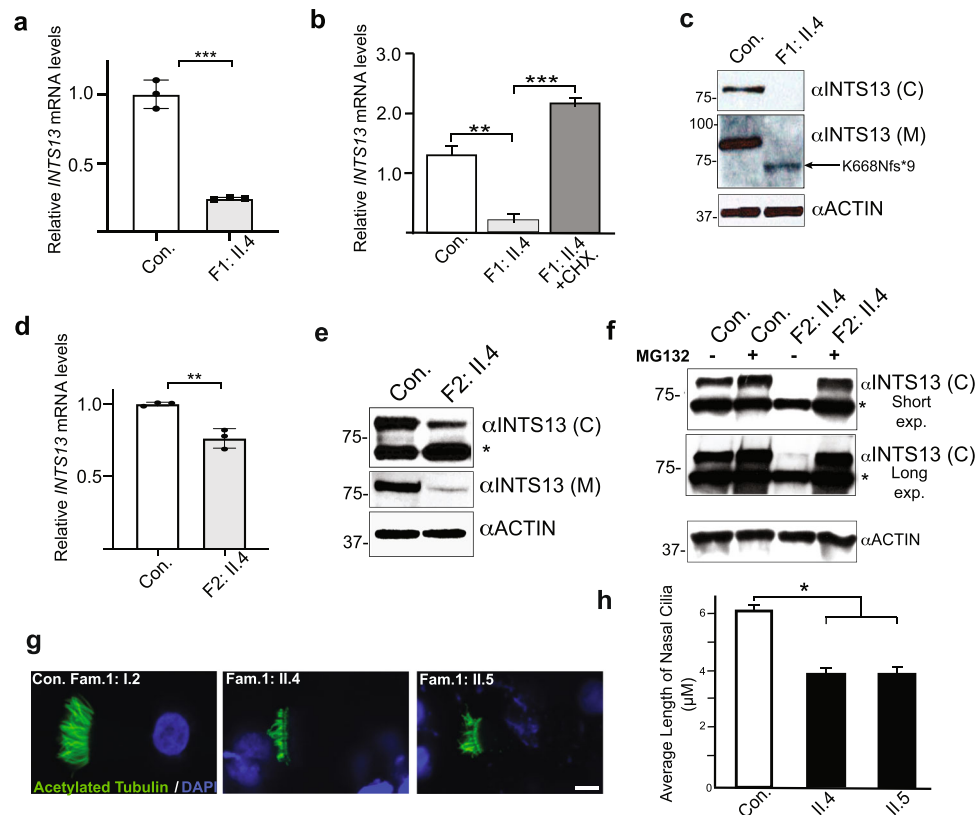
Fig. 2 | Characterization of INTS13 variants using patients' cells. a RT-qPCR shows that endogenous *INTS13* transcript levels were significantly reduced relative to control cells in primary dermal fibroblasts of affected individual II.4 from Family 1. Data are mean +/- SD, $n = 3$ replicate cell samples. Statistical significance was calculated using a two-tailed unpaired *t*-test. $p = 0.000223$. **b** Primary dermal fibroblasts were treated with cycloheximide (CHX) (100 g/ml). Following treatment, qRT-PCR was done at 8 h and shows that *INTS13* mRNA level was increased in treated cells indicating a non-sense mediated decay. Data are mean ± s.e.m., *$P < 0.05$; **$P < 0.005$; ***$P < 0.001$; one-tailed Student's *t*-test. **c** Western blot analysis of primary dermal fibroblasts of affected individual II.4 from Family 1 using two polyclonal antibodies with antibody directed against the C-terminus of INTS13 (C) and the second antibody directed against the center of INTS13 (M). Actin was used as a loading control and molecular weight markers shown are in kilodaltons. **d** RT-qPCR shows that *INTS13* transcript levels were significantly reduced in primary dermal fibroblast cells of affected individual II.4 from Family 2 relative to control

cells. Data are mean +/- SD, $n = 3$ replicate cell samples. Statistical significance was calculated using a two-tailed unpaired *t*-test. $p = 0.003833$. **e** Western blot analysis of primary dermal fibroblasts of affected individual II.4 from Family 2 shows reduced levels of INTS13 protein compared to control cells. **f** Western blot analysis of primary dermal fibroblasts from individual II.4 from Family 2 treated with MG132. Treated cells show a level of INTS13 protein comparable to control cells. * indicates a non-specific signal. Multiple western blots were developed (> 3), and the results shown are representative of the data obtained. **g** Multiciliated airway cells were obtained from nasal biopsies of three members of Family 1: the unaffected carrier mother (Con.) and the two affected children (II.4 and II.5). Probing for acetylated α-tubulin shows significantly shorter, less dense, and disorganized cilia in the affected individuals compared to the mother's cells. scale bar is 5 μm. **h** The average length of nasal cilia is significanty reduced in the two affected sisters compared to that of their mother. Data are mean ± s.e.m., ***$P < 4 \times 10^{-9}$, $n \geq 31$ cells for each patient in one biological replicate. Source data are provided as a Source Data file.

compared mRNA expression levels by RT-qPCR of *INTS13* in Family 1 cells (F1) containing the INTS13 truncation mutation. We found the level of *INTS13* transcripts to be significantly decreased (Fig. 2a), which we hypothesized may be due to the presence of the premature termination codon triggering nonsense-mediated decay (NMD). Therefore, we isolated total RNA from patient cells that were either mock-treated or treated with cycloheximide, a translational inhibitor that blocks NMD, and RT-qPCR was repeated. The results showed that the level of mutant mRNA was increased following cycloheximide treatment, suggesting that the mutant *INTS13* mRNA was indeed targeted for NMD (Fig. 2b). To assess the levels of INTS13 protein in F1 cells, we raised two custom antibodies against antigens from the central portion of INTS13 (α-M) and the C-terminal region (α-C), which is downstream of the truncation. These antibodies were affinity-purified and used to probe cell lysates from patient fibroblasts using western blot analysis. We observed that the C-terminal antibody could detect endogenous INTS13 protein in control cells, but not in F1 patient fibroblasts, likely due to the loss of the epitope in the truncated mutant (Fig. 2c). Consistent with this notion, we could detect low levels of truncated INTS13 in F1 patient fibroblasts using the α-M antibody

(Fig. 2c). Overall, these results reveal that the *INTS13* mutation observed in Family 1 creates an mRNA subject to NMD that in turn leads to low levels of a truncated INTS13 protein.

We next analyzed the expression of INTS13 in fibroblasts isolated from a Family 2 (F2) patient harboring the homozygous p.S652L mutation. We observed a slight but significant reduction in the level of endogenous *INTS13* mRNA in these mutant cells (Fig. 2d) and a pronounced reduction in the level of INTS13 protein as assessed using our custom antibodies (Fig. 2e). We posited that the basis for the reduction in INTS13 protein, beyond the minor reduction in mRNA, may be due to increased protein degradation in F2 mutant cells. To test this, primary fibroblasts were treated with MG132 to inhibit proteasome activity, and we observed an increased level of INTS13 protein comparable to control cells by western blot analysis (Fig. 2f). This suggests that the INTS13 mutation observed in Family 2 generates a protein which is unstable and targeted for degradation via the proteasome.

The epithelia lining the human respiratory system are covered with multiple cell types dominated by ciliated cells. These ciliated cells have hundreds of motile cilia to generate fluid flow that keeps mucus moving to protect the airway from inhaled irritants, pathogens and

foreign particles. Abnormalities of mucociliary clearance caused by either aberrant ciliated cells or cilia dysfunction often result in recurrent respiratory tract infections. Reduction or impairment of mucociliary clearance in humans results in an array of pathologies such as primary ciliary dyskinesia, chronic obstructive pulmonary disease and asthma[51–55]. In this study, both probands of Family 1 suffered from recurrent respiratory tract infection and chronic cough. Therefore, we obtained respiratory epithelial cells from the two patients and the mother to assess cilia formation and performance. Immunofluorescence staining was performed using anti-acetylated α-Tubulin antibody to visualize cilia. The average length of nasal cilia was then scored. The results showed that cilia were shorter, fewer in number and occasionally disorganized compared to those of the heterozygous asymptomatic mother, which may explain the recurrent respiratory tract infections in the two affected subjects (Fig. 2g, h). Importantly, these results also indicate that ciliogenesis was indeed disrupted in patient cells.

The patient mutations within the INTS13 C-terminus disrupt association with the Integrator Cleavage Module

The structure of INTS13 associated with INTS14 has been recently elucidated[56]. However, the INTS13 C-terminus was unable to be resolved in that structure, thereby obscuring our ability to ascertain the impact of OFD2 mutations[56]. The INTS13/14 heterodimer was shown to interact with INTS10 as well as the Integrator CM, the latter of which was found to be mediated independently through the INTS13 C-terminus[56]. With this information in hand, and our previous success mapping Integrator interactions using yeast two-hybrid analysis[33,50,57,58], we fused INTS13 to the Gal4 DNA binding domain and screened for interactions with the other Integrator subunits that were each expressed individually as Gal4 activation domain fusions. This initial approach failed to detect interaction between INTS13 and other members of the complex (Fig. 3a, left panels). Given the structural information for INTS13, the most likely explanation for the inability to detect binding partners is that INTS13 may interact with a
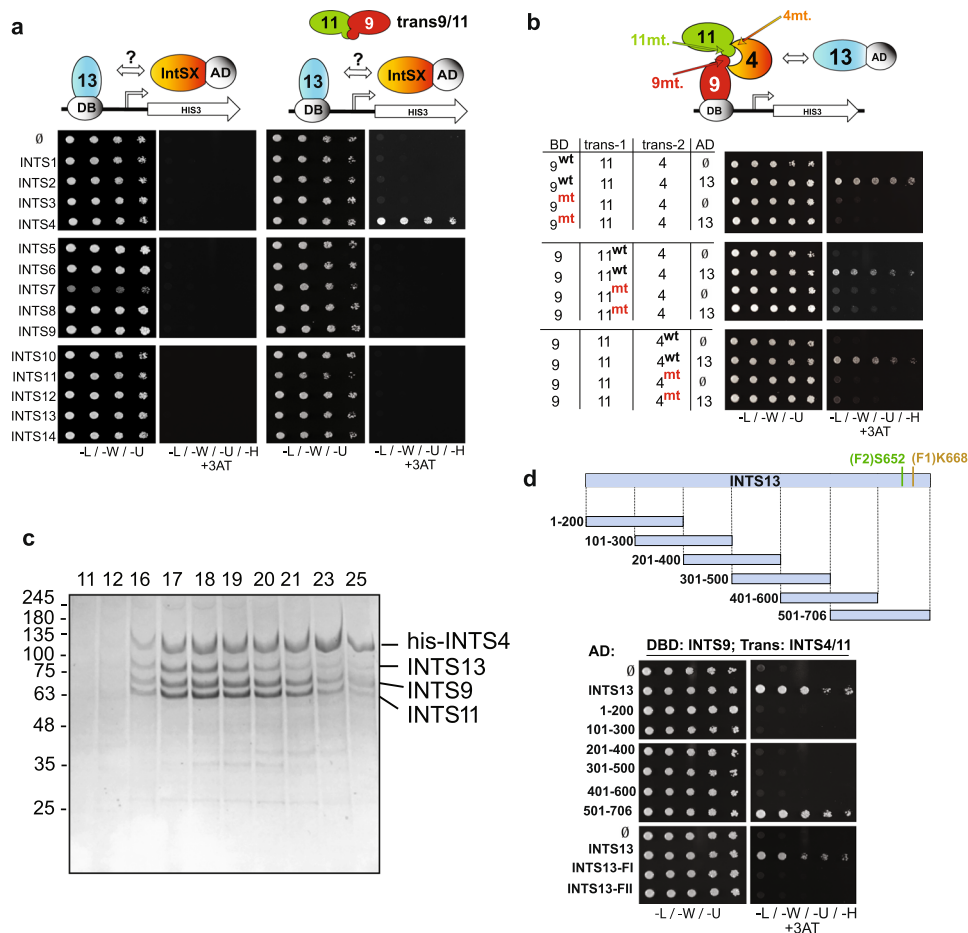

Fig. 3 | INTS13 associates with the INTS4/9/11 heterotrimer. a Yeast two hybrid screening for interactors of INTS13. Haploid strains of yeast expressing each individual Integrator subunit fused to the Gal4 activating domain (AD) were mated with strains expressing INTS13 fused to the DNA binding domain (BD). The first and third panels show mated yeast grown on selective media as a positive control (media lacking leucine, tryptophan, and uracil). The remaining panels are of yeast plated on restrictive media additionally lacking histidine to test for protein-protein interactions and with 3AT (3-amino-1,2,4-triazole) to test the strength of the interaction. No interaction was observed between INTS13 and any other individual subunit (second panel). When INTS9 and INTS11 were expressed in *trans* with INTS13, there was a positive interaction with INTS4 (fourth panel). **b** Point mutants of INTS4, INTS9, and INTS11 known to disrupt their interactions were used in yeast two hybrid to demonstrate that the formation of the INTS4/9/11 heterotrimer is necessary for INTS13 association. General locations of the mutations are indicated with arrows on the schematic. INTS13 is fused to the activating domain in the rest of the yeast two hybrid assays shown in this study. **c** SDS PAGE gel of fractions from gel filtration column of co-expressed full-length INTS4, INTS9, INTS11 and INTS13. The sample was first purified by nickel affinity chromatography. INTS4 carried an N-terminal His-tag, while the other three proteins were untagged. The positions of the fractions are indicated at the top of the gel and molecular weight markers shown are in kilodaltons. Multiple purifications were conducted (> 3), and the gel shown is representative of the results attained. **d** Consecutive 200-amino acid truncations of INTS13 were tested for interaction with INTS4/9/11. In modified yeast two-hybrid assays, the only positive interactions seen are full length INTS13 and the last ~200 residues, 501-706. The patient mutations are marked on the schematic and occur in this region. Loss of interaction with INTS4/9/11 was observed using the mutant versions of INTS13. Source data are provided as a Source Data file.

domain created by an interface between multiple Integrator subunits. We previously demonstrated this to be the case for the INTS4/9/11 Cleavage Module where INTS4 interaction with INTS9 or INTS11 was only detected with yeast two-hybrid if all three subunits were co-expressed[33]. We therefore decided to focus on whether INTS13 could interact with members of the cleavage module and thus repeated the screen with INTS9 and INTS11 expressed simultaneously in *trans*. Using this approach, we could detect a robust interaction between INTS13 and INTS4 only when INTS9 and INTS11 were co-expressed in the same yeast strain (Fig. 3a, right panels and Supplementary Fig. 2a). To confirm this interaction, we altered the position of each subunit in the screen and could detect an interaction between INTS9 and INTS13 only when both INTS4 and INTS11 were simultaneously expressed in *trans* (Supplementary Fig. 2b). These data expand on recent structural and biochemical analyses to indicate that INTS13 can, in the absence of INTS10 and INTS14, associate with subunits of the ICM.

The binding interfaces of the Integrator Cleavage Module have been previously well-characterized using both biochemical and structural approaches[33,58]. Based upon these prior analyses, amino acid mutations within each of the three ICM subunits are known to disrupt formation of the heterotrimer. Therefore, we repeated our modified two-hybrid assay with each of these mutants tested individually and discovered that any disruption of the INTS4/9/11 Cleavage Module also results in loss of INTS13 interaction (Fig. 3b). Further, in order to confirm the interaction between these four subunits using an independent approach, we co-infected High Five cells with baculoviruses individually expressing His-tagged-INTS4, or untagged INTS9, INTS11, and INTS13. The complex was purified from cell lysate using nickel affinity resin and subjected to gel filtration analysis. We observed that all four Integrator subunits remained intact as a complex and could be detected in the same fractions at near stoichiometric levels (Fig. 3c). Collectively, these results indicate that INTS13 interacts with the three components of the ICM only once the module is fully assembled.

We next wanted to determine which region of INTS13 is responsible for this interaction and whether the identified germline mutations may impact it. We tested a series of INTS13 fragments in modified yeast two-hybrid assays to determine whether any of these regions were sufficient to mediate interaction with INTS4/9/11 (Fig. 3d schematic). The INTS13 C-terminal region, composed of amino acids 501-706, was as effective as the full-length INTS13 to support growth on selective media (Fig. 3d). We further dissected the INTS13 C-terminus through a series of smaller deletion mutants, which ultimately defined that the last 130 residues, amino acids 577-706, were sufficient to bind to the Integrator CM (Supplementary Fig. 2c). This result is consistent with biochemical approaches recently described where a GST fusion protein containing the C-terminus of INTS13 can pull down subunits of the Integrator CM[56]. Because both mutations observed in Family 1 (p.K668Nfs*9) and Family 2 (p.S652L) are within the region of INTS13 found to be sufficient to bind INTS4/9/11, we tested whether these mutations could disrupt binding. Indeed, we found that either mutation in the context of the full-length INTS13 protein resulted in a loss of interaction between INTS13 and INTS4/9/11, indicating that this interaction may be specifically disrupted in the identified subjects (Fig. 3d).

The INTS13 C-terminus connects an INTS10/13/14 submodule to the Integrator complex

We next sought to determine if the C-terminal INTS13 mutations have an impact on the assembly of the entire Integrator complex. To do so, we created individual 293 T cell lines harboring three different versions of the *INTS13* cDNA: wild type (wt), Family 1 mutant (F1), or Family 2 mutant (F2). Each of these lines expresses exogenous INTS13 with an N-terminal FLAG tag under the control of the doxycycline-inducible promoter. After 48 h of doxycycline treatment, nuclear fractions from each of the cell lines as well as an untransfected naïve line serving as a negative control were extracted. Purified Integrator complexes

associated with each INTS13 protein, and the degree of co-associated INT subunits was assessed by LC/MS and western blot analysis. We observed that both patient-derived INTS13 mutants not only disrupted association with INTS4/9/11, confirming our yeast two-hybrid experiments, but also resulted in reduced binding to nearly all other members of the complex (Fig. 4a). The exception to this observation was the binding of INTS14 and to a lesser extent INTS10, which were not as dramatically impacted by the mutations (Fig. 4a).

We have shown previously that Integrator constituency and function is highly conserved in *Drosophila*[50]. Therefore, we also examined if the patient mutations could also have disruptive effects on the orthologous fly complex. Supporting this approach, the C-terminal region of INTS13 is conserved in the *Drosophila* IntS13 orthologue (Asunder or dIntS13). We introduced the analogous patient mutations into dIntS13 (Supplementary Figure 3a) and created individual S2 cell lines expressing copper-inducible FLAG-dIntS13-wild type (wt), Family 1 mutant (dF1), or Family 2 mutant (dF2). We generated nuclear extract from these cell lines as well as a naïve S2 control and purified *Drosophila* Integrator complexes using anti-FLAG affinity resin. Similar to the human Integrator purifications, we used both LC/MS and western blot analysis to assess complex integrity. In accordance with the human data, we found that patient mutation 'equivalents' within *Drosophila* IntS13 reduced association with most other Integrator subunits (Fig. 4b). We also observed that association of dIntS10 and dIntS14 were only marginally affected by the patient mutations.

These results confirm that the identified mutations disrupt INTS13 assembly into Integrator, but also suggest that INTS13 mutants do not affect INTS14 and possibly INTS10 binding. Next we set out to determine whether an interaction of INTS13 with INTS10 and INTS14 could be documented using our modified two-hybrid assay. We initially expressed INTS14 fused to the Gal4 DNA binding domain and screened for interactions individually with each of the other Integrator subunits fused to the activating domain. This approach did not detect any pairwise interactions (Supplementary Fig. 3b, left panel). Therefore, we expressed each Integrator subunit individually in trans with INTS14 fused to the DNA binding domain and INTS13 fused to the activating domain. In this context, we found that the additional presence of INTS10 could allow for an interaction between INTS14 and INTS13, supporting the model that these three subunits form a module as has been recently shown biochemically[56] (Supplementary Fig. 3b, right panel). Finally, we tested whether our patient-derived mutations within INTS13 affect the interaction of INTS10/13/14 and observed that both INTS13 mutants could interact with INTS10/14 comparable to the wild-type INTS13 (Fig. 4c). Overall, the results conveyed here provide a molecular basis for OFD2 where INTS13 mutations disrupt interaction between the INTS10/13/14 submodule and the INTS4/9/11 Cleavage Module (Fig. 4d).

Reduced INTS13 expression in RPE cells leads to altered ciliogenesis and disruption of transcriptional regulation of ciliary genes

Results presented thus far reveal that OFD2 patient cells exhibit disrupted ciliogenesis and that these same mutations also disrupt the INTS13 module's interaction with the ICM. As such, we next determined whether reduced expression of INTS13 in cultured human cells was sufficient to inhibit ciliogenesis. We employed human retinal pigment epithelial (RPE) cells as a model, because these cells harbor clearly visible primary cilia upon serum starvation[41]. We treated RPE cells with either control siRNA or one of two distinct siRNA targeting *INTS13* and discovered that either siRNA could effectively deplete endogenous INTS13 (Fig. 5a). To visualize primary cilia, we probed fixed cells using antibodies raised to ADP-ribosylation factor-like protein 13B (ARL13B)[59], which stains the ciliary axoneme, and γ-tubulin, which stains the basal body[39]. We observed that the basal body and axoneme components were still present and appropriately juxtaposed
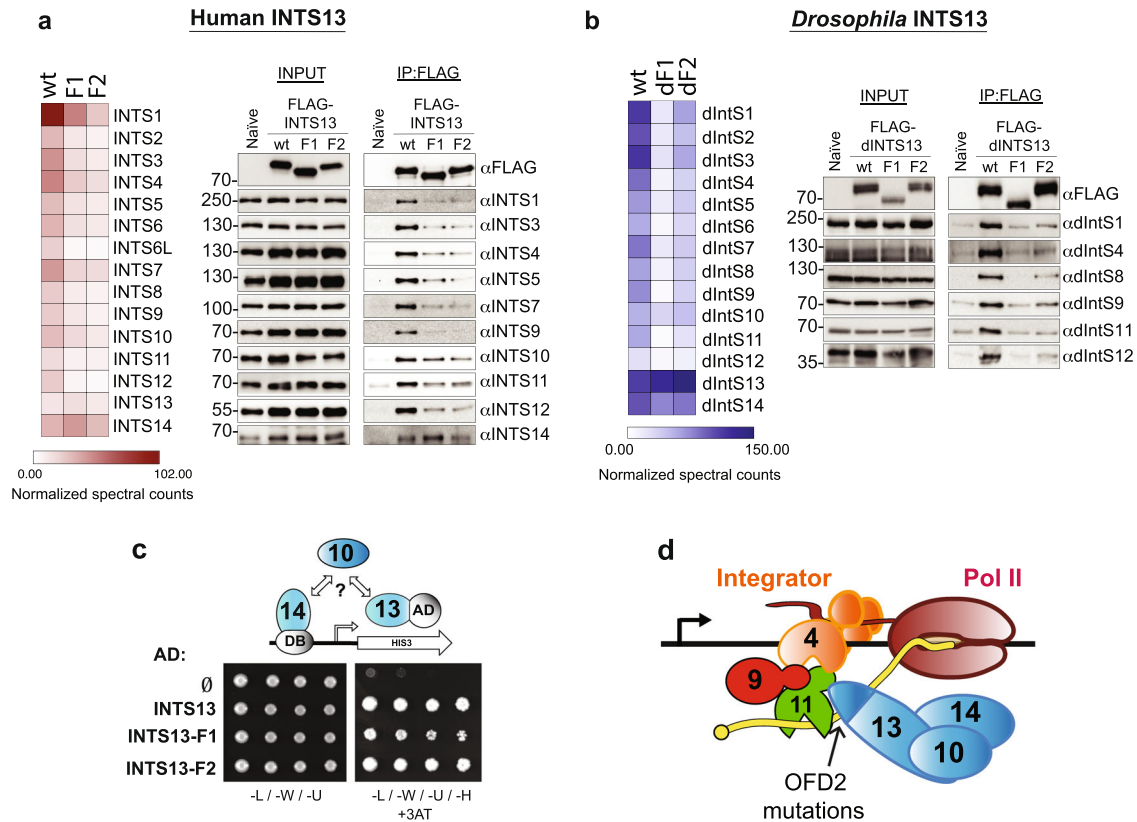
Fig. 4 | INTS13 patient variants disrupt interactions between the INTS10/13/14 module and the Integrator complex. a Immunoprecipitation of FLAG-INTS13 from nuclear extract derived from HEK-293T cells stably expressing FLAG epitope tagged human INTS13-wild type (wt), Family 1 INTS13 mutant (F1), or Family 2 INTS13 mutant (F2). Nuclear extract derived from naïve HEK-293T cells not expressing any FLAG epitope was used as a control. Left panel: IP-LC/MS was done for each sample, and quantitation is shown as a heat map. Results from the MS experiments were quantified from triplicate experiments, while the IP/WB was conducted in duplicate with representative images shown in the panel. The color scale shown denotes the normalized spectral counts for each protein (see methods). Right panel: western blot analysis of either nuclear extract input or anti-FLAG immunoprecipitation using human Integrator antibodies with molecular weight markers shown are in kilodaltons. **b** Results shown are from parallel experiments conducted in a manner identical with panel (**a**) with the exception that nuclear

extract was derived from *Drosophila* S2 cells expressing FLAG-dINTS13-wt or FLAG-tagged dINTS13 proteins where patient mutations were introduced into homologous regions of the *Drosophila* INTS13 cDNA. Results from the MS experiments were quantified from triplicate experiments, while the IP/WB was conducted in duplicate with representative images shown in the panel. **c** Modified yeast two hybrid analysis where human INTS14 is expressed as a Gal4-DNA binding domain fusion, human INTS13 wt or mutant is expressed as a Gal4-Activation domain fusion, and human INTS10 is expressed in *trans*. Dilutions of yeast are plated on both permissive (left) and restrictive (right) media. Empty activation domain vector is used as a negative control. **d** Schematic of the Integrator CM associating with the remainder of Integrator subunits that is bridged to the INTS10/13/14 module through the C-terminus of INTS13, which is disrupted by OFD2 mutations. Source data are provided as a Source Data file.

in INTS13-depleted cells but were clearly not utilized to form a normal primary cilium (Fig. 5b and Supplementary Figure 4a). We found that depletion of INTS13 with either siRNAs reduced the number of properly assembled primary cilia by ~50% relative to control siRNA (Fig. 5c).

To understand the underlying changes in gene expression that occurred after *INTS13* depletion, we subjected control siRNA-treated and each of the *INTS13* siRNA-treated RPE cells to RNA-seq analysis. We conducted RNA-seq using three biological replicates from cells treated with each siRNA and used DESeq to determine differential expression (Supplementary Data 2). Taking advantage of two distinct *INTS13* siRNAs, we only considered genes observed to be significantly increased/decreased (> 2-fold, $p < 0.0001$) in both knockdown conditions (Supplementary Fig. 4b, c). Using these filters, we identified 254 genes that were increased and 156 genes that were decreased upon *INTS13* knockdown (Fig. 5d). Gene ontology analysis revealed a striking enrichment in genes involved in cilium biogenesis uniquely within the increased gene set, suggesting that knockdown of *INTS13* leads to de-repression of a collection of genes involved in cilium biogenesis (Fig. 5e, Supplementary Data 2). We validated several of these gene changes using RT-qPCR including well-characterized ciliary genes such

as *BBOF1*, *CFAP206*, and *CFAP53* (Fig. 5f). Moreover, we observed that protein levels of CFAP206, CFAP53, and to a lesser extent BBOF1 also increase in cells depleted of INTS13 (Supplementary Fig. 4d).

Finally, to address the possibility of an 'Integrator-independent' function for INTS13 in ciliogenesis, we analyzed localization of endogenous INTS13 and associated proteins in RPE cells. Using antibodies raised to INTS13, we conducted indirect immunofluorescence in RPE cells under serum starvation. Similar to what was reported previously in *Drosophila* S2 cells, we found INTS13 to be predominantly in the nucleus without any ciliary localization (Supplementary Fig. 5a, b). Moreover, when we conjugated wild-type INTS13 or either of the two mutant INTS13 to eGFP, the transiently expressed proteins were always nuclear localized (Supplementary Fig. 5c–e). Finally, we purified both nuclear and total cell extracts from serum-starved RPE1 cells, subjected them to immunoprecipitation using anti-INTS13 antibodies, and then analyzed the associated proteins using mass spectrometry. Using this approach, we observed co-immunoprecipitation of nearly all Integrator subunits from the nuclear extracts and total cell extracts, but did not detect any other associating factors with known roles in ciliogenesis (Supplementary Data 3). Altogether, these results indicate that depletion of INTS13 in human cells leads to disrupted ciliogenesis
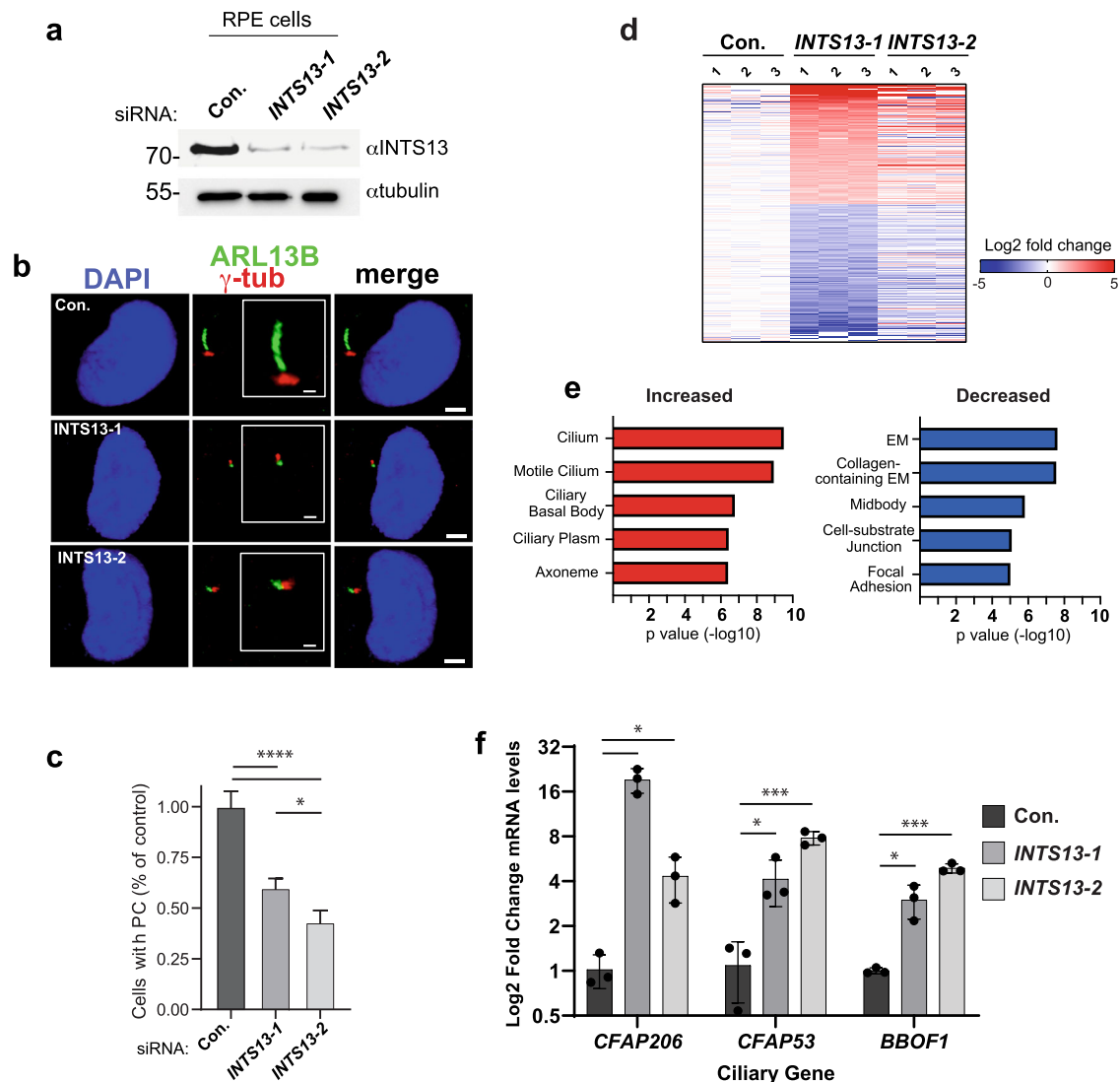
Fig. 5 | INTS13 depletion leads to a loss of primary cilia and broad transcriptional perturbation of ciliary genes. a Western blot of whole cell lysates from RPE cells treated with non-targeting control siRNA (Con.) or either of two siRNA targeting *INTS13* (*INTS13-1*, *INTS13-2*). Lysates were probed for either INTS13 or tubulin as a loading control and molecular weight markers shown are in kilodaltons. RNAi-depletion and Western blotting was repeated in two independent experiments with a representative image shown. **b** Immunofluorescence imaging of RPE cells treated with siRNAs as in (**a**) to visualize primary cilia (PC). DAPI staining in blue indicates the nuclei, γ-tubulin staining is shown in red to visualize the basal body, and ARL13B staining is shown in green to visualize the ciliary axoneme. Representative images show disrupted primary cilia that occurred when INTS13 was targeted by RNAi. The scale bar for the merged image is set at 3 μm while the scale bar in the figure inset present in the center is 1 μm. **c** Images from (**b**) were quantified for each condition and graphed as a percentage of the number of control-treated cells with PC (data are mean +/- SD, con. $n = 3785$ cells, 13-1 $n = 2969$ cells, 13-2 $n = 2675$ cells. Raw calculations for each condition were divided by the control mean to adjust control to 1). Statistical significance was calculated using a one-way ANOVA with Tukey's multiple comparisons test. Control vs INTS13-1 and INTS13-2 $p < 0.0001$, and INTS13-1 vs INTS13-2 $p = 0.0024$. **d** RNA sequencing was done on siRNA treated RPE cells in triplicate. Genes with significant changes in expression are shown by color indicating the log2 fold change. **e** Results from (**d**) were used for gene ontology analysis: the top five significant results are shown for upregulated and downregulated genes. All five GO terms for upregulated genes are related to cilia. **f** Quantitative RT-PCR validating three ciliary genes with expression changes identified using RNA-seq. Results are quantified from independent biological replicates ($n = 3$, data are mean +/- SD). Statistical significance was calculated using multiple unpaired t-tests with a Holm-Sidak correction. For 13-1 compared to control, CFAP206 $p = 0.00287$, CFAP53 $p = 0.02511$, and BBOF1 $p = 0.0221$. For 13-2 compared to control, CFAP206 $p = 0.01878$, CFAP53 $p = 0.00049$, and BBOF1 $p = 0.00013$ *$p < 0.05$ **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$ respectively. Source data are provided as a Source Data file.

that is likely driven by broad transcriptional deregulation of genes involved in primary cilium biogenesis.

Human wild-type INTS13, but not a truncated mutant form, rescues *Drosophila asun* spermatogenesis phenotypes

Previous studies showed that *IntS13* (also known as *asunder* or *asun*), the *Drosophila* homolog of human *INTS13*, plays a critical role in regulating dynein localization and nucleus-centrosome coupling during spermatogenesis. *Drosophila* males homozygous for a null allele of

asun (*piggyBac* transposon insertion within intron 2 of the gene; *asun^f02815^* allele) display characteristic phenotypes in their germ cells, including severe reduction of perinuclear dynein and loss of nucleus-centrosome coupling[39,49]. Given the conserved nature of OFD-like mutants in *Drosophila* (Fig. 4b), we examined whether expression of human wild-type, full-length INTS13 or a truncated form, INTS13-F1 (encoded by *INTS13* c.2004delA), could rescue *Drosophila asun* mutant phenotypes (Fig. 6). We generated transgenic *Drosophila* lines with expression of either wild-type or truncated human INTS13 cDNA under
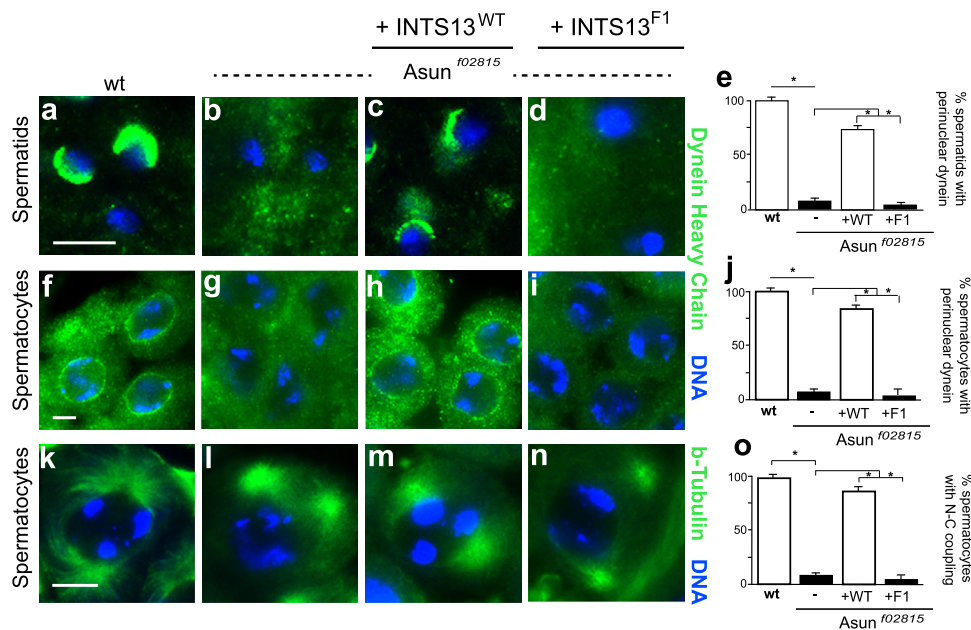
Fig. 6 | Human wild-type INTS13, but not mutant INTS13-F1, rescues spermatogenesis defects of *Drosophila* mutant ints13 males. a–j Male germline expression of wild-type, full-length human INTS13 (INTS13-wt), but not a truncated form (INTS13-F1), restored perinuclear dynein in *Drosophila asun* primary spermatids (**a–d**; quantified in **e**) and spermatocytes (**f–i**; quantified in **j**). Representative G2 spermatocytes and immature spermatids stained for dynein heavy chain (green) and DNA (blue) are shown. Bar graphs depict percentages of spermatids and spermatocytes showing properly localized dynein heavy chain. **k–o** Male germline expression of wild-type, full-length human INTS13, but not INTS13-F1, restored nucleus-centrosome coupling in *asun* primary spermatocytes (**k–n**; quantified in **o**). Representative prophase I spermatocytes stained for β-tubulin (green) and DNA (blue) are shown. Bar graphs depict percentages of spermatocytes showing normal nucleus-centrosome coupling. Data are mean ± s.e.m. *$P < 0.0001$, one-tailed Student's *t*-test. **e, j** $n > 10$ flies were used with $n > 584$ spermatids or spermatocytes scored. **o** $n > 10$ flies were used with $n > 120$ spermatocytes scored. INTS13-F1 is encoded by *INTS13* c.2004delA. Scale bars, 10 μm. Source data are provided as a Source Data file.

control of the male germline-specific *βtub85D* promoter and then introduced these transgenes into *asun* mutant male flies. Germline cells from transgenic *asun* flies expressing either wild-type or truncated human INTS13 were then assessed for perinuclear dynein and nucleus-centrosome coupling. We found that the wild-type, full-length human INTS13 almost fully rescued the spermatogenesis defects of *asun* male flies, whereas the truncated mutant form of human INTS13 failed to rescue these dynein localization and nucleus-centrosome coupling phenotypes (Fig. 6). These findings are highly supportive that INTS13 mutations observed in OFD-like patients that disrupt association with the Integrator cleavage module are likely disease-causative variants.

Knockdown of *ints13* causes developmental defects and ciliary phenotypes in *Xenopus*

To further examine the role of INTS13 in vivo and in an organism with readily observable ciliary phenotypes, we first examined *ints13* expression in *Xenopus laevis* embryos using whole mount in situ hybridization (WISH). *ints13* showed both maternal and zygotic contributions with relatively ubiquitous expression. At late gastrula (stage 12), the expression of *ints13* was enriched in ectodermal lineages, with no apparent expression in the endoderm. At neurula (stage 17), its expression was enriched in the neural tube and neural folds (Supplementary Fig. 6a). From tadpole stages onwards, *ints13* transcripts were abundantly observed in ciliated tissues and organs including the brain, branchial arches, pronephros, eye and otic vesicles. Remarkably *ints13* was found to follow a punctate pattern of expression in the skin, which coincided with that of multiciliated cells (MCCs). This was validated by whole mount antibody staining, using the α-M antibody highlighting that endogenous INTS13 was enriched in epidermal cells endowed with multicilia (Supplementary Fig. 6b).

In order to gain insight into the precise function of ints13 during *Xenopus* embryonic development, we performed loss-of-function

analysis using a morpholino oligonucleotide (MO) targeted against the *ints13* start codon (IntS13-MO). This IntS13-MO was previously established to be specific[47], and we confirmed the translational block efficiency achieved by this MO by demonstrating that *ints13* morphants were nearly devoid of endogenous ints13 protein (Fig. 7a). The Ints13-MO resulted in developmental delay, a small head, shortened body axis, bent tail, cardiac oedema, polyploidy and early developmental arrest in *Xenopus* morphants (Fig. 7b). In addition to the previously described gastrulation defects, we also observed prominent deficiencies in ciliogenesis and craniofacial development. Based on the enrichment of ints13 in MCCs, from which protrude hundreds of beating cilia[60,61], we examined the possibility that *ints13* may affect multi-cilia formation and function. In situ hybridization using a specific MCC marker *ccdc19*[62] showed that morphants had significantly less *ccdc19* than did control embryos (Fig. 7c), which was validated by whole mount antibody staining using the anti-acetylated α-tubulin antibody marking cilia protruding from MCCs (Fig. 7d). We noted increased distances between puncta and irregular staining in *ints13*-depleted embryos, suggesting that the epidermis of *ints13*-depleted tadpoles could have fewer differentiated MCCs and/or that these cells harbor fewer cilia. Scanning electron microscopy (SEM) performed on both *ints13* morphants and control embryos revealed that MCCs were defective, with a significantly reduced number of cilia per cell (Fig. 7e, f). While the length of cilia seemed to be normal, inconsistency was observed within and between MCCs, suggestive of incomplete penetrance (Fig. 7e). Cilia exhibited abnormalities in morphology, protrusion and direction (Fig. 7f). Of note, we were unable to detect overt abnormalities in the ultrastructure of the cilium by transmission electron microscopy (TEM) (Fig. 7g). These results prompted us to examine the function of the cilia, which normally generate an anterior-to-posterior directional flow over the tadpole's skin. Using high-resolution imaging, time-lapsed movies revealed that morphants were completely unable to move liquid around their bodies compared to
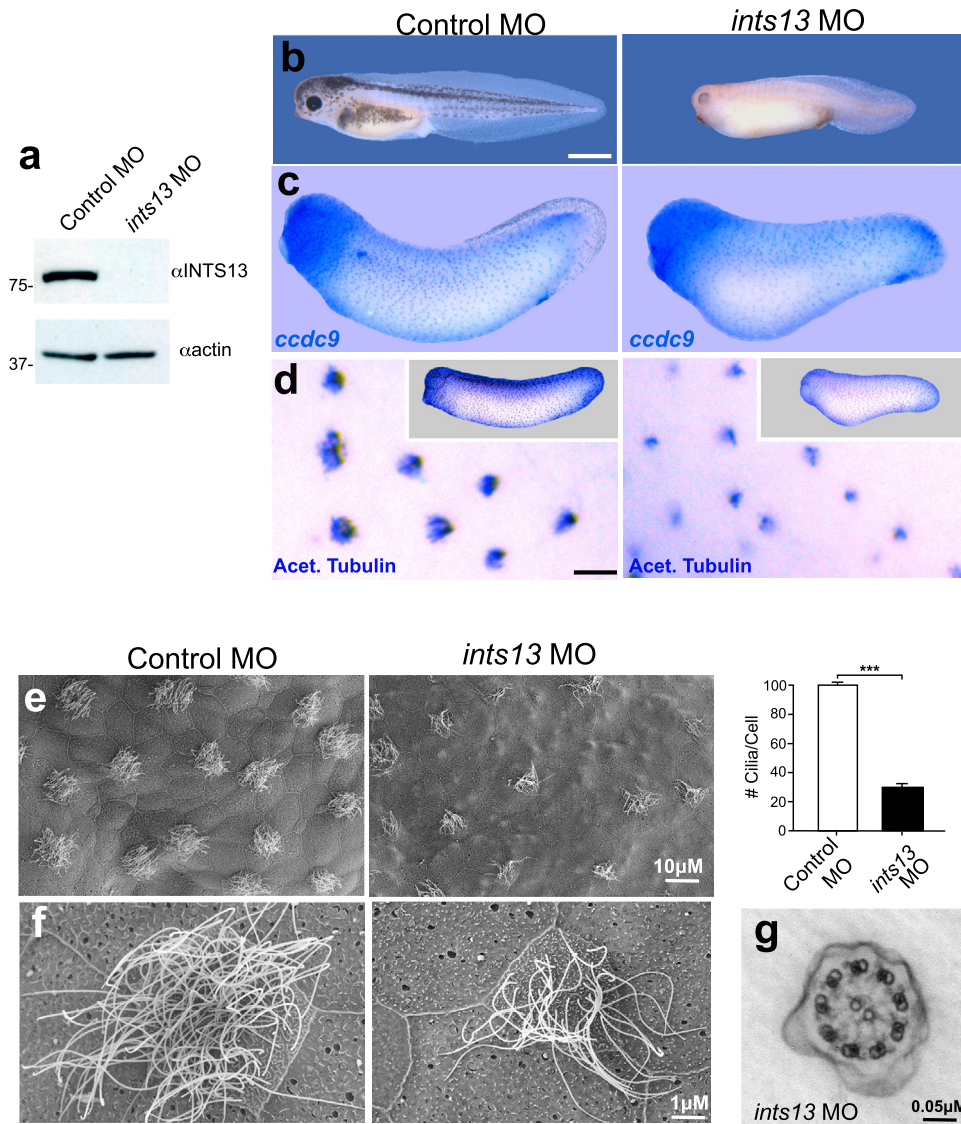
Fig. 7 | INTS13 is required for embryonic development and ciliogenesis in Xenopus laevis. a Western blot showing that the INTS13 MO was able to deplete INTS13 protein in *Xenopus* embryos and molecular weight markers shown are in kilodaltons. **b** In comparison to control embryos, morpholino (MO)-mediated knock down of INTS13 in Xenopus leads to severe developmental defects including a small head, big belly, short and down curved body axes, cardiac edema and massive developmental delay. Scale bar represents 500 μM. **c** In situ hybridization using a specific probe for the multi-ciliated cell (MCC) marker ccdc9 shows a reduction in ccdc9 expression in MCC of morphants compared to control embryos. **d** Whole mount antibody staining for acetylated α-tubulin shows a punctate pattern similar to ccdc9. The puncta have a more condensed and stronger signal in control embryos. A higher magnification view of the MCC shows marked reduction in cilia compared to controls. **b**–**d** $n > 60$ control or morphants embryos were used for three sets of independent injections. Scale bar represents 50 μM. **e** Low magnification SEM imaging of Xenopus MCC at embryonic stage 28. MCCs of a control embryo possess multiple, long cilia, whereas the morphant MCCs display fewer cilia per cell which appear to be disorganized. **f** High magnification SEM imaging of *Xenopus* MCC at embryonic stage 28. Ciliary length seems unperturbed in morphant embryos, while the average number of cilia on individual MCCs is clearly reduced by INTS13 knockdown. The average MCC in morphants have less than half of the number of cilia per cell compared to control, as shown in the bar graph. **g** The ultrastructure of the cilia in morphants appears to show a normal 9 + 2 structure as imaged by TEM. **e**, **f**, $n = 3$ control or morphant tadpoles were processed for SEM and $n > 30$ multiciliated cells were examined for the number of cilia present. **g** $n = 5$ cilia were examined by TEM. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, ****$p < 0.0001$, respectively. Source data are provided as a Source Data file.

control embryos (Supplementary Movie #1). Moreover, we found that ciliary beating in *ints13*-depleted embryos was slow, erratic and uncoordinated compared to control embryos (Supplementary Movie #2). These data indicate that INTS13 is necessary for MCC differentiation, as well as cilia formation and beating, thus revealing its critical role in ciliogenesis during development.

Discussion

Oral facial digital syndromes are an expanding group of genetic conditions characterized by malformations of the face, oral cavity, and digits. These disorders exhibit extensive variability of clinical features with intra-familial and inter-familial variability. Here we describe two homozygous germline variants in *INTS13* in two independent families, with Family 1 fulfilling clinical criteria for an OFD-like syndrome. Through molecular and biochemical approaches, we determined that INTS13 associates with INTS10/14 to form a heterotrimer that interacts with the Integrator Cleavage Module through the C-terminus of INTS13. The two disease-causing mutations both reduce steady-state levels of endogenous INTS13 protein in patient cells and disrupt INTS13 interaction with the ICM, likely causing transcriptional defects in a broad collection of ciliary genes. The binding of INTS10/13/14 to the ICM is clearly vital, as depletion of INTS13 was sufficient to disrupt

ciliogenesis in developing *Xenopus* embryos. These results underscore both the importance of Integrator assembly during development and reveal the importance of Integrator transcriptional regulation in ciliogenesis.

INTS13 and INTS14 were not initially described as components of Integrator[25], but were subsequently discovered through a genome-wide RNAi screen for factors regulating UsnRNA 3' end formation[50]. Since that initial identification, INTS13 has been shown to play a key role in the differentiation of monocytes[63]. A study in which the structure of INTS13 and INTS14 as a heterodimer was recently reported[56]. Using elegant biochemical approaches, Sabath and colleagues determined that INTS13/14 associates into a heterotrimer with INTS10 and the C-terminus of INTS13 binds to the Integrator Cleavage Module[33]. These structural and biochemical findings are remarkably concordant with results obtained here. Interestingly, it was also discovered that the INTS10/13/14 heterotrimer has an overall architecture related to the Ku70-Ku80 DNA repair complex[64] and is capable of binding DNA or RNA in vitro with the highest affinity for RNA hairpins. Thus, our work here and the structural data in that study establish that Integrator has at least two submodules: the previously defined INTS4/9/11 Cleavage Module containing the RNA endonucleolytic activity and an Auxiliary Module containing nucleic acid binding capability composed of INTS10/13/14. Interaction between these two modules is mediated in part through the C-terminus of INTS13. The structure of the INTS13 C-terminus was not visualized by crystallography, thus obscuring our ability to understand how the germline mutations identified herein disrupt the folding and/or binding of that region[56]. Nevertheless, the collective implication from these two studies is that the Integrator Auxiliary Module could bind to RNA hairpins to position the ICM for RNA cleavage and that this process is perturbed in congenital syndromes affecting the development of multiple organs.

Our results also provide evidence that *INTS13* is a novel ciliopathy gene, and phenotypes associated with *INTS13* mutation are likely due to a dysfunction in Integrator-mediated transcriptional regulation. This finding is consistent with previous studies showing that depletion of multiple Integrator subunits can also lead to disrupted ciliogenesis[41], and that mutations within INTS1 or INTS8 also create ciliopathy-related symptoms[42,65]. Herein, we discovered that depletion of INTS13 causes a ~2- to 10-fold upregulation in a broad collection of ciliary genes in RPE cells, which is consistent with recent work demonstrating that depletion of Integrator subunits in *Drosophila* cells primarily causes an increase in transcription in a subset of genes[34,38,66]. We posit that an increase in transcription of a broad collection of ciliary genes is sufficient to disrupt ciliogenesis by potentially over-producing specific components. The concept of dosage imbalance disrupting ciliogenesis has been well documented through reports where gene duplication of centrosome- or cilia-encoding genes can cause ciliopathies[67–69]. Even small increases in ciliary gene expression can prove disruptive to ciliogenesis, as evidenced by Down Syndrome patients who present ciliopathy symptoms and, due to its location on chromosome 21, overproduce *Pericentrin* a mere ~1.5-fold[70]. Overall, our results illustrate how sensitive the process of ciliogenesis is to the dosage of components and underscores the need for coordinate transcriptional regulation of genes critical to ciliogenesis. It will be interesting in future studies to determine the mechanism by which Integrator coordinately regulates ciliary genes and how disruption of the Auxiliary Module and Integrator Cleavage Modules gives rise to loss of this transcriptional control.

Methods
Clinical work
The index cases of the Jordanian family were initially diagnosed at the National Center for Diabetes, Endocrinology & Genetics in Jordan by Prof. Hanan Hamamy. Clinical assessments were undertaken at different clinics at the University hospital of Jordan. Saliva samples were collected from the whole family members (I.1, I.2, and II.1 to II.6) and skin biopsies were obtained from Family 1 (individuals I.2 and II.4). Family 2 was diagnosed at the Imagine Institute in France by Prof. Jeanne Amiel. Blood and skin biopsy were collected from age-matched health individuals and patient II.4 from Family 2. Genomic DNA was extracted (Puregene DNA Purification Kit, GentraSystems) following manufacturer's instructions. Nasal epithelial cells were also obtained (Family 1: I.2, II.4 and II.5) by nasal-brush biopsy (Cytobrush Plus, Medscand, Sweden). These investigations were conducted following ethical approval by the local IRBs in France and Singapore (A*STAR, 2019-087). Patients and their families have given informed consent for publication of photographs.

Genomic loci capture
Briefly, custom arrays (Agilent 244 K) were designed to target every exonic sequence of the genes present in the five homozygous IBD regions in Family 1 (chr6:167,901,317-170,287,795); (chr8:106,743,988-122,186,235); (chr12:12,129,350-28,091,505); (chr13:21,704,404-26,337,017); (chr16:58,296,974-65,631,053). A DNA library from patient (II:4) was prepared according to the Illumina library generation protocol version 2.3 and hybridized to custom arrays according to the Agilent CGH 244 K array protocol, and washed, eluted and amplified. The sample was submitted to one channel of Illumina flow cell and sequenced by Illumina Genome Analyzer (GAII) using standard manufacturer's protocol. The image data was processed by the provided GA Pipeline (Firecrest version 1.3.4 and Bustard version 1.3.4) and all sequences were aligned to the human reference genome (UCSC, build 18) by Blat-like Fast Accurate Search Tool (BFAST)[71]. Mismatches were further filtered to identify variants seen ten or more times, in which the variant was called as homozygous and did not overlap with a known dbSNP129 entry mismatch. Nonsynonymous mutations were identified with additional SeqWare tools and the "knownGene" gene model from the UCSC hg18. The open-source SeqWare project that provides a LIMS tool for tracking samples (SeqWare LIMS) and a pipeline for sequence analysis (SeqWare Pipeline) was used throughout this work.

For exome sequencing in Family 2, sequences were aligned to the human genome reference sequence (UCSC Genome Browser, GRCh38 build) by Burrows-Wheeler Aligner. Downstream processing was carried out with the Genome Analysis Toolkit (GATK), SAMtools and Picard Tools. Substitution calls were made with GATK Unified Genotyper, whereas indel calls were made with a GATK IndelGenotyperV2. All calls with a read coverage ≤ 2X and a Phred-scaled single-nucleotide polymorphism (SNP) quality of ≤ 20 were filtered out. Variants were annotated using an in-house-developed annotation software (Poly-Web) allowing filtering of variants according to relevant genetic models.

Mutation analysis
Positional candidate genes were obtained from GenBank and Ensembl databases. Segregation analysis was done for candidate mutations by Sanger sequencing with the BigDye Terminator cycle sequencing kit (Applied Biosystems) using primers flanking each mutation. The primer sequences for *INTS13* (NM_018164) and *SACS* (NM_014363) mutations screening are given in Supplementary Data 4.

Cell culture and transfections
Mouse fibroblast cells and hTERT transformed RPE1 cells were cultured in DMEM and Ham's F12/DMEM (Invitrogen), respectively. Lymphocytes cells were grown in cell culture medium (RPMI) (Invitrogen). All cultured media were supplemented with 10% Fetal Bovine Serum (FBS) (HyClone), 1X Penicillin / Streptomycin and 2 mM L-glutamine under standard conditions (37 °C, 5% CO_2). To induce ciliogenesis, cells were grown in serum starvation medium (0.5% FBS) for 72 h. Transient transfection was done using Lipofectamine 2000 (Life Technologies) according to manufacturer's recommended protocol.

HEK 293 T (ATCC, Manassas VA, CRL-3216) cells were grown at 37 °C with 5% CO_2 in DMEM (Gibco, Waltham MA, #11965-092), supplemented with 10% (v/v) FBS and 1% (v/v) penicillin-streptomycin (Gibco, #15070-063). RPE cells (ATCC, ARPE-19, CRL-2302, lot #70004873) were grown at 37 °C with 5% CO_2 in DMEM/F-12 50/50 (Corning, Manassas VA, 10-092-CV), supplemented with 10% (v/v) FBS and 1% (v/v) penicillin-streptomycin. *Drosophila* S2 cells were grown at 27 °C in Sf-900 II SFM (Gibco, #10902-088), supplemented with 1% (v/v) Antibiotic-Antimycotic (Gibco, 15240-062).

293 T cells were transfected using CRISPR/Cas9 to make inducible stable lines. 300,000 cells per well were seeded in a 6-well plate. The following day, 2 μl of Lipofectamine 2000 (Invitrogen, Carlsbad CA, 11668-019) was combined with 100 μl of Opti-MEM I (Gibco, 31985-070) per well and incubated at room temperature for 7 min. 100 μl of this solution was added to 500 ng of the Cas9 AAVS1 sgRNA plasmid and 500 ng of a pTet-3xFLAG-INTS13 plasmid per well and incubated at room temperature for 20 min. The transfection solution was then added to wells containing 1 ml of media. Each insert gene was done in duplicate wells. The next day, duplicate wells were combined, and cells were expanded to a 10 cm dish. 48 h after transfection, stable cell selection began with 1.5 μg/ml puromycin. Once stable lines were created, cells were maintained at 1 μg/ml puromycin. Human *INTS13* cDNA was mutated by site-directed PCR to create the genes for Family 1 and Family 2. All three versions of the *INTS13* gene, including wild type, were cloned individually into a plasmid containing the tetracycline-inducible Tet-On 3 G promotor with a 3x-FLAG tag for insertion into the AAVS1 locus. This was accomplished by co-transfection with the Cas9 plasmid containing the AAVS1 guide RNA (ACCCCACAGTGGGGCCACTA). Protein expression for nuclear extract was induced for 48 h with 0.4 μg/ml doxycycline for INTS13 wild type and Family 1 and 12 μg/ml for Family 2 in 30 15 cm diameter dishes per line.

Stable lines were also made with S2 cells. *Drosophila* INTS13 cDNA was mutated using site-directed PCR to make orthologous versions of the Family 1 and Family 2 mutations indicated in Supplementary Fig 3. The Family I cDNA also had an early stop codon added to recapitulate the truncation seen in the human gene. These modified genes, and wild type *INTS13*, were cloned into a pMT-3xFLAG-puro plasmid[34] following the metallothionein promotor and 3x-FLAG tag. 2×10^6 cells were plated in Schneider's media supplemented with 10% FBS in a 6-well plate overnight. 2 μg of plasmid was transfected using Fugene HD (Promega, Madison WI, #E2311). Plasmid DNA was mixed with 8 μL Fugene and 100 μL media and incubated at room temperature for 15 min before being added to cells. After 24 h, 2.5 μg/mL puromycin was added to the media to select and maintain the cell population. Cells were transitioned to SFX media without serum for large scale growth. Protein expression for nuclear extract was induced by adding 500 mM copper sulfate for 48 h to 1l of each cell line grown to approximately 1×10^7 cells/mL.

Immunofluorescence staining

For patient cells: cells were seeded on 12 mm cover slips. Cells were grown to around 80% confluency and then incubated with 0.5% serum media for 72 h to induce cilia formation, washed and kept at 4 °C until use. Nasal epithelial cells were suspended in RPMI without supplements to avoid artefacts caused by serum. The cytobrush was moved gently up and down for 5 min to detach ciliary cells from the brush. Samples were then spread onto non-coated glass slides, air dried, and stored at −80 °C until use. (See our protocol for immunofluorescence staining, URL).

For RNAi-treated RPE cells (Fig. 5b): After RNAi treatment (described in the RNA interference section) and serum starvation of cells grown on cover slips, cells were washed with DPBS and each cover slip was placed in an individual well of a 24-well plate. Cells were fixed to the cover slips by incubating in ice cold methanol for 7–8 min and washed thoroughly with PBS. Then samples were blocked for 30 min

(all incubation steps done at 37 °C with humidity). The blocking buffer used was 1% bovine serum albumin (Sigma-Aldrich, St. Louis MO, A7906-50G) (w/v) in PBS + 0.1% Tween 20 (Fisher Scientific, Fair Lawn NJ, BP337-500). Probing was performed for 2 h with the following primary antibodies at 1:1000 in blocking buffer: rabbit anti-gamma tubulin (Proteintech, Rosemont IL, 15176-1-AP) and mouse anti-ARL13B (Proteintech, 66739-1-Ig), followed by three washes with PBS for 5 min each. Secondary antibody probing was for 1 h in the dark with the following antibodies at 1:1000 in blocking buffer: goat anti-rabbit Alexa Fluor 594 (Life Technologies, Carlsbad CA, A11037) and goat anti-mouse Alexa Fluor 488 (Life Technologies, A11029), followed by three more washes with PBS. Cover slips were mounted on slides using Fluoroshield Mounting Medium with DAPI (Abcam, Cambridge, UK, ab104139). Confocal imaging was performed on a Leica TCS-SP8 3X system (Leica Microsystems, Wetzlar Germany) equipped with a Leica HC PL APO CS2 63×/1.4 oil immersion objective. A tunable (470–670 nm) pulsed white light laser (Leica WLL) and a UV 405 nm diode laser was used for excitation. Images were processed with the image analysis software, Imaris Pro 9.3.0.

Reverse transcription (RT-PCR) and Quantitative PCR

Figure 2: Total RNA was isolated from cells or embryos using Trizol (Ambion by Life Technologies, Carlsbad CA, 15596018) followed by extraction using the RNeasy Mini Kit (Qiagen, Hilden Germany). RNA (1 μg) was reverse transcribed using Iscript™ cDNA Synthesis Kit (Biotranscript lead) and the mis-splicing of the transcript was assessed using pair of primers designed in the exons upstream and downstream of the morpholino binding site. Quantitative real-time PCR was performed with power SYBR green master mix (Applied Biosystems) using a 7300 real-time polymerase chain reaction machine (Applied Biosystems) and normalized to the housekeeping gene β-actin. Primer sequences for RT-PCR and quantitative real-time PCR are listed in the Supplementary Data 4.

Figure 5: Total RNA was isolated using Trizol for phenol-chloroform extraction. cDNA was reverse transcribed using Superscript III Reverse Transcriptase (Thermo Scientific, Rockford IL, #18080085) according to the manufacturer's instructions. Random hexamers were used for cDNA synthesis primers and qRT-PCR was then carried out in triplicate using Bio-Rad iTaq Universal SYBR Green Supermix (Bio-Rad, Hercules CA, #1725120) and measured with the CFX Connect Real-Time System (Bio-Rad). Data was analyzed using the ΔΔCt method to calculate fold change, with *7SK* as the control gene.

Drosophila rescue

The *Drosophila IntS13* mutant line used herein ($asun^{f02815}$) was previously described[49]. The human *INTS13* coding sequence was subcloned into pCS2+ and a mutated construct (pCS2-INTS13$^{c.2004delA}$) was generated by introducing a frameshift mutation into wild-type cpenreading frame using site-specific mutagenesis (Stratagene). Sequence-verified wild-type and mutant constructs were subcloned into testis expression vector tv3 tagged with mCherry (CHY) at the N-terminal end[39]. These transgenes were introduced into the $asun^{f02815}$ background using standard genetic crosses. Immunostaining of *Drosophila* testes was performed as previously described[39].

Morpholino, RNA injection and embryological methods

The *ints13* morpholino (MO) purchased from Gene Tools were resuspended in sterile water to a concentration of 1 mM according to the manufacturer's instructions. RNA was then extracted and converted to cDNA. To target non-neural ectoderm, 1:4 of diluted *ints13*-MO was injected at the two- to four- cell stage. Morpholino sequences are given in Supplementary Data 4. Protocols for *Xenopus* fertilization, MO-injections, antibody staining, whole-mount in situ hybridization and immunofluorescence staining are at our protocol website (https://sites.google.com/a/reversade.com/www/protocols/). All experiments

with *Xenopus* embryos were approved by the Singapore National Advisory on Laboratory Animal Research. A stereomicroscope equipped with an ICD digital camera (Leica M205 FA) was used to capture images of embryos in successive focal planes. Images were then combined into one picture with Photoshop CS6 (Adobe).

Western blot

For Figs. 2 and 7: Cells or embryos were lysed in RIPA buffer supplemented with a cocktail of protease inhibitor (Roche). Total protein was resolved on SDS polyacrylamide gels (Bio-Rad) with DTT, followed by transfer on polyvinylidene difluoride (PVDF) membranes. The raw data for Fig. 2c, e, f is unavailable. The included blots are historical data, which could not be located. The original samples were historically collected and are not suitable to repeat the experiments. Orthogonal data is provided in Fig. 2a, b. The raw data for Fig. 7a is also unavailable as the included blot is historical data, which could not be located. Due the covid19 pandemic, the Xenopus facility is closed and the experiment could not be repeated. The following antibodies were used for HRP-mediated chemiluminescence at 1:1000 dilutions unless noted otherwise: polyclonal Rabbit antibody to INTS13 synthetic peptides (C & M), monoclonal Mouse antibody to actin (Chemicon; Mab1501R), β-tubulin (E7, Developmental Studies Hybridoma Bank, University of Iowa, Iowa City, IA), dynein heavy chain (P1H4, gift from T. Hays, University of Minnesota, Minneapolis, MN) anti-HA antibody (Sigma, H 6533 (clone HA-7), hIntS13 (Proteintech, #19892-1-AP), hArl13b (Proteintech, #66739-1-Ig), γ-tubulin (Proteintech, #15176-1-AP), hINTS1 (Bethyl, #A300-361A), hINTS3 (Proteintech, #16620-1-AP), hINTS4 (Bethyl, #A301-296A), hINTS5 (Proteintech, #14069-1-AP), hINTS7 (Bethyl, #A300-271A), hINTS9 (Atlas, #HPA066822), hINTS10 (Proteintech, #15271-1-AP), hINTS11 (Bethyl, #A301-274A), hINTS12 (Proteintech, #16455-1-AP), hINTS14/C15orf44 (Bethyl, #A303-577A), hCFAP53 (RayBiotech, #144-64537-50), hC6orf165 (CFAP206) (Biorbyt, #orb156105), hCCDCl76 (BBOFl) (lnvitrogen, #PA5-70942) anti-rabbit donkey ECL Horseradish Peroxidase-linked secondary (GE Healthcare, Sigma Aldrich, #NA934V), Alexa Fluor 594 goat anti-rabbit secondary (Life Technologies, #A11037), Alexa Fluor 488 goat anti-mouse secondary (Life Technologies, #A11029), anti-FLAG M2-Peroxidase (HRP) mouse monoclonal (Sigma, #A8592). Antibodies to *Drosophila* proteins used in the Fig. 3b western blot were made in the Wagner laboratory from recombinant proteins expressed in E. coli and all were raised in guinea pig. Antibodies used in Fig. 3b and citations of initial use: dIntS1[35], dIntS4[66], dIntS8[66], dIntS9[35], dIntS11[35], dIntS12[57].

For Figs. 4 and 5: Protein was extracted directly from RPE cells treated with siRNA by adding 2X SDS loading buffer (120 mM Tris pH6.8, 4% SDS, 200 mM DTT, 20% Glycerol, and 0.02% Bromophenol blue) to cells while on the plate. Samples were then incubated at room temperature while on the plate with periodic swirling prior to a 10 min boiling at 95 °C and a short sonication. Denatured protein samples were then resolved on a 10% polyacrylamide gel and transferred to a PVDF membrane (Thermo Scientific, 88518). Blots were probed as previously described[34]. Samples from nuclear extract immunoprecipitation are described in the anti-FLAG affinity purification.

Electron microscopy (EM)

Embryos at the desired stages were fixed in 3% PBS-glutaraldehyde, pH 7.4, for 4 h with agitation. After 3 washes in fresh phosphate buffer (NaH2PO4,2H2O / Na2HPO4,12H2O) for 5 min each with vigorous agitation, embryos were rinsed 3 times/5 min in distilled water, dehydrated in ethanol with vigorous agitation, dried in Leica EM CPD030 and sputter coated in gold coating, 7 nm thick layer using Leica SCD050. Embryos were viewed using a scanning electron microscopy (FESEM, JEOL, JSM-6701F). For Transmission electron microscopy (TEM) dehydrated embryos were embedded in Epon 812, prepared sections (50–60 nm) and contrasted with 4% uranyl acetate in 50% ethanol and 2.6% lead nitrate in 1 M NaOH. Embryos were scanned using a Hitachi 7000 electron microscope (Hitachi, Tokyo, Japan). EM scanning was performed at electron microscopy facility (Institute of Molecular and Cell Biology (IMCB), A*STAR, Singapore).

Plasmid construction and Yeast two-hybrid analysis

Yeast two-hybrid analysis was carried out in PJ69-4α and PJ49-4a strains as described[33]. In brief, human INTS1 through INTS14, and INTS13 variations, were cloned into pGAD, pOBD, pTEF, or pCEV-G1Ura3/TEFdual vectors using conventional cloning. pOBD plasmids were then transformed into PJ69-4α yeast and were selected on tryptophan-dropout medium (Clontech, Mountain View CA, #630413); pGAD plasmids were transformed into PJ49-4a yeast and were selected on leucine-dropout medium (Clontech, #630414). pTEF or pTEFdual plasmids were included to express proteins in trans and were transformed with pOBD plasmids and selected on media lacking tryptophan and uracil (Sunrise Science Products, San Diego CA, 1316-030). Yeast strains were then mated and subsequently selected on medium lacking tryptophan, leucine, and uracil (Sunrise Science Products, 1328-030). Interactions were tested through serial dilution (one to five) of diploid yeast followed by plating on medium lacking tryptophan, leucine, and uracil or on medium lacking tryptophan, leucine, uracil, and histidine (Sunrise Science Products, 1334-030) that also was supplemented with 1 mM 3-amino-1,2,4-triazole (MP Biomedicals, Irvine CA, 4061-722).

Expression and purification of INTS4-INTS9-INTS11-INTS13 complex

Full-length human INTS4, INTS9, INTS11, and INTS13 were co-expressed in insect cells using Multibac technology (Geneva Biotech)[72]. INTS4 carried an N-terminal hexa-histidine tag while the other three proteins were untagged. High Five cells were grown in ESF 921 medium (Expression Systems) by shaking at 120 rpm at 27 °C until the density reached 1.5-2×10^6 cells mL^{-1}. Cells were infected by P1 viruses and harvested after 48 h. The cells were lysed by sonication in a buffer containing 20 mM Tris (pH 8.0), 200 mM NaCl, 5% (v/v) glycerol and 0.1 mM PMSF. The supernatant of cell lysates was incubated with nickel beads (Qiagen) and the eluate was then subjected to gel filtration chromatography (Superdex 200, GE Healthcare), in a running buffer of 20 mM Tris (pH 8.0), 250 mM NaCl, and 5 mM DTT.

Nuclear extract

Cells were collected and washed in cold PBS. 293 T cells were detached from dishes while S2 cells were grown in suspension and pelleted by centrifugation. Cells were then resuspended in five times the cell pellet volume of Buffer A (10 mM Tris pH8, 1.5 mM MgCl2, 10 mM KCl, 0.5 mM DTT, and 0.2 mM PMSF). Resuspended cells were allowed to swell during a 15 min rotation at 4 °C. After pelleting down at 1000 g for 10 min, two volumes of the original cell pellet of Buffer A was added and cells were homogenized with a dounce pestle B for 20 strokes on ice. Nuclear and cytosolic fractions were then separated by centrifugation at 2000 g for 10 min. To attain a nuclear fraction, the pellet was washed once with Buffer A before resuspending in an equal amount of the original cell pellet volume of Buffer C (20 mM Tris pH8, 420 mM NaCl, 1.5 mM MgCl2, 25% glycerol, 0.2 mM EDTA, 0.5 mM PMSF, and 0.5 mM DTT). The sample was then homogenized with a dounce pestle B for 20 strokes on ice and rotated for 30 min at 4 °C before centrifuging at 15,000 g for 30 min at 4 °C. Finally, supernatants were collected and subjected to dialysis in Buffer D (20 mM HEPES, 100 mM KCl, 0.2 mM EDTA, 0.5 mM DTT, and 20% glycerol) overnight at 4 °C. Prior to any downstream applications, nuclear extracts were centrifuged again at 15,000 g for 3 min at 4 °C to remove any precipitate.

Anti-FLAG affinity purification

To purify FLAG-tagged Integrator complexes for mass spectrometry, generally between 8 and 10 mg of nuclear extract (approximately

1.9 mL of extract depending on the concentration) was mixed with 100 µL anti-Flag M2 affinity agarose slurry (Sigma-Aldrich, #A2220) washed with 0.1 M glycine then equilibrated in binding buffer (20 mM HEPES pH7.4, 150 mM KCl, 10% Glycerol, 0.1% NP-40). This mixture was rotated for four hours at 4 °C. Following the 4 h incubation/rotation, five sequential washes were carried out in binding buffer with a 10 min rotation at 4 °C followed by a 1000 g centrifugation at 4 °C. After a final wash with 20 mM HEPES buffer, the supernatant was removed using a pipette and the beads were kept cold and submitted to the mass spectrometry core where the protein complexes were eluted by digestion (described below). For immunoprecipitation samples intended for western blot, a similar protocol was used. 25 µL of bead slurry and 200 µL of extract sample were rotated for two hours at 4 °C. After the fifth wash with binding buffer, protein complexes were eluted from the anti-FLAG resin by adding 50 µL of 2X SDS loading buffer and boiled at 95 °C for 5 min. For western blots, input samples were generated by adding equal volume of 2X SDS loading buffer to nuclear extract and 1/10 of the immunoprecipitation was loaded as estimated by protein mass.

Mass spectrometry sample digestion

The samples were prepared in a similar manner as described previously[73]. Briefly, the agarose bead-bound proteins were washed several times with 50 mM Triethylammonium bicarbonate (TEAB) pH 7.1, before being solubilized with 40 µL of 5% SDS, 50 mM TEAB, pH 7.55 followed by a room temperature incubation for 30 min. The supernatant containing the proteins of interest was then transferred to a new tube, reduced by making the solution 10 mM Tris(2-carboxyethyl)phosphine (TCEP) (Thermo, #77720), and further incubated at 65 °C for 10 min. The sample was then cooled to room temperature and 1 µL of 1 M iodoacetamide acid was added and allowed to react for 20 min in the dark. Then, 5 µL of 12% phosphoric acid was added to the 50 µL protein solution followed by 350 µL of binding buffer (90% Methanol, 100 mM TEAB final; pH 7.1). The resulting solution was administered to an S-Trap spin column (Protifi, Farmingdale NY) and passed through the column using a bench top centrifuge (30 s spin at 4000 g). The spin column was then washed three times with 400 µL of binding buffer and centrifuged (4000 g, 1 min). Trypsin (Promega, #V5280) was then added to the protein mixture in a ratio of 1:25 in 50 mM TEAB, pH = 8, and incubated at 37 °C for 4 h. Peptides were eluted with 80 µL of 50 mM TEAB, followed by 80 µL of 0.2% formic acid, and finally 80 µL of 50% acetonitrile, 0.2% formic acid. The combined peptide solution was then dried in a speed vacuum (room temperature, 1.5 h) and resuspended in 2% acetonitrile, 0.1% formic acid, 97.9% water and aliquoted into an autosampler vial.

NanoLC MS/MS Analysis

Peptide mixtures were analyzed by nanoflow liquid chromatography-tandem mass spectrometry (nanoLC-MS/MS) using a nano-LC chromatography system (UltiMate 3000 RSLCnano, Dionex, Thermo Fisher Scientific, San Jose, CA). The nano-LC-MS/MS system was coupled on-line to a Thermo Orbitrap Fusion mass spectrometer (Thermo Fisher Scientific, San Jose, CA) through a nanospray ion source (Thermo Scientific). A trap and elute method was used to desalt and concentrate the sample, while preserving the analytical column. The trap column (Thermo Scientific) was a C18 PepMap100 (300 µm X 5 mm, 5 µm particle size) while the analytical column was an Acclaim PepMap 100 (75 µm X 25 cm) (Thermo Scientific). After equilibrating the column in 98% solvent A (0.1% formic acid in water) and 2% solvent B (0.1% formic acid in acetonitrile (ACN)), the samples (2 µL in solvent A) were injected onto the trap column and subsequently eluted (400 nL/min) by gradient elution onto the C18 column as follows: isocratic at 2% B, 0–5 min; 2% to 32% B, 5–39 min; 32% to 70% B, 39–49 min; 70% to 90% B, 49–50 min; isocratic at 90% B, 50–54 min; 90% to 2%, 54–55 min; and isocratic at 2% B, until the 65 min mark.

All LC-MS/MS data were acquired using XCalibur, version 2.1.0 (Thermo Fisher Scientific) in positive ion mode using a top speed data-dependent acquisition (DDA) method with a 3 s cycle time. The survey scans (m/z 350-1500) were acquired in the Orbitrap at 120,000 resolution (at m/z = 400) in profile mode, with a maximum injection time of 100 m s and an AGC target of 400,000 ions. The S-lens RF level was set to 60. Isolation was performed in the quadrupole with a 1.6 Da isolation window, and CID MS/MS acquisition was performed in profile mode using rapid scan rate with detection in the ion-trap using the following settings: parent threshold = 5000; collision energy = 32%; maximum injection time 56 ms; AGC target 500,000 ions. Mono-isotopic precursor selection (MIPS) and charge state filtering were on, with charge states 2–6 included. Dynamic exclusion was used to remove selected precursor ions, with a +/− 10 ppm mass tolerance, for 15 s after acquisition of one MS/MS spectrum.

Database Searching

Tandem mass spectra were extracted and charge state deconvoluted using Proteome Discoverer (Thermo Fisher, version 2.2.0388). Deisotoping was not performed. All MS/MS spectra were searched against the appropriate database, either Uniprot *Drosophila* database (version 04-04-2018) or Uniprot Human database (reviewed 06272018), using Sequest. Searches were performed with a parent ion tolerance of 5 ppm and a fragment ion tolerance of 0.60 Da. Trypsin was specified as the enzyme, allowing for two missed cleavages. Fixed modification of carbamidomethyl (C) and variable modifications of oxidation (M) and deamidation were specified in Sequest. Heat maps in Fig. 4 were made using Morpheus from the Broad Institute, https://software.broadinstitute.org/morpheus.

RNA interference

siRNA used were Universal Negative Control #2 (Sigma Aldrich, #SIC002), INTS13-1 (Sigma Aldrich, SASI_Hs01_00134940/ASUN), and INTS13-2 (Sigma Aldrich, SASI_Hs01_00134941/ASUN). RPE cells were seeded at 50,000 cells/well in a 24-well plate on Day 1. On Day 2, 3 µL of 20 µM siRNA was added to 50 µL of Opti-MEM I per well and incubated at room temperature for 5 min. 2 µL of Lipofectamine RNAiMax (Invitrogen, 13778-150) was also added to 50 µL of Opti-MEM per well and incubated at room temperature for 5 min. The siRNA and RNAi Max solutions were combined and incubated at room temperature for 20 min, then added to the cells. Cells were expanded to a 6-well plate on Day 3. On Day 4, cells were washed gently with DPBS and 900 µL of DMEM/F-12 media was added to each well. The siRNA treatment from Day 2 was repeated. Cells were further incubated and then harvested on Day 6. For RNA extraction, cells were harvested with 500 µL of Trizol. For protein extraction, cells were harvested with SDS loading buffer detailed previously.

A modified version of the above protocol was followed for cells intended for immunofluorescence staining of primary cilia: RPE cells were seeded onto three cover slips in each well of the 6-well plate on Day 3. On Day 4, cells were washed twice with DPBS and grown in serum free (0% FBS) DMEM/F-12 + 1% (v/v) penicillin-streptomycin for 48 h before fixation.

Library preparation and RNA sequencing

RNAi-treated RPE cells were harvested in 500 µL of Trizol, and total RNA was purified by phenol-chloroform extraction. 1 µg total RNA per sample was used for sequencing library preparation. The NEBNext Poly(A) mRNA magnetic isolation kit (New England Biolabs, Ipswich MA, E7490) was used to purify mRNA from the total RNA. The Click-Seq protocol was used following this step[74]. RT-PCR to make cDNA was done with 1:35 azido-NTP:dNTP for random strand termination by SuperScript III reverse transcriptase. Following Click-Seq preparations, libraries were PCR amplified for 17 cycles and products from 400–600 bp were collected for sequencing. Samples were sequenced

Table 1 | RNA-sequencing statistics

Sample	Total Reads	Mappable Fragments (Percent of total)	Agreement between replicates (Spearman's rho)	Normalization Factor
Control	131,005,450	94.33%	> 0.99	1.142296033
INTS13-1	149,803,846	94.27%	> 0.99	0.999334293
INTS13-2	168,253,066	94.13%	> 0.99	0.889755677

on a NextSeq500 (Illumina) using a single read 75 bp run. Sequencing adapters removed and reads with a Qscore < 30 were discarded using fastp (ver 0.14.1). Reads were aligned using Hisat2 (ver 2.1.0) to the hg38 genome using default parameters. Differential expression changes were determined using DESeq2 (ver 1.23.10) and assigned using the comprehensive gene annotation in GENCODE32 and tests for significance were using the nbinomialWald test. Alignment information and lowest correlation between samples in each data set is given in Table 1.

Statistics and Reproducibility

Figure 5a: siRNA treatment of RPE cells to knock down INTS13 was done multiple times for various experiments. This blot is a representative image of those results. We saw slight variations in the level of protein knock down between experiments but reduced expression was always seen.

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The RNA-seq data generated in this study have been deposited in the GEO database under accession GSE154197. Proteomics data submitted to the ProteomeXchane Consortium via the PRIDE under accession PXD020593 [http://proteomecentral.proteomexchange.org/cgi/GetDataset]. Source data are provided with this paper. Full blots for Figs. 4, 5, and Supplementary Figure 4 are provided in the source data. Full blots for Figs. 2 and 7 are not in the source data as they were stored on hard drives that were lost. Source data are provided with this paper.

References

1. Lee, J. E. & Gleeson, J. G. A systems-biology approach to understanding the ciliopathy disorders. *Genome Med.* **3**, 59 (2011).
2. Lee, J. E. & Gleeson, J. G. Cilia in the nervous system: linking cilia function and neurodevelopmental disorders. *Curr. Opin. Neurol.* **24**, 98–105 (2011).
3. Raidt, J. et al. Ciliary beat pattern and frequency in genetic variants of primary ciliary dyskinesia. *Eur. Respir. J.* **44**, 1579–1588 (2014).
4. Waters, A. M. & Beales, P. L. Ciliopathies: An expanding disease spectrum. *Pediatr. Nephrol.* **26**, 1039–1056 (2011).
5. Arts, H. H. & Knoers, N. V. Current insights into renal ciliopathies: What can genetics teach us? *Pediatr. Nephrol.* **28**, 863–874 (2013).
6. Estrada-Cuzcano, A., Roepman, R., Cremers, F. P., den Hollander, A. I. & Mans, D. A. Non-syndromic retinal ciliopathies: Translating gene discovery into therapy. *Hum. Mol. Genet* **21**, R111–124 (2012).
7. Forsythe, E. & Beales, P. L. Bardet-Biedl syndrome. *Eur. J. Hum. Genet.* **21**, 8–13 (2013).
8. Alexiev, B. A., Lin, X., Sun, C. C. & Brenner, D. S. Meckel-Gruber syndrome: Pathologic manifestations, minimal diagnostic criteria, and differential diagnosis. *Arch. Pathol. Lab Med* **130**, 1236–1238 (2006).
9. Badano, J. L., Mitsuma, N., Beales, P. L. & Katsanis, N. The ciliopathies: An emerging class of human genetic disorders. *Annu Rev. Genomics Hum. Genet.* **7**, 125–148 (2006).
10. Baker, K. & Beales, P. L. Making sense of cilia in disease: The human ciliopathies. *Am. J. Med Genet C. Semin Med Genet.* **151C**, 281–295 (2009).
11. Davis, E. E. & Katsanis, N. The ciliopathies: A transitional model into systems biology of human genetic disease. *Curr. Opin. Genet. Dev.* **22**, 290–303 (2012).
12. Ferkol, T. W. & Leigh, M. W. Ciliopathies: The central role of cilia in a spectrum of pediatric disorders. *J. Pediatr.* **160**, 366–371 (2012).
13. Leigh, M. W. et al. Clinical and genetic aspects of primary ciliary dyskinesia/Kartagener syndrome. *Genet. Med.* **11**, 473–487 (2009).
14. Gurrieri, F., Franco, B., Toriello, H. & Neri, G. Oral-facial-digital syndromes: review and diagnostic guidelines. *Am. J. Med Genet A* **143A**, 3314–3323 (2007).
15. Thauvin-Robinet, C. et al. The oral-facial-digital syndrome gene C2CD3 encodes a positive regulator of centriole elongation. *Nat. Genet* **46**, 905–911 (2014).
16. Adly, N., Alhashem, A., Ammari, A. & Alkuraya, F. S. Ciliary genes TBC1D32/C6orf170 and SCLT1 are mutated in patients with OFD type IX. *Hum. Mutat.* **35**, 36–40 (2014).
17. Ferrante, M. I. et al. Convergent extension movements and ciliary function are mediated by ofd1, a zebrafish orthologue of the human oral-facial-digital type 1 syndrome gene. *Hum. Mol. Genet* **18**, 289–303 (2009).
18. Franco, B. & Thauvin-Robinet, C. Update on oral-facial-digital syndromes (OFDS). *Cilia* **5**, 12 (2016).
19. Lopez, E. et al. C5orf42 is the major gene responsible for OFD syndrome type VI. *Hum. Genet.* **133**, 367–377 (2014).
20. Shamseldin, H. E. et al. Mutations in DDX59 implicate RNA helicase in the pathogenesis of orofaciodigital syndrome. *Am. J. Hum. Genet.* **93**, 555–560 (2013).
21. Thauvin-Robinet, C. et al. OFD1 mutations in males: phenotypic spectrum and ciliary basal body docking impairment. *Clin. Genet.* **84**, 86–90 (2013).
22. Thomas, S. et al. TCTN3 mutations cause Mohr-Majewski syndrome. *Am. J. Hum. Genet.* **91**, 372–378 (2012).
23. Ferrante, M. I. et al. Identification of the gene for oral-facial-digital type I syndrome. *Am. J. Hum. Genet.* **68**, 569–576 (2001).
24. Roberson, E. C. et al. TMEM231, mutated in orofaciodigital and Meckel syndromes, organizes the ciliary transition zone. *J. Cell Biol.* **209**, 129–142 (2015).
25. Baillat, D. et al. Integrator, a multiprotein mediator of small nuclear RNA processing, associates with the C-terminal repeat of RNA polymerase II. *Cell* **123**, 265–276 (2005).
26. Baillat, D. & Wagner, E. J. Integrator: surprisingly diverse functions in gene expression. *Trends Biochem. Sci.* **40**, 257–264 (2015).
27. Gardini, A. et al. Integrator regulates transcriptional initiation and pause release following activation. *Mol. Cell* **56**, 128–139 (2014).
28. Skaar, J. R. et al. The Integrator complex controls the termination of transcription at diverse classes of gene targets. *Cell Res.* **25**, 288–305 (2015).
29. Stadelmayer, B. et al. Integrator complex regulates NELF-mediated RNA polymerase II pause/release and processivity at coding genes. *Nat. Commun.* **5**, 5531 (2014).
30. Albrecht, T. R. & Wagner, E. J. snRNA 3' end formation requires heterodimeric association of integrator subunits. *Mol. Cell Biol.* **32**, 1112–1123 (2012).

31. Callebaut, I., Moshous, D., Mornon, J. P. & de Villartay, J. P. Metallo-beta-lactamase fold within nucleic acids processing enzymes: the beta-CASP family. *Nucleic Acids Res.* **30**, 3592–3601 (2002).

32. Mandel, C. R. et al. Polyadenylation factor CPSF-73 is the pre-mRNA 3'-end-processing endonuclease. *Nature* **444**, 953–956 (2006).

33. Albrecht, T. R. et al. Integrator subunit 4 is a 'Symplekin-like' scaffold that associates with INTS9/11 to form the Integrator cleavage module. *Nucleic Acids Res.* **46**, 4241–4255 (2018).

34. Elrod, N. D. et al. The integrator complex attenuates promoter-proximal transcription at protein-coding genes. *Mol. Cell* **76**, 738–752 e737 (2019).

35. Ezzeddine, N. et al. A subset of Drosophila integrator proteins is essential for efficient U7 snRNA and spliceosomal snRNA 3'-end formation. *Mol. Cell Biol.* **31**, 328–341 (2011).

36. Lai, F., Gardini, A., Zhang, A. & Shiekhattar, R. Integrator mediates the biogenesis of enhancer RNAs. *Nature* **525**, 399–403 (2015).

37. Rubtsova, M. P. et al. Integrator is a key component of human telomerase RNA biogenesis. *Sci. Rep.* **9**, 1701 (2019).

38. Tatomer, D. C. et al. The Integrator complex cleaves nascent mRNAs to attenuate transcription. *Genes Dev.* **33**, 1525–1538 (2019).

39. Jodoin, J. N. et al. Nuclear-localized Asunder regulates cytoplasmic dynein localization via its role in the integrator complex. *Mol. Biol. Cell* **24**, 2954–2965 (2013).

40. Sorokin, S. Centrioles and the formation of rudimentary cilia by fibroblasts and smooth muscle cells. *J. Cell Biol.* **15**, 363–377 (1962).

41. Jodoin, J. N. et al. The snRNA-processing complex, Integrator, is required for ciliogenesis and dynein recruitment to the nuclear envelope via distinct mechanisms. *Biol. Open* **2**, 1390–1396 (2013).

42. Breslow, D. K. et al. A CRISPR-based screen for Hedgehog signaling provides insights into ciliary function and ciliopathies. *Nat. Genet.* **50**, 460–471 (2018).

43. Sakai, N., Nakakita, N., Yamazaki, Y., Ui, K. & Uchinuma, E. Oral-facial-digital syndrome type II (Mohr syndrome): clinical and genetic manifestations. *J. Craniofac Surg.* **13**, 321–326 (2002).

44. Bonnard, C. et al. Mutations in IRX5 impair craniofacial development and germ cell migration via SDF1. *Nat. Genet* **44**, 709–713 (2012).

45. Duquette, A., Brais, B., Bouchard, J. P. & Mathieu, J. Clinical presentation and early evolution of spastic ataxia of Charlevoix-Saguenay. *Mov. Disord.* **28**, 2011–2014 (2013).

46. Synofzik, M. et al. Autosomal recessive spastic ataxia of Charlevoix Saguenay (ARSACS): expanding the genetic, clinical and imaging spectrum. *Orphanet J. Rare Dis.* **8**, 41 (2013).

47. Lee, L. A. et al. Drosophila genome-scale screen for PAN GU kinase substrates identifies Mat89Bb as a cell cycle regulator. *Dev. Cell* **8**, 435–442 (2005).

48. Stebbings, L., Grimes, B. R. & Bownes, M. A testis-specifically expressed gene is embedded within a cluster of maternally expressed genes at 89B in Drosophila melanogaster. *Dev. Genes Evol.* **208**, 523–530 (1998).

49. Anderson, M. A. et al. Asunder is a critical regulator of dynein-dynactin localization during Drosophila spermatogenesis. *Mol. Biol. Cell* **20**, 2709–2721 (2009).

50. Chen, J. et al. An RNAi screen identifies additional members of the Drosophila Integrator complex and a requirement for cyclin C/Cdk8 in snRNA 3'-end formation. *RNA* **18**, 2148–2156 (2012).

51. Boon, M. et al. Primary ciliary dyskinesia: critical evaluation of clinical symptoms and diagnosis in patients with normal and abnormal ultrastructure. *Orphanet J. Rare Dis.* **9**, 11 (2014).

52. Erle, D. J. & Sheppard, D. The cell biology of asthma. *J. Cell Biol.* **205**, 621–631 (2014).

53. Livraghi, A. & Randell, S. H. Cystic fibrosis and other respiratory diseases of impaired mucus clearance. *Toxicol. Pathol.* **35**, 116–129 (2007).

54. Turner, J. & Jones, C. E. Regulation of mucin expression in respiratory diseases. *Biochem. Soc. Trans.* **37**, 877–881 (2009).

55. Wallmeier, J. et al. Mutations in CCNO result in congenital mucociliary clearance disorder with reduced generation of multiple motile cilia. *Nat. Genet.* **46**, 646–651 (2014).

56. Sabath, K. et al. INTS10-INTS13-INTS14 form a functional module of Integrator that binds nucleic acids and the cleavage module. *Nat. Commun.* **11**, 3422 (2020).

57. Chen, J., Waltenspiel, B., Warren, W. D. & Wagner, E. J. Functional analysis of the integrator subunit 12 identifies a microdomain that mediates activation of the Drosophila integrator complex. *J. Biol. Chem.* **288**, 4867–4877 (2013).

58. Wu, Y., Albrecht, T. R., Baillat, D., Wagner, E. J. & Tong, L. Molecular basis for the interaction between Integrator subunits IntS9 and IntS11 and its functional importance. *Proc. Natl Acad. Sci. USA* **114**, 4394–4399 (2017).

59. Blacque, O. E. et al. Functional genomics of the cilium, a sensory organelle. *Curr. Biol.* **15**, 935–941 (2005).

60. Brooks, E. R. & Wallingford, J. B. Multiciliated cells. *Curr. Biol.* **24**, R973–982 (2014).

61. Wallingford, J. B. Planar cell polarity signaling, cilia and polarized ciliary beating. *Curr. Opin. Cell Biol.* **22**, 597–604 (2010).

62. Hayes, J. M. et al. Identification of novel ciliogenesis factors using a new in vivo model for mucociliary epithelial development. *Dev. Biol.* **312**, 115–130 (2007).

63. Barbieri, E. et al. Targeted Enhancer Activation by a Subunit of the Integrator Complex. *Mol. Cell* **71**, 103–116 e107 (2018).

64. Walker, J. R., Corpina, R. A. & Goldberg, J. Structure of the Ku heterodimer bound to DNA and its implications for double-strand break repair. *Nature* **412**, 607–614 (2001).

65. Oegema, R. et al. Human mutations in integrator complex subunits link transcriptome integrity to brain development. *PLoS Genet.* **13**, e1006809 (2017).

66. Huang K. L. et al. Integrator Recruits Protein Phosphatase 2A to Prevent Pause Release and Facilitate Transcription Termination. *Mol. Cell* **80**, 345–358.e9 (2020).

67. Del Viso, F. et al. Congenital Heart Disease Genetics Uncovers Context-Dependent Organization and Function of Nucleoporins at Cilia. *Dev. Cell* **38**, 478–492 (2016).

68. Fakhro, K. A. et al. Rare copy number variations in congenital heart disease patients identify unique genes in left-right patterning. *Proc. Natl Acad. Sci. USA* **108**, 2915–2920 (2011).

69. Schmidt, K. N. et al. Cep164 mediates vesicular docking to the mother centriole during early steps of ciliogenesis. *J. Cell Biol.* **199**, 1083–1101 (2012).

70. Galati, D. F., Sullivan, K. D., Pham, A. T., Espinosa, J. M. & Pearson, C. G. Trisomy 21 Represses Cilia Formation and Function. *Dev. Cell* **46**, 641–650 e646 (2018).

71. Homer, N., Merriman, B. & Nelson, S. F. BFAST: an alignment tool for large scale genome resequencing. *PLoS One* **4**, e7767 (2009).

72. Fitzgerald, D. J. et al. Protein complex expression by using multi-gene baculoviral vectors. *Nat. Methods* **3**, 1021–1032 (2006).

73. Anderson, A. P., Luo, X., Russell, W. & Yin, Y. W. Oxidative damage diminishes mitochondrial DNA polymerase replication fidelity. *Nucleic Acids Res.* **48**, 817–829 (2020).

74. Jaworski, E. & Routh, A. ClickSeq: Replacing Fragmentation and Enzymatic Ligation with Click-Chemistry to Prevent Sequence Chimeras. *Methods Mol. Biol.* **1712**, 71–85 (2018).

Author contributions

L.G.M. and M.Shboul contributed equally to this work. B.R. initiated this project with H.H. who clinically ascertained Family 1 with the support of B.A.R., O.A., R.F. and M.E.K. Family 2 was ascertained by J.A. and L.C. L.G.M. conducted RPE analysis, biochemical purifications, and yeast two-hybrid analysis. M.Shboul performed all experiments on patients' primary cells and in *Xenopus* with the help of N.E.B. and S.R. N.D.E. prepared samples for and analyzed RNA-seq. K-L.H. performed western blot analysis. N.P. built the pTEFdual vector. H.L., S.F.N., and B.M. conducted linkage/sequence analysis. Y.W. and L.T. conducted recombinant protein production and gel filtration. M.Singh and L.J.K. imaged and analyzed PC immunofluorescence. J.N.J, P.S. and L.A.L. performed *Drosophila* rescue experiments. W.K.R. prepared proteomic samples and analyzed LC/MS data. L.G.M., M.Shboul, B.R., and E.J.W. wrote and revised the manuscript. All authors read and reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-022-33547-8.

Correspondence and requests for materials should be addressed to Bruno Reversade or Eric J. Wagner.

Peer review information *Nature Communications* thanks Anita Corbett, Miriam Schmidts, Yongsheng Shi and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

[1]Department of Biochemistry and Molecular Biology, The University of Texas Medical Branch at Galveston, Galveston, TX 77550, USA. [2]Department of Medical Laboratory Sciences, Jordan University of Science and Technology, Irbid, Jordan. [3]Inserm UMR 1163, Institut Imagine, 24 Boulevard du Montparnasse, 75015 Paris, France. [4]Department of Genetic Medicine and Development, University Hospital, Geneva, Switzerland. [5]Department of Biochemistry and Biophysics, University of Rochester School of Medicine Dentistry, Rochester, NY 14642, USA. [6]Center for RNA Biology, University of Rochester School of Medicine Dentistry, Rochester, NY 14642, USA. [7]Department of Pathology and Laboratory Medicine, Department of Human Genetics, David Geffen School of Medicine, University of California Los Angeles, Los Angeles, CA 90095, USA. [8]Department of Cell and Developmental Biology, Vanderbilt University Medical Center, Nashville, TN 37232, USA. [9]Department of Pathology, Microbiology and Immunology, Vanderbilt University Medical Center, Nashville, TN 37232, USA. [10]National Center for Diabetes, Endocrinology and Genetics, Amman, Jordan. [11]Faculty of Medicine, Hospital of the University of Jordan, University of Jordan, Amman, Jordan. [12]Department of Medical Genetics, KOÇ University, Istanbul, Turkey. [13]Institute of Molecular and Cell Biology, A*STAR, Singapore, Singapore. [14]Department of Biological Sciences, Columbia University, New York, NY 10027, USA. [15]Department of Paediatrics, School of Medicine, NUS, Singapore, Singapore. [16]Department of Biological Sciences, Faculty of Science, NUS, Singapore, Singapore. [17]Service de Génétique, Institut Imagine, 24 Boulevard du Montparnasse, 75015 Paris, France. [18]Smart-Health Initiative, Biological and Environmental Sciences and Engineering Division, King Abdullah University of Science and Technology (KAUST), Thuwal, Saudi Arabia. [19]Laboratory of Human Genetics & Therapeutics, Genome Institute of Singapore, A*STAR, Singapore 137673, Singapore. [20]Present address: 3billion, Inc., Seoul, South Korea. [21]These authors contributed equally: Lauren G. Mascibroda, Mohammad Shboul. ✉e-mail: bruno@reversade.com; Eric_Wagner@URMC.Rochester.edu