INTS13 Mutations Causing a Developmental Ciliopathy Disrupt Integrator Complex Assembly

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

[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 Pathology and Laboratory Medicine, Department of Human Genetics, David Geffen School of Medicine, University of California Los Angeles, Los Angeles CA 90095, USA

[6] Department of Cell and Developmental Biology, Vanderbilt University Medical Center, Nashville, TN 37232, USA

[7]Department of Pathology, Microbiology and Immunology, Vanderbilt University Medical Center, Nashville, TN 37232, USA

[8] Faculty of Medicine, Hospital of the University of Jordan, University of Jordan, Jordan

[9] National Center for Diabetes, Endocrinology and Genetics, Jordan

[10] Department of Biological Sciences, Columbia University, New York, NY 10027, USA

[11] Service de Génétique, Institut Imagine, 24 Boulevard du Montparnasse 75015, Paris, France

[12] Department of Paediatrics, School of Medicine, NUS, Singapore

[13] Department of Medical Genetics, KOÇ University, Istanbul, Turkey

[14] Genome Institute of Singapore, A*STAR, 137673, Singapore

[15] Institute of Molecular and Cell Biology, A*STAR, Singapore

* These authors contributed equally to this work

** Corresponding authors: bruno@reversade.com and ejwagner@utmb.edu

ABSTRACT

Oral-facial-digital syndromes (OFD) are a heterogeneous group of congenital disorders characterized by malformations of the face and oral cavity, and digit anomalies. To date, mutations in 12 ciliary-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 mutations in *INTS13,* a subunit of the Integrator complex. This 14-component complex associates with RNAPII and can cleave nascent RNA to modulate gene expression. We determined that INTS13 utilizes a discrete domain within its C-terminus to bind the Integrator cleavage module, which is disrupted by the identified germline *INTS13* mutations. 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 lead to motile cilia anomalies. Altogether, we show that mutations in *INTS13* cause an autosomal recessive ciliopathy, which reveals key interactions within Integrator components.

63 **INTRODUCTION**

64

65 Ciliopathies encompass a large and expanding group of human genetic disorders caused
66 by mutations in genes either encoding ciliary proteins or that participate indirectly in the
67 formation and function of cilia[1, 2, 3, 4]. There have been several hundred diseases identified
68 as ciliopathies that range from renal cystic disease[5] and isolated blindness[6] to multiple
69 organ disorders such as a Bardet-Biedl syndrome (BBS)[7] or Meckel Gruber syndrome
70 (MKS)[8]. The clinical manifestations associated with ciliopathies are broad and highly
71 variable and include: retinal degeneration, specific nervous system malformations,
72 hepatic, kidney and pancreatic cysts, heterotaxia, polydactyly, encephalocele,
73 hydrocephalus, hearing loss, anosmia, intellectual disability, developmental delay,
74 skeletal malformation, craniofacial anomalies, respiratory function impairment, infertility
75 and obesity[4, 9, 10, 11, 12, 13].

76 Among ciliopathies, oral-facial-digital syndromes (OFDs) are a group of clinically
77 and genetically heterogeneous developmental disorders characterized by defects in the
78 development of the face and oral cavity along with digit anomalies. Symptoms associated
79 with OFDs can also be found in other organs, which has led to the classification of at least
80 14 different types of OFD[14, 15]. To date, 12 genes have been identified to be mutated in
81 OFD syndromes and the encoded proteins have been implicated in primary cilia
82 biogenesis and function. The phenotypic spectrum of OFD syndromes can overlap with
83 other known ciliopathies, and mutations in OFD genes can also cause other types of
84 ciliopathies[16, 17, 18, 19, 20, 21, 22]. Of the 12 known genes involved in OFDs, the majority
85 encode structural components of the primary cilium, such as *OFD1*[23] and *TMEM231*[24].
86 However, not all OFD-causing gene mutations such as those in the RNA helicase
87 DDX59[19] localize to the primary cilium, and are thus thought to contribute indirectly to
88 disease pathology. Despite significant progress, seven OFD syndromes are still without
89 a clear genetic aetiology and at least three unclassified forms need further clinical and
90 molecular characterization.

91 The Integrator complex (INT), which is unique to metazoans, is composed of at
92 least 14 subunits, and it associates with RNA polymerase II (RNAPII) where it modulates
93 transcription[25, 26, 27, 28, 29]. Critical among its protein components is Integrator subunit 11

94 (INTS11)[30], which is a member of the metallo-β-lactamase/β-CASP family of RNA
95 endonucleases[31] and is a paralog of the cleavage and polyadenylation specificity factor
96 of 73 kDa (CPSF73), which cleaves pre-mRNA[32]. INTS11 has been found to associate
97 with INTS4 and INTS9 to form what is called the Integrator Cleavage Module (CM), which
98 is thought to be the key enzymatic factor within the complex[33]. The Integrator CM, through
99 the activity of INTS11, has been shown to be essential to cleave the 3' ends of a variety
100 of noncoding RNA and nascent mRNA of protein-coding genes with paused RNAPII. This
101 catalytic activity thus serves as a key functional component of Integrator-mediated
102 transcriptional repression and for the biogenesis of noncoding RNA[34, 35, 36, 37, 38].

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

116 Here, we used a combination of homozygosity mapping and deep sequencing of
117 patient samples to identify two distinct recessive germline mutations in the *INTS13*
118 (MIM615079) gene causing a variable form of OFD syndrome type 2 (OFD2), also known
119 as Mohr Syndrome (MIM252100)[43]. Using multiple molecular and biochemical
120 approaches, we demonstrate that the C-terminus of INTS13 interacts with the Integrator
121 CM and that mutations observed in OFD2 patients specifically disrupt this interaction. We
122 also find that human cells depleted of INTS13 not only have reduced primary cilia, but
123 also exhibit broad transcriptional dysregulation that impacts numerous known genes
124 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 human ciliogenesis and identify a vital interaction between two Integrator subcomplexes.

RESULTS

Characterization of 4 patients with recessive loss-of-function *INTS13* mutations

In Family 1, two affected sisters (individuals II.4 and II.5 in Figure 1a) born to a consanguineous Jordanian couple presented with a congenital disorder consistent with OFD2 (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 shape 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 Supplemental Figure 1a-m, and a summary is provided in Table 1. Inter- and intra-familial variations have been reported in OFD and were clearly observed between the two sisters. Many clinical features are unique to this

156 family and overlap with OFD2 or may be a new form of OFD, with a distinct genetic
157 aetiology.

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

174 The second family had two affected siblings born to French parents not known to
175 be endogamous. The affected 27-year-old woman (II.2) and her affected 20-year-old
176 brother (II.4) (Family 2, Figure 1a) presented with some overlapping phenotypes, but also
177 distinct features from those reported in the Jordanian patients. Both affected sibs were
178 born at term with low birth parameters for weight, length and occipitofrontal circumference
179 (OFC). II.4 presented with hypospadias. Both remained short and became overweight
180 from the age of 3 years, while OFC was in the low normal range at this age. II.2 walked
181 unaided before 18 months and II.4 walked at 22 months. For both, language was delayed
182 and remained limited to a few spoken words, although they could understand simple
183 conversational topics. II.2 was treated for early-onset puberty for 18 months starting at 8
184 years. Both presented with short extremities with brachydactyly of fingers and toes,
185 hyperlordosis when standing, flessum of the knees and pes valgus. Facial features
186 include upslanting palpebral fissures, macrostomia, gum 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 Supplemental Figure 1o-w, and a summary is provided in Table 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 (Figure 1c). This mutation converts Serine652 to a Leucine and is located 16 amino acids upstream of where the frameshift mutation was observed in Family 1.

INTS13 mRNA consists of 17 exons (Figure 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]. Mat89b was later changed to 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 (Figure 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 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 down-regulated (Figure 2a), which we

218 hypothesized may be due to the presence of the premature termination codon triggering
219 nonsense-mediated decay (NMD). Therefore, we isolated total RNA from patient cells
220 that were either mock-treated or treated with cycloheximide, a translational inhibitor that
221 blocks NMD, and RT-qPCR was repeated. The results showed that the level of mutant
222 mRNA was increased following cycloheximide treatment, suggesting that the mutant
223 *INTS13* mRNA was indeed targeted for NMD (Figure 2b). To assess the levels of INTS13
224 protein in F1 cells, we raised two custom antibodies against antigens from the central
225 portion of INTS13 (α-M) and the C-terminal region (α-C), which is downstream of the
226 truncation. These antibodies were affinity-purified and used to probe cell lysates from
227 patient fibroblasts using Western blot analysis. We observed that the C-terminal antibody
228 could detect endogenous INTS13 protein in control cells, but not in F1 patient fibroblasts,
229 likely due to the loss of the epitope in the truncated mutant (Figure 2c). Consistent with
230 this notion, we could detect low levels of truncated INTS13 in F1 patient fibroblasts using
231 the α-M antibody (Figure 2c). Overall, these results reveal that the *INTS13* mutation
232 observed in Family 1 creates an mRNA subject to NMD that is translated at low levels,
233 thereby producing a truncated INTS13 protein.

234 We next analyzed the expression of INTS13 in fibroblasts isolated from a Family
235 2 (F2) patient harboring the homozygous p.S652L mutation. We observed a slight but
236 significant reduction in the level of endogenous *INTS13* mRNA in these mutant cells
237 (Figure 2d) and a pronounced reduction in the level of INTS13 protein as assessed using
238 either antibody (Figure 2e). We posited that the basis for the reduction in INTS13 protein,
239 beyond the minor reduction in mRNA, may be due to increased protein degradation of the
240 INTS13 F2 mutant. To test this, primary fibroblasts were treated with MG132 to inhibit
241 proteasome activity, and we observed an increased level of INTS13 protein comparable
242 to control cells by Western blot analysis (Figure 2f). This suggests that the INTS13
243 mutation observed in Family 2 generates a protein which may be unstable and targeted
244 for degradation via the proteasome.

245 The epithelia lining the human respiratory system are covered with multiple cell
246 types dominated by ciliated cells. These ciliated cells have hundreds of motile cilia to
247 generate fluid flow that keeps mucus moving to protect the airway from inhaled irritants,
248 pathogens and foreign particles. Abnormalities of mucociliary clearance caused by either

249 aberrant ciliated cells or cilia dysfunction often result in recurrent respiratory tract
250 infections. Reduction or impairment of mucociliary clearance in humans results in an array
251 of pathologies such as primary ciliary dyskinesia, chronic obstructive pulmonary disease
252 and asthma [51, 52, 53, 54, 55]. In this study, both probands of Family 1 suffered from recurrent
253 respiratory tract infection and chronic cough. Therefore, we obtained respiratory epithelial
254 cells from the two patients and the mother to assess cilia formation and performance.
255 Immunofluorescence staining was performed using anti-acetylated α-Tubulin antibody to
256 visualize cilia. The average length of nasal cilia was then scored. The results showed that
257 cilia were shorter, fewer in number and occasionally disorganized compared to those of
258 the heterozygous but unaffected mother, which may explain the recurrent respiratory tract
259 infections in the two affected subjects (Figure 2g-h). Importantly, these results also
260 indicate that ciliogenesis was indeed disrupted in patient cells.
261

262 **OFD2 patient mutations within the INTS13 C-terminus disrupt association with the**
263 **Integrator Cleavage Module**

264 Recently, the structure of INTS13 associated with INTS14 has been elucidated.
265 However, the INTS13 C-terminus was unable to be resolved in that structure, thereby
266 obscuring our ability to ascertain the impact of OFD2 mutations[56]. The INTS13/14
267 heterodimer was shown to interact with INTS10 as well as the Integrator CM, the latter of
268 which was found to be mediated independently through the INTS13 C-terminus[56]. With
269 this information in hand, and our previous success mapping Integrator interactions using
270 yeast two-hybrid analysis[33, 50, 57, 58], we fused INTS13 to the Gal4 DNA binding domain
271 and screened for interactions with the other Integrator subunits that were each expressed
272 individually as Gal4 activation domain fusions. This initial approach failed to detect
273 interaction between INTS13 and other members of the complex (Figure 3a, left panels).
274 Given the structural information for INTS13, the most likely explanation for the inability to
275 detect binding partners is that INTS13 may interact with a domain created by an interface
276 between multiple Integrator subunits. We previously demonstrated this to be the case for
277 the INTS4/9/11 Cleavage Module where INTS4 interaction with INTS9 or INTS11 was
278 only detected with yeast two-hybrid if all three subunits were co-expressed[33]. We
279 therefore decided to focus on whether INTS13 could interact with members of the

280 cleavage module and thus repeated the screen with INTS9 and INTS11 expressed
281 simultaneously in *trans*. Using this approach, we could detect a robust interaction
282 between INTS13 and INTS4 only when INTS9 and INTS11 were co-expressed in the
283 same yeast strain (Figure 3a, right panels and Supplemental Figure 2a). To confirm this
284 interaction, we altered the position of each subunit in the screen and could detect an
285 interaction between INTS9 and INTS13 only when both INTS4 and INTS11 were
286 simultaneously expressed in *trans* (Supplemental Figure 2b). These data extend on
287 recent structural and biochemical analyses to indicate that INTS13 can, in the absence
288 of INTS10 and INTS14, associate with subunits of the Integrator CM.

289 The binding interfaces of the Integrator CM have been previously well-
290 characterized using both biochemical and structural approaches[33, 57]. Based upon these
291 prior analyses, amino acid mutations within each of the three CM subunits are known to
292 disrupt formation of the heterotrimer. Therefore, we repeated our modified two-hybrid
293 assay with each of these mutants tested individually and discovered that any disruption
294 of the INTS4/9/11 CM also results in loss of INTS13 interaction (Figure 3b). Further, in
295 order to confirm the interaction between these four subunits using an independent
296 approach, we co-infected High Five cells with baculoviruses individually expressing His-
297 tagged-INTS4, or untagged INTS9, INTS11, and INTS13. The complex was purified from
298 cell lysate using nickel affinity resin and subjected to gel filtration analysis. We observed
299 that all four Integrator subunits remained intact as a complex and could be detected in
300 the same fractions at near stoichiometric levels (Figure 3c). Collectively, these results
301 indicate that INTS13 interacts with the three components of the Integrator CM only once
302 the module is fully assembled.

303 We next wanted to determine which region of INTS13 is responsible for this
304 interaction and whether OFD2 mutations may impact it. We tested a series of INTS13
305 fragments in modified yeast two-hybrid assays to determine whether any of these regions
306 were sufficient to mediate interaction with INTS4/9/11 (Figure 3d schematic). The INTS13
307 C-terminal region, composed of amino acids 501-706, was as effective as the full-length
308 INTS13 to support growth on selective media (Figure 3d). We further dissected the
309 INTS13 C-terminus through a series of smaller deletion mutants, which ultimately defined
310 that the last 130 residues, amino acids 577-706, was sufficient to bind to the Integrator

311 CM (Supplemental Figure 2c). This result is consistent with biochemical approaches
312 recently described where a GST fusion protein containing the C-terminus of INTS13 can
313 pull down subunits of the Integrator CM[56]. Because both mutations observed in Family 1
314 (p.K668Nfs*9) and Family 2 (p.S652L) are within the region of INTS13 found to be
315 sufficient to bind INTS4/9/11, we tested whether these mutations could disrupt binding.
316 Indeed, we found that either mutation in the context of the full-length INTS13 protein
317 resulted in a loss of interaction between INTS13 and INTS4/9/11, indicating that this
318 interaction may be specifically disrupted in the four affected patients (Figure 3d).
319

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

322 We next sought to determine if the C-terminal INTS13 mutations have an impact on the
323 assembly of the entire Integrator complex. To do so, we created individual 293T cell lines
324 harboring three different versions of the *INTS13* cDNA: wild type (wt), Family 1 mutant
325 (F1), or Family 2 mutant (F2). Each of these lines expresses exogenous INTS13 with an
326 N-terminal FLAG tag under the control of the doxycycline-inducible promoter. After 48
327 hours of doxycycline treatment, nuclear fractions from each of the cell lines as well as an
328 untransfected naïve line serving as a negative control were extracted. Purified Integrator
329 complexes associated with each INTS13 protein, and the degree of co-associated INT
330 subunits was assessed by LC/MS and Western blot analysis. We observed that both
331 patient-derived INTS13 mutants not only disrupted association with INTS4/9/11,
332 confirming our yeast two-hybrid experiments, but also resulted in reduced binding to
333 nearly all other members of the complex (Figure 4a). The exception to this observation
334 was the binding of INTS14 and to a lesser extent INTS10, which were not as dramatically
335 impacted by the mutations (Figure 4a).

336 We have shown previously that Integrator constituency and function is highly
337 conserved in *Drosophila*[50]. Therefore, we wanted to test if patient mutations have the
338 same disruptive effect in the orthologous fly complex. Supporting this approach, the C-
339 terminal region of INTS13 is conserved in the *Drosophila* IntS13 orthologue (Asunder or
340 dIntS13) and so we introduced the analogous patient mutations into dIntS13
341 (Supplemental Figure 3a). We then created individual S2 cell lines expressing copper-

342 inducible FLAG-dIntS13-wild type (wt), Family 1 mutant (dF1), or Family 2 mutant (dF2).
343 We generated nuclear extract from these cell lines as well as a naïve S2 control and
344 purified *Drosophila* Integrator complexes using anti-FLAG affinity resin. Similar to the
345 human Integrator purifications, we used both LC/MS and Western blot analysis to assess
346 complex integrity. In accordance with the human data, we found that patient mutation
347 'equivalents' within *Drosophila* IntS13 reduced association with most other Integrator
348 subunits (Figure 4b). We also observed that association of dIntS10 and dIntS14 were
349 only marginally affected by the patient mutations.

350 These results confirm that OFD2-causing mutations disrupt INTS13 assembly into
351 Integrator, but also suggest that INTS13 mutants do not affect INTS14 and possibly
352 INTS10 binding. Therefore, we wanted to determine whether we could detect an
353 interaction of INTS13 with INTS10 and INTS14 using our modified two-hybrid assay and
354 if INTS13 mutations affected it. We initially expressed INTS14 fused to the Gal4 DNA
355 binding domain and screened for interactions individually with each of the other Integrator
356 subunits fused to the activating domain. This approach did not detect any pairwise
357 interactions (Supplemental Figure 3b, left panel). Therefore, we expressed each
358 Integrator subunit individually *in trans* with INTS14 fused to the DNA binding domain and
359 INTS13 fused to the activating domain, and we found that the additional presence of
360 INTS10 can allow for an interaction between INTS14 and INTS13, supporting the model
361 that these three subunits form a module as has been recently shown biochemically[56]
362 (Supplemental Figure 3b, right panel). Finally, we tested whether our patient-derived
363 mutations within INTS13 affect the interaction of INTS10/13/14 and observed that both
364 INTS13 mutants could interact with INTS10/14 comparable to the wild-type INTS13
365 (Figure 4c). Overall, the results presented here provide a molecular basis for OFD2 where
366 INTS13 mutations disrupt interaction between the INTS10/13/14 submodule and the
367 INTS4/9/11 CM, along with the rest of the Integrator complex (Figure 4d).
368

369 **Reduced INTS13 expression in RPE cells leads to loss of ciliogenesis and broad**
370 **disruption of transcriptional regulation of ciliary genes.**

371 Results presented thus far reveal that OFD2 patient cells exhibit disrupted ciliogenesis
372 and that these same mutations also disrupt the INTS13 module's interaction with the CM.

373 We next wanted to determine whether reduced expression of INTS13 in cultured cells
374 was sufficient to inhibit ciliogenesis. We utilized human retinal pigment epithelial (RPE)
375 cells as a model, because these cells exhibit clearly visible primary cilium upon serum
376 starvation. We treated RPE cells with either control siRNA or one of two distinct siRNA
377 targeting *INTS13* and discovered that either siRNA could effectively deplete endogenous
378 INTS13 (Figure 5a). To visualize primary cilia, we probed fixed cells using antibodies
379 raised to ADP-ribosylation factor-like protein 13B (ARL13B)[59], which stains the ciliary
380 axoneme, and γ-tubulin, which stains the basal body[39]. We observed that the basal body
381 and axoneme components were still present and appropriately juxtaposed in INTS13-
382 depleted cells but were clearly not utilized to form a normal primary cilium (Figure 5b).
383 We found that depletion of INTS13 with either siRNA reduced the number of properly
384 assembled primary cilium by ~50% relative to control siRNA (Figure 5c).
385 To understand the underlying changes in gene expression that occurred after
386 *INTS13* depletion, we subjected control siRNA-treated and each of the *INTS13* siRNA-
387 treated RPE cells to RNA-seq analysis. We conducted RNA-seq using three biological
388 replicates from cells treated with each siRNA and used DESeq to determine differential
389 expression (Supplemental Table 1). Taking advantage of two distinct *INTS13* siRNAs, we
390 only considered genes observed to be significantly upregulated/downregulated (>2-fold,
391 $p<0.0001$) in both knockdown conditions. Using these filters, we identified 254 genes that
392 were upregulated and 156 genes that were downregulated upon *INTS13* knockdown
393 (Figure 5d). Gene ontology analysis revealed a striking enrichment in genes involved in
394 cilium biogenesis uniquely within the upregulated gene set, suggesting that knockdown
395 of *INTS13* leads to de-repression of a collection of genes involved in cilium biogenesis
396 (Figure 5e). We validated several of these gene changes using RT-qPCR including well-
397 characterized ciliary genes such as *BBOF1*, *CFAP206*, and *CFAP53* (Figure 5f).
398 Altogether, these results indicate that depletion of INTS13 in human cells leads to
399 disrupted ciliogenesis that is likely caused by broad transcriptional deregulation of genes
400 involved in primary cilium biogenesis.
401
402 **Knockdown of *ints13* causes developmental defects and ciliary phenotypes in**
403 ***Xenopus*.**

404 To examine the role of INTS13 *in vivo*, we first examined its expression in *Xenopus laevis*
405 embryos using whole mount in situ hybridization (WISH). *ints13* showed both maternal
406 and zygotic contributions with relatively ubiquitous expression. At late gastrula (stage 12),
407 the expression of *ints13* was enriched in ectodermal lineages, with no apparent
408 expression in the endoderm. At neurula (stage 17), its expression was enriched in the
409 neural tube and neural folds (Supplemental Figure 5a). From tadpole stages onwards,
410 *ints13* transcripts were abundantly observed in ciliated tissues and organs including the
411 brain, branchial arches, pronephros, eye and otic vesicles. Remarkably *ints13* was found
412 to follow a punctate pattern of expression in the skin, which coincided with that of
413 multiciliated cells (MCCs). This was validated by whole mount antibody staining, using
414 the α-M antibody highlighting that endogenous INTS13 was enriched in epidermal cells
415 endowed with multicilia (Supplemental Figure 5b).

416 In order to gain insight into the precise function of ints13 during *Xenopus*
417 embryonic development, we performed loss-of-function analysis using a morpholino
418 oligonucleotide (MO) targeted against the *ints13* start codon (IntS13-MO). This IntS13-
419 MO was previously established to be specific[47], and we confirmed the translational block
420 efficiency achieved by this MO by demonstrating that *ints13* morphants were nearly
421 devoid of endogenous ints13 protein (Figure 6d). The Ints13-MO resulted in
422 developmental delay, a small head, shortened body axis, bent tail, cardiac oedema,
423 polyploidy and early developmental arrest in *Xenopus* morphants (Figure 6a). In addition
424 to the previously described gastrulation defects, we also observed prominent deficiencies
425 in ciliogenesis and craniofacial development. Based on the enrichment of ints13 in MCCs,
426 from which protrude hundreds of beating cilia[60, 61], we examined the possibility that *ints13*
427 may affect multi-cilia formation and function. *In situ* hybridization using a specific MCC
428 marker *ccdc19*[62] showed that morphants had significantly less *ccdc19* than did control
429 embryos (Figure 6b), which was validated by whole mount antibody staining using the
430 anti-acetylated α-tubulin antibody marking cilia protruding from MCCs (Figure 6c). We
431 noted increased distances between puncta and irregular staining in *ints13*-depleted
432 embryos, suggesting that the epidermis of *ints13*-depleted tadpoles could have fewer
433 differentiated MCCs and/or that these cells harbor fewer cilia. Scanning electron
434 microscopy (SEM) performed on both *ints13* morphants and control embryos revealed

435 that MCCs were defective, with a significantly reduced number of cilia per cell (Figures
436 6e, f). While the length of cilia seemed to be normal, inconsistency was observed within
437 and between MCCs, suggestive of incomplete penetrance (Figure 6e). Cilia exhibited
438 abnormalities in morphology, protrusion and direction (Figure 6f). Of note, we were unable
439 to detect overt abnormalities in the ultrastructure of the cilium by transmission electron
440 microscopy (TEM) (Figure 6g). These results prompted us to examine the function of the
441 cilia, which normally generate an anterior-to-posterior directional flow over the tadpole's
442 skin. Using high-resolution imaging, time-lapsed movies revealed that morphants were
443 completely unable to move liquid around their bodies compared to control embryos
444 (Supplemental Movie 1). Moreover, we found that ciliary beating in *ints13*-depleted
445 embryos was slow, erratic and uncoordinated compared to control embryos
446 (Supplemental Movie 2). These data indicate that INTS13 is necessary for MCC
447 differentiation, as well as cilia formation and beating, thus revealing its critical role in
448 ciliogenesis during development.
449
450 **Discussion**
451
452 Oral facial digital syndromes are an expanding group of genetic conditions characterized
453 by malformations of the face, oral cavity, and digits. These disorders exhibit extensive
454 variability of clinical features with intra-familial and inter-familial variability. Here we
455 describe two homozygous germline mutations in the *INTS13* gene in two independent
456 families, with Family 1 fulfilling clinical criteria for OFD. Through molecular and
457 biochemical approaches, we determined that INTS13 associates with INTS10/14 to form
458 a heterotrimer that interacts with the Integrator CM through the C-terminus of INTS13.
459 The two disease-causing mutations both reduce steady-state levels of endogenous
460 INTS13 protein in patient cells and disrupt INTS13 interaction with the CM, likely causing
461 transcriptional defects in a broad collection of ciliary genes. The binding of INTS10/13/14
462 to the Integrator CM is clearly vital, as depletion of INTS13 was sufficient to disrupt
463 ciliogenesis in developing *Xenopus* embryos. These results underscore both the
464 importance of Integrator assembly during development and reveal the importance of
465 Integrator transcriptional regulation in ciliogenesis.

466 INTS13 and INTS14 were not initially described as components of Integrator[25], but
467 were subsequently discovered through a genome-wide RNAi screen for factors regulating
468 UsnRNA 3' end formation[50]. Since that initial identification, INTS13 has been shown to
469 play a key role in the differentiation of monocytes[63]. During the course of this manuscript
470 preparation, we became aware of another study in which the structure of INTS13 and
471 INTS14 as a heterodimer was solved[56]. Further, using elegant biochemical approaches,
472 Sabath and colleagues determined that INTS13/14 associates into a heterotrimer with
473 INTS10 and the C-terminus of INTS13 binds to the Integrator CM[56]. These structural and
474 biochemical findings are remarkably concordant with results attained here. Interestingly,
475 it was also discovered that the INTS10/13/14 heterotrimer has an overall architecture
476 related to the Ku70-Ku80 DNA repair complex[64] and is capable of binding DNA or RNA
477 *in vitro* with the highest affinity for RNA hairpins. Thus, our work here and the structural
478 data in that study establish that Integrator has at least two submodules: the previously
479 defined INTS4/9/11 Cleavage Module containing the RNA endonucleolytic activity and a
480 Nucleic Acid Binding (NAB) module composed of INTS10/13/14. Interaction between
481 these two modules is mediated in part through the C-terminus of INTS13. The structure
482 of the INTS13 C-terminus was not visualized by crystallography, thus obscuring our ability
483 to understand how the germline mutations identified herein disrupt the folding and/or
484 binding of that region[56]. Nevertheless, the collective implication from these two studies is
485 that the Integrator NAB module can bind to RNA hairpins to position the Integrator CM for
486 RNA cleavage and that this process is perturbed in OFD 2/Mohr syndrome patients.
487 Our results also provide evidence that *INTS13* is a novel ciliopathy gene, and
488 phenotypes associated with *INTS13* mutation are likely due to a dysfunction in Integrator-
489 mediated transcriptional regulation. This finding is consistent with previous studies
490 showing that depletion of multiple Integrator subunits can also lead to disrupted
491 ciliogenesis[41], and that mutations within INTS1 or INTS8 also create ciliopathy-related
492 symptoms[42, 65]. Herein, we discovered that depletion of INTS13 causes a ~2- to 10-fold
493 upregulation in a broad collection of ciliary genes in RPE cells, which is consistent with
494 recent work demonstrating that depletion of Integrator subunits in *Drosophila* cells
495 primarily causes an increase in transcription in a subset of genes[37, 38]. We posit that an
496 increase in transcription of a broad collection of ciliary genes is sufficient to disrupt

497 ciliogenesis by potentially overproducing specific components. The concept of dosage
498 imbalance disrupting ciliogenesis has been well documented through reports where gene
499 duplication of centrosome- or cilia-encoding genes can cause ciliopathies[66, 67, 68]. Even
500 small increases in ciliary gene expression can prove disruptive to ciliogenesis, as
501 evidenced by Down Syndrome patients who present ciliopathy symptoms and, due to its
502 location on chromosome 21, overproduce *Pericentrin* a mere ~1.5-fold[69]. Overall, our
503 results illustrate how sensitive the process of ciliogenesis is to the dosage of components
504 and underscores the need for coordinate transcriptional regulation of genes critical to
505 ciliogenesis. It will be interesting in future studies to determine the mechanism by which
506 Integrator coordinately regulates ciliary genes and how disruption of the NAB and
507 Cleavage Modules gives rise to loss of this transcriptional control.

508

509

510 **METHODS**

511 **Clinical work.** The index cases of the Jordanian family were initially diagnosed at the
512 National Center for Diabetes, Endocrinology & Genetics in Jordan by Prof. Hanan
513 Hamamy. Clinical assessments were undertaken at different clinics at the University
514 hospital of Jordan. Saliva samples were collected from the whole family members (I.1,
515 I.2, and II.1 to II.6) and skin biopsies were obtained Family 1 (individuals I.2 and II.4).
516 Family 2 was diagnosed at the Imagine Institute in France by Prof. Jeanne Amiel. Blood
517 and skin biopsy were collected from age-matched healthy individuals and patient II:4 from
518 Family 2. Genomic DNA was extracted using (Puregene DNA Purification Kit, Gentra
519 Systems) following the manufacturer's instructions. Nasal epithelial cells were also
520 obtained from (Family 1: I.2, II.4 and II.5) by nasal-brush biopsy (Cytobrush Plus,
521 Medscand, Sweden).

522

523 **Genomic loci capture.** Briefly, custom arrays (Agilent 244K) were designed to target
524 every exonic sequence of the genes present in the five homozygous IBD regions in Family
525 1 (chr6:167,901,317-170,287,795); (chr8:106,743,988-122,186,235); (chr12:12,129,350-
526 28,091,505); (chr13:21,704,404-26,337,017); (chr16:58,296,974-65,631,053). A DNA

527 library from patient (II:4) was prepared according to the Illumina library generation
528 protocol version 2.3 and hybridized to custom arrays according to the Agilent CGH 244K
529 array protocol, and washed, eluted and amplified. The sample was submitted to one
530 channel of Illumina flow cell and sequenced by Illumina Genome Analyzer (GAII) using
531 standard manufacturer's protocol. The image data was processed by the provided GA
532 Pipeline (Firecrest version 1.3.4 and Bustard version 1.3.4) and all sequences were
533 aligned to the human reference genome (UCSC, build 18) by Blat-like Fast Accurate
534 Search Tool (BFAST)[70]. Mismatches were further filtered to identify variants seen ten or
535 more times, in which the variant was called as homozygous and did not overlap with a
536 known dbSNP129 entry mismatch. Nonsynonymous mutations were identified with
537 additional SeqWare tools and the ''knownGene'' gene model from the UCSC hg18. The
538 open-source SeqWare project that provides a LIMS tool for tracking samples (SeqWare
539 LIMS) and a pipeline for sequence analysis (SeqWare Pipeline) was used throughout this
540 work.

541 For exome sequencing in Family 2, sequences were aligned to the human genome
542 reference sequence (UCSC Genome Browser, GRCh38 build) by Burrows-Wheeler
543 Aligner. Downstream processing was carried out with the Genome Analysis Toolkit
544 (GATK), SAMtools and Picard Tools. Substitution calls were made with GATK Unified
545 Genotyper, whereas indel calls were made with a GATK IndelGenotyperV2. All calls with
546 a read coverage ≤2X and a Phred-scaled single-nucleotide polymorphism (SNP) quality
547 of ≤20 were filtered out. Variants were annotated using an in-house-developed annotation
548 software (PolyWeb) allowing filtering of variants according to relevant genetic models.
549

550 **Mutation analysis.** Positional candidate genes were obtained from GenBank and
551 Ensembl databases. Segregation analysis was done for candidate mutations by Sanger
552 sequencing with the BigDye Terminator cycle sequencing kit (Applied Biosystems) using
553 primers flanking each mutation. The primer sequences for *INTS13* (NM_018164) and
554 *SACS* (NM_014363) mutations screening are given in Supplemental Table 2.

555 **Cell culture and transfections.** Mouse fibroblast cells and hTERT transformed RPE1
556 cells were cultured in DMEM and Ham's F12/DMEM (Invitrogen), respectively.
557 Lymphocytes cells were grown in cell culture medium (RPMI) (Invitrogen). All cultured

558 media were supplemented with 10% Fetal Bovine Serum (FBS) (HyClone), 1X Penicillin
559 / Streptomycin and 2 mM L-glutamine under standard conditions (37°C, 5% CO2). To
560 induce ciliogenesis, cells were grown in serum starvation medium (0.5% FBS) for 72
561 hours. Transient transfection was done using Lipofectamine 2000 (Life Technologies)
562 according to manufacturer's recommended protocol.

563 HEK 293T (ATCC, Manassas VA, CRL-3216) cells were grown at 37°C with 5%
564 CO_2 in DMEM (Gibco, Waltham MA, #11965-092), supplemented with 10% (v/v) FBS and
565 1% (v/v) penicillin-streptomycin (Gibco, #15070-063). RPE cells (ATCC, ARPE-19, CRL-
566 2302, lot #70004873) were grown at 37°C with 5% CO_2 in DMEM/F-12 50/50 (Corning,
567 Manassas VA, 10-092-CV), supplemented with 10% (v/v) FBS and 1% (v/v) penicillin-
568 streptomycin. *Drosophila* S2 cells were grown at 27°C in Sf-900 II SFM (Gibco, #10902-
569 088), supplemented with 1% (v/v) Antibiotic-Antimycotic (Gibco, 15240-062).

570 293T cells were transfected using CRISPR/Cas9 to make inducible stable lines.
571 300,000 cells per well were seeded in a 6-well plate. The following day, 2 µl of
572 Lipofectamine 2000 (Invitrogen, Carlsbad CA, 11668-019) was combined with 100 µl of
573 Opti-MEM I (Gibco, 31985-070) per well and incubated at room temperature for 7
574 minutes. 100 µl of this solution was added to 500 ng of the Cas9 AAVS1 sgRNA plasmid
575 and 500 ng of a pTet-3xFLAG-INTS13 plasmid per well and incubated at room
576 temperature for 20 minutes. The transfection solution was then added to wells containing
577 1 ml of media. Each insert gene was done in duplicate wells. The next day, duplicate wells
578 were combined, and cells were expanded to a 10-cm dish. 48 hours after transfection,
579 stable cell selection began with 1.5 µg/ml puromycin. Once stable lines were created,
580 cells were maintained at 1 µg/ml puromycin. Human *INTS13* cDNA was mutated by site-
581 directed PCR to create the genes for Family 1 and Family 2. All three versions of the
582 *INTS13* gene, including wild type, were cloned individually into a plasmid containing the
583 tetracycline-inducible Tet-On 3G promotor with a 3x-FLAG tag for insertion into the
584 AAVS1 locus. This was accomplished by co-transfection with the Cas9 plasmid
585 containing the AAVS1 guide RNA (ACCCCACAGTGGGGCCACTA). Protein expression
586 for nuclear extract was induced for 48 hours with 0.4 µg/ml doxycycline for INTS13 wild
587 type and Family 1 and 12 µg/ml for Family 2 in 30 15-cm diameter dishes per line.

588 Stable lines were also made with S2 cells. *Drosophila* INTS13 cDNA was mutated
589 using site-directed PCR to make orthologous versions of the Family 1 and Family 2
590 mutations indicated in Supplemental Fig 3. The Family I cDNA also had an early stop
591 codon added to recapitulate the truncation seen in the human gene. These modified
592 genes, and wild type *INTS13*, were cloned into a pMT-3xFLAG-puro plasmid [37] following
593 the metallothionein promotor and 3x-FLAG tag. $2x10^6$ cells were plated in Schneider's
594 media supplemented with 10% FBS in a 6-well plate overnight. 2 µg of plasmid was
595 transfected using Fugene HD (Promega, Madison WI, #E2311). Plasmid DNA was mixed
596 with 8 µL Fugene and 100 µL media and incubated at room temperature for 15 minutes
597 before being added to cells. After 24 hours, 2.5 µg/mL puromycin was added to the media
598 to select and maintain the cell population. Cells were transitioned to SFX media without
599 serum for large scale growth. Protein expression for nuclear extract was induced by
600 adding 500 mM copper sulfate for 48 hours to 1 liter of each cell line grown to
601 approximately $1x10^7$ cells/mL.
602
603 **Immunofluorescence staining.** For patient cells: cells were seeded on 12 mm cover
604 slips. Cells were grown to around 80% confluency and then incubated with 0.5% serum
605 media for 72 hours to induce cilia formation, washed and kept at 4°C until use. Nasal
606 epithelial cells were suspended in RPMI without supplements to avoid artefacts caused
607 by serum. The cytobrush was moved gently up and down for 5 min to detach ciliary cells
608 from the brush. Samples were then spread onto non-coated glass slides, air dried, and
609 stored at -80°C until use. (See our protocol for immunofluorescence staining, URL)
610 For RNAi-treated RPE cells (Figure 5b): After RNAi treatment (described in the
611 RNA interference section) and serum starvation of cells grown on cover slips, cells were
612 washed with DPBS and each cover slip was placed in an individual well of a 24-well plate.
613 Cells were fixed to the cover slips by incubating in ice cold methanol for 7-8 minutes and
614 washed thoroughly with PBS. Then samples were blocked for 30 min (all incubation steps
615 done at 37°C with humidity). The blocking buffer used was 1% bovine serum albumin
616 (Sigma-Aldrich, St. Louis MO, A7906-50G) (w/v) in PBS + 0.1% Tween 20 (Fisher
617 Scientific, Fair Lawn NJ, BP337-500). Probing was performed for 2 hours with the
618 following primary antibodies at 1:1000 in blocking buffer: rabbit anti-gamma tubulin

619 (Proteintech, Rosemont IL, 15176-1-AP) and mouse anti-ARL13B (Proteintech, 66739-1-
620 Ig), followed by three washes with PBS for 5 minutes each. Secondary antibody probing
621 was for 1 hour in the dark with the following antibodies at 1:1000 in blocking buffer: goat
622 anti-rabbit Alexa Fluor 594 (Life Technologies, Carlsbad CA, A11037) and goat anti-
623 mouse Alexa Fluor 488 (Life Technologies, A11029), followed by three more washes with
624 PBS. Cover slips were mounted on slides using Fluoroshield Mounting Medium with DAPI
625 (Abcam, Cambridge, UK, ab104139). Confocal imaging was performed on a Leica TCS-
626 SP8 3X system (Leica Microsystems, Wetzlar Germany) equipped with a Leica HC PL
627 APO CS2 63×/1.4 oil immersion objective. A tunable (470-670 nm) pulsed white light laser
628 (Leica WLL) and a UV 405 nm diode laser was used for excitation. Images were
629 processed with the image analysis software, Imaris Pro 9.3.0.
630
631 **Reverse transcription (RT-PCR) and Quantitative PCR.** Figure 2: Total RNA was
632 isolated from cells or embryos using Trizol (Ambion by Life Technologies, Carlsbad CA,
633 15596018) followed by extraction using the RNeasy Mini Kit (Qiagen, Hilden Germany).
634 RNA (1 µg) was reverse transcribed using Iscript™ cDNA Synthesis Kit (Bio-transcript
635 lead) and the mis-splicing of the transcript was assessed using pair of primers designed
636 in the exons upstream and downstream of the morpholino binding site. Quantitative real-
637 time PCR was performed with power SYBR green master mix (Applied Biosystems) using
638 a 7300 real-time polymerase chain reaction machine (Applied Biosystems) and
639 normalized to the housekeeping gene *β-actin*. Primer sequences for RT-PCR and
640 quantitative real-time PCR are listed in the Supplemental Table 2.

641 Figure 5: Total RNA was isolated using Trizol for phenol-chloroform extraction.
642 cDNA was reverse transcribed using Superscript III Reverse Transcriptase (Thermo
643 Scientific, Rockford IL, #18080085) according to the manufacturer's instructions. Random
644 hexamers were used for cDNA synthesis primers and qRT-PCR was then carried out in
645 triplicate using Bio-Rad iTaq Universal SYBR Green Supermix (Bio-Rad, Hercules CA,
646 #1725120) and measured with the CFX Connect Real-Time System (Bio-Rad). Data was
647 analyzed using the ΔΔCt method to calculate fold change, with *7SK* as the control gene.

648 **Morpholino, RNA injection and embryological methods.** The *ints13* morpholino (MO)
649 purchased from Gene Tools were resuspended in sterile water to a concentration of 1
650 mM according to the manufacturer's instructions. RNA was then extracted and converted
651 to cDNA. To target non-neural ectoderm, 1:4 of diluted *ints13*-MO was injected at the two-
652 to four- cell stage. Morpholino sequences are given in Supplemental Table 2. Protocols
653 for *Xenopus* fertilization, MO-injections, antibody staining, whole-mount in situ
654 hybridization and immunofluorescence staining are at our protocol website
655 (https://sites.google.com/a/reversade.com/www/protocols/). All experiments with
656 *Xenopus* embryos were approved by the Singapore National Advisory on Laboratory
657 Animal Research. A stereomicroscope equipped with an ICD digital camera (Leica M205
658 FA) was used to capture images of embryos in successive focal planes. Images were
659 then combined into one picture with Photoshop CS6 (Adobe).

660 **Western Blot.** For Figures 2 and 6: Cells or embryos were lysed in RIPA buffer
661 supplemented with a cocktail of protease inhibitor (Roche). Total protein was resolved on
662 SDS polyacrylamide gels (Bio-Rad) with DTT, followed by transfer on polyvinylidene
663 difluoride (PVDF) membranes. The following antibodies were used for HRP-mediated
664 chemiluminescent: polyclonal Rabbit antibody to INTS13 (C & M), monoclonal Mouse
665 antibody to actin (Chemicon; Mab1501R), β-tubulin (E7, Developmental Studies
666 Hybridoma Bank, University of Iowa, Iowa City, IA), dynein heavy chain (P1H4, gift from
667 T. Hays, University of Minnesota, Minneapolis, MN) and anti-HA antibody (Sigma, H 6533
668 (clone HA-7).

669 For Figures 4 and 5: Protein was extracted directly from RPE cells treated with
670 siRNA by adding 2X SDS loading buffer (120mM Tris pH6.8, 4% SDS, 200mM DTT, 20%
671 Glycerol, and 0.02% Bromophenol blue) to cells while on the plate. Samples were then
672 incubated at room temperature while on the plate with periodic swirling prior to a 10-
673 minute boiling at 95°C and a short sonication. Denatured protein samples were then
674 resolved on a 10% polyacrylamide gel and transferred to a PVDF membrane (Thermo
675 Scientific, 88518). Blots were probed as previously described [37]. Samples from nuclear
676 extract immunoprecipitation are described in the anti-FLAG affinity purification.
677

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

690 **Plasmid construction and Yeast two-hybrid analysis**. Yeast two-hybrid analysis was
691 carried out in PJ69-4α and PJ49-4a strains as described[33]. In brief, human INTS1 through
692 INTS14, and INTS13 variations, were cloned into pGAD, pOBD, pTEF, or pCEV-
693 G1Ura3/TEFdual vectors using conventional cloning. pOBD plasmids were then
694 transformed into PJ69-4α yeast and were selected on tryptophan-dropout medium
695 (Clontech, Mountain View CA, #630413); pGAD plasmids were transformed into PJ49-4a
696 yeast and were selected on leucine-dropout medium (Clontech, #630414). pTEF or
697 pTEFdual plasmids were included to express proteins in trans and were transformed with
698 pOBD plasmids and selected on media lacking tryptophan and uracil (Sunrise Science
699 Products, San Diego CA, 1316-030). Yeast strains were then mated and subsequently
700 selected on medium lacking tryptophan, leucine, and uracil (Sunrise Science Products,
701 1328-030). Interactions were tested through serial dilution (one to five) of diploid yeast
702 followed by plating on medium lacking tryptophan, leucine, and uracil or on medium
703 lacking tryptophan, leucine, uracil, and histidine (Sunrise Science Products, 1334-030)
704 that also was supplemented with 1 mM 3-amino-1,2,4-triazole (MP Biomedicals, Irvine
705 CA, 4061-722).

706 **Expression and purification of INTS4-INTS9-INTS11-INTS13 complex**. Full-length
707 human INTS4, INTS9, INTS11, and INTS13 were co-expressed in insect cells using

708 Multibac technology (Geneva Biotech)[71]. INTS4 carried an N-terminal hexa-histidine tag
709 while the other three proteins were untagged. High Five cells were grown in ESF 921
710 medium (Expression Systems) by shaking at 120 rpm at 27 °C until the density reached
711 1.5-2x10^6 cells mL^{-1}. Cells were infected by P1 viruses and harvested after 48 h. The
712 cells were lysed by sonication in a buffer containing 20 mM Tris (pH 8.0), 200 mM NaCl,
713 5% (v/v) glycerol and 0.1 mM PMSF. The supernatant of cell lysates was incubated with
714 nickel beads (Qiagen) and the eluate was then subjected to gel filtration chromatography
715 (Superdex 200, GE Healthcare), in a running buffer of 20 mM Tris (pH 8.0), 250 mM NaCl,
716 and 5 mM DTT.

717 **Nuclear extract.** Cells were collected and washed in cold PBS. 293T cells were detached
718 from dishes while S2 cells were grown in suspension and pelleted by centrifugation. Cells
719 were then resuspended in five times the cell pellet volume of Buffer A (10mM Tris pH8,
720 1.5 mM $MgCl_2$, 10 mM KCl, 0.5mM DTT, and 0.2mM PMSF). Resuspended cells were
721 allowed to swell during a 15-minute rotation at 4°C. After pelleting down at 1,000g for 10
722 minutes, two volumes of the original cell pellet of Buffer A was added and cells were
723 homogenized with a dounce pestle B for 20 strokes on ice. Nuclear and cytosolic fractions
724 were then separated by centrifugation at 2,000g for 10 minutes. To attain a nuclear
725 fraction, the pellet was washed once with Buffer A before resuspending in an equal
726 amount of the original cell pellet volume of Buffer C (20 mM Tris pH8, 420mM NaCl, 1.5
727 mM $MgCl_2$, 25% glycerol, 0.2 mM EDTA, 0.5 mM PMSF, and 0.5 mM DTT). The sample
728 was then homogenized with a dounce pestle B for 20 strokes on ice and rotated for 30
729 minutes at 4°C before centrifuging at 15,000g for 30 minutes at 4°C. Finally, supernatants
730 were collected and subjected to dialysis in Buffer D (20 mM HEPES, 100 mM KCl, 0.2
731 mM EDTA, 0.5 mM DTT, and 20% glycerol) overnight at 4°C. Prior to any downstream
732 applications, nuclear extracts were centrifuged again at 15,000g for 3 minutes at 4°C to
733 remove any precipitate.

734 **Anti-FLAG affinity purification.** To purify FLAG-tagged Integrator complexes for mass
735 spectrometry, generally between 8 and 10 mg of nuclear extract (approximately 1.9 mL
736 of extract depending on the concentration) was mixed with 100 µL anti-Flag M2 affinity
737 agarose slurry (Sigma-Aldrich, #A2220) washed with 0.1 M glycine then equilibrated in

738 binding buffer (20 mM HEPES pH7.4, 150 mM KCl, 10% Glycerol, 0.1% NP-40). This
739 mixture was rotated for four hours at 4°C. Following the four-hour incubation/rotation, five
740 sequential washes were carried out in binding buffer with a 10-minute rotation at 4°C
741 followed by a 1,000g centrifugation at 4°C. After a final wash with 20 mM HEPES buffer,
742 the supernatant was removed using a pipette and the beads were kept cold and submitted
743 to the mass spectrometry core where the protein complexes were eluted by digestion
744 (described below). For immunoprecipitation samples intended for Western blot, a similar
745 protocol was used. 25 µL of bead slurry and 200 µL of extract sample were rotated for
746 two hours at 4°C. After the fifth wash with binding buffer, protein complexes were eluted
747 from the anti-FLAG resin by adding 50 µL of 2X SDS loading buffer and boiled at 95°C
748 for five minutes. For Western blots, input samples were generated by adding equal
749 volume of 2X SDS loading buffer to nuclear extract and 1/10 of the immunoprecipitation
750 was loaded as estimated by protein mass.
751
752 **Mass spectrometry sample digestion.** The samples were prepared in a similar manner
753 as described previously [72]. Briefly, the agarose bead-bound proteins were washed several
754 times with 50 mM Triethylammonium bicarbonate (TEAB) pH 7.1, before being solubilized
755 with 40 µL of 5% SDS, 50 mM TEAB, pH 7.55 followed by a room temperature incubation
756 for 30 minutes. The supernatant containing the proteins of interest was then transferred
757 to a new tube, reduced by making the solution 10 mM Tris(2-carboxyethyl)phosphine
758 (TCEP) (Thermo, #77720), and further incubated at 65°C for 10 minutes. The sample was
759 then cooled to room temperature and 1 µL of 1M iodoacetamide acid was added and
760 allowed to react for 20 minutes in the dark. Then, 5 µL of 12% phosphoric acid was added
761 to the 50 µL protein solution followed by 350 µL of binding buffer (90% Methanol, 100 mM
762 TEAB final; pH 7.1). The resulting solution was administered to an S-Trap spin column
763 (Protifi, Farmingdale NY) and passed through the column using a bench top centrifuge
764 (30 second spin at 4,000g). The spin column was then washed three times with 400 µL
765 of binding buffer and centrifuged (1200 rpm, 1 min). Trypsin (Promega, #V5280) was then
766 added to the protein mixture in a ratio of 1:25 in 50 mM TEAB, pH=8, and incubated at
767 37°C for 4 hours. Peptides were eluted with 80 µL of 50 mM TEAB, followed by 80 µL of
768 0.2% formic acid, and finally 80 µL of 50% acetonitrile, 0.2% formic acid. The combined

769 peptide solution was then dried in a speed vacuum (room temperature, 1.5 hours) and
770 resuspended in 2% acetonitrile, 0.1% formic acid, 97.9% water and aliquoted into an
771 autosampler vial.

772 **NanoLC MS/MS Analysis**. Peptide mixtures were analyzed by nanoflow liquid
773 chromatography-tandem mass spectrometry (nanoLC-MS/MS) using a nano-LC
774 chromatography system (UltiMate 3000 RSLCnano, Dionex, Thermo Fisher Scientific,
775 San Jose, CA). The nano-LC-MS/MS system was coupled on-line to a Thermo Orbitrap
776 Fusion mass spectrometer (Thermo Fisher Scientific, San Jose, CA) through a nanospray
777 ion source (Thermo Scientific). A trap and elute method was used to desalt and
778 concentrate the sample, while preserving the analytical column. The trap column (Thermo
779 Scientific) was a C18 PepMap100 (300 µm X 5 mm, 5 µm particle size) while the analytical
780 column was an Acclaim PepMap 100 (75 µm X 25 cm) (Thermo Scientific). After
781 equilibrating the column in 98% solvent A (0.1% formic acid in water) and 2% solvent B
782 (0.1% formic acid in acetonitrile (ACN)), the samples (2 µL in solvent A) were injected
783 onto the trap column and subsequently eluted (400 nL/min) by gradient elution onto the
784 C18 column as follows: isocratic at 2% B, 0-5 min; 2% to 32% B, 5-39 min; 32% to 70%
785 B, 39-49 min; 70% to 90% B, 49-50 min; isocratic at 90% B, 50-54 min; 90% to 2%, 54-
786 55 min; and isocratic at 2% B, until the 65 minute mark.

787 All LC-MS/MS data were acquired using XCalibur, version 2.1.0 (Thermo Fisher
788 Scientific) in positive ion mode using a top speed data-dependent acquisition (DDA)
789 method with a 3 second cycle time. The survey scans (m/z 350-1500) were acquired in
790 the Orbitrap at 120,000 resolution (at m/z = 400) in profile mode, with a maximum injection
791 time of 100 m s and an AGC target of 400,000 ions. The S-lens RF level was set to 60.
792 Isolation was performed in the quadrupole with a 1.6 Da isolation window, and CID
793 MS/MS acquisition was performed in profile mode using rapid scan rate with detection in
794 the ion-trap using the following settings: parent threshold = 5,000; collision energy = 32%;
795 maximum injection time 56 msec; AGC target 500,000 ions. Monoisotopic precursor
796 selection (MIPS) and charge state filtering were on, with charge states 2-6 included.
797 Dynamic exclusion was used to remove selected precursor ions, with a +/- 10 ppm mass
798 tolerance, for 15 seconds after acquisition of one MS/MS spectrum.
799

800 **Database Searching**. Tandem mass spectra were extracted and charge state
801 deconvoluted using Proteome Discoverer (Thermo Fisher, version 2.2.0388). Deisotoping
802 was not performed. All MS/MS spectra were searched against the appropriate database,
803 either Uniprot *Drosophila* database (version 04-04-2018) or Uniprot Human database
804 (reviewed 06272018), using Sequest. Searches were performed with a parent ion
805 tolerance of 5 ppm and a fragment ion tolerance of 0.60 Da. Trypsin was specified as the
806 enzyme, allowing for two missed cleavages. Fixed modification of carbamidomethyl (C)
807 and variable modifications of oxidation (M) and deamidation were specified in Sequest.
808 Heat maps in Figure 4 were made using Morpheus from the Broad Institute,
809 https://software.broadinstitute.org/morpheus.
810
811 **RNA interference:** siRNA used were Universal Negative Control #2 (Sigma Aldrich,
812 #SIC002), INTS13-1 (Sigma Aldrich, SASI_Hs01_00134940/ASUN), and INTS13-2
813 (Sigma Aldrich, SASI_Hs01_00134941/ASUN). RPE cells were seeded at 50,000
814 cells/well in a 24-well plate on Day 1. On Day 2, 3 µL of 20 µM siRNA was added to 50
815 µL of Opti-MEM I per well and incubated at room temperature for 5 minutes. 2 µL of
816 Lipofectamine RNAiMax (Invitrogen, 13778-150) was also added to 50 µL of Opti-MEM
817 per well and incubated at room temperature for 5 minutes. The siRNA and RNAi Max
818 solutions were combined and incubated at room temperature for 20 minutes, then added
819 to the cells. Cells were expanded to a 6-well plate on Day 3. On Day 4, cells were washed
820 gently with DPBS and 900 µL of DMEM/F-12 media was added to each well. The siRNA
821 treatment from Day 2 was repeated. Cells were further incubated and then harvested on
822 Day 6. For RNA extraction, cells were harvested with 500 µL of Trizol. For protein
823 extraction, cells were harvested with SDS loading buffer detailed previously.
824 A modified version of the above protocol was followed for cells intended for
825 immunofluorescence staining of primary cilia: RPE cells were seeded onto three cover
826 slips in each well of the 6-well plate on Day 3. On Day 4, cells were washed twice with
827 DPBS and grown in serum free (0% FBS) DMEM/F-12 + 1% (v/v) penicillin-streptomycin
828 for 48 hours before fixation.
829

830 **Library preparation and RNA sequencing:** RNAi-treated RPE cells were harvested in
831 500 µL of Trizol, and total RNA was purified by phenol-chloroform extraction. 1 µg total
832 RNA per sample was used for sequencing library preparation. The NEBNext Poly(A)
833 mRNA magnetic isolation kit (New England Biolabs, Ipswich MA, E7490) was used to
834 purify mRNA from the total RNA. The Click-Seq protocol was used following this step[73].
835 RT-PCR to make cDNA was done with 1:35 azido-NTP:dNTP for random strand
836 termination by SuperScript III reverse transcriptase. Following Click-Seq preparations,
837 libraries were PCR amplified for 17 cycles and products from 400-600 bp were collected
838 for sequencing. Samples were sequenced on a NextSeq500 (Illumina) using a single read
839 75bp run. Sequencing adapters removed and reads with a Qscore<30 were discarded
840 using fastp (ver 0.14.1). Reads were aligned using Hisat2 (ver 2.1.0) to the hg38 genome
841 using default parameters. Differential expression changes were determined using
842 DESeq2 (ver 1.23.10) and assigned using the comprehensive gene annotation in
843 GENCODE32 and tests for significance were using the nbinomialWald test. Alignment
844 information and lowest correlation between samples in each data set is given in table
845 below.

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

846

847

848 **ACKNOWLEDGMENTS**

849 We thank all families for partaking in this study. The authors would also like to thank
850 members of their team for technical assistance and fruitful discussions. This work was
851 supported by a grant from the French Ministry of Health to J.A., Strategic Positioning Fund
852 on Genetic Orphan Diseases from the Biomedical Research Council, A*STAR, Singapore
853 to B.R., and by the UCLA California Center for Rare Diseases to H.L., S.F.N. B.R. is a
854 fellow of the Branco Weiss Foundation, an A*STAR Investigator and Young EMBO
855 Investigator. This work was supported by National Institutes of Health grant R01-

GM134539 (E.J.W.), and the Welch Foundation grant H-1889 to The University of Texas Medical Branch at Galveston (E.J.W.). The UTMB Mass Spectrometry Facility is supported in part by The Cancer Prevention Research Institute of Texas (CPRIT) grant number RP190682 (W.K.R). M.K.S. was supported by CPRIT grant number P71709-B01.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests

AUTHOR CONTRIBUTIONS

L.G.M. and M.S. 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. 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.S. performed all experiments on patients' primary cells and in *Xenopus* with the help of N.E.B. 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.X. and L.T. conducted recombinant protein production and gel filtration. M.K.S. and L.J.K. imaged and analyzed PC immunofluorescence. W.K.R. prepared proteomic samples and analyzed LC/MS data. L.G.M., M.S., B.R., and E.J.W. wrote the manuscript. All authors critically reviewed the manuscript.

WEB RESOURCES

The URLs for data presented herein are as follows:

BLAST, http://blast.ncbi.nlm.nih.gov/Blast.cgi

dbSNP146, http://www.ncbi.nlm.nih.gov/snp

1000 Genomes, http://www.1000genomes.org/

Online Mendelian Inheritance in Man (OMIM), http://www/omim.org/

883 ExAC, http://exac.broadinstitute.org/

884 ANNOVAR, http://www.openbioinformatics.org/annovar/

885 PolyPhen-2, http://genetics.bwh.harvard.edu/pph2/

886 Human Splicing Finder, http://www.umd.be/HSF3/HSF.html

887 ClinVar, ftp://ftp.ncbi.nlm.nih.gov/pub/clinvar/vcf_GRCh37/

888 Exome Sequencing Project, http://evs.gs.washington.edu/EVS/

889

890

891 **FIGURE LEGENDS**

892

893 **Fig 1. Identification of two families segregating recessive mutations in *INTS13*. a**
894 Pedigree of Family 1 (top): affected individuals (II.4 and II.5) were born to first-cousin
895 parents (I.1 and I.2) with three unaffected children (II.1-II.3). Double lines indicate a
896 consanguineous marriage. Pedigree of Family 2 (bottom): affected individuals (II.2 and
897 II.4) were born to parents with two unaffected children (II.1 and II.3). **b** Homozygosity
898 mapping of Family 1 delineated five candidate loci totaling 46 cM on chromosomes 6, 8,
899 12, 13, and 16. **c** Schematic representation of human *INTS13* located on chromosome
900 12 (ch12p13.2-p11.22) consisting of 17 exons. Locus capture followed by massive
901 parallel sequencing identified a homozygous frameshift mutation as a disease-causing
902 mutation in Family 1. The c.2004delA (p.K668Nfs*9) mutation is caused by a single base
903 pair deletion in exon 16 (marked in brown), resulting in a frame shift and premature
904 termination codon (PTC) which alters nine amino acids (marked in red) and deletes the
905 last 31. In Family 2, the c.1955C>T (S652L) mutation is caused by a single base pair
906 substitution in exon 16 (marked in green). **d** Amino acid alignment of the C-terminus of
907 the INTS13 protein shows that both mutations occur at highly conserved residues. Red
908 and yellow shading indicates regions of conservation.

909 **Fig 2. Characterization of Loss-of-function INTS13 mutations in patient cells. a** RT-
910 qPCR shows that endogenous *INTS13* transcript levels were significantly reduced relative
911 to control cells in primary dermal fibroblasts of affected individual II.4 from Family 1. Data
912 are mean +/- SD, n = 3. Statistical significance was calculated using a two-tailed unpaired
913 t-test. **b** Primary dermal fibroblasts from individual II.4 from Family 1 were treated with
914 100 g/ml cycloheximide (CHX). Following treatment, RT-qPCR was done which shows
915 that the *INTS13* mRNA level was increased in treated cells. Data are mean +/- SEM, n =
916 3. Statistical significance was calculated using a one-tailed Student's t-test. **c** Western
917 blot analysis of primary dermal fibroblasts of affected individual II.4 from Family 1 using
918 two polyclonal antibodies designed against distinct regions of INTS13. No signal was
919 detected in patient cells with the antibody directed against the C-terminus of INTS13 (C).
920 A truncated protein was detected in patient cells by the antibody directed against the
921 center of the protein (M). Actin was used as a loading control. **d** RT-qPCR shows that
922 *INTS13* transcript levels were significantly reduced in primary dermal fibroblast cells of
923 affected individual II.4 from Family 2 relative to control cells. Data are mean +/- SD, n =
924 3. Statistical significance was calculated using a two-tailed unpaired t-test. **e** Western blot
925 analysis of primary dermal fibroblasts of affected individual II.4 from Family 2 shows
926 reduced levels of INTS13 protein compared to control cells. **f** Western blot analysis of
927 primary dermal fibroblasts from individual II.4 from Family 2 treated with MG132. Treated
928 cells show a level of INTS13 protein comparable to control cells. * indicates a non-specific
929 signal. **g** Multiciliated airway cells were obtained from nasal biopsies of three members
930 of Family 1: the unaffected carrier mother (Con.) and the two affected children (II.4 and
931 II.5). Probing for acetylated α-tubulin shows significantly shorter, less dense, and
932 disorganized cilia in the affected individuals compared to the mother's cells. **h** Bar graph
933 representing the average length of nasal cilia. Data are mean +/- SEM, n = 200 cells for
934 each individual. Statistical significance was calculated using a one-tailed t-test. *, **, ***,
935 **** correspond to p-values < 0.05, 0.01, 0.001, and 0.0001, respectively.

936

937 **Fig 3. INTS13 associates with the INTS4/9/11 heterotrimer. a** Yeast two hybrid
938 screening for interactors of INTS13. Haploid strains of yeast expressing each individual

939 Integrator subunit fused to the Gal4 activating domain (AD) were mated with strains
940 expressing INTS13 fused to the DNA binding domain (BD). The first and third panels
941 show mated yeast grown on selective media as a positive control (media lacking leucine,
942 tryptophan, and uracil). The remaining panels are of yeast plated on restrictive media
943 additionally lacking histidine to test for protein-protein interactions and with 3AT (3-amino-
944 1,2,4-triazole) to test the strength of the interaction. No interaction was observed between
945 INTS13 and any other individual subunit (second panel). When INTS9 and INTS11 were
946 expressed in *trans* with INTS13, there was a positive interaction with INTS4 (fourth panel).
947 **b** Point mutants of INTS4, INTS9, and INTS11 known to disrupt their interactions were
948 used in yeast two hybrid to demonstrate that the formation of the INTS4/9/11 heterotrimer
949 is necessary for INTS13 association. General locations of the mutations are indicated
950 with arrows on the schematic. INTS13 is fused to the activating domain in the rest of the
951 yeast two hybrid assays shown in this study. **c** SDS PAGE gel of fractions from gel
952 filtration column of co-expressed full-length INTS4, INTS9, INTS11 and INTS13. The
953 sample was first purified by nickel affinity chromatography. INTS4 carried an N-terminal
954 His-tag, while the other three proteins were untagged. The positions of the fractions are
955 indicated at the top of the gel. **d** Consecutive 200-amino acid truncations of INTS13 were
956 tested for interaction with INTS4/9/11. In modified yeast two-hybrid assays, the only
957 positive interactions seen are full length INTS13 and the last ~200 residues, 501-706.
958 The patient mutations are marked on the schematic and occur in this region. Loss of
959 interaction with INTS4/9/11 was observed using the mutant versions of INTS13.

960

961 **Fig 4. OFD2 mutations lead to conserved disruption of interactions between the**
962 **INTS10/13/14 module and the remainder of the Integrator complex**. **a**
963 Immunoprecipitation of FLAG-INTS13 from nuclear extract derived from HEK-293T cells
964 stably expressing FLAG epitope tagged human INTS13-wild type (wt), Family 1 INTS13
965 mutant (F1), or Family 2 INTS13 mutant (F2). Nuclear extract derived from naïve HEK-
966 293T cells not expressing any FLAG epitope was used as a control. Left panel: IP-LC/MS
967 was done for each sample, and quantitation is shown as a heat map. The color scale
968 shown denotes the normalized spectral counts for each protein (see methods). Right

969 panel: Western blot analysis of either nuclear extract input or anti-FLAG
970 immunoprecipitation using human Integrator antibodies. **b** Results shown are from
971 parallel experiments conducted in a manner identical with panel (a) with the exception
972 that nuclear extract was derived from *Drosophila* S2 cells expressing FLAG-dINTS13-wt
973 or FLAG-tagged dINTS13 proteins where patient mutations were introduced into
974 homologous regions of the *Drosophila INTS13* cDNA. **c** Modified yeast two hybrid
975 analysis where human INTS14 is expressed as a Gal4-DNA binding domain fusion,
976 human INTS13 wt or mutant is expressed as a Gal4-Activation domain fusion, and human
977 INTS10 is expressed in *trans*. Dilutions of yeast are plated on both permissive (left) and
978 restrictive (right) media. Empty activation domain vector is used as a negative control. **d**
979 Schematic of the Integrator CM associating with the remainder of Integrator subunits that
980 is bridged to the INTS10/13/14 module through the C-terminus of INTS13, which is
981 disrupted by OFD2 mutations.

982

983 **Fig 5. INTS13 depletion leads to a loss of primary cilia and broad transcriptional**
984 **perturbation of ciliary genes. a** Western blot of whole cell lysates from RPE cells treated
985 with non-targeting control siRNA (Con.) or either of two siRNA targeting *INTS13* (*INTS13-*
986 *1*, *INTS13-2*). Lysates were probed for either INTS13 or tubulin as a loading control. **b**
987 Immunofluorescence imaging of RPE cells treated with siRNAs as in panel (a) to visualize
988 primary cilia (PC). DAPI staining in blue indicates the nuclei, γ-tubulin staining is shown
989 in red to visualize the basal body, and ARL13B staining is shown in green to visualize the
990 ciliary axoneme. Representative images show disrupted primary cilia that occurred when
991 INTS13 was targeted by RNAi. **c** Images from panel (b) were quantified for each condition
992 and graphed as a percentage of the number of control-treated cells with PC (data are
993 mean +/- SD, con. n = 3785, 13-1 n = 2969, 13-2 n = 2675. Raw calculations for each
994 condition were divided by the control mean to adjust control to 1). Statistical significance
995 was calculated using a one-way ANOVA with multiple comparisons. **d** RNA sequencing
996 was done on siRNA treated RPE cells in triplicate. Genes with significant changes in
997 expression are shown by color indicating the log2 fold change. **e** Results from panel (d)
998 were used for gene ontology analysis: the top five significant results are shown for

999 upregulated and downregulated genes. All five GO terms for upregulated genes are
1000 related to cilia. **f** Quantitative RT-PCR validating three ciliary genes with expression
1001 changes identified using RNA-seq. Results are quantified from independent biological
1002 replicates (n = 3, data are mean +/- SD). Statistical significance was calculated using
1003 multiple Student t-tests with a Bonferroni correction. *, **, ***, **** correspond to p-values
1004 < 0.05, 0.01, 0.001, and 0.0001, respectively.

1005 **Fig 6. INTS13 is required for embryonic development and ciliogenesis in *Xenopus***
1006 ***laevis*. a** In comparison to control embryos, morpholino (MO)-mediated knock down of
1007 INTS13 in *Xenopus* leads to severe developmental defects including a small head, big
1008 belly, short and down curved body axes, cardiac edema and massive developmental
1009 delay. **b** In situ hybridization using a specific probe for the multi-ciliated cell (MCC) marker
1010 *ccdc9* shows a reduction in *ccdc9* expression in MCC of morphants compared to control
1011 embryos. **c** Whole mount antibody staining for acetylated α-tubulin shows a punctate
1012 pattern similar to *ccdc9*. The puncta have a more condensed and stronger signal in control
1013 embryos. A higher magnification view of the MCC shows marked reduction in cilia
1014 compared to controls. **d** Western blot showing that the *INTS13* MO was able to deplete
1015 INTS13 protein in *Xenopus* embryos. **e** Low magnification SEM imaging of *Xenopus*
1016 MCC at embryonic stage 28. MCCs of a control embryo possess multiple, long cilia,
1017 whereas the morphant MCCs display fewer cilia per cell which appear to be disorganized.
1018 **f** High magnification SEM imaging of *Xenopus* MCC at embryonic stage 28. Ciliary length
1019 seems unperturbed in morphant embryos, while the average number of cilia on individual
1020 MCCs is clearly reduced by INTS13 knockdown. The average MCC in morphants have
1021 less than half of the number of cilia per cell compared to control, as shown in the bar
1022 graph. **g** The ultrastructure of the cilia in morphants appears to show a normal 9+2
1023 structure as imaged by TEM. *, **, ***, **** correspond to p-values < 0.05, 0.01, 0.001,
1024 and 0.0001, respectively.

1025

1026 **SUPPLEMENTAL FIGURE LEGENDS**

1027 **Supplemental Fig 1. Presentation of OFD2 Syndrome patients from Family 1 and**
1028 **Family 2. a-d, m** Photographs of affected individual II.4 of Family 1 with common

1029 phenotype and orofacial dysmorphisms at different ages including: **a-c** Bilateral cleft lip
1030 and palate (corrected after multiple surgical operations), hypertelorism, broad nasal
1031 bridge, flat philtrum, dental abnormalities, low set ear and rough and sparse hair. **d**
1032 Crowded optic disc. **m** Clinodactyly of toes. **e-l** Photographs of affected individual II.5 of
1033 Family 1. **e-g** Bilateral cleft lip and palate, hypertelorism, broad nasal bridge, flat philtrum,
1034 dental abnormalities. **h** Crowded optic disc. **i** Low set ear and rough and sparse hair. **j, h**
1035 Broad and mild brachydactyly and single palmar creases. **l** Clinodactyly of toes. **n**
1036 Pedigree of Family I. The genotypes for the corresponding mutation are indicated below
1037 each individual. The c.2004delA mutation segregated with the disease in this family. (+/-
1038) denotes the heterozygous alleles, and (-/-) denotes the homozygous mutant alleles. **o-**
1039 **v** Photographs of affected individuals II.2 and II.4 of Family 2 at 27 and 20 years old,
1040 respectively. **w** Pedigree of Family 2. The genotypes for the corresponding mutation are
1041 shown below each individual which confirm the segregation of the c.1955C>T mutation
1042 in this family. (+/+) denotes the homozygous wildtype alleles.

1043 **Supplemental Fig 2. The C-terminus of INTS13 interacts with the Integrator**
1044 **Cleavage Module. a** Expanded modified yeast two hybrid assay from Fig 2a. Included
1045 here are tests where INTS11 or INTS9 were expressed in *trans* along with INTS13-
1046 binding domain and crossed with each Integrator subunit. No positive interactions were
1047 seen. **b** Left: Modified yeast two hybrid with INTS9 fused to the binding domain and
1048 INTS11 expressed in *trans* mated with yeast expressing each Integrator subunit fused to
1049 the activating domain to reconfirm that INTS4 interacts with the INTS9/11 heterodimer.
1050 Right: The same experiment with the addition of INTS4 expressed in *trans* to show that
1051 INTS13 interacts with the INTS4/9/11 heterotrimer. **c** Smaller portions (approximately 20
1052 amino acids) of the INTS13 protein were removed from the necessary region for
1053 interaction shown in Fig 2d to determine that residues 577-706 were sufficient to interact
1054 with INTS4/9/11.
1055

1056 **Supplemental Fig 3. INTS13 interacts with INTS10/14. a** Partial protein alignment of
1057 the C-terminus of human and *Drosophila* INTS13 to show the orthologous mutations
1058 made to recapitulate the patient mutations in *Drosophila* cells. The location and altered
1059 residues are marked in brown for Family 1 and in green for Family 2. **b** Modified yeast

1060 two hybrid testing the interaction between INTS10/13/14. Left: INTS14 fused to the
1061 binding domain was crossed with each Integrator subunit fused to the activating domain,
1062 and no positive interaction was seen on selective media. Right: INTS13 fused to the
1063 activating domain was crossed with yeast strains expressing INTS14-binding domain with
1064 each Integrator subunit in *trans*. A positive interaction was seen when INTS10 was
1065 expressed in *trans*.
1066
1067 **Supplemental Fig 4. Significant overlap in gene expression changes between two**
1068 ***INTS13*-targeting siRNAs. a** Venn diagrams show the number of significantly
1069 upregulated or **b** downregulated genes for siRNA 13-1 and 13-2, and the genes in
1070 common between the two siRNAs as determined by RNA-seq.
1071
1072 **Supplemental Fig 5. Expression of *ints13* during *Xenopus* development. a** Detailed
1073 expression pattern of *ints13* in *Xenopus* embryos analyzed by WISH and whole mount
1074 immunostaining. *ints13* has both maternal and zygotic contribution. At 4 cell stage, *ints13*
1075 transcripts are localized within the animal half. At stage 10, its expression covers the
1076 whole embryo except the blastopore. At stage 17, *ints13* is observed in neural tube and
1077 neural folds. **b** Lateral view, anterior is right stage 22 *ints13* starts to be enriched in brain,
1078 branchial arches (ba), pronephros (pn), eye vesicles (ev), ear (e), skin cells (sk) and
1079 somites (so). Lateral view, anterior is right of the embryo at stage 28, *ints13* expression
1080 is also detected in ear (e), skin cells (sk) and pronephros (pn). Correlation of transcripts
1081 localization of ints13 protein at stage 28 by wholemount embryo antibody staining using
1082 a custom polyclonal INTS13 (α-M) antibody, Lateral view, anterior is left.
1083
1084 **Table 1. Clinical phenotypes of the affected individuals.** (n.d) not determined, (+)
1085 affirmative, (-) negative, (M) male and (F) female.
1086
1087 **Supplemental Table 1. RNA-seq differential expression analysis results.**
1088
1089
1090 **Supplemental Table 2. Primer sequences.**

1091

1092

1093 ## References

1094

1095 1. Lee JE, Gleeson JG. A systems-biology approach to understanding the ciliopathy disorders.
1096 *Genome Med* **3**, 59 (2011).

1097
1098 2. Lee JE, Gleeson JG. Cilia in the nervous system: linking cilia function and neurodevelopmental
1099 disorders. *Curr Opin Neurol* **24**, 98-105 (2011).

1100
1101 3. Raidt J*, et al.* Ciliary beat pattern and frequency in genetic variants of primary ciliary dyskinesia.
1102 *Eur Respir J* **44**, 1579-1588 (2014).

1103
1104 4. Waters AM, Beales PL. Ciliopathies: an expanding disease spectrum. *Pediatr Nephrol* **26**, 1039-
1105 1056 (2011).

1106
1107 5. Arts HH, Knoers NV. Current insights into renal ciliopathies: what can genetics teach us? *Pediatr*
1108 *Nephrol* **28**, 863-874 (2013).

1109
1110 6. Estrada-Cuzcano A, Roepman R, Cremers FP, den Hollander AI, Mans DA. Non-syndromic retinal
1111 ciliopathies: translating gene discovery into therapy. *Hum Mol Genet* **21**, R111-124 (2012).

1112
1113 7. Forsythe E, Beales PL. Bardet-Biedl syndrome. *Eur J Hum Genet* **21**, 8-13 (2013).

1114
1115 8. Alexiev BA, Lin X, Sun CC, Brenner DS. Meckel-Gruber syndrome: pathologic manifestations,
1116 minimal diagnostic criteria, and differential diagnosis. *Arch Pathol Lab Med* **130**, 1236-1238
1117 (2006).

1118
1119 9. Badano JL, Mitsuma N, Beales PL, Katsanis N. The ciliopathies: an emerging class of human
1120 genetic disorders. *Annu Rev Genomics Hum Genet* **7**, 125-148 (2006).

1121
1122 10. Baker K, Beales PL. Making sense of cilia in disease: the human ciliopathies. *Am J Med Genet C*
1123 *Semin Med Genet* **151C**, 281-295 (2009).

1124
1125 11. Ferkol TW, Leigh MW. Ciliopathies: the central role of cilia in a spectrum of pediatric disorders. *J*
1126 *Pediatr* **160**, 366-371 (2012).

1127
1128 12. Davis EE, Katsanis N. The ciliopathies: a transitional model into systems biology of human
1129 genetic disease. *Curr Opin Genet Dev* **22**, 290-303 (2012).

1130
1131 13. Leigh MW*, et al.* Clinical and genetic aspects of primary ciliary dyskinesia/Kartagener syndrome.
1132 *Genet Med* **11**, 473-487 (2009).

1133
1134 14. Thauvin-Robinet C*, et al.* The oral-facial-digital syndrome gene C2CD3 encodes a positive
1135 regulator of centriole elongation. *Nat Genet* **46**, 905-911 (2014).

1136
1137 15. Gurrieri F, Franco B, Toriello H, Neri G. Oral-facial-digital syndromes: review and diagnostic
1138 guidelines. *Am J Med Genet A* **143A**, 3314-3323 (2007).

1139
1140 16. Ferrante MI*, et al.* Convergent extension movements and ciliary function are mediated by ofd1,
1141 a zebrafish orthologue of the human oral-facial-digital type 1 syndrome gene. *Hum Mol Genet*
1142 **18**, 289-303 (2009).

1143
1144 17. Thauvin-Robinet C*, et al.* OFD1 mutations in males: phenotypic spectrum and ciliary basal body
1145 docking impairment. *Clin Genet* **84**, 86-90 (2013).

1146
1147 18. Thomas S*, et al.* TCTN3 mutations cause Mohr-Majewski syndrome. *Am J Hum Genet* **91**, 372-
1148 378 (2012).

1149
1150 19. Shamseldin HE*, et al.* Mutations in DDX59 implicate RNA helicase in the pathogenesis of
1151 orofaciodigital syndrome. *Am J Hum Genet* **93**, 555-560 (2013).

1152
1153 20. Adly N, Alhashem A, Ammari A, Alkuraya FS. Ciliary genes TBC1D32/C6orf170 and SCLT1 are
1154 mutated in patients with OFD type IX. *Hum Mutat* **35**, 36-40 (2014).

1155
1156 21. Franco B, Thauvin-Robinet C. Update on oral-facial-digital syndromes (OFDS). *Cilia* **5**, 12 (2016).

1157
1158 22. Lopez E*, et al.* C5orf42 is the major gene responsible for OFD syndrome type VI. *Hum Genet* **133**,
1159 367-377 (2014).

1160
1161 23. Ferrante MI*, et al.* Identification of the gene for oral-facial-digital type I syndrome. *Am J Hum*
1162 *Genet* **68**, 569-576 (2001).

1163
1164 24. Roberson EC*, et al.* TMEM231, mutated in orofaciodigital and Meckel syndromes, organizes the
1165 ciliary transition zone. *J Cell Biol* **209**, 129-142 (2015).

1166
1167 25. Baillat D, Hakimi MA, Naar AM, Shilatifard A, Cooch N, Shiekhattar R. Integrator, a multiprotein
1168 mediator of small nuclear RNA processing, associates with the C-terminal repeat of RNA
1169 polymerase II. *Cell* **123**, 265-276 (2005).

1170
1171 26. Baillat D, Wagner EJ. Integrator: surprisingly diverse functions in gene expression. *Trends*
1172 *Biochem Sci* **40**, 257-264 (2015).

1173
1174 27. Gardini A*, et al.* Integrator Regulates Transcriptional Initiation and Pause Release following
1175 Activation. *Mol Cell* **56**, 128-139 (2014).

1176
1177 28. Skaar JR*, et al.* The Integrator complex controls the termination of transcription at diverse
1178 classes of gene targets. *Cell Res* **25**, 288-305 (2015).

1179
1180 29. Stadelmayer B*, et al.* Integrator complex regulates NELF-mediated RNA polymerase II
1181 pause/release and processivity at coding genes. *Nat Commun* **5**, 5531 (2014).

1182
1183 30. Albrecht TR, Wagner EJ. snRNA 3' end formation requires heterodimeric association of
1184 integrator subunits. *Mol Cell Biol* **32**, 1112-1123 (2012).

1185
1186 31. Callebaut I, Moshous D, Mornon JP, de Villartay JP. Metallo-beta-lactamase fold within nucleic
1187 acids processing enzymes: the beta-CASP family. *Nucleic Acids Res* **30**, 3592-3601 (2002).

1188
1189 32. Mandel CR*, et al.* Polyadenylation factor CPSF-73 is the pre-mRNA 3'-end-processing
1190 endonuclease. *Nature* **444**, 953-956 (2006).

1191
1192 33. Albrecht TR*, et al.* Integrator subunit 4 is a 'Symplekin-like' scaffold that associates with
1193 INTS9/11 to form the Integrator cleavage module. *Nucleic Acids Res* **46**, 4241-4255 (2018).

1194
1195 34. Ezzeddine N*, et al.* A subset of Drosophila integrator proteins is essential for efficient U7 snRNA
1196 and spliceosomal snRNA 3'-end formation. *Mol Cell Biol* **31**, 328-341 (2011).

1197
1198 35. Lai F, Gardini A, Zhang A, Shiekhattar R. Integrator mediates the biogenesis of enhancer RNAs.
1199 *Nature* **525**, 399-403 (2015).

1200
1201 36. Rubtsova MP*, et al.* Integrator is a key component of human telomerase RNA biogenesis. *Sci Rep*
1202 **9**, 1701 (2019).

1203

1204 37. Elrod ND*, et al.* The Integrator Complex Attenuates Promoter-Proximal Transcription at Protein-
1205 Coding Genes. *Mol Cell* **76**, 738-752 e737 (2019).

1206
1207 38. Tatomer DC*, et al.* The Integrator complex cleaves nascent mRNAs to attenuate transcription.
1208 *Genes Dev* **33**, 1525-1538 (2019).

1209
1210 39. Jodoin JN*, et al.* Nuclear-localized Asunder regulates cytoplasmic dynein localization via its role
1211 in the integrator complex. *Mol Biol Cell* **24**, 2954-2965 (2013).

1212
1213 40. Sorokin S. Centrioles and the formation of rudimentary cilia by fibroblasts and smooth muscle
1214 cells. *J Cell Biol* **15**, 363-377 (1962).

1215
1216 41. Jodoin JN*, et al.* The snRNA-processing complex, Integrator, is required for ciliogenesis and
1217 dynein recruitment to the nuclear envelope via distinct mechanisms. *Biol Open* **2**, 1390-1396
1218 (2013).

1219
1220 42. Breslow DK*, et al.* A CRISPR-based screen for Hedgehog signaling provides insights into ciliary
1221 function and ciliopathies. *Nat Genet* **50**, 460-471 (2018).

1222
1223 43. Sakai N, Nakakita N, Yamazaki Y, Ui K, Uchinuma E. Oral-facial-digital syndrome type II (Mohr
1224 syndrome): clinical and genetic manifestations. *J Craniofac Surg* **13**, 321-326 (2002).

1225
1226 44. Bonnard C*, et al.* Mutations in IRX5 impair craniofacial development and germ cell migration via
1227 SDF1. *Nat Genet* **44**, 709-713 (2012).

1228
1229 45. Duquette A, Brais B, Bouchard JP, Mathieu J. Clinical presentation and early evolution of spastic
1230 ataxia of Charlevoix-Saguenay. *Mov Disord* **28**, 2011-2014 (2013).

1231
1232 46. Synofzik M*, et al.* Autosomal recessive spastic ataxia of Charlevoix Saguenay (ARSACS):
1233 expanding the genetic, clinical and imaging spectrum. *Orphanet J Rare Dis* **8**, 41 (2013).

1234
1235 47. Lee LA*, et al.* Drosophila genome-scale screen for PAN GU kinase substrates identifies Mat89Bb
1236 as a cell cycle regulator. *Dev Cell* **8**, 435-442 (2005).

1237
1238 48. Stebbings L, Grimes BR, Bownes M. A testis-specifically expressed gene is embedded within a
1239 cluster of maternally expressed genes at 89B in Drosophila melanogaster. *Dev Genes Evol* **208**,
1240 523-530 (1998).

1241

1242 49. Anderson MA, Jodoin JN, Lee E, Hales KG, Hays TS, Lee LA. Asunder is a critical regulator of
1243 dynein-dynactin localization during Drosophila spermatogenesis. *Mol Biol Cell* **20**, 2709-2721
1244 (2009).

1245
1246 50. Chen J*, et al.* An RNAi screen identifies additional members of the Drosophila Integrator complex
1247 and a requirement for cyclin C/Cdk8 in snRNA 3'-end formation. *RNA* **18**, 2148-2156 (2012).

1248
1249 51. Boon M*, et al.* Primary ciliary dyskinesia: critical evaluation of clinical symptoms and diagnosis in
1250 patients with normal and abnormal ultrastructure. *Orphanet J Rare Dis* **9**, 11 (2014).

1251
1252 52. Erle DJ, Sheppard D. The cell biology of asthma. *J Cell Biol* **205**, 621-631 (2014).

1253
1254 53. Livraghi A, Randell SH. Cystic fibrosis and other respiratory diseases of impaired mucus
1255 clearance. *Toxicol Pathol* **35**, 116-129 (2007).

1256
1257 54. Turner J, Jones CE. Regulation of mucin expression in respiratory diseases. *Biochem Soc Trans*
1258 **37**, 877-881 (2009).

1259
1260 55. Wallmeier J*, et al.* Mutations in CCNO result in congenital mucociliary clearance disorder with
1261 reduced generation of multiple motile cilia. *Nat Genet* **46**, 646-651 (2014).

1262
1263 56. Sabath K, Staubli ML, Marti S, Leitner A, Moes M, Jonas S. INTS10-INTS13-INTS14 form a
1264 functional module of Integrator that binds nucleic acids and the cleavage module. *Nat Commun*
1265 **11**, 3422 (2020).

1266
1267 57. Wu Y, Albrecht TR, Baillat D, Wagner EJ, Tong L. Molecular basis for the interaction between
1268 Integrator subunits IntS9 and IntS11 and its functional importance. *Proc Natl Acad Sci U S A* **114**,
1269 4394-4399 (2017).

1270
1271 58. Chen J, Waltenspiel B, Warren WD, Wagner EJ. Functional analysis of the integrator subunit 12
1272 identifies a microdomain that mediates activation of the Drosophila integrator complex. *J Biol*
1273 *Chem* **288**, 4867-4877 (2013).

1274
1275 59. Blacque OE*, et al.* Functional genomics of the cilium, a sensory organelle. *Curr Biol* **15**, 935-941
1276 (2005).

1277
1278 60. Brooks ER, Wallingford JB. Multiciliated cells. *Curr Biol* **24**, R973-982 (2014).

1279

1280 61. Wallingford JB. Planar cell polarity signaling, cilia and polarized ciliary beating. *Curr Opin Cell Biol*
1281 **22**, 597-604 (2010).

1282
1283 62. Hayes JM*, et al.* Identification of novel ciliogenesis factors using a new in vivo model for
1284 mucociliary epithelial development. *Dev Biol* **312**, 115-130 (2007).

1285
1286 63. Barbieri E*, et al.* Targeted Enhancer Activation by a Subunit of the Integrator Complex. *Mol Cell*
1287 **71**, 103-116 e107 (2018).

1288
1289 64. Walker JR, Corpina RA, Goldberg J. Structure of the Ku heterodimer bound to DNA and its
1290 implications for double-strand break repair. *Nature* **412**, 607-614 (2001).

1291
1292 65. Oegema R*, et al.* Human mutations in integrator complex subunits link transcriptome integrity to
1293 brain development. *PLoS Genet* **13**, e1006809 (2017).

1294
1295 66. Del Viso F*, et al.* Congenital Heart Disease Genetics Uncovers Context-Dependent Organization
1296 and Function of Nucleoporins at Cilia. *Dev Cell* **38**, 478-492 (2016).

1297
1298 67. Fakhro KA*, et al.* Rare copy number variations in congenital heart disease patients identify
1299 unique genes in left-right patterning. *Proc Natl Acad Sci U S A* **108**, 2915-2920 (2011).

1300
1301 68. Schmidt KN, Kuhns S, Neuner A, Hub B, Zentgraf H, Pereira G. Cep164 mediates vesicular
1302 docking to the mother centriole during early steps of ciliogenesis. *J Cell Biol* **199**, 1083-1101
1303 (2012).

1304
1305 69. Galati DF, Sullivan KD, Pham AT, Espinosa JM, Pearson CG. Trisomy 21 Represses Cilia Formation
1306 and Function. *Dev Cell* **46**, 641-650 e646 (2018).

1307
1308 70. Homer N, Merriman B, Nelson SF. BFAST: an alignment tool for large scale genome
1309 resequencing. *PLoS One* **4**, e7767 (2009).

1310
1311 71. Fitzgerald DJ, Berger P, Schaffitzel C, Yamada K, Richmond TJ, Berger I. Protein complex
1312 expression by using multigene baculoviral vectors. *Nat Methods* **3**, 1021-1032 (2006).

1313
1314 72. Anderson AP, Luo X, Russell W, Yin YW. Oxidative damage diminishes mitochondrial DNA
1315 polymerase replication fidelity. *Nucleic Acids Res* **48**, 817-829 (2020).

1316
1317 73. Jaworski E, Routh A. ClickSeq: Replacing Fragmentation and Enzymatic Ligation with Click-
1318 Chemistry to Prevent Sequence Chimeras. *Methods Mol Biol* **1712**, 71-85 (2018).

1319

1320

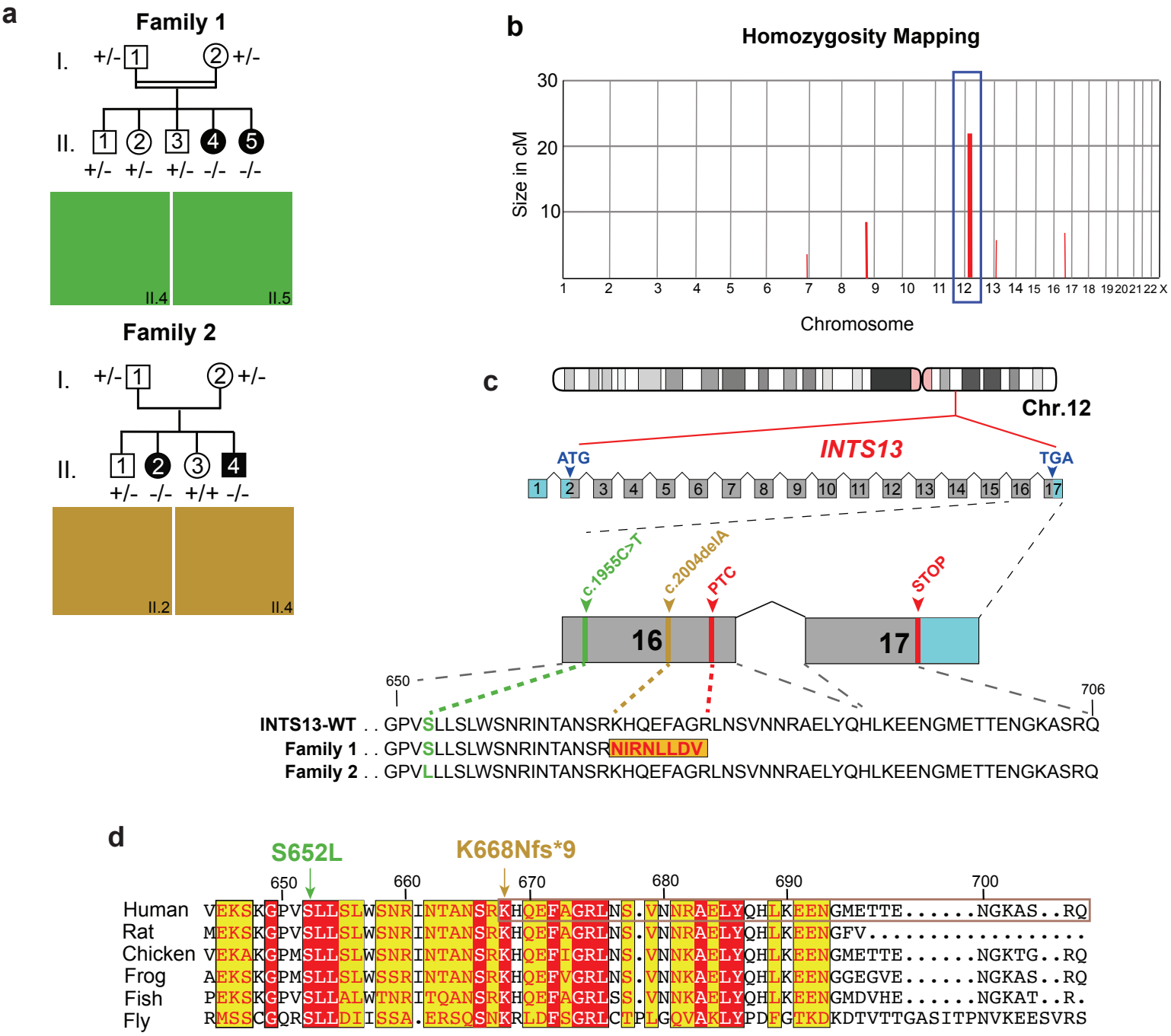

Fig 1. Identification of two families segregating recessive mutations in INTS13. **a** Pedigree of Family 1 (top): affected individuals (II.4 and II.5) were born to first-cousin parents (I.1 and I.2) with three unaffected children (II.1-II.3). Double lines indicate a consanguineous marriage. Pedigree 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 frame shift 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 (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.

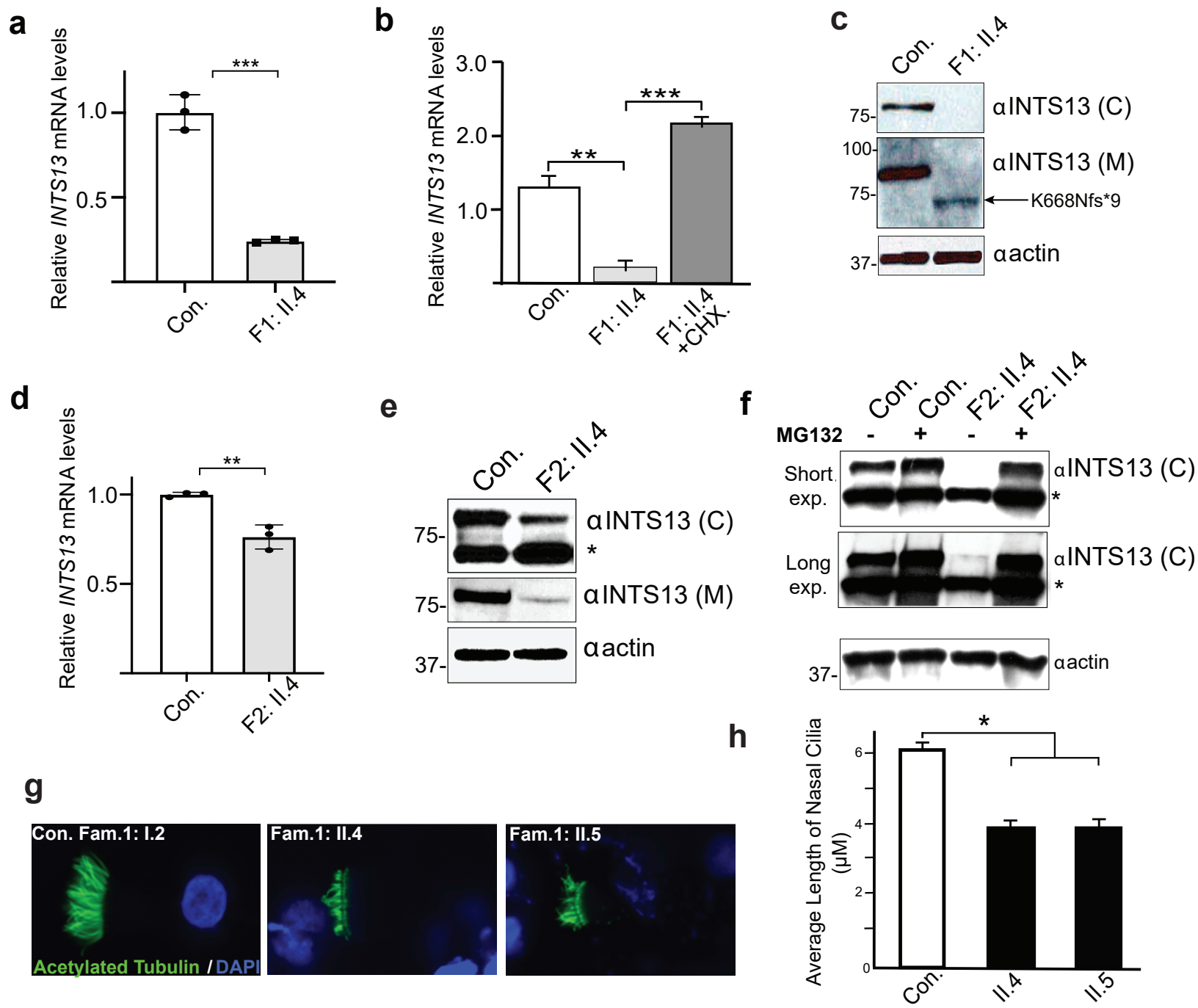

Fig 2. Characterization of Loss-of-function INTS13 mutations in patient cells. a RT-qPCR shows that endogenous INTS13 transcript levels were significantly reduced relative to control cells in primary dermal fibroblasts of affected individual II.4 from Family 1. Data are mean +/- SD, n=3, p<0.001, two-tailed unpaired t-test. **b** Primary dermal fibroblasts from individual II.4 from Family 1 were treated with 100 g/ml cycloheximide (CHX). Following treatment, RT-qPCR was done which shows that the INTS13 mRNA level was increased in treated cells. **c** Western blot analysis of primary dermal fibroblasts of affected individual II.4 from Family 1 using two polyclonal antibodies designed against distinct regions of INTS13. No signal was detected in patient cells with the antibody directed against the C-terminus of INTS13 (C). A truncated protein was detected in patient cells by the antibody directed against the center of the protein (M). Actin was used as a loading control. **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, p<0.01, two-tailed unpaired t-test. **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. **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. **h** Bar graph representing the average length of nasal cilia. Data are mean ± s.e.m., n ≥ 31 cells for each individual. *, **, ***, **** correspond to p-values < 0.05, 0.01, 0.001, and 0.0001, respectively.

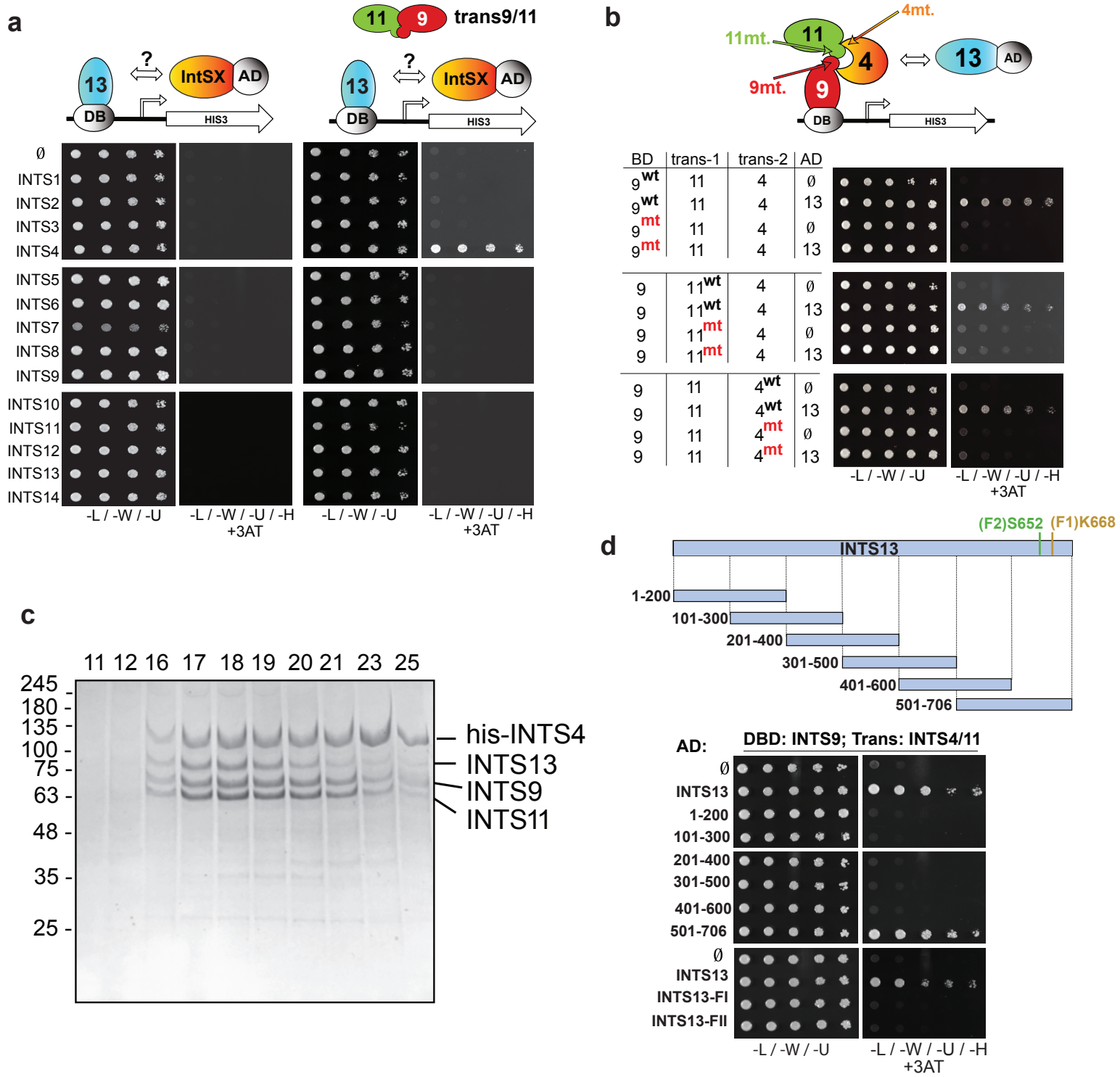

Fig 3. INTS13 associates with the INTS4/9/11 heterotrimer. a Yeast two hybrid screening for interactors of INTS13. Haploid strains of yeast expressing each individual Integrator subunit fused to the Gal4 activating domain (AD) were mated with strains expressing INTS13 fused to the DNA binding domain (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. **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.

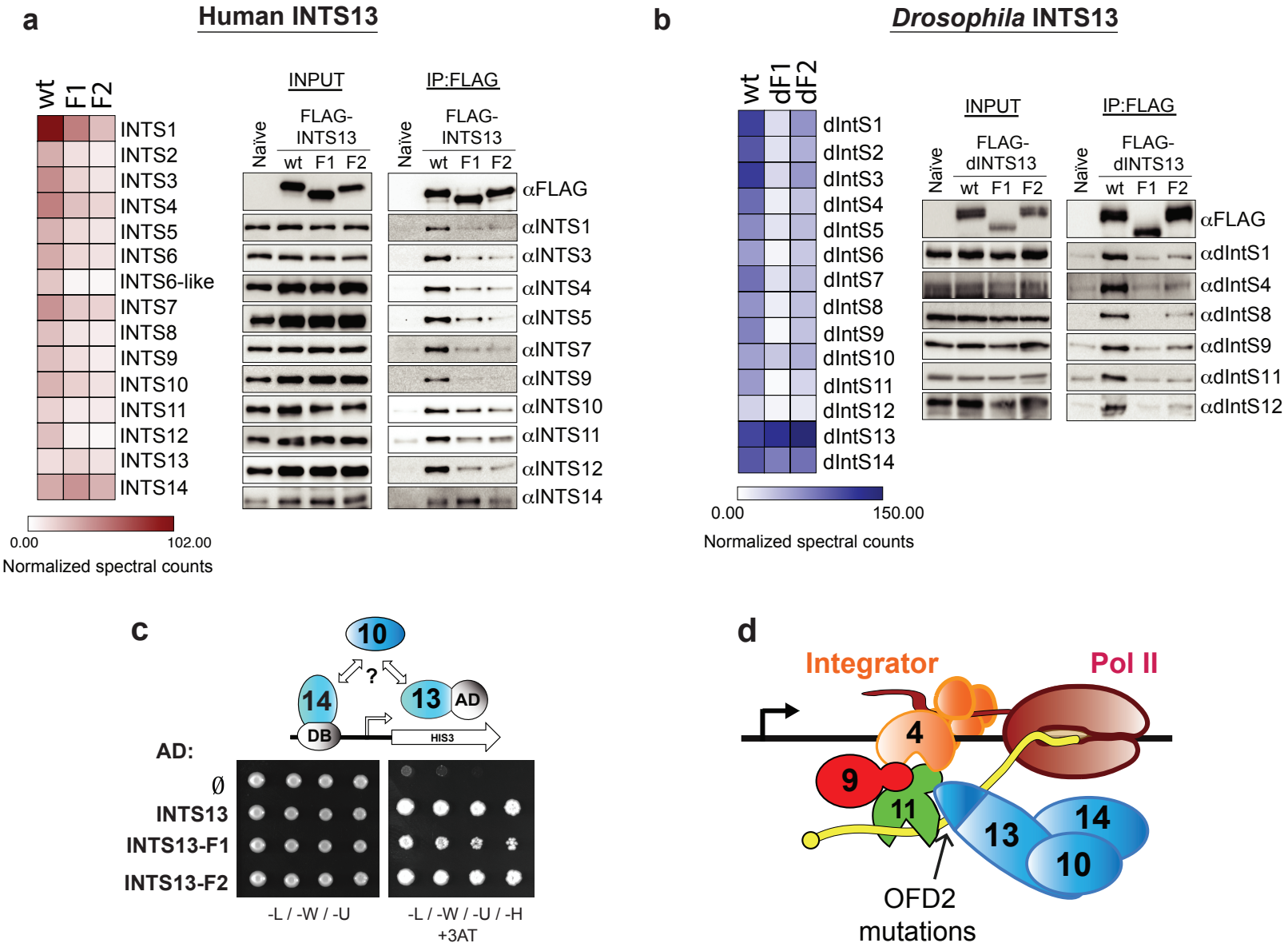

Fig 4. OFD2 mutations lead to conserved disruption of interactions between the INTS10/13/14 module and the remainder of the Integrator complex. a Immunoprecipitation of FLAG-INTS13 from nuclear extract derived from HEK-293T cells stably expressing FLAG epitope tagged human INTS13-wild type (wt), Family 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. 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 immuno-precipitation using human Integrator antibodies. **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. **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.

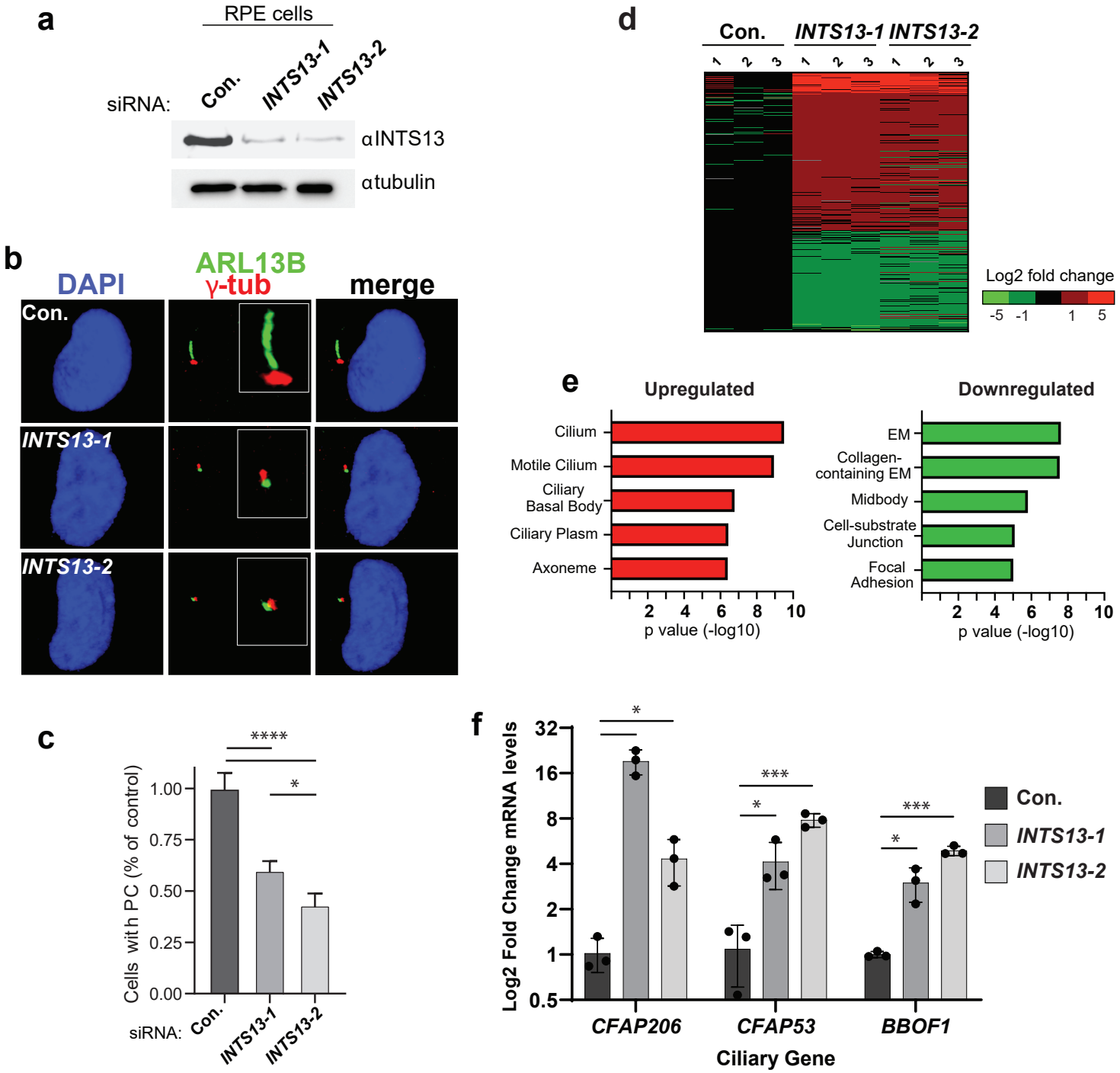

Fig 5. INTS13 depletion leads to a loss of primary cilia and broad transcriptional perturbation of ciliary genes. a Western blot of whole cell lysates from RPE cells treated with non-targeting control siRNA (Con.) or either of two siRNA targeting INTS13 (INTS13-1, INTS13-2). Lysates were probed for either INTS13 or tubulin as a loading control. **b** Immunofluorescence imaging of RPE cells treated with siRNAs as in panel (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. **c** Images from panel (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, 13-1 n = 2969, 13-2 n = 2675. 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 multiple comparisons with $p<0.05$(*) and $p<0.0001$(**). **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 panel (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 Student T-tests with a Bonferroni correction, with $p<0.05$(*) and $p<0.001$(**). *, **, ***, **** correspond to p-values < 0.05, 0.01, 0.001, and 0.0001, respectively.

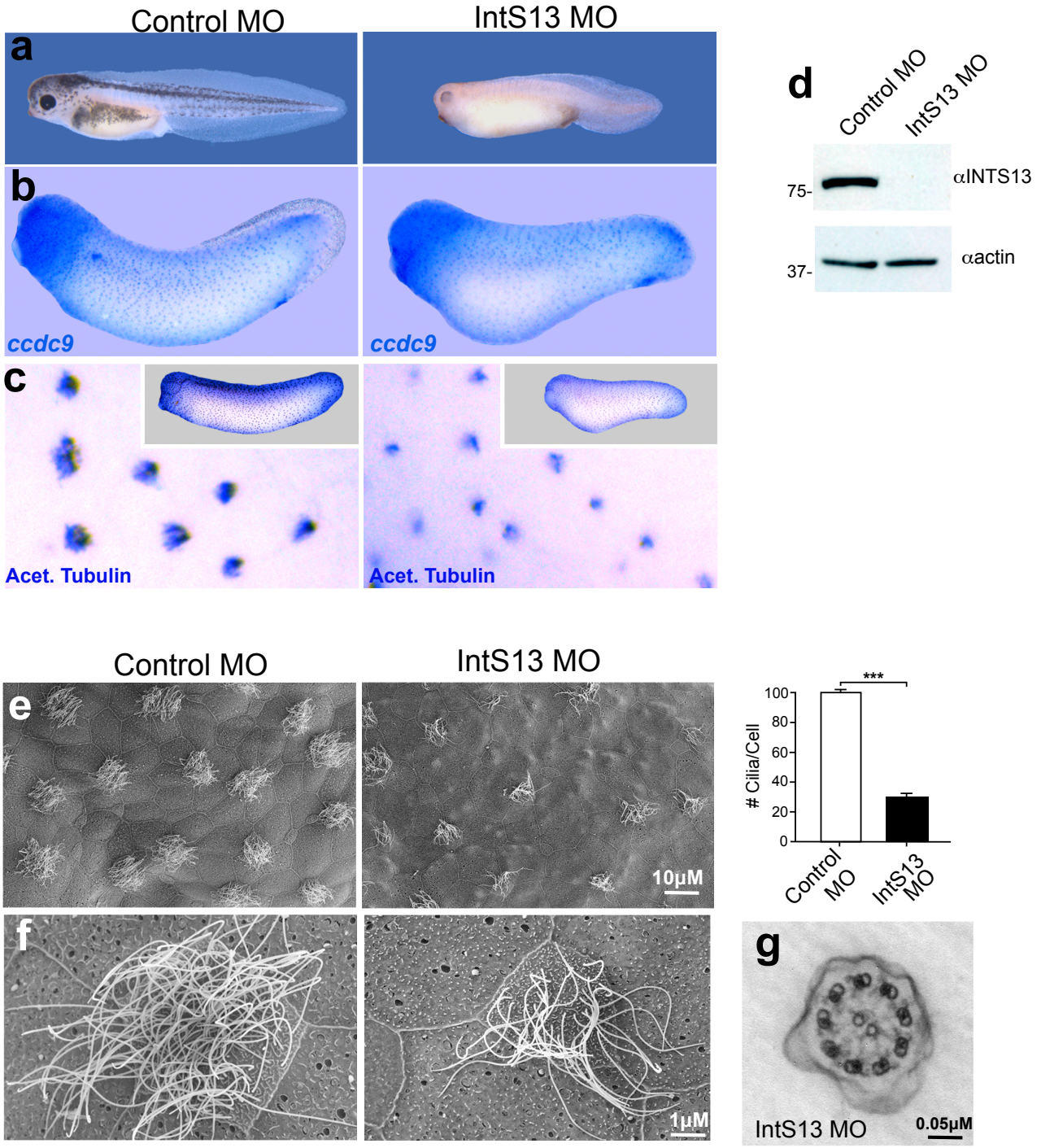

Fig 6. INTS13 is required for embryonic development and ciliogenesis in Xenopus laevis. a In comparison to control embryos, morpholino (MO)-mediated knock down of INTS13 in Xenopus leads to severe developmental defects including a small head, big belly, short and down curved body axes, cardiac edema and massive developmental delay. **b** 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. **c** 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. **d** Western blot showing that the INTS13 MO was able to deplete INTS13 protein in Xenopus embryos. **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. *, **, ***, **** correspond to p-values < 0.05, 0.01, 0.001, and 0.0001, respectively.

OFDII: MIM 252100

Clinical synopsis	**Observations in probands**			
Origin	Jordan	Jordan	France	France
Consanguinity	+	+	+	+
Patient number (refer to pedigree)	II.4	II.5	II.2	II.4
Year of birth	2002	2005	1987	1994
Gender	F	F	F	M
Oral				
Cleft / high arched palate	+	+	-	-
Tongue anomalies / hamartoma/ lipoma	-	-	-	-
Teeth abnormalities	+	+	-	-
Micrognathia	+	+	-	-
Craniofacial				
Microcephaly	+	+	-	-
Ocular hyperletorism	+	+	-	-
Median cleft upper lip	Bilateral	Bilateral	-	-
Broad nasal bridge	+	+	-	-
Nasal tip anomalies	+	+	+	+
Low set ears	+	+	-	-
Hands				
Syndactyly	-	-	-	-
Clinodactyly	+	-	-	-
Pre-Polydactyly	+	-	-	-
Bifid/duplicated thumb	-	-	-	-
Feet				
Syndactyly	-	-	-	-
Clinodactyly	+	-	-	-
Polydactyly (pre-, post-axial)	-	-	-	-
Bilateral hallux duplication	-	-	-	-
Variable Findings				
Lower stature/short humerus	-	-	+	+
Pectus escavatum	-	-	-	-
Absence of central incisors	-	-	-	-
Supernumerary teeth	+	+	-	-
Oral frenulae	-	-	-	-
Long philtrum	+	+	-	-
Thick septum	+	+	-	-
Epiglottis anomaly	+	+	-	-
Central Y-shaped metacarpa	-	-	-	-
CNS abnormalities	-	-	+	+
Speech abnormality	+	+	+	+
Congenital heart defects	+	-	-	-
Conductive hearing loss	+	+	-	-
Epicanthal fold	+	+	-	-
Respiratory tract infections and cough	+	+	-	-
Unique features seen in present propositae				
Alveolar pulmonary stenosis	-	+	-	-
Omega shaped epiglottis	+	+	-	-
Underdeveloped aryepiglottic folds	+	+	-	-
Hypertrophied false vocal cords	+	+	nd	nd
Presence of single central upper incisors	+	+	-	-
Racemose angioma	+	+	-	-
Crowded optic disc	+	+	nd	nd
Myringosclerosis	+	+	nd	nd
Adenoid hypertrophy	n.d.	+	nd	nd
Opacification of mastoid and ethmoid air cells	+	n.d.	nd	nd
Rough and sparse hair	+	+	+	+
Bilateral simian creases	+	+	-	-
Short & broad fingers	+	+	+	+

(n.d) not determined, (+) affirmative, (-) negative,
(M) male and (F) female.