The American Journal of Human Genetics, Volume *93*

Supplemental Data

Mutations in *ZMYND10*, a Gene Essential for Proper Axonemal Assembly of Inner and Outer Dynein Arms in Humans and Flies, Cause Primary Ciliary Dyskinesia

Daniel J. Moore, Alexandros Onoufriadis, Amelia Shoemark, Michael A. Simpson, Petra I. zur Lage, Sandra C. de Castro, Lucia Bartoloni, Giuseppe Gallone, Stavroula Petridi, Wesley J. Woollard, Dinu Antony, Miriam Schmidts, Teresa Didonna, Periklis Makrythanasis, Jeremy Bevillard, Nigel P. Mongan, Jana Djakow, Gerard Pals, Jane S. Lucas, June K. Marthin, Kim G. Nielsen, Federico Santoni, Michel Guiponni, Claire Hogg, Stylianos E. Antonarakis, Richard D. Emes, Eddie M.K. Chung, Nicholas D.E. Greene, Jean-Louis Blouin, Andrew P. Jarman, and Hannah M. Mitchison

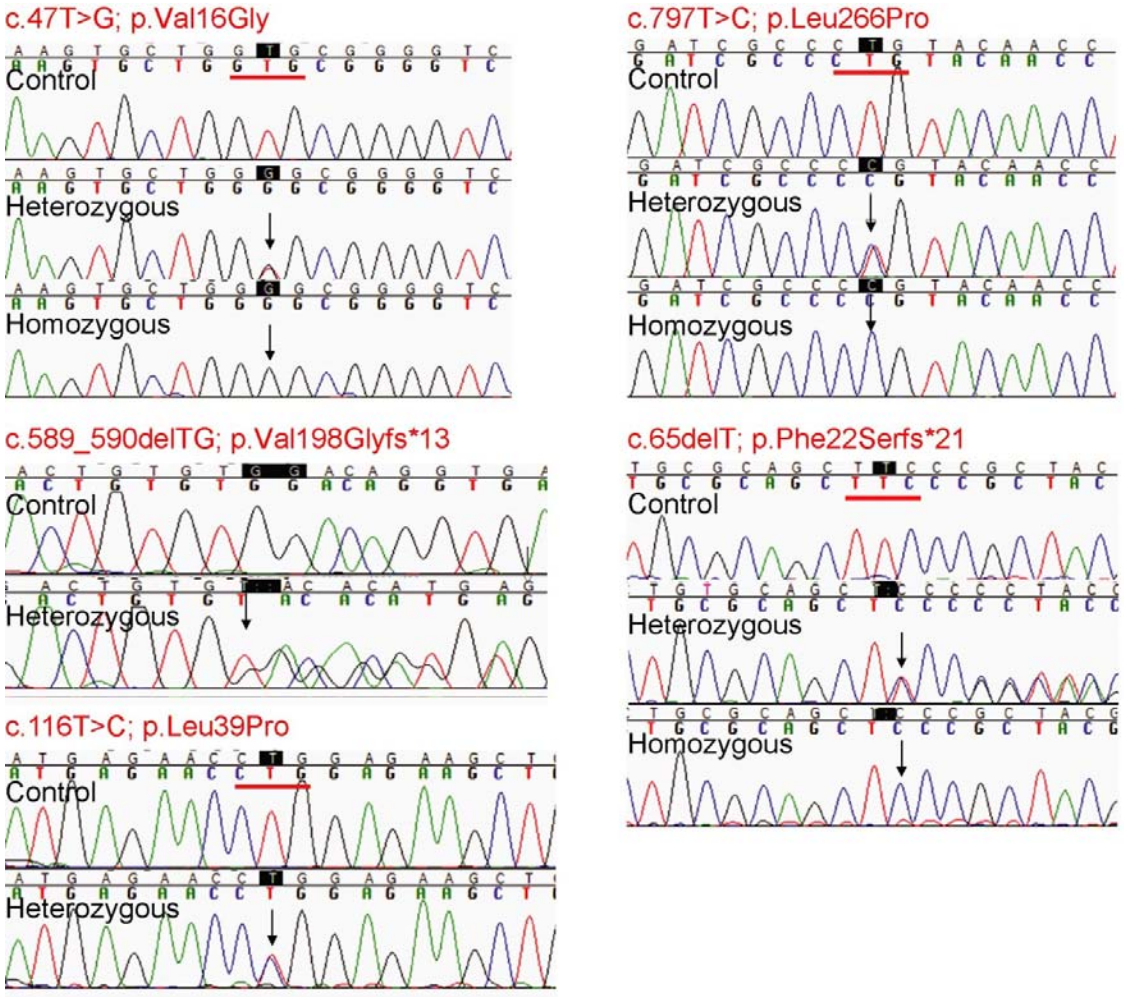

Figure S1. Sequence Chromatograms of *ZMYND10* Mutations

Reference genomic sequence amplified from a control individual is shown in comparison to sequence from carrier heterozygote and homozygote affected members representing all five *ZMYND10* mutations identified, with the variants indicated by an arrow.

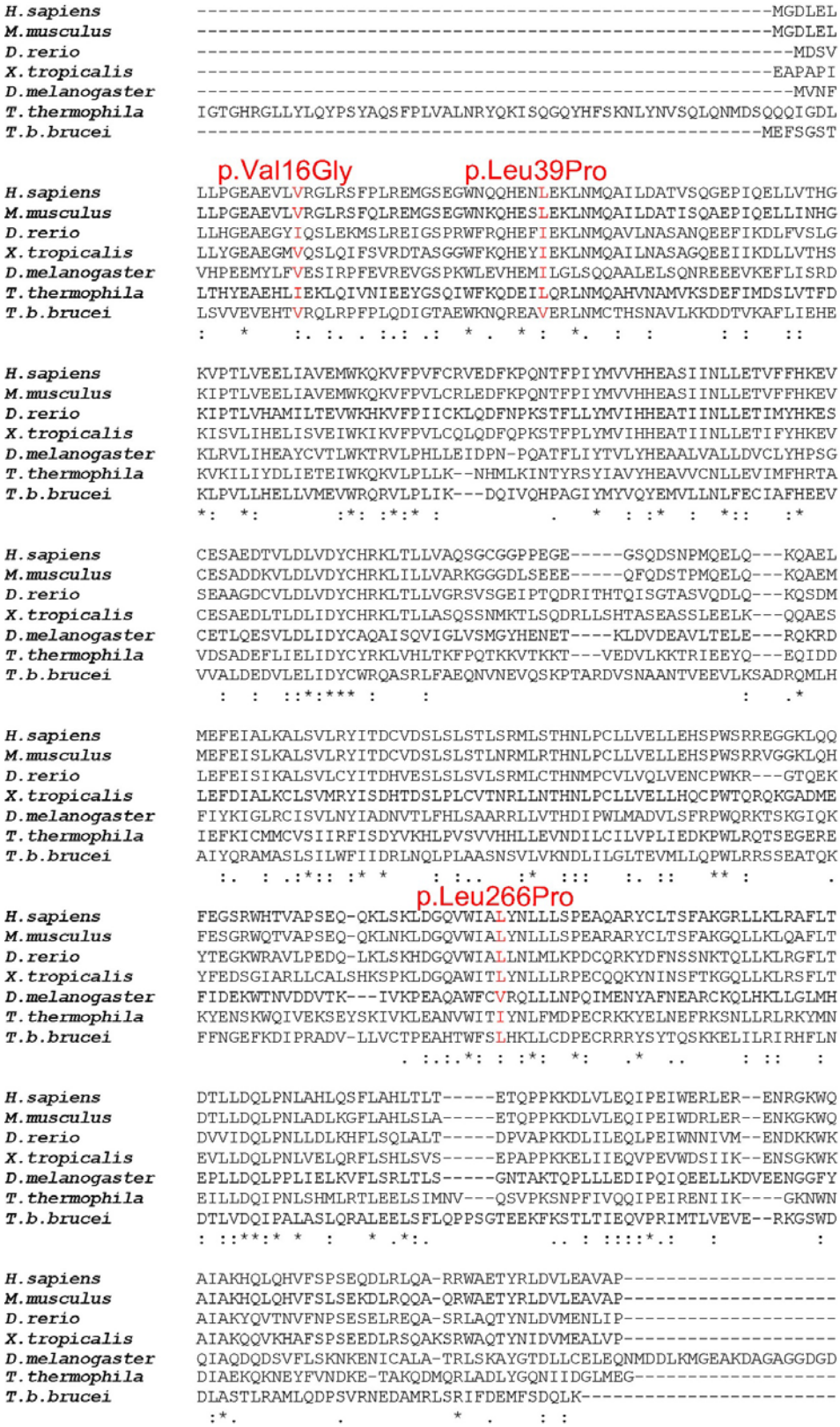

Figure S2. ZMYND10 Cross-Species Protein Homology

Alignment of the protein sequences encoded by a selected number of *ZMYND10* orthologs is shown indicating (in red) the position of the three amino acids affected by missense variants in patients.

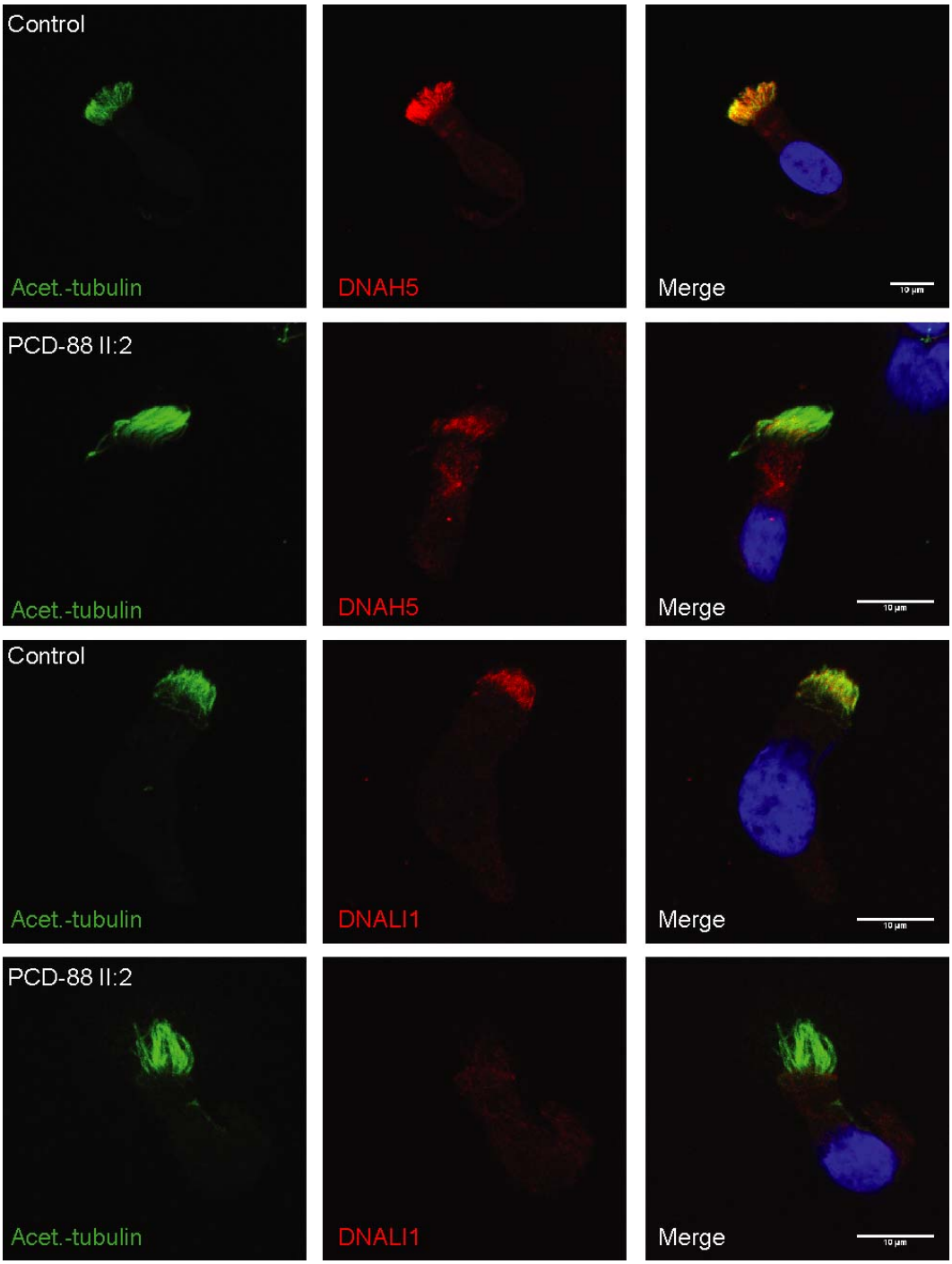

Figure S3. Loss of Inner and Outer Dynein Arms in UCL-88 II:2

Cilia axonemal immunostaining was performed using antisera against the ODA protein DNAH5 (Sigma HPA037470) and IDA protein DNALI1 (Sigma HPA028305) in nasal epithelial cells from the UCL-88 II:2 patient. Subcellular localization shows a reduction of

both DNAH5 and DNALI1 in the cilia compared to control (in red). Axoneme-specific staining with anti-acetylated α-tubulin antibodies (Sigma T-7451) (green) was used as a control, in addition to DNA (DAPI, blue).

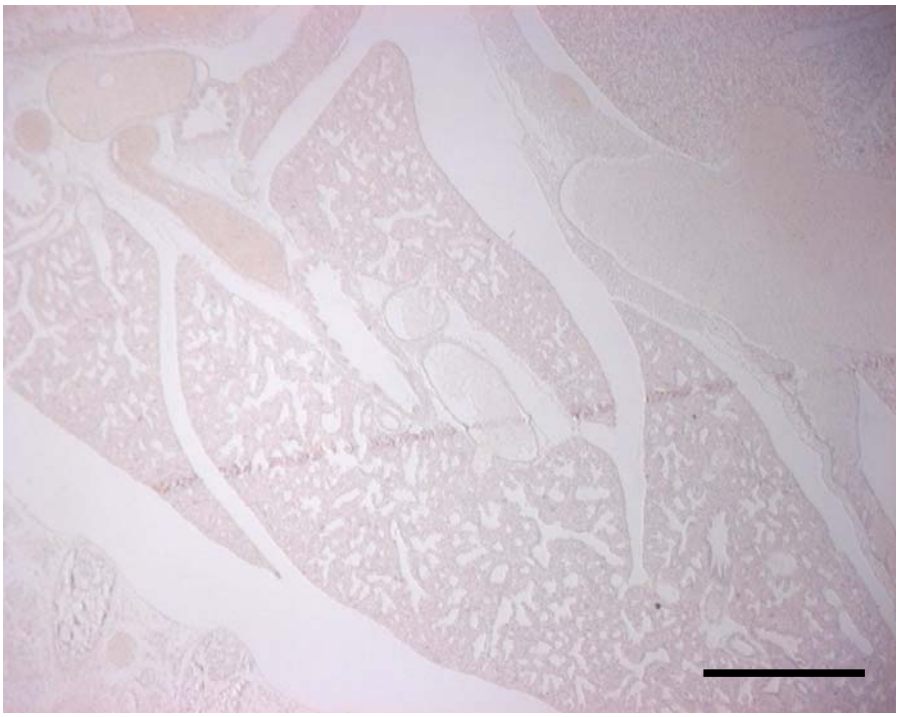

Figure S4. Mouse Lung Tissue Stained with *Zmynd10* Sense Probe

Scale bar, 0.5 mm.

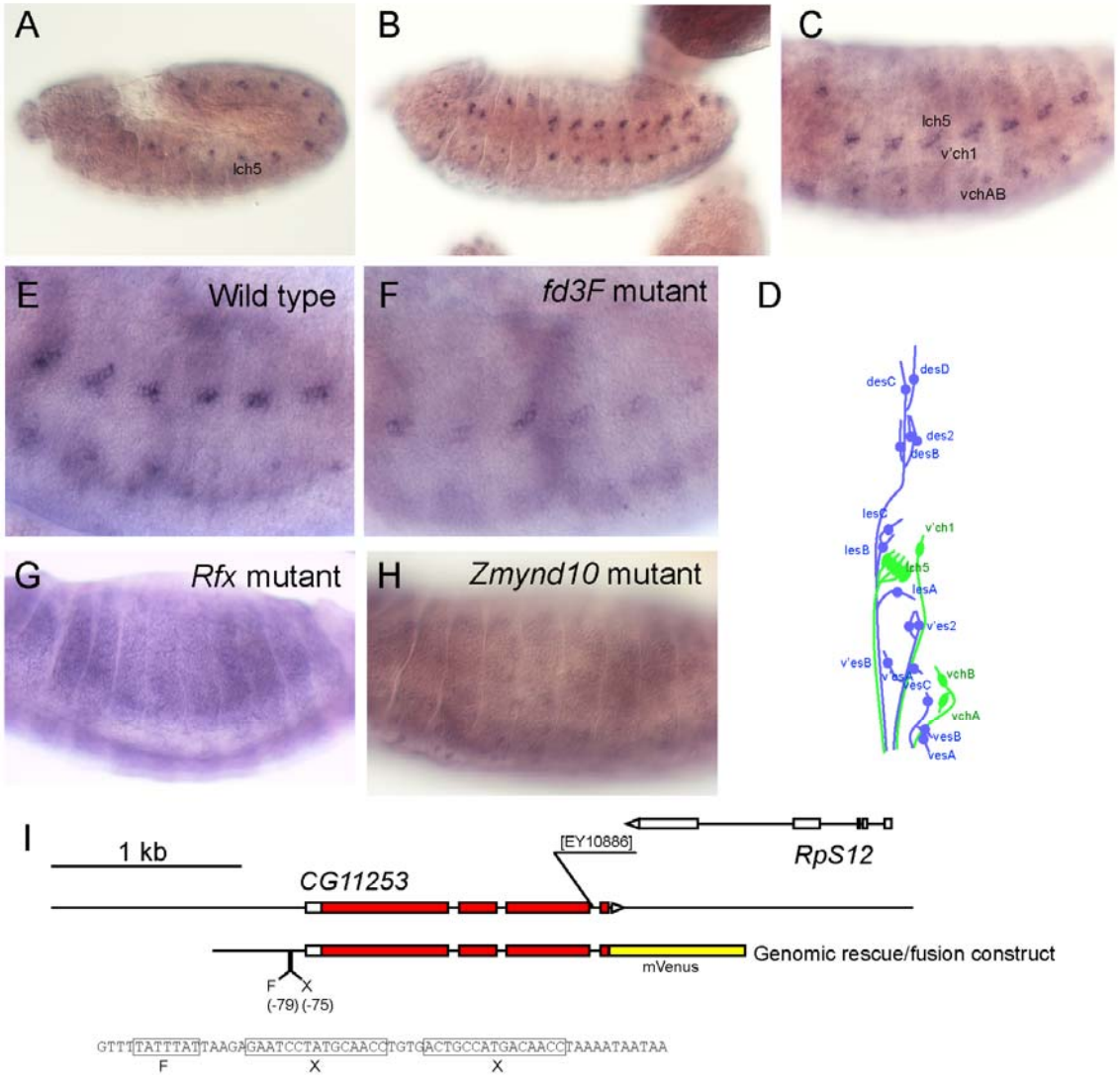

Figure S5. Expression of *Drosophila Zmynd10* during normal development and in the mutant embryos described in the study

(A–C) Developmental time-course of *Zmynd10* mRNA expression in stage 11 (A) and 14

(B and C) embryos, indicating positively stained sensory neurons.

(D) Schematic of the sensory neurons in an embryonic abdominal segment. Ch neurons are green.

(E–G) *Zmynd10* expression is reduced in *Rfx* and *fd3F* mutant embryos at stage 14.

(H) Expression is undetectable in embryos homozygous for the P element insertion, EY10886.

(I) Schematic of the *Zmynd10* locus showing gene structure, insertion site of the P element ('EY10886' insertion), and region used to construct the *Zmynd10-mVenus* fusion gene. The location of putative binding motifs for Rfx (X) and Fd3F (F) is shown. To create the fusion gene, primers 5'-

GGGGACAAGTTTGTACAAAAAAGCAGGCTCCTATAACACCTATTAGTTCTCAG-3’ and 5’-GGGGACCACTTTGTACAAGAAAGCTGGGTCAGTTTTAAGGCAGACGAGCT-3’ were used to amplify a fragment from fly genomic DNA that was cloned into pDonr221 then transferred to pBID-UASC-GV (Gateway technology, Invitrogen). Transformants were made by microinjection into syncytial blastoderm embryos.

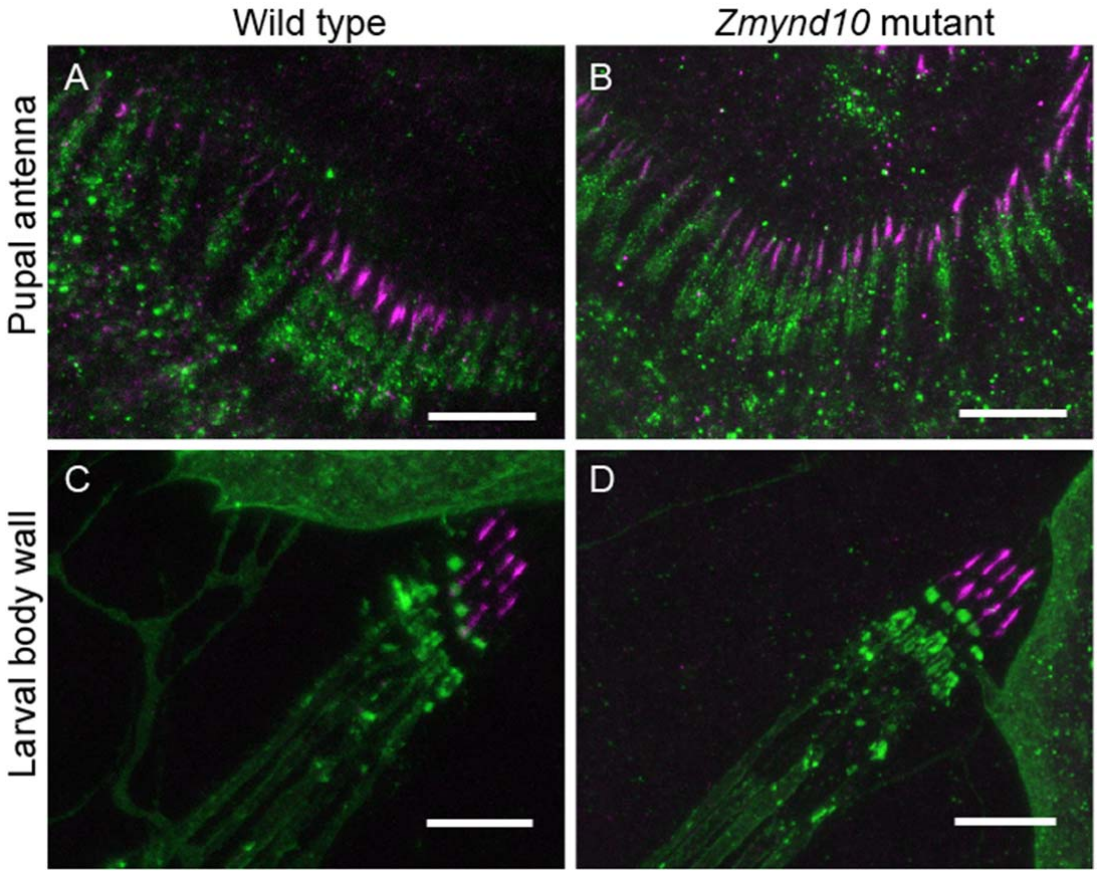

Figure S6. Localization of NOMPC in *Zmynd10* Mutant Ch Neurons

Localization of distal ciliary protein NOMPC (magenta) relative to anti-HRP immunoreactivity (green, labeling the neuronal dendritic outer segments) in Ch neurons in pupal antennae (A and B) and larval body wall (C and D). There is no gross difference in NOMPC distribution between wild-type (A and C) and *Zmynd10* mutant (B and D) tissues. Scale bars: 10 μm

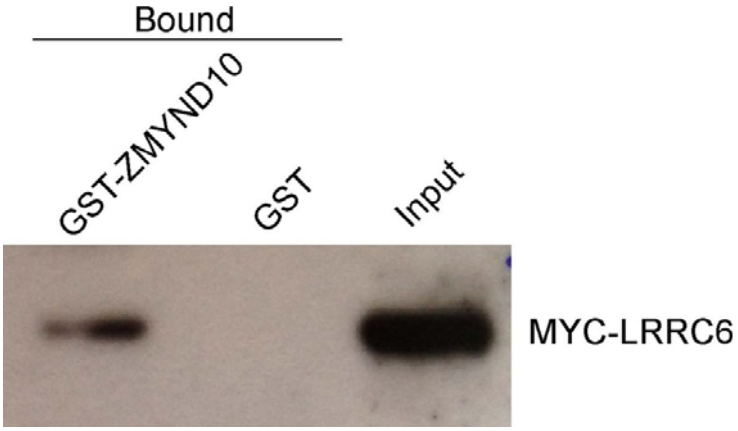

Figure S7. LRRC6 Binds Specifically to ZMYND10 in a GST Pull-Down Assay

MYC-tagged in vitro translated LRRC6 was subjected to GST pull-down assay using GST or GST-ZMYND10. A GST-ZMYND10 fusion protein (ORF cloned into the *Eco*RI site of pGEX-2T) was induced in *BL21(pLysS)* cells using 1 mM IPTG, and LRRC6 protein expressed from the yeast two-hybrid vector pGBKT7 using the TNT Quick Coupled Transcription/Translation System (Promega), resulting in LRRC6 protein with an N-terminal c-Myc epitope tag. The GST pulldown was carried out according to Nguyen, T.N. and Goodrich, J.A. (2006). Protein-protein interaction assays: eliminating false positive interactions. *Nature Methods* 3, 135–139. After electrophoresis of the washed GST samples alongside a sample of the MYC-LRR6 protein (input), the Western blot was probed with anti-c-MYC antibody (1:4000, Sigma) as a primary antibody and anti-rabbit-HRP (1:10000, Jackson Laboratories) as a secondary. Detection performed using Amersham ECL Select Western blotting detection reagents (GE Healthcare).

Table S1. Exome Sequencing Coverage and Mapping Statistics

	UCL-142 II:1	**UCL-88 II:2**	**UCL-157 II:5**
total_reads	60061451	51249050	94486779
mapped_to_target_reads	47443940	27636951	65894306
percentage	78.99	53.93	69.74
mapped_to_target_reads_plus_150bp	53550783	32502610	73656490
percentage	89.16	63.42	77.95
mean_coverage	81.25	49.46	122.23
accessible_target_bases	33323618	33323618	33323618
accessible_target_bases_1x	32934659	32643148	32765248
percentage	98.83	97.96	98.32
accessible_target_bases_5x	31832799	31546980	32195792
percentage	95.53	94.67	96.62
accessible_target_bases_10x	30717588	29846450	31598540
percentage	92.18	89.57	94.82
target_bases_20x	28784969	25414499	30102792
percentage	86.38	76.27	90.34

Table S2. Variant Calling for Exome Sequenced Individuals

	UCL-142 II:1		**UCL-88 II:2**		**UCL-157 II:5**	
variant_type	known	novel	known	novel	known	novel
variants	23625	104	23135	153	24098	141
het_variants	14357	101	14046	153	14368	138
hom_variants	9268	3	9089	0	9730	3
coding_variants	20914	95	20532	139	21256	120
het_coding_variants	12745	92	12471	139	12690	118
hom_coding_variants	8169	3	8061	0	8566	2
splice_variants	2711	9	2603	14	2842	21
het_splice_variants	1612	9	1575	14	1678	20
hom_splice_variants	1099	0	1028	0	1164	1
nonsynonymous_SNVs	9702	48	9435	86	9451	73
het_nonsynonymous_SNVs	5938	47	5723	86	5658	73
hom_nonsynonymous_SNVs	3764	1	3712	0	3793	0
synonymous_SNVs	10644	35	10500	39	10833	33
het_synonymous_SNVs	6498	35	6448	39	6480	32
hom_synonymous_SNVs	4146	0	4052	0	4353	1
stoploss_SNVs	51	0	42	1	12	0
het_stoploss_SNVs	21	0	16	1	7	0
hom_stoploss_SNVs	30	0	26	0	5	0
stopgain_SNVs	86	2	74	3	70	0
het_stopgain_SNVs	70	2	57	3	49	0
hom_stopgain_SNVs	16	0	17	0	21	0
deletions	184	5	204	5	223	8
het_deletions	111	4	104	5	135	7
hom_deletions	73	1	100	0	88	1
insertions	218	5	234	4	194	5
het_insertions	80	4	102	4	100	5
hom_insertions	138	1	132	0	94	0
frameshift_deletions	65	4	67	3	82	3
het_frameshift_deletions	37	3	24	3	46	3
hom_frameshift_deletions	28	1	43	0	36	0
frameshift_insertions	106	1	119	2	54	2
het_frameshift_insertions	24	1	30	2	22	2
hom_frameshift_insertions	82	0	89	0	32	0
ts_tv_ratio	3.04	2.06	2.91	2.08	2.97	3.03
het_ts_tv_ratio	3.06	2.03	2.95	2.08	2.98	3
hom_ts_tv_ratio	3.01	n/a	2.85	n/a	2.95	n/a

Table S3. Summary of Whole Exome Filtering Process

	UCL-142 II:1	**UCL-88 II:2**	**UCL-157 II:5**
Total variants	23729	23288	24228
Variants in MAF<0.01 in 1000 Genomes	2698	2309	2411
Heterozygous	2085	1849	1849
Heterozygous nonsynonymous, splice-site, or insertion/deletion variants	1477	1363	1372
Heterozygous variants MAF<0.01 in 1000 Genomes, and 700 in-house non-PCD exomes	308	396	365
Genes with compound heterozygous variants	13	16	6
Genes in CILBD (Ross *et al*., 2007)	0	1 *(ZMYND10)*	0
Homozygous	613	460	562
Homozygous nonsynonymous, splice-site, or insertion/deletion variants	527	426	508
Homozygous variants MAF<0.01 in 1000 Genomes, and 700 in-house non-PCD exomes	10	5	18
Genes in CILDB (Ross *et al*., 2007)	1 *(ZMYND10)*	0	1 *(ZMYND10)*

Table S4. Exome Sequencing Data for Family GVA-09

	GVA-09 II:1 (affected)				GVA-09 II:2 (affected)				GVA-09 II:3 (unaffected)			
	variants	% in exons	present in dbSNP	% in dbSNP	variants	% in exons	present in dbSNP	% in dbSNP	variants	% in exons	present in dbSNP	% in dbSNP
synonymous SNV	10155	50.98%	10095	99.41%	10468	50.92%	10405	99.40%	10306	51.22%	10242	99.38%
missense SNV	8947	44.92%	8851	98.93%	9186	44.68%	9092	98.98%	8968	44.57%	8870	98.91%
nonsense SNV	73	0.37%	68	93.15%	73	0.36%	65	89.04%	65	0.32%	61	93.85%
stoploss SNV	18	0.09%	16	88.89%	22	0.11%	20	90.91%	19	0.09%	18	94.74%
frameshift indels	118	0.59%	30	66.90%	145	0.71%	39	72.60%	138	0.69%	35	77.05%
nonframeshift indels	175	0.88%	72	86.29%	201	0.98%	88	94.47%	171	0.85%	67	83.06%
uncertain annotation	433	2.17%	407	94.00%	463	2.25%	435	93.95%	453	2.25%	425	93.82%
total exonic	19919	100.00%	19539		20558	100.00%	20144		20120	100.00%	19718	
splicing (±10bp)	2812		2393	85.10%	3022		2498	82.66%	2846		2402	84.40%
total (exonic+splicing)	22731		21932		23580		22642		22966		22120	

	GVA-09 II:1	GVA-09 II:2	GVA-09 II:3
Number of reads			
total	157304542	190680200	169255446
rmdup	123009073	157720954	144383524
on target	67897700	82965171	76829774
coverage			
average	115.8	155.1	145
% of reads covered			
8x	92.92%	94.96%	94.14%
10x	91.57%	93.96%	93.01%

Table S5. Summary of Whole Exome Filtering Process for GVA-09 Family

Filtering steps	**Number of genes remaining**
Genes after filtration by VariantMaster (respecting autosomal recessive heredity, filtering out low quality variants, synonymous variants, variants with frequency > 0.01 and variants in segmental duplications)	10
Genes not found in the laboratory's database of exome variants	5
Filtering out genes causing known disorders with phenotypes unrelated to the patient's clinical description	3
Genes without false positive calls	2
Genes with variants predicted pathogenic by SIFT, PolyPhen2, MutationTaster	1 (*ZMYND10*)

The annotated list of variants and the mapped reads files (.bam) were input for VariantMaster (submitted), a software testing for heredity and call accuracy. Variants selected were consistent for recessive inheritance (homozygous, compound heterozygotes) in the three siblings (two affected, one unaffected), and met the selected quality control threshold (score of 50 for SNVs called by SAMtools and score of 600 for indels called by Pindel). Variants in segmental duplications, synonymous variants and variants with MAF > 0.01 in dbSNP, Exome Variant Server or 1000Genomes project were excluded.

Table S6. Mutational Analysis and Clinical Symptoms of *ZMYND10* Patients

Patient (Method)	Mutations	Age at diagnosis	Sex	Origin	EM findings	Situs	Perinatal respiratory symptoms	Nasal NO	Light microscopy	Respiratory disease	Documented bronchiectais	Rhino-sino disease	Otitis media	Subfertility
UCL-142 II:1 (WES)	c.47T>G (p.Val16Gly)	8 y	F	N. European	Reduced IDA+ODA	Situs solitus	Recurrent respiratory infections from birth	23 ppb	Slowed frequency, asynchronous (stiff)	Chronic cough with purulent sputum	Multiple on HRCT, incl. chronic right middle lung lobe atelectasis	Persistant rhinitis and rhinorrhea	Frequent in childhood, repeated grommet insertions, alternating hearing loss during infections	Yes, IVF program planned
UCL-88 II:2 (WES)	c.47T>G (p.Val16Gly); c.589_590delTG (p.Val198Glyfs*13)	2 y	F	N. European	Absent IDA+ODA	Dextrocardia	Respiratory distress at birth	NA	Static cilia	Chronic cough with purulent sputum	Yes	Persistant	Glue ear, grommet insertions	Unknown
GVA-09 II:1 (SS, WES)	c.797T>C (p.Leu266Pro)	NA	F	N. European	Absent IDA+ODA	Complete situs inversus	NA	NA	Motility defect document	Recurrent infections and pneumonia	NA	Nasal polyps	NA	Had IVF
GVA-09 II:2 (SS, WES)	c.797T>C (p.Leu266Pro)	NA	F	N. European	Absent IDA+ODA	Situs solitus	NA	NA	Motility defect document	Recurrent infections and pneumonia	NA	NA	NA	NA
UCL-157 II:3 (ND)	ND	NA	M	N. European (1st cousin marriage)	NA	Complete situs inversus	NA	NA	NA	NA	NA	NA	NA	NA
UCL-157 II:4 (SS)	c.47T>G (p.Val16Gly)	NA	F	N. European (1st cousin marriage)	Reduced IDA+ODA, with reduced arm lengths	Situs solitus	NA	NA	Static cilia documented	NA	NA	NA	NA	NA
UCL-157 II:5 (WES)	c.47T>G (p.Val16Gly)	NA	F	N. European (1st cousin marriage)	Reduced IDA+ODA, with reduced arm lengths	Situs solitus	NA	NA	Motility defect, labelled 'normal' after culture	NA	NA	NA	NA	NA
UCL226 II:2 (SS)	c.65delT (p.Phe22Serfs*21)	4 mo	F	Pakistani (1st cousin marriage)	Absent IDA+ODA	Complete situs inversus	Respiratory distress from birth	NA (too young when diagnosed)	Static cilia	Chronic cough with purulent sputum	NA	Persistant	Glue ear and mild hearing loss	Unknown (paediatric)
UCL226 II:4 (ND)	ND	3 mo	F	Pakistani (1st cousin marriage)	Absent IDA+ODA	Complete situs inversus	Respiratory distress from birth	NA -Too young	Static cilia	Cough	NA	Persistant rhinitis	neonatal hearing test normal	Unknown (paediatric patient)
UCL-233 II:1 (Sanger sequencing)	c.47T>G (p.Val16Gly); c.116T>C (p.Leu39Pro)	3 y	M	N. European	Absent or short IDA+ODA	Dextrocardia with unusual atrial arrangement, and complex cyanotic heart disease	Complex cyanotic cardiac disease-respiratory status unclear	32 ppb	Static cilia	Chronic cough with purulent sputum	No, but no HRCT performed	Persistant	Bilateral effusions with hearing impairment	Unknown

WES, whole exome sequencing; SS, Sanger sequencing; ND, not sequenced; NA, information not available; IVF, in vitro fertilisation assistance used.

Table S7. Frequency of Visible Dynein Arms in *Drosophila Zmynd10* Mutant

Presence of:	Revertant (n=45)	Mutant (n=81)	*P* (Fisher exact test)
ODA	93.3%	53.1%	<0.0001
IDA	86.7%	21.0%	<0.0001

Key: n = number of doublets scored.

Table S8. WES-Derived Haplotype Analysis across the *ZMYND10* p.Val16Gly Mutation

Chrom	Position	dbSNP ID	1000G MAF	Gene	Reference	Family haplotypes					
						UCL-157		UCL-142		UCL-88	
						Mat	Pat	Mat	Pat	Mat	Pat
chr3	49949071	*rs868891*	0.2784	*MON1A*	A	G	G	G	G	G	G
chr3	50251835	-	-	*SLC38A3*	-	insG	insG	insG	insG	-	-
chr3	50306757	-	-	*SEMA3B*	-	insC	insC	-	-	insC	insC
chr3	50329826	*rs1076872*	0.7801	*IFRD2*	A	G	G	G	G	G	G
chr3	50332697	*rs13100173*	0.2679	*HYAL3*	G	A	A	A	A	A	A
chr3	50357869	*rs709210*	0.7736	*HYAL2*	A	C	C	C	C	C	C
chr3	50380558	-	-	*ZMYND10*	CA	CA	CA	CA	CA	CA	delCA
chr3	50382964	*rs138815960*	-	*ZMYND10*	A	C	C	C	C	C	A
chr3	50597092	*rs1034405*	0.8252	*C3orf18*	G	A	A	A	A	A	A
chr3	50609624	*rs2232248*	0.7257	*HEMK1*	T	G	G	G	G	G	G
chr3	51273773	*rs1480361*	0.994	*DOCK3*	T	C	C	C	C	C	C
chr3	51812952	*rs11130296*	0.162	*IQCF6*	C	T	T	T	T	T	C
chr3	51929096	*rs114559032*	0.0051	*IQCF1*	C	C	C	C	C	C	T
chr3	51929183	*rs11927897*	0.0197	*IQCF1*	C	C	C	C	C	C	T
chr3	51978220	*rs28547534*	0.9951	*PARP3*	A	G	G	G	G	G	G
chr3	51990119	-	0.9956	*GPR62*	A	C	C	G	G	C	C

Haplotype information on common SNPs across the ZMYND10 was derived from whole exome sequencing data in three available families carrying the p.Val16Gly missense variant, UCL-157 and UCL-142 homozygous for the variant, and UCL-88 compound heterozygous for p.Val16Gly missense (maternally inherited) and c.589_590delTG (paternally inherited, called 'delCA' here since calls for *ZMYND10* are from the reverse strand). Parental phase is defined in Figure 1, main paper. A common 1.6Mb haplotype carrying the p.Val16Gly missense change (c.47T>G, rs138815960) can be defined, shared across the three patients (shown by red border). However, the markers within this haplotype are mostly uninformative, as seen by comparison to the paternal haplotype carrying the c.589_590delTG *ZMYND10* mutation in UCL-88 (far right column). Five SNP markers surrounding the two *ZMYND10* mutations (indicated in grey) are uninformative.