Family	Consanguinity	# of homozygous peaks	Cumulative Homozygosity	Hypothesis from mapping	Total sequence reads (Millions)	Matched Reads	Total DIPs	Exonic DIPs	DIPs not SNP129	DIPs in linked region	DIPs after inspection and not SNP132>1%	Sanger Confirmation/Segregation	Total SNPs	Exonic SNPs	SNPs not SNP129	SNPs in linked regions	SNPs after inspection and not SNP132 (>1%)	Sanger Confirmation/Segregation	Causative Gene
F438	Yes	8	175 Mb	Hom	164	97.5%	81,744	311	251	33	1	0	358,148	5,581	518	59	14	4	*SLC41A1*

Supplemental Table 1 – Filtering process for variants from normal reference sequence following whole exome capture. Mean read coverage in homozygous regions was 302 with 96% of homozygous regions covered by 1 or more reads. Hom – homozygous; Mb – Megabases; DIPs – Deletion Insertion Polymorphisms; SNPs – Single Nucleotide Polymorphisms.

Gene	hg18 Position	Transcript	Nucleotide Change	AA Change	SNP	1000 Gen	Exome Variant Server EA Allele#	Poly Phen 1	Poly Phen 2	SIFT	Mutation Taster
SLC41A1	Chr1:204,034,566	ENST00000367137	c.698G>T	p.G233V (CSA Site)	np	np	np	2.273	1	Damaging	Disease Causing
ZNF224	Chr19:49,303,179	ENST00000336976	c.1026A>C	p.K342N	rs143243662	Yes	C=3/A=8597	1.486	1	?	polymorphism
PDZD3	Chr11:118,564,698	ENST00000355547	c.1199G>A	p.R400Q	rs190935754	Yes	A=2/G=8588	1.717	0.983	?	polymorphism
CNTRL	Chr9:122,969,741	ENST00000373855	c.5809G>A	p.E1937K	rs145379701	Yes	A=58/G=8542	1.297	0.154	?	polymorphism

Supplemental Table 2 – Analysis of variant pathogenicity for 4 surviving genes. AA – Amino Acid; EA – European American; np – not present; SNP – single nucleotide polymorphism; 1000 Gen – 1000 Genomes.

Gene	AA Change	Mouse	Chicken	Frog	Zebrafish	Ciona	Drosophila	C. elegans
SLC41A1	p.G233V (CSA Site)	G	G	G	G	G	S	D
ZNF224	p.K342N	K	K	K	K	K	-	T
PDZD3	p.R400Q	C	-	-	R	-	-	-
CNTRL	p.E1937K	R	E	E	Q	-	K	A

Supplemental Table 3 – Amino Acid Conservation across species at variant position. Amino acids in different species are denoted using the single letter amino acid codes. AA – amino acid.

Phenotype	Control MO	SLC41A1 MO
Ventral body curvature	2/156 (1.2%)	124/171 (72.5%)
Hydrocephalus	1/156 (0.6%)	97/171 (56.7%)
Cysts	0/131 (0%)	54/92 (58.7%)

Supplemental Table 4 – Phenotype of control versus *slc41a1* morphants. 1st number denotes number of embryos with indicated phenotype and 2nd number denotes total number of embryos examined. Number in parentheses indicates percentage of embryos with indicated phenotype. MO – morpholino.

Gene	UCSC hg18 Position	Nt change	AA Change	Status in III5	Status in III6	Status in IV1	Status in IV2	Status in III7	Status in III8	Status in IV3
DNAH14	chr1:223,447,185	c.5769_5772delAAAG	p.1925fsX1930	het	np	het	het	-	np	het
DNAH14	chr1:223,635,765	c.12632G>A	p.R4211Q	het	np	het	het	-	np	het
DNAH5	chr5:13,935,851	c.3248A>T	p.E1083V	het	np	het	het	-	np	np
DNAH5	chr5:13,839,358	c.8750A>G	p.Y2917C	het	np	het	het	-	np	np
DNALI1	Chr1:37,800,402	c.796A>G	p.Q259R							
DNAH9	Chr17:11,442,742	c.202C>T	p.P68S							

Supplemental Table 5 – Variants from Exome data found in known primary cilia dyskinesia genes. np – not present; het – heterozygous; Nt – nucleotide; AA – amino acid.



Supplemental Figure 1 – Sequence alignment of human and zebrafish SLC41A1 proteins. Asterisks denotes residues that are identical between sequences. Numbering on right hand side denotes sequence position.