Supplementary Information

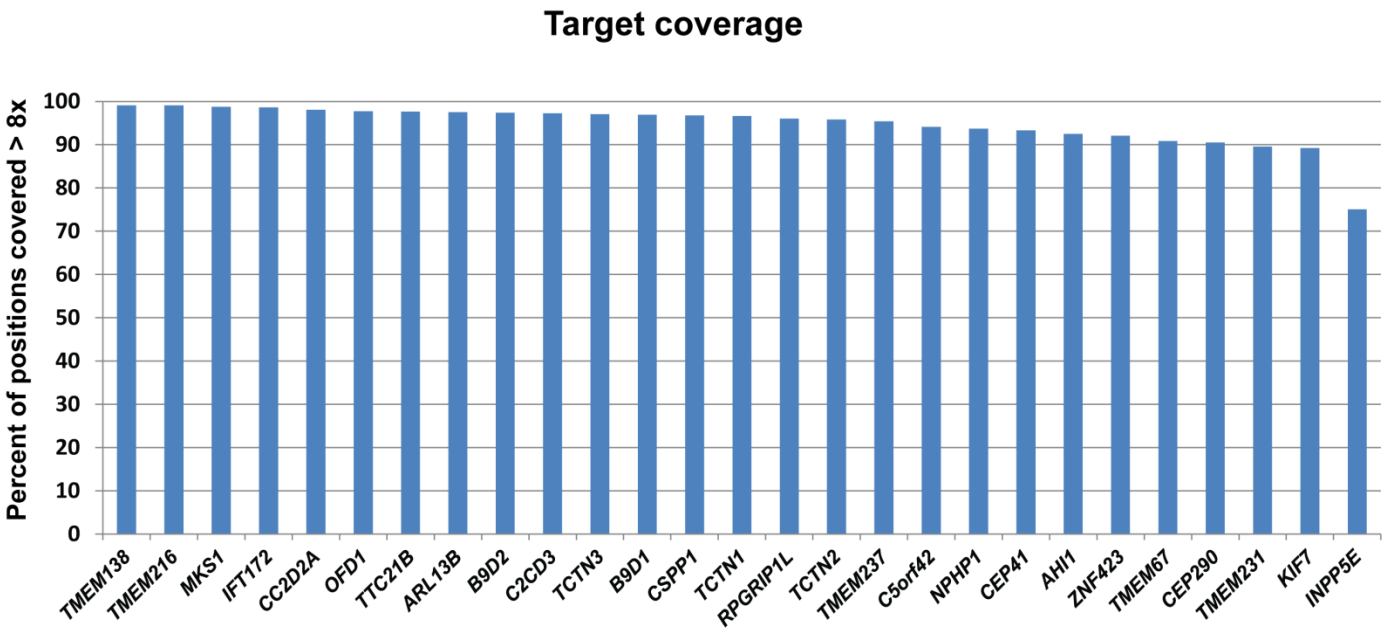
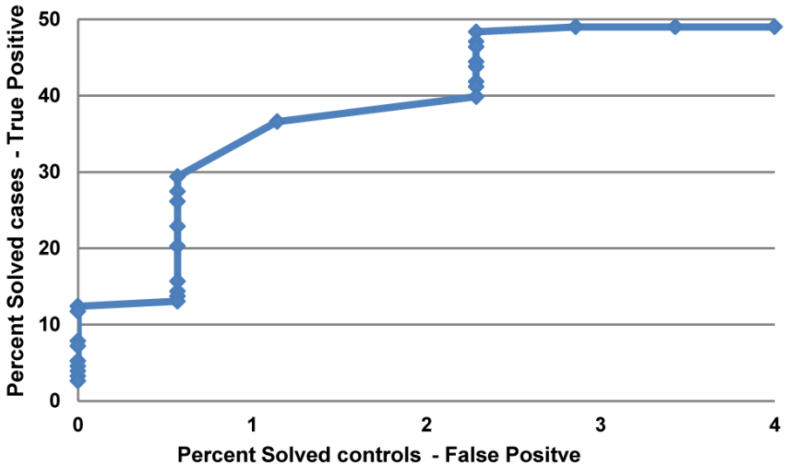


Figure S1: Coverage efficiency of the MIPs target by gene

A



CADD score cut-off	**Controls "solved"** (n)	**Cases "solved"** (n)	**False positive rate** (%)	**True positive rate** (%)	**solve rate (%)** (true positives-false positives)
0	7	75	4	49	45
11	5	75	2.9	49	46.1
12	4	74	2.9	48	45.1
13	4	72	2.9	47	44.1

B

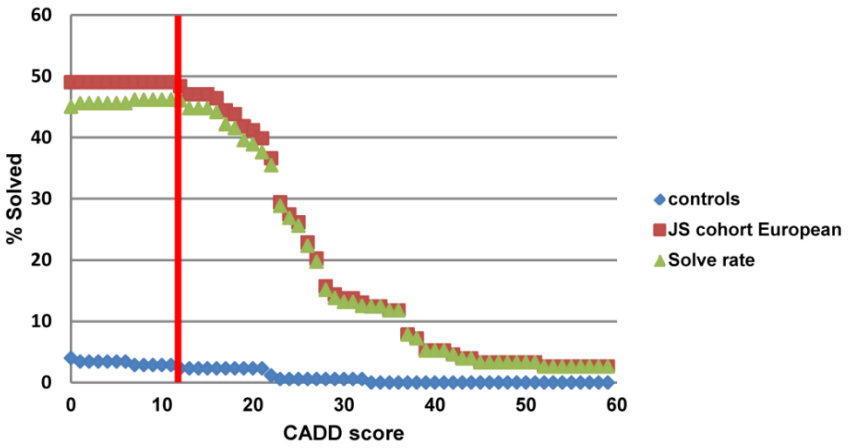


Figure S2. Receiver-Operating-Characteristic (ROC) curve method to determine CADD score cut off for causal mutations (A) ROC curve plotting the proportion of affected individuals with two RDVs against the proportion of control individuals with two RDVs. The table beneath the graph indicates the values for affected and control individuals at representative CADD score cutoffs, as well as the "solve rate" = proportion of affected individuals with two RDVs minus proportion of control individuals with two RDVs (presumed-false positives). Given that the control individuals are all European, the ROC curves were calculated using only the affected individuals of European descent. **(B)** Proportion of samples with two RDVs plotted against CADD score cutoff. The red curve represents affected individuals, the blue curve indicates the control individuals and the green curve indicates the solve rate. The vertical red bar indicates the CADD value of 11 that was used to define causal variants.

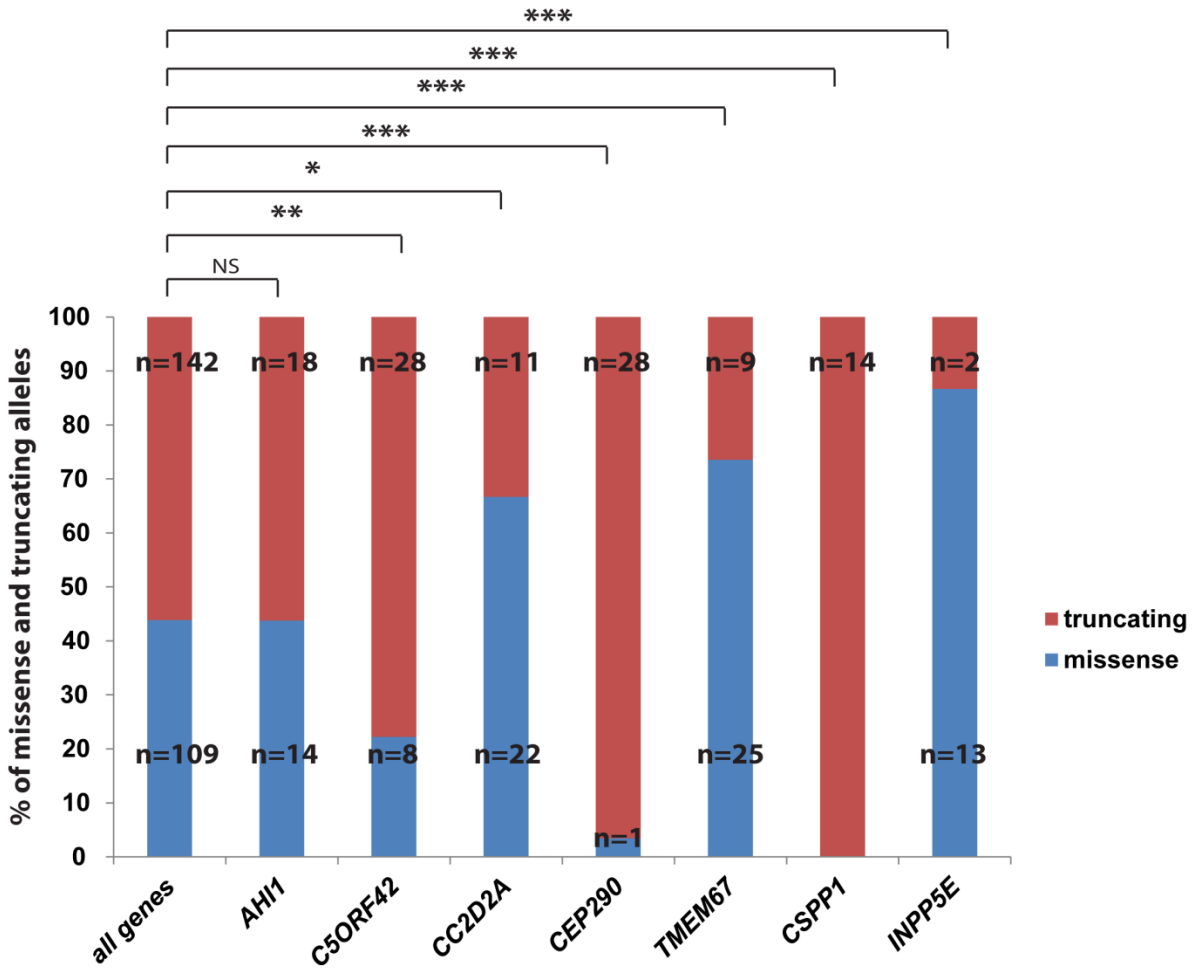


Figure S3: Differential distribution of mutation types (missense vs truncating) across the 7 genes most commonly associated with JS. The distribution of mutation types is not significantly different for *AHI1* compared with the entire cohort, but *C5ORF42*, *CEP290* and *CSPP1* have significantly fewer missense mutations than average while *CC2D2A*, *INPP5E* and *TMEM67* have significantly fewer truncating mutations (***$p<0.0001$, **$p<0.001$, *$p<0.01$; Fisher's exact test). Each mutation was only counted once, even if it occurred multiple times (homozygous or in >1 affected individual).

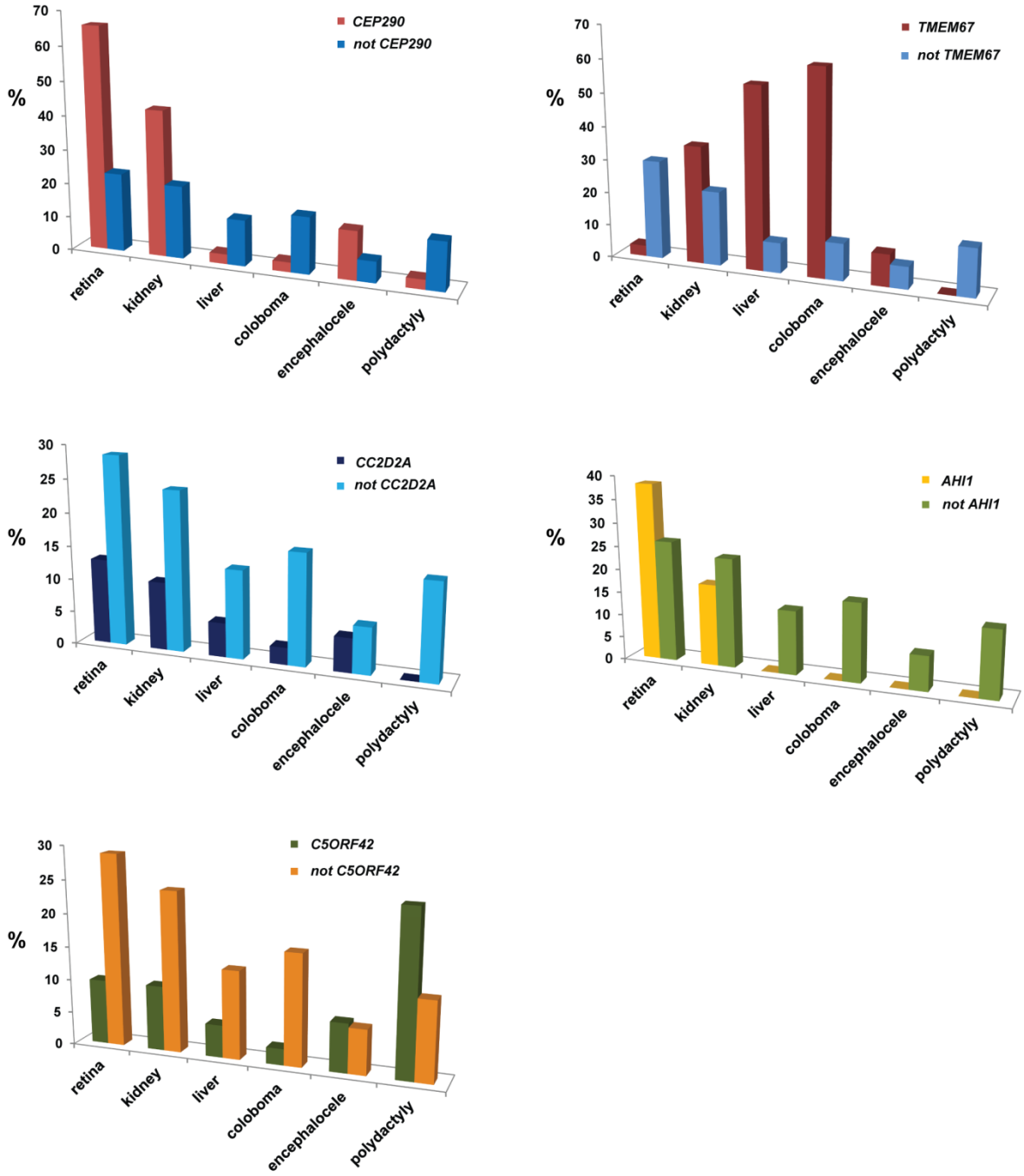


Figure S4. Frequency of variable clinical features in affected individuals with causal mutations in *CEP290*, *TMEM67*, *AHI1*, *CC2D2A*, or *C5ORF42* compared to the rest of the JS cohort.

Table S1. MIPS target

No	Gene	Gene name	Accession#	hg19 position	Exons	Additionally sequenced
1	***AHI1***	Abelson helper integration site 1	NM_001134831.1	chr6:135,285,296-135,497,776	29	chr6:135709110-135709181
2	***ARL13B***	ADP-ribosylation factor-like 13B	NM_182896.2	chr3: 93,980,155-94,054,069	11	
3	***B9D1***	B9 protein domain 1	NM_015681.3	chr17:19,343,170-19,362,713	7	
4	***B9D2***	B9 protein domain 2	NM_030578.3	chr19: 41,354,421-41,364,173	4	
5	***C2CD3***	C2 calcium-dependent domain containing 3	NM_001286577.1	chr11:74,012,714-74,171,019	33	chr11:73745633-73745663
6	***C5ORF42***	chromosome 5 open reading frame 42	NM_023073.3	chr5:37,106,228-37,249,428	52	chr5:37157405-37157532
7	***CC2D2A***	coiled-coil and C2 domain containing 2A	NM_001080522.2	chr4:15,469,930-15,601,557	38	
8	***CEP290***	centrosomal protein 290kDa	NM_025114.3	chr12:88,049,020-88,142,216	54	chr12:88509247-88509357
9	***CEP41***	centrosomal protein 41kDa	NM_018718.2	chr7:130,393,771-130,441,237	11	
10	***CSPP1***	centrosome and spindle pole associated protein 1	NM_024790.6	chr8:67,064,368-67,196,263	29	chr8:67999048-67999091; chr8:68004028-68004128
11	***IFT172***	intraflageller transport 172 homolog	NM_015662.2	chr2:27667240-27712678	48	
12	***INPP5E***	inositol polyphosphate-5-phosphatase, 72kDa	NM_019892.4	chr9:136,428,619-136,439,822	10	
13	***KIF7***	kinesin family member 7	NM_198525.2	chr15:89,627,977-89,655,451	19	
14	***MKS1***	Meckel syndrome, type 1	NM_017777.3	chr17:58,205,437-58,219,305	18	
15	***NPHP1***	nephrocystin-1 (juvenile)	NM_000272.3	chr2:110,122,311-110,205,042	20	
16	***OFD1***	oral-facial-digital syndrome 1	NM_003611.2	chrX:13,734,745-13,769,353	23	
17	***RPGRIP1L***	RPGRIP1-like / FTM	NM_015272.2	chr16:53,600,778-53,703,846	27	
18	***TCTN1***	tectonic family member 1	NM_001082538.2	chr12:110,614,164-110,649,128	15	chr12:111054109-111054167
19	***TCTN2***	tectonic family member 2	NM_024809.4	chr12:123,671,113-123,708,403	18	
20	***TCTN3***	tectonic family member 3	NM_015631.5	chr10:95,663,396-95,694,143	14	
21	***TMEM138***	transmembrane protein 138	NM_016464.4	chr11:61,362,001-61,369,509	5	
22	***TMEM216***	transmembrane protein 216	NM_001173990.2	chr11:11: 61,391,687-61,398,851	5	
23	***TMEM231***	transmembrane protein 231	NM_001077418.2	chr16:75,536,744-75,556,286	7	chr16:75,589,872-75,589,968
24	***TMEM237***	transmembrane protein 237	NM_001044385.2	chr2:201,620,184-201,643,517	13	chr2:202504975-202504998
25	***TMEM67***	transmembrane protein 67	NM_153704.5	chr8:93,754,857-93,818,057	28	chr8:94772079-94772231
26	***TTC21B***	tetratricopeptide repeat domain 21B	NM_024753.4	chr2:165,873,362-165,953,843	29	
27	***ZNF423***	zinc finger protein 423	NM_015069.3	chr16:49,490,605-49,822,738	8	

Table S2. Phenotypic assessment criteria

Major Feature	Positive criteria to determine the presence of this feature (one or more)
Retinal disease	Pigmentary abnormality on fundal examination, abnormal electroretinogram recording, retinal dystrophy, diagnosis of Leber congenital amaurosis or retinitis pigmentosa
Renal disease	Increased echogenicity or cystic kidneys on ultrasound, positive kidney biopsy (microcysts, fibrosis), diagnosis of nephronophthisis, chronic renal failure, dialysis, renal transplant
Liver disease	Persistently elevated transaminases or GGT, increased echogenicity on ultrasound, hepatomegaly, splenomegaly, upper gastrointestinal bleeding, positive liver biopsy (liver fibrosis, ductal plate malformation), portal hypertension, variceal bleeding, spleen removal, medication treatment of liver disease, liver transplant
Encephalocele	Based on clinical examination or head imaging
Coloboma	Based on clinical examination
Polydactyly	Single or multiple, pre-, meso-, or postaxial polydactyly

Supplementary Table S3: detailed phenotypic combinations

Phenotypes present	n	%	*AHI1*	*ARL13B*	*B9D1*	*C9D2*	*C2CD3*	*C5ORF42*	*CC2D2A*	*CEP290*	*CSPP1*	*IFT172*	*INPP5E*	*KIF7*	*MKS1*	*NPHP1*	*OFD1*	*RPGRIP1L*	*TCTN1*	*TCTN2*	*TCTN3*	*TMEM138*	*TMEM216*	*TMEM237*	*TMEM67*	*unkown*
pure JS	**68**	**33.83**	8	1	0	0	0	9	9	0	6	0	2	0	1	1	0	0	0	0	0	0	1	0	0	30
retina only	**26**	**12.93**	**4**	0	0	0	0	1	0	**3**	1	0	1	0	2	0	0	0	0	0	0	0	0	0	0	14
retina+kidney	**19**	**9.45**	**4**	0	0	0	0	0	1	**8**	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	4
kidney only	**12**	**5.97**	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	1	0	0	0	0	1	1	0	7
polydactyly only	**11**	**5.47**	0	0	0	0	0	**5**	0	0	0	0	0	1	0	0	0	1	0	0	0	0	1	0	0	3
coloboma only	**8**	**3.98**	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	**2**	4
liver only	**7**	**3.48**	0	0	0	0	0	1	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4
coloboma+liver	**6**	**2.98**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**3**	3
retina+liver+kidney	**5**	**2.48**	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	3
encephalocele only	**5**	**2.48**	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4
coloboma+liver+kidney	**5**	**2.48**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**5**	0
encephalocele+retina only	**4**	**1.99**	0	0	0	0	0	1	0	2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
coloboma+retina	**4**	**1.99**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4
coloboma+retina+kidney	**3**	**1.49**	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
kidney+liver only	**2**	**0.99**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1
polydactyly+kidney only	**2**	**0.99**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1
coloboma+retina+liver	**2**	**0.99**	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
coloboma+polydactyly	**2**	**0.99**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1
polydactyly+liver only	**1**	**0.49**	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
polydactyly+retina only	**1**	**0.49**	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
polydactyly+retina + kidney	**1**	**0.49**	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Polydactyly+retina+liver	**1**	**0.49**	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
polydactyly+encephalocele	**1**	**0.49**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
polydactyly+encephalocele+retina+kidney	**1**	**0.49**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
coloboma+kidney	**1**	**0.49**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
coloboma+encephalocele+liver	**1**	**0.49**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
coloboma+encephalocele+retina	**1**	**0.49**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
coloboma+polydactyly+encephalocele+retina	**1**	**0.49**	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
total	**201**	**100**	16	1	0	0	0	21	11	16	10	1	5	1	4	3	1	4	0	2	1	0	4	1	12	87

Only individuals for whom definite positive or negative information for all 6 major associated features was available (retinal disease, renal disease, liver disease, coloboma, polydactyly, encephalocele) were considered for this analysis (n=201).

Table S4. Correlations between common features seen in individuals with JS

Positive associations			
	Odds ratio	**95% CI**	**p-value**
Liver disease and coloboma	6.5	3.2, 13.4*	<0.0001
Kidney disease and retinal dystrophy	3.0	1.7, 5.2*	<0.0001
Liver disease and kidney disease	3.0	1.6, 5.5*	0.0004
Polydactyly and encephalocele	2.8	1.03, 7.8*	0.03
Retinal dystrophy and encephalocele	2.2	0.8, 5.5	0.1
Coloboma and encephalocele	1.6	0.5, 5.0	0.18
Coloboma and kidney disease	1.5	0.8, 2.9	0.22
Liver disease and encephalocele	1.2	0.3, 4.3	0.24
Kidney disease and polydactyly	1.2	0.6, 2.3	0.62
Polydactyly and coloboma	1.1	0.4, 2.9	0.87
Negative associations			
	Odds ratio	**95% CI**	**p-value**
Kidney disease and encephalocele	0.9	0.4, 2.5	0.91
Retinal dystrophy and liver disease	0.7	0.3, 1.6	0.39
Retinal dystrophy and polydactyly	0.7	0.3, 1.6	0.39
Retinal dystrophy and coloboma	0.6	0.3, 1.3	0.2
Polydactyly and liver disease	0.4	0.1, 1.5	0.08

CI = Confidence Interval, * statistically significant (p-value <0.05, Fisher's exact test if <5 observations are present in one category or Chi-Square test in all other situations)

Supplementary Table S7: Detailed phenotypic description for individuals with *B9D2* mutations

ID	Mutations	MTS	retina	kidney	liver	Coloboma	Encephalocele	PD	Other
UW284-3	p.Gly155Ser (mat) p.Pro74Ser (pat)	+	NA	-	NA	-	+ (foramen magnum encephalocele)	+	**Oral features**: Cleft palate, tongue tumors, **Skeletal features**: tibial and fibular mesomelic dysplasia, **Neurological features**: shunted hydrocephalus, seizures, interpeduncular heterotopia **Dysmorphic features**: small palpebral fissures **Visual features**: poor pupillary response to light, optic disks normal **Other**: micropenis, hearing loss, patent ductus arteriosus
UW309-3	p.Leu36Pro hmz	+	NA	-	-	-	-	+	**Neurological features**: seizures and abnormal EEG **Dysmorphic features**: frontal bossing, epicanthus, dysplastic ears, down turned corners of mouth, retrognathia, ptosis, right eye exotropia **Other**: Hypospadias

MTS Molar Tooth Sign*, PD* polydactyly, *NA* not available, *EEG* Electro-encephalogram, *mat* maternal, *pat* paternal, *hmz* homozygous

Supplementary Table S8: Recurrent alleles

Gene	allele	n unrelated* families	Reported country of origin or ethnicity (*n families*)
AHI1	**Q423X**	**3**	Armenian (1), Australian (1), Native American/mixed European (EU) (1)
C5ORF42	**p.Gly2663Alafs*40**	**4**	**Saudi Arabian (4)**
C5ORF42	**L595X**	**3**	mixed EU (3)
C5ORF42	**p.Thr2755Asnfs*8**	**3**	mixed EU (2), mixed EU /Native American (1)
C5ORF42	**R2493X**	**3**	Native American/ mixed EU (1), French Canadian (1), Native American/ mixed EU (1)
C5ORF42	**R2904X**	**4**	mixed EU/Canadian (1), mixed EU (2), African American/Korean/mixed EU (1)
C5ORF42	**W2593X**	**3**	Saudi Arabian (2), India (1)
CC2D2A	**D1556V**	**7**	Mixed EU (5), Australian (1), Native American/ mixed EU /Filipino (1)
CC2D2A	**P1122S**	**3**	Saudi Arabian (3)
CC2D2A	**R1528C**	**4**	Turkish (1), mixed EU (2), Trinidadian/ mixed EU (1)
CC2D2A	**T1116M**	**3**	Brazilian (1), mixed EU (2)
CC2D2A	**Val1097Phefs*2**	**4**	mixed EU (4)
CEP290	**G1890X**	**10**	India (3), mixed EU (3), Saudi Arabian (1), Iraqi (1), Canadian/ mixed EU (1), mixed EU /Indian (1)
CEP290	**K1575X**	**3**	mixed EU /Thai (1), mixed EU (2)
CSPP1	**NM_024790.6:c.2953+1G>A**	**3**	Brazilian (3)
CSPP1	**p.Tyr1071***	**3**	Brazilian (3)
NPHP1	**deletion**	**5**	Turkish (1), Native American/ mixed EU /French Canadian (1), Peruvian/ mixed EU (1), mixed EU (2)
TMEM216	**R73L**	**10**	**Ashkenazi (10)**
TMEM67	**I833T**	**5**	mixed EU (4), Japanese/mixed EU (1)

* not reported to be related; alleles previously described in specific ethnic groups are highlighted. The number in parenthesis indicates the number of families for each country of origin or ethnicity.

Supplementary Table S9. Gene-Phenotype Correlations in a large JS cohort

	Retina	Kidney	Liver	Polydactyly	Coloboma	Encephalocele
AHI1	**2.3** (1.1-5.2) p=0.03	**0.9** (0.4-2.4) p=0.90	p=0.15	p=0.06	p=0.02	p=0.24
C5ORF42	**0.4** (0.1-1.1) p=0.08	**0.4** (0.12-1.1) p=0.06	**0.4** (0.1-1.7) p=0.28	**2.7** (1.2-5.9) p=0.01	**0.15** (0.02-1.1) p=0.04	**1.2** (0.3-4.2) p=0.73
CC2D2A	**0.5** (0.3-2.0) p=0.52	**0.4** (0.1-1.2) p=0.13	**0.4** (0.1-1.9) p=0.40	p=0.01	**0.2** (0.02-1.4) p=0.1	**0.9** (0.2-3.9) p=1.0
CEP290	**22.9** (6.7-78.4) p<0.0001	**3.3** (1.6-7.1) P=0.001	0.2 (0.03-1.9) p=0.23	0.3 (0.04-2.0) p=0.22	0.2 (0.02-1.5) p=0.1	**3.5** (1.2-10.2) p=0.01
CSPP1	**0.6** (0.1-2.7) p=0.73	p=0.03	**1.0** (0.2-4.8) p=1.0	p=0.23	p=0.22	**1.1** (0.1-9.0) p=1.0
INPP5E	**1.6** (0.4-5.7) p=0.50	**1.1** (0.3-4.3) p=1.0	**0.8** (0.1-6.3) p=1.0	p=0.6	p=0.61	p=1.0
MKS1	**3.6** (0.6-21.7) p=0.16	**0.5** (0.06-4.2) p=0.7	**2.6** (0.5-13.6) p=0.25	**1.5** (0.2-13.5) p=0.54	p=0.59	p=1.0
NPHP1	p=0.56	3.0 (0.4-21.) p=0.26	p=1.0	p=1.0	p=1.0	p=1.0
OFD1	p=1.0	**1.5** (0.1-16.7) p=1.0	**2.1** (0.2-20.6) p=0.45	**6.1** (0.8-44.2) p=0.10	p=1.0	**13.1** (1.8-97.0) p=0.03
RPGRIP1L	**0.7** (0.1-3.2) p=0.38	**4.7** (1.3-17.0) p=0.02	**1.3** (0.1-11.0) p=0.59	**4.2** (1.1-15.3) p=0.04	p=1.0	**1.4** (0.2-11.3) p=0.55
TCTN2	**1.2** (0.1-13.0) p=1.0	p=0.58	p=1.0	**18.7** (1.9-182.9) p=0.01	**10.1** (0.9-113.5) p=0.08	**13.6** (2.6-70.8) p=0.007
TMEM216	p=0.06	**2.2** (0.7-7.1) p=0.18	**0.6** (0.1-4.4) p=1.0	3.1 (0.9-10.7) p=0.08	p=1.0	p=0.61
TMEM67	**0.1** (0.01-0.8) p=0.006	**2.0** (0.9-4.5) p=0.07	**17.3** (7.2-42.0) p<0.0001	P=0.05	**22.9** (8.6-61.1) p<0.0001	**2.3** (0.6-8.4) p=0.18

Odds ratios are indicated in **bold** followed by the 95% Confidence Interval in brackets. Statistical significance (Fisher's exact test if <5 observations are present in one category or Chi-Square test in all other situations) is indicated below with the respective p-value. Cells with significant positive correlations are shaded in yellow, and those remaining significant after Bonferroni correction are shaded in green. Cells shaded in orange indicate negative correlations. In cells without odds ratios, no individuals with the relevant clinical feature were observed to have that genetic cause. In those cells, the p-value indicates the likelihood that there is a correlation even in the absence of an individual with the feature and the genetic cause.