Neuron, Volume *84*

Supplemental Information

Katanin p80 Regulates Human Cortical Development by Limiting Centriole and Cilia Number

Wen F. Hu, Oz Pomp, Tawfeg Ben-Omran, Andrew Kodani, Katrin Henke, Ganeshwaran H. Mochida, Timothy W. Yu, Mollie B. Woodworth, Carine Bonnard, Grace Selva Raj, Thong Teck Tan, Hanan Hamamy, Amira Masri, Mohammad Shboul, Muna Al Saffar, Jennifer N. Partlow, Mohammed Al-Dosari, Anas Alazami, Mohammed Alowain, Fowzan S. Alkuraya, Jeremy F. Reiter, Matthew P. Harris, Bruno Reversade, and Christopher A. Walsh

Supplemental Figures and Figure Legends

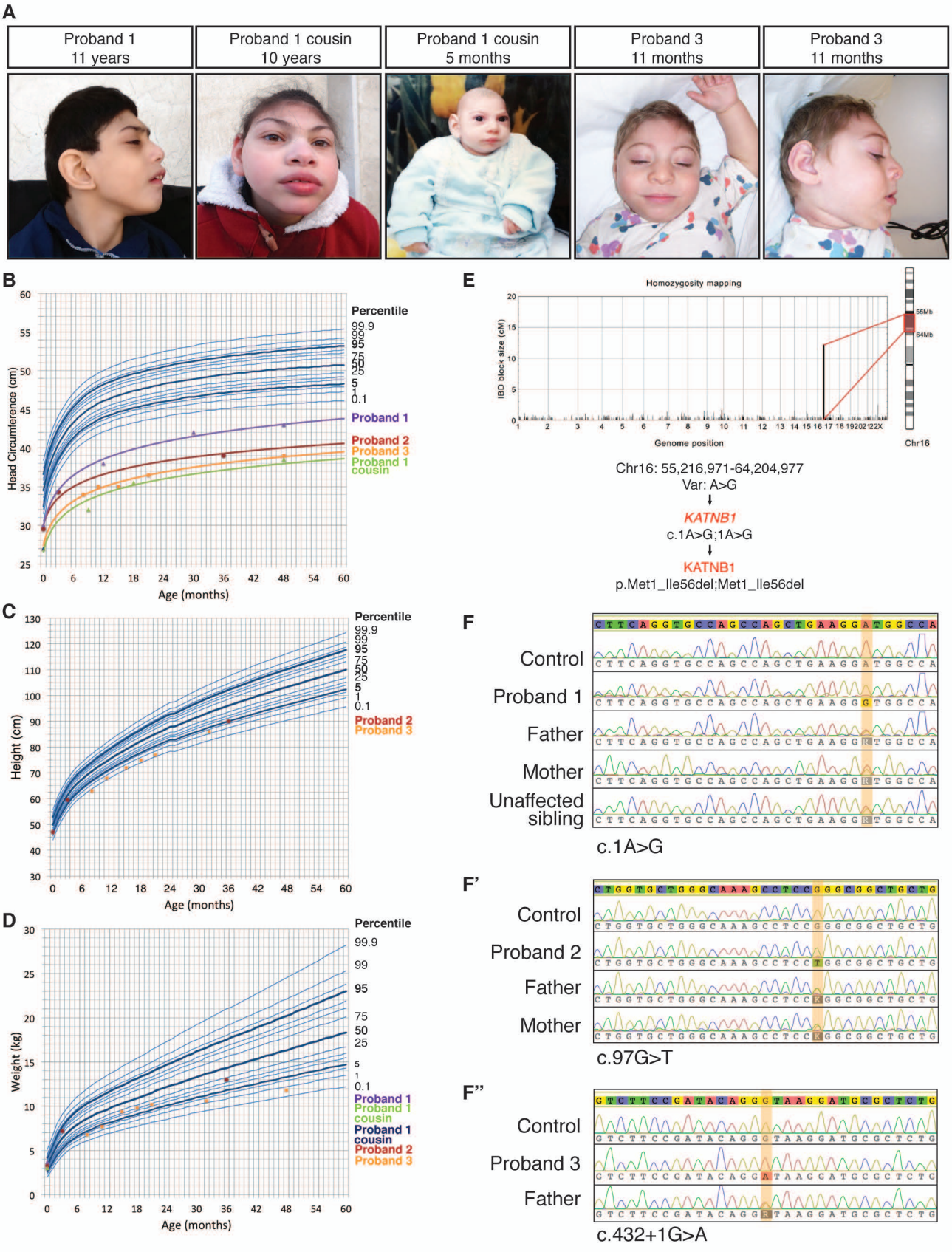
Supplemental Figure S1, related to Figure 1. *KATNB1* affected individuals have small head circumference relative to height and weight.

(A) Affected *KATNB1* individuals show sloping foreheads indicative of reduced cranial size, and facial dysmorphisms.

(B-D) *KATNB1* affected individuals display head circumference well below 0.1st percentile for ages between birth and five years, disproportionate to height and weight. Proband data fitted with logarithmic regression line. Proband 1, purple; Proband 1 cousin (female), green; Proband 1 cousin (male), dark blue; Proband 2, red; Proband 3, orange. Standard growth curves from World Health Organization data(WHO Multicentre Growth Reference Study Group, 2006)

(E) Homozygosity mapping in Family 1 identifies a single shared identical-by-descent (IBD) locus on Chromosome 16.

(F-F'') Sanger sequencing traces showing homozygous mutations in probands and heterozygous changes in the unaffected parents and siblings.

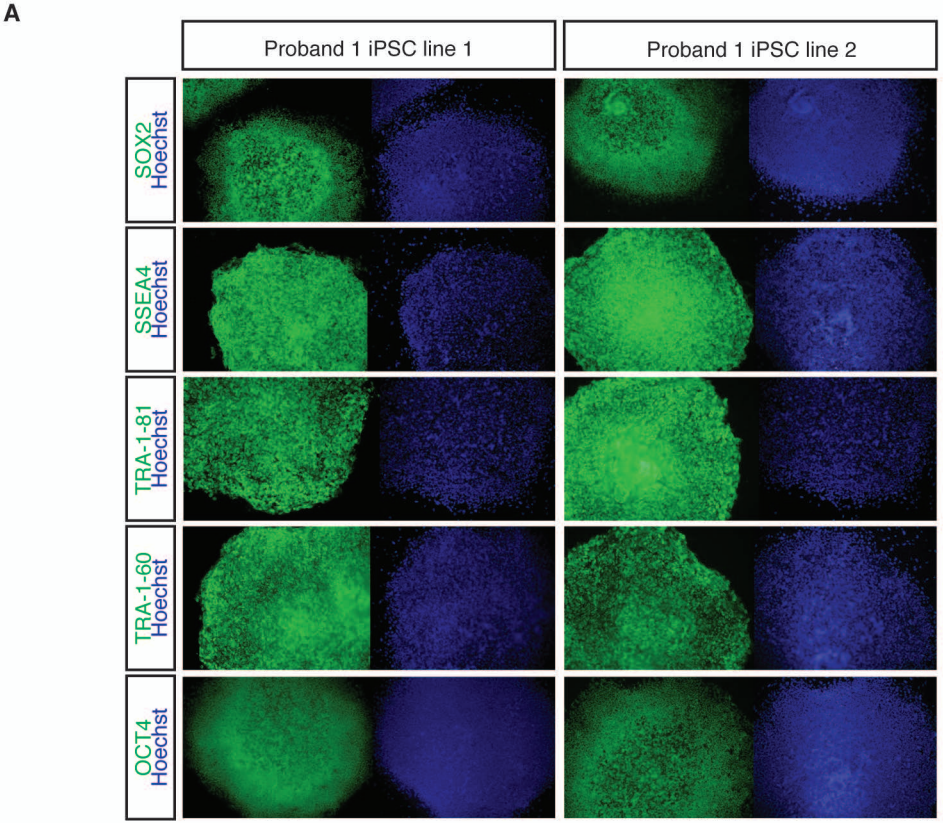

A
Proband 1
11 years
Proband 1 cousin
10 years
Proband 1 cousin
5 months
Proband 3
11 months
Proband 3
11 months
B
Head Circumference (cm)
Age (months)
Percentile
99.9
99
95
75
50
25
5
1
0.1
Proband 1
Proband 2
Proband 3
Proband 1
cousin
C
Height (cm)
Age (months)
Percentile
99.9
99
95
75
50
25
5
1
0.1
Proband 2
Proband 3
D
Weight (kg)
Age (months)
Percentile
99.9
99
95
75
50
25
5
1
0.1
Proband 1
Proband 1
cousin
Proband 1
cousin
Proband 2
Proband 3
E
Homozygosity mapping
IBD block size (cM)
Genome position
55Mb
64Mb
Chr16
Chr16: 55,216,971-64,204,977
Var: A>G
KATNB1
c.1A>G;1A>G
KATNB1
p.Met1_Ile56del;Met1_Ile56del
F
Control
Proband 1
Father
Mother
Unaffected sibling
c.1A>G
F'
Control
Proband 2
Father
Mother
c.97G>T
F''
Control
Proband 3
Father
c.432+1G>A

Supplemental Movie S1, related to Figure 1. MRI scans of Proband 2. Full sagittal, axial and coronal inversion recovery MRI sequences of Proband 2 show reduced brain size, simplification of gyral folding pattern, enlarged lateral ventricles posteriorly and thinning of the corpus callosum, with sparing of the cerebellum, basal ganglia, thalamus and brainstem.

Supplemental Figure S2, related to Figure 2. Characterization of iPSC colonies and Family 1 mutation.

(A) Characterization of representative human iPSC colonies derived from Proband 1. Colonies with ES-like morphology were manually picked 4-5 weeks post-infection. Several representative clones were characterized. These clones expressed pluripotency markers, including surface antigens TRA1-60, TRA-1-81 and SSEA4, as well as the nuclear transcription factors OCT4 and SOX2.

(B) Initiator codon mutation in Family 1 is predicted to result in N-terminal truncation of 56 amino acids. Asterisks, conserved residues.



B

```
                1         10        20        30        40        50        60        70
KATNB1          |         |         |         |         |         |         |         |
                **    * * *********   *** ***   *** **** **** *   **** ********
Zebrafish       MALTNTTITSWKLQEIVAHSSNVSSLVLGKSSGRLLATGGEDCRVNIWAVSKPNCIMSLTGHTSAVGCIQ..
Xenopus         MAAAVVTKTAWKLQEIVAHSSNVSSLVLGKSTGRLLATGGDDCRVNVWSVNKPNCVMSLTGHTTPIESLQ..
Chicken         MAAPSPTKTTWKLQEIVAHGCSVSSVVLGRSSGRLVATGGDDCRVHLWSVNKPNCIMSLTGHTTPVESVR..
Mouse           MATPVVTKTAWKLQEIVAHASNVSSLVLGKASGRLLATGGDDCRVNLWSINKPNCIMSLTGHTSPVESVR..
Human           MATPVVTKTAWKLQEIVAHASNVSSLVLGKASGRLLATGGDDCRVNLWSINKPNCIMSLTGHTSPVESVR..
p.Met1_Ile56del .............     ....................................MSLTGHTSPVESVR..
```

Supplemental Figure S3, related to Figure 4. Sequences and characterization of zebrafish *katnb1* mutant alleles.

(A) Multiple sequence alignment of the human, mouse and zebrafish *Katnb1*, showing a high level of conservation of the protein sequence between the species.

(B) Summary of identified TALEN-induced mutations in zebrafish *katnb1*. Genomic region surrounding the exon 6 - intron 6 boundary in zebrafish is highlighted. TALEN recognition sites are marked with blue boxes; exonic sequence (capital letters); intronic sequence (lowercase). Wild-type sequence is listed above and the sequences of identified alleles are detailed below; deleted base pairs (red dashes); inserted base pairs (red letters). The effect on the genomic sequence is summarized to the right of each sequence: - deletion; + insertion; presence (+) or absence (Δ) of the splice site; allele names; frequency of the allele in 288 fish analyzed.

(C) Progeny from homozygous female carriers of *katnb1* mutations develop normally through epiboly. Summary of phenotypes appearing at 24 hpf ranging from mild microcephaly, anencephaly and disrupted embryos.

A CLUSTAL W multiple sequence alignment

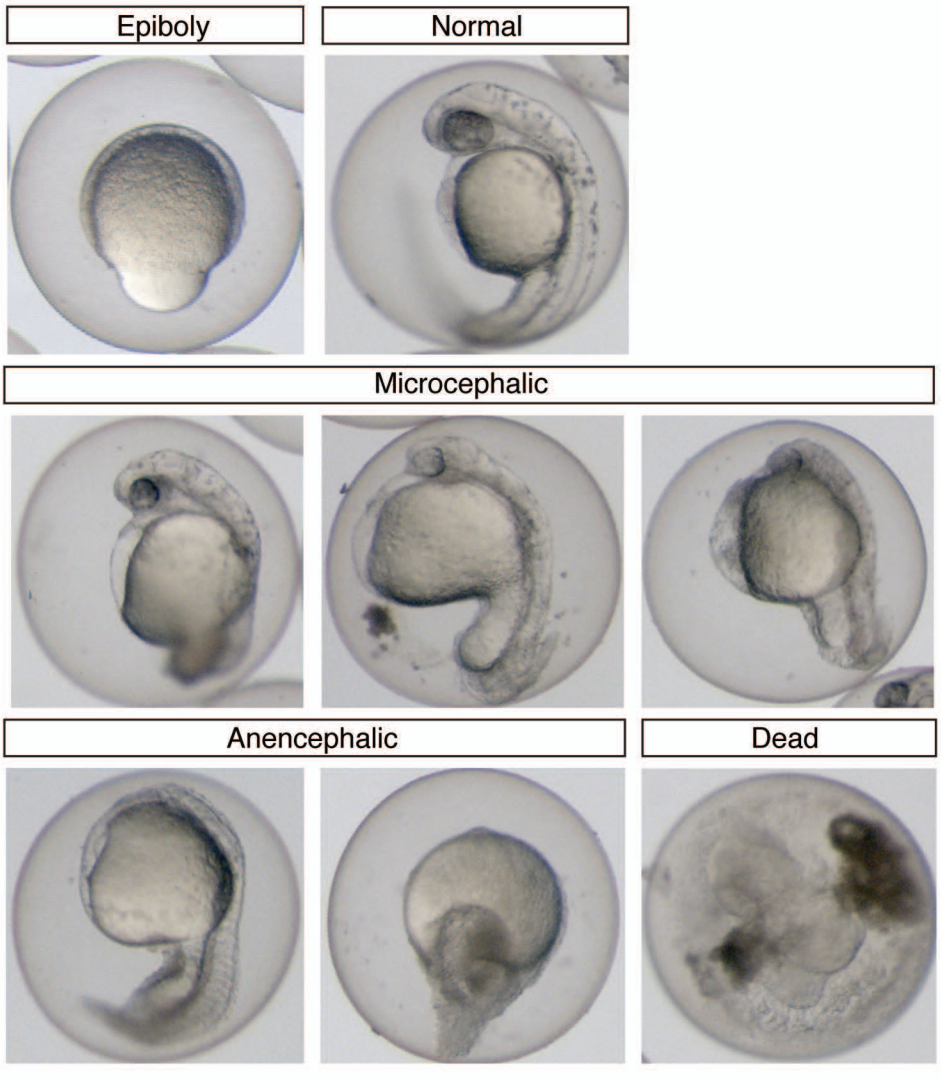
```
hs_KATNB1   MATPVVTKTAWKLQEIVAHASNVSSLVLGKASGRLLATGGDDCRVNLWSINKPNCIMSLTGHTSPVESVRLNTPEELIVAGSQSGSIRVWDLEAAKILRT
mm_KATNB1   MATPVVTKTAWKLQEIVAHASNVSSLVLGKASGRLLATGGDDCRVNLWSINKPNCIMSLTGHTSPVESVRLNTPEELIVAGSQSGSIRVWDLEAAKILRT
dr_Katnb1   MALTNTTITSWKLQEIVAHSSNVSSLVLGKSSGRLLATGGEDCRVNIWAVSKPNCIMSLTGHTSAVGCIQFNSSEERVVAGSLSGSLRLWDLEAAKILRT

hs_KATNB1   LMGHKANICSLDFHPYGEFVASGSQDTNIKLWDIRRKGCVFRYRGHSQAVRCLRFSPDGKWLASAADDHTVKLWDLTAGKMMSEFPGHTGPVNVVEFHPN
mm_KATNB1   LMGHKANICSLDFHPYGEFVASGSQDTNIKLWDIRRKGCVFRYRGHSQAVRCLRFSPDGKWLASAADDHTVKLWDLTAGKMMSEFPGHTGPVNVVEFHPN
dr_Katnb1   LMGHKASISSLDFHPMGEYLASGSVDSNIKLWDVRRKGCVFRYKGHTQAVRCLAFSPDGKWLASASDDSTVKLWDLIAGKMITEFTSHTSAVNVVQFHPN
                                                  Exon6

hs_KATNB1   EYLLASGSSDRTIRFWDLEKFQVVSCIEGEPGPVRSVLFNPDGCCLYSGCQDSLRVYGWEPERCFDVVLVNWGKVADLAICNDQLIGVAFSQSNVSSYVV
mm_KATNB1   EYLLASGSSDRTIRFWDLEKFQVVSCIEGEPGPVRSVLFNPDGCCLYSGCQDSLRVYGWEPERCFDVVLVNWGKVADLAICNDQLIGVAFSQSNVSSYVV
dr_Katnb1   EYLLASGSADRTVKLWDLEKFNMIGSSEGETGVVRSVLFNPDGSCLYSGSENTLRVYGWEPDRCFDVVHVGWGKVSDLAISNNQMIAVSYSHTNVSWYVV

hs_KATNB1   DLTRVTRTGTVARDPVQDHRPLAQPLPNPSAPLRRIYERPSTTCSKPQRVKQNSES-------ERRSPSSEDDRDERESRAEIQNAEDYNEIFQPKNSIS
mm_KATNB1   DLTRVTRTGTVTQDPVQANQPLTQQTPNPGVSLRRIYERPSTTCSKPQRVKHNSES-------ERRSPSSEDDRDERESRAEIQNAEDYNEIFQPKNSIS
dr_Katnb1   DLNRVKKSGSVIQGLIQD-KPIPAPSSALGTTLRRNYERPTTSCTG-QEMKQSSEADRRSPEGERRSPSSEDEKEDKESSAEITNPEDYKEIFQPRSVIS

hs_KATNB1   RTPPRRSEPFPAPPEDDAATAKEAAKPSPAMDVQFP---VPNLEVLPRPPVVASTPA----------PKAEPAIIP
mm_KATNB1   RTPPRRSEPFPAPPEDDAATVKEVSKPSPAMDVQLPQLPVPNLEVPARPSVMTSTPA----------PKGEPDIIP
dr_Katnb1   RTPPKTTEPFPAPLEHSFS-ESVLEKPGPVVKIVTP---VIDRAGQLKGPITSSTPVQRVEPTVIAAAPRPVAVVTTSASSPSRPVVNTTKPKPSTGIIL

hs_KATNB1   ATRNEPIGLKASDFLPAVKIPQQAELVDEDAMSQIRKGHDTMCVVLTSRHKNLDTVRAVWTMGDIKTSVDSAVAINDLSVVVDLLNIVNQKASLWKLDLC
mm_KATNB1   ATRNEPIGLKASDFLPAVKVPQQAELVDEDAMSQIRKGHDTMFVVLTSRHKNLDTVRAVWTTGDIKTSVDSAVAINDLSVVVDLLNIVNQKASLWKLDLC
dr_Katnb1   STRNEPIGLNAGDFLSHARNAKASAMGDEEALAQIRKGHDTMCVMLSSRSKNLDSVRSVWASGDVKTSLDSAVSMNDLSIVVDVLNIINLKPSLWKLDLC

hs_KATNB1   TTVLPQIEKLLQSKYESYVQTGCTSLKLILQRFLPLITDMLAAPPSVGVDISREERLHKCRLCYKQLKSISGLVKSKSGLSGRHGSTFRELHLLLMASL-D
mm_KATNB1   TTVLPQIEKLLQSKYESYVQTGCTSLKLILQRFLPLITDILAAPPSVGVDISREERLHKCRLCFKQLKSISGLVKSKSGLSGRHGSAFRELHLLLMASL-D
dr_Katnb1   TSILPQIEELLQSRYESYVQTGCMSLKLILKRFWPLISDTLNAPPSVGVDITREERHQKCKACYKQLKNLSNVVKNRAEQVGRHGSTFRELQLLLMAPLDY
```

B

L1- recognition sequence — R1- recognition sequence

TTGTGGGATGTGAGAAGGAAAGGATGTGTGTTCAGATACAAGgtaagcactaagctgactttgtgt	Reference		Splice Site	Allele	Frequency
TTGTGGGATGTGAGAAGGAAAGGATGTGTGTT------ACAAGgtaagcactaagctgactttgtgt	- 5		+		8x
TTGTGGGATGTGAGAAGGAAAGGATGTGT---------- aagcactaagctgactttgtgt	- 15		Δ		4x
TTGTGGGATGTGAGAAGGAAAGGATGTGTGT------ACAAGgtaagcactaagctgactttgtgt	- 6		+	*mh101*	7x
TTGTGGGATGTGAGAAGGAAA---------- agctgactttgtgt	- 31		Δ		1x
TTGTGGGATGTGAGAAGGAAAGGATGTGTGTTAAGCACTAAGGTAAGCACTAAGGTAC	+ 19	(+ 26, - 7)	+		1x
TTGTGGGATGTGAGAAGGAAAGGATGCGTGTTACAAGGTACACAAGgtaagcactaagctga	+ 4	(+ 9, - 5)	+		6x
TTGTGGGATGTGAGAAGGAAAGGATGTGTGTAAG----CAAGgtaagcactaagctgactttgtgt	- 4	(+ 3, - 7)	+	*mh102*	32x
TTGTGGGATGTGAGAAGGAAAGGATGTGTT---------- taagctgactttgtgt	- 20	(+ 1, - 21)	Δ	*mh103*	14x
TTGTGGGATGTGAGAAGGAAAGGATGTGTGTT----TACAAGgtaagcactaagctgactttgtgt	- 4		+		5x
TTGTGGGATGTGAGAAGGAAAGGATGTGTGTTAATATTATACAAGgtaagcactaagctgacttt	+ 3	(+ 6, - 3)	+		6x
TTGTGGGATGTGAGAAGGAAAGGATGTGTGTT---------- agcactaagctgactttgtgt	- 13		Δ		4x
TTGTGGGATGTGAGAAGGAAAGGATGT---------- CAAGgtaagcactaagctgactttgtgt	- 11		+		4x
TTGTGGGATGTGAGAAGGAAAGGATGTGTGTTCAAGATACAAGgtaagcactaagctgactttgtg	+ 1		+		1x
TTGTGGGATGTGAGAAGGAAAGGATGTGTGTTAAGCACTAAGGTAAGCACTAAGGTAC	+ 19	(+ 23, - 4)	+		1x
TTGTGGGATGTGAGAAGGAAAGGATGTGTGT---------- AGgtaagcactaagctgactttgtgt	- 9		+		1x

C



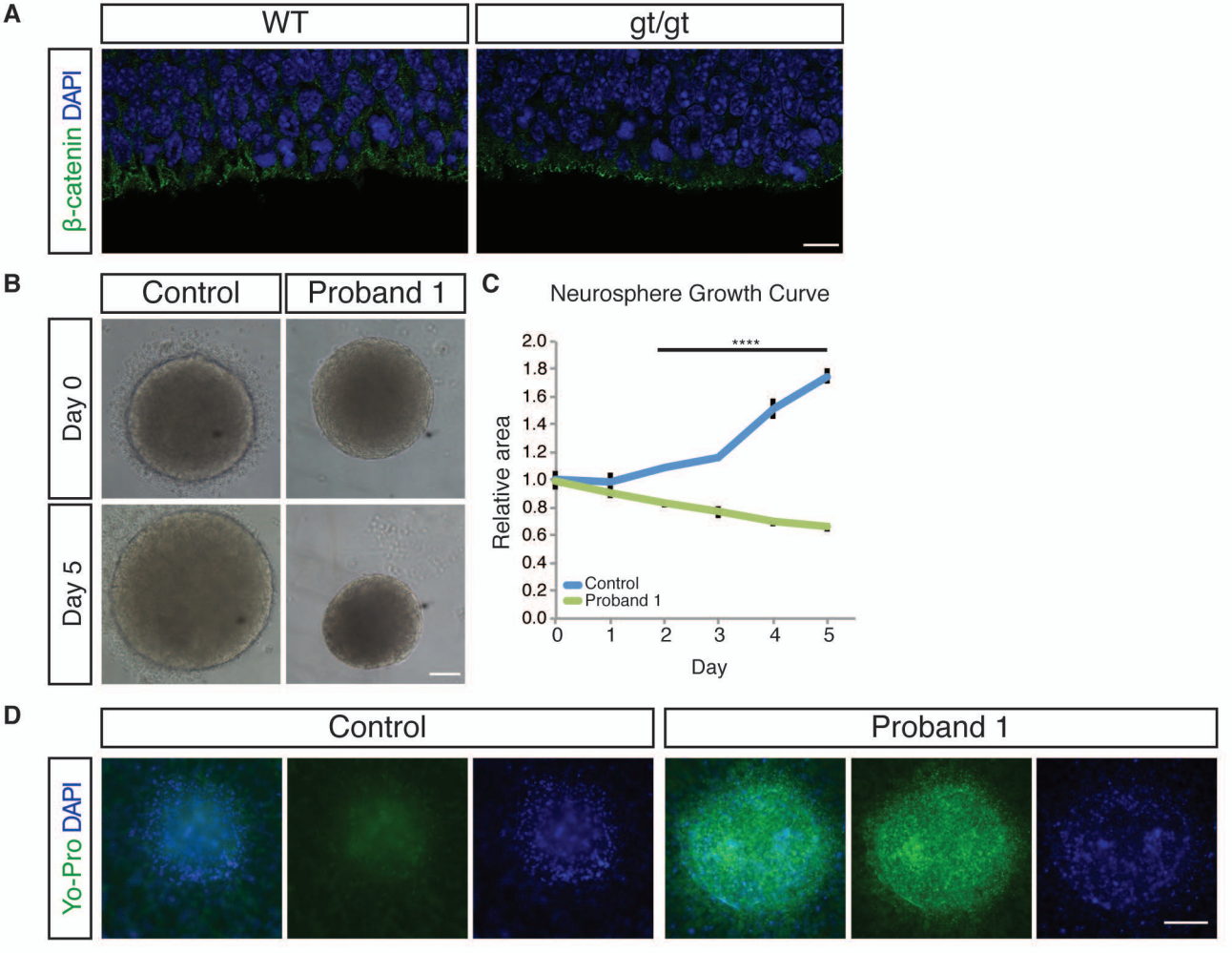
Supplemental Figure S4, related to Figure 5. Katanin p80 in neural progenitor cells.

(A) Apical complex marker, beta-catenin, localizes to the ventricular surface in both wild-type and gene-trap developing mouse cortex, indicating intact cellular polarity. Scale bar, 10 um.

(B) Neurospheres derived from Proband 1 iPSCs fail to grow. Scale bar, 100 um.

(C) Growth curves for neurospheres derived from Proband 1 and control. Two-way ANOVA, $p<0.0001$.

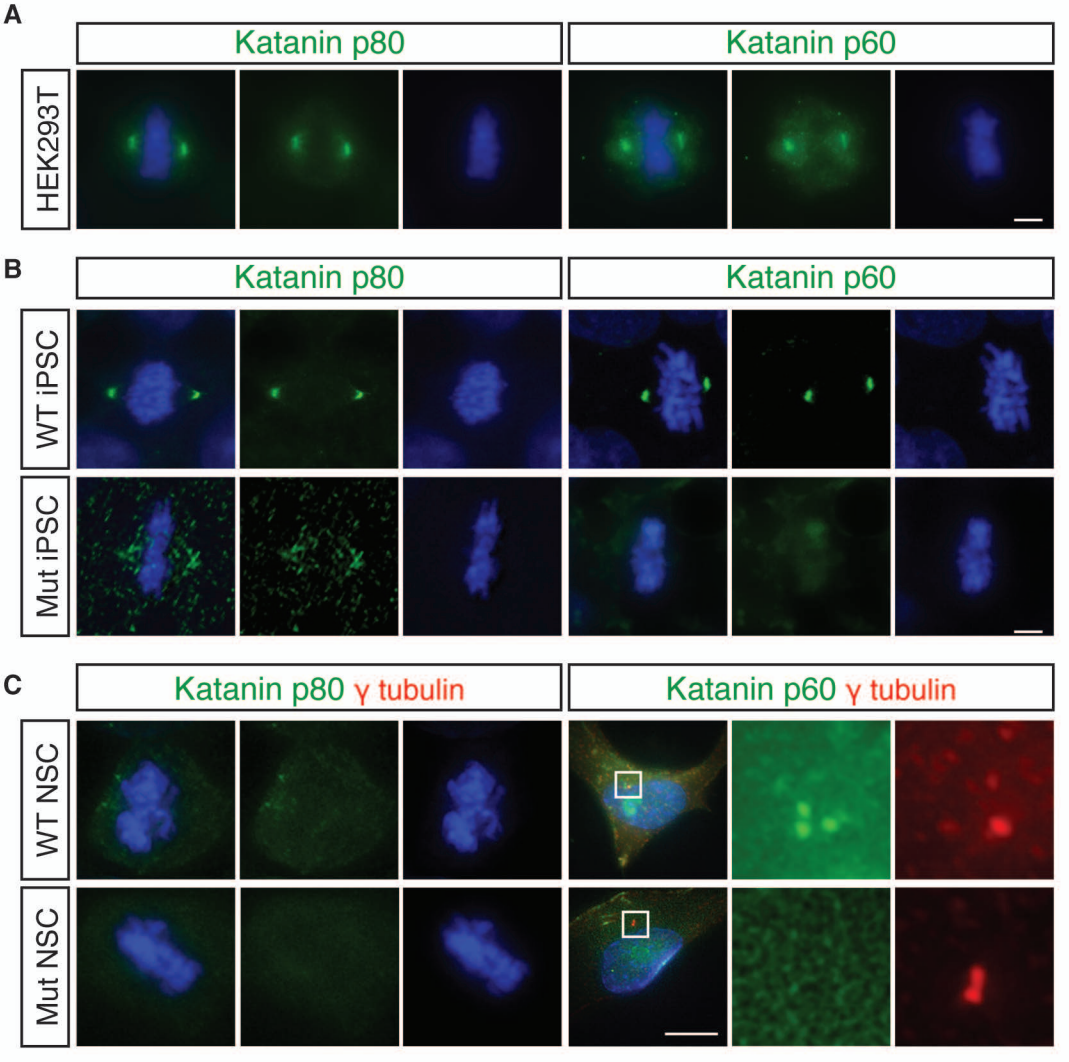
(D) Neurospheres derived from Proband 1 show abundant apoptosis compared to control cells. Scale bar, 100 um.


A
WT
gt/gt
β-catenin DAPI
B
Control
Proband 1
Day 0
Day 5
C
Neurosphere Growth Curve
****
Relative area
2.0
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
Control
Proband 1
0
1
2
3
4
5
Day
D
Control
Proband 1
Yo-Pro DAPI

Supplemental Figure S5, related to Figure 6. Endogenous katanin complex localizes to the centrosome and spindle pole.

(A) Endogenous katanin p80 and p60 localize to the spindle poles of mitotic HEK293T cells. Scale bar, 10 um.

(B-C) Endogenous katanin p80 and p60 localize to the spindle poles of mitotic iPS cells (B), and to the centrosome of interphase neural stem cells (NSCs) differentiated from iPSCs (C). Localization is lost in mutant iPSCs and NSCs derived from Proband 1. Scale bar, 10 um.


A
Katanin p80
Katanin p60
HEK293T
B
Katanin p80
Katanin p60
WT iPSC
Mut iPSC
C
Katanin p80 γ tubulin
Katanin p60 γ tubulin
WT NSC
Mut NSC

Supplemental Figure S6, related to Figure 7. Ultrastructural analysis of centrioles and cytoskeletal immunohistochemistry.

(A) Gene-trap MEF centrosomes contain multiple unpaired centrioles by transmission electron microscopy. Arrowhead, centriole. Scale bar, 200 nm.

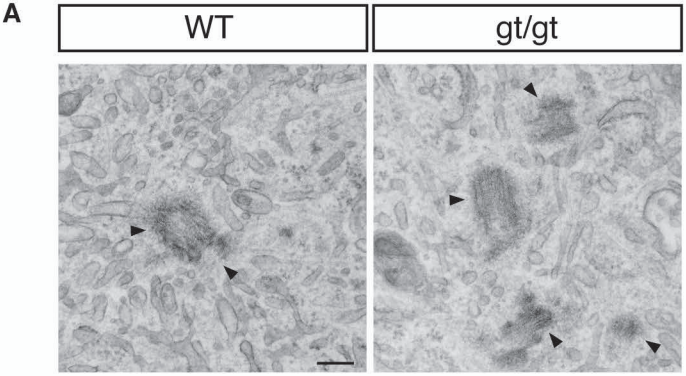
(B) Cross-section of gene-trap MEF centriole shows normal 9-fold symmetry. Arrowhead, centriole. Scale bar, 200 nm.

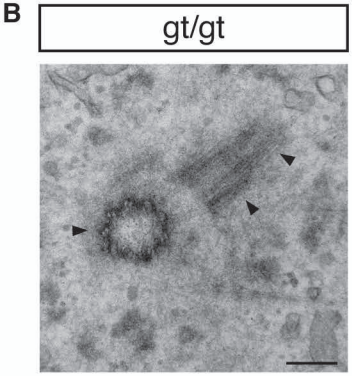
(C) Homozygous gt/gt MEFs fail to proliferate. Fisher's exact test, p=0.03.

(D) Gene-trap MEF cytoskeleton is comprised of increased acetylated tubulin. Scale bar, 10 um.

(E) Gene-trap MEFs exhibit increased staining for CAMSAP2, a microtubule minus-end stabilizing protein. Scale bar, 10 um.

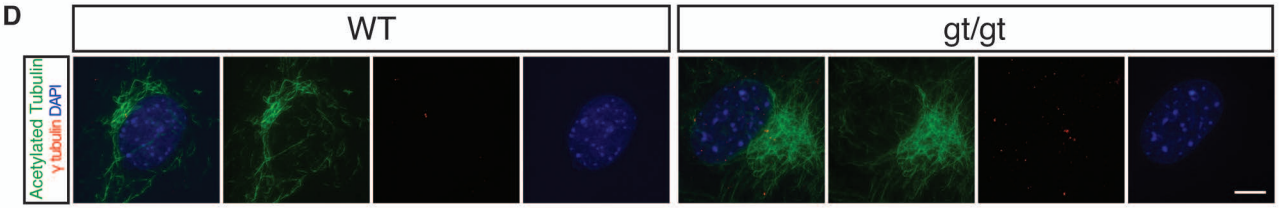
(F) Gene-trap MEFs show increased linear staining for EB1, a microtubule plus-end protein, compared to the shorter, more puncta-like staining in wild-type MEFs. Scale bar, 10 um.

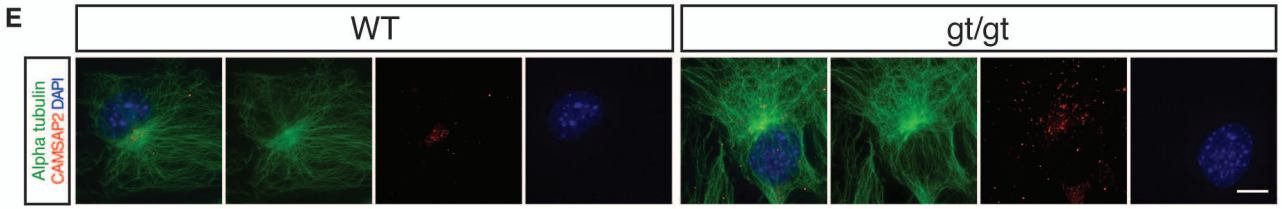


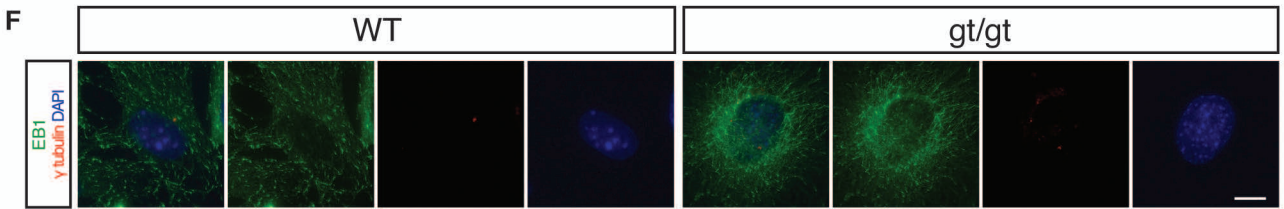


C

	WT	gt/+	gt/gt	
MEF line grew	11	15	4	
MEF line failed to grow	0	0	2	*

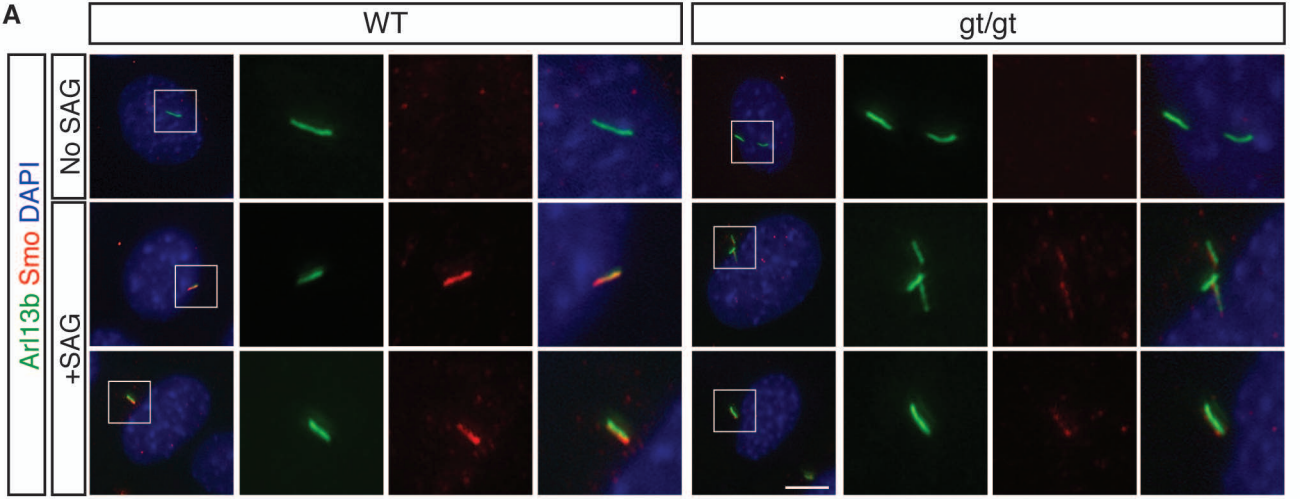






Supplemental Figure S7, related to Figure 8. Cilia in neural progenitor cells.

(A) After stimulation by Smoothened agonist (SAG), gt/gt MEFs show reduced Smoothened (Smo) localization to the cilium, indicating a deficit in Sonic hedgehog signaling (right). Wild-type MEFs robustly relocate Smo to the cilium after SAG stimulation (left). Scale bar, 10 um.


A
WT
gt/gt
Arl13b Smo DAPI
No SAG
+SAG

Supplemental Table S1, related to Figure 1. Clinical summary of affected individuals.

Individual	Proband 1	Proband 1 cousin	Proband 1 cousin	Proband 2	Proband 3
Gender	Male	Female	Male, deceased at 5 months	Male	Male
Parental consanguinity	Yes	Yes	Yes	Yes	None reported, distant relation suspected based on homozygosity mapping
Prenatal and birth history	Full term normal spontaneous vaginal delivery	Born by Caesarian section, no NICU admission	Unknown	Unremarkable, full term normal spontaneous vaginal delivery	Microcephaly noted in 2nd trimester, full term normal spontaneous vaginal delivery
Head circumference at birth	30 cm (-3.5 SD)	27 cm (-5.9 SD)	27 cm (-5.9 SD)	29.5 cm (-3.9 SD)	Unknown
Weight at birth	3.7 kg (75th percentile)	3 kg (25th percentile)	Unknown	3.3 kg (50th percentile)	3.2 kg (40th percentile)
Length at birth	Unknown	Unknown	Unknown	47 cm (-1.5 SD)	Unknown
Age at first evaluation	1 year old	9 months	Unknown	3 months	8 months
Head circumference at first evaluation	38 cm (-6.3 SD)	32 cm (-10.3 SD)	Unknown	34.3 cm (-5.3 SD)	34 cm (-8.5 SD)
Weight at first evaluation	Unknown	Unknown	Unknown	7.2 kg (85th percentile)	6.8 kg (2nd percentile)
Length at first evaluation	Unknown	Unknown	Unknown	59.5 cm (-0.9 SD)	63 cm (-3.5 SD)
Age at most recent evaluation	10 years	10 years	Not applicable	3 years	4 years
Head circumference at most recent evaluation	42.5 cm	41 cm	Not applicable	39 cm (-7.4 SD)	39 cm (-7.7 SD)
Weight at most recent evaluation	22 kg	24 kg	Not applicable	13 kg (20th percentile)	11.8 kg (0.5th percentile)
Length at most recent evaluation	118.5 cm	120 cm	Not applicable	90 cm (-1.6 SD)	Unknown
Neurologic findings	Spasticity in achilles tendon bilaterally	Increased tone in achilles tendon bilaterally with exaggerated deep tendon reflexes (+3) in knees and ankles, no clonus, babinski is downwards	Not applicable	Spastic quadriplegia	Mild hypertonia, increased deep tendon reflexes
Developmental history	10 years old: sitting alone, does not walk, no words (global delay)	10 years old: sitting alone, does not walk, no words	Not applicable	Severe and global delay, no speech	Head control at 1 year, rolled at 1.25 years, unable to sit unsupported at 4 years, no speech
Seizures	9 years old: seizures in the form of uprolling of eyes with decreased tone all over. Repeated after 6 months, treated with valproic acid	Developed seizures only once at the age of 7 years with cyanosis , uprolling of eyes and tonic contraction of upper limbs and lower limbs	Not applicable	Yes, generalized tonic-clonic, onset 1 year, treated with Keppra	Yes, onset at 21 months, treated with phenobarbital
Hearing	Grossly normal	Grossly normal	Not applicable	Grossly normal	Mild hearing loss at 45dB
Vision	Grossly normal	Grossly normal	Not applicable	Grossly normal (examined in a private clinic, report not available)	Bilateral vision loss with disc pallor and macular lesions
Dysmorphic features	Microcephaly with decreased cranio-facial ratio and small cranial vault, large ears, synephrous	Microcephaly with decreased cranio-facial ratio and small cranial vault, large ears, synephrous, bulbous nose, almond eyes, big mouth	Not applicable	Facial features consistent with microcephaly, single palmar creases, secondary craniosynostosis	Sloping forehead, round face, arched eyebrows, small nose with anteverted nares, thin upper lip, widely spaced teeth, increased internipplary distance, tapered fingers, small testes
Brain imaging	MRI at 1 year and 11 years	MRI at 10 years old	Not applicable	MRI at 17 months	MRI at 11 months
cerebral cortex	Simplification of the gyral pattern with thickened cortex (packygyria), broad flat gyri, smooth grey-white matter interface, brain atrophy	Simplification of the gyral pattern with thickened cortex (packygyria), broad flat gyri, smooth grey-white matter interface, shallow sylvian fissure		Moderately to severely simplified gyral pattern, normal cortical thickness	Severely simplified gyral pattern, normal cortical thickness
basal ganglia and thalami	Grossly normal	Grossly normal		Grossly normal	Grossly normal basal ganglia, small thalami
corpus callosum	Partial agenesis of body and splenium (genu present)	Agenesis of the body and splenium (genu present), hypomyelination is noted		Partial agenesis (genu and body present)	Partial agenesis (only genu present)
ventricles	Prominent ventricles enlarged ventricles (colpocephaly)	Enlarged ventricle (colpocephaly), the 4th ventricle is communicating with an enlarged cisterna magna, through a slit-like opening finding representing Dandy-Walker variant		Enlarged posteriorly	Enlarged posteriorly, periventricular heterotopia
cerebellum and brain stem	Atrophy involving the cerebellum and cerebellar vermis	Atrophy involving the cerebellum and cerebellar vermis		Grossly normal, enlarged cisterna magna	Grossly normal, enlarged cisterna magna
Other findings	None	None	Not applicable	None	Oral pharyngeal dysphagia, elevated liver enzymes
Normal screening tests	Blood	Not available	Not applicable	Abdominal ultrasound	Karyotype, toxoplasma, rubella, cytomegalovirus, herpes simplex, HIV, complete blood count, serum electrolytes, renal and thyroid function tests, creatine kinase, abdominal ultrasound

Supplemental Table S2, related to Figure 4. Plasmid and protein sequences of TALENs used to generate zebrafish *katnb1* mutants.

katnb1-L	Plasmid sequence: gctgcttcgcgatgtacgggccagatatacgcgtggctccggtgcccgtcagtgggcagagcgcacatcgcccacagtccccgagaagttgggggggagggtcggcaattga accggtgcctagagaaggtggcgcggggtaaactgggaaagtgatgtcgtgtactggctccgccttttcccgagggtgggggagaaccgtatataagtgcagtagtcgccgtg aacgttcttttttcgcaacgggtttgccgccagaacacaatttaggtgacactatagaagctagcgtttaaacttaagctgatccactagtccagtgtggtggaattcgccatggacta caaagaccatgacggtgattataaagatcatgacatcgattacaaggatgacgatgacaagatggcccccaagaagaagaggaaggtgggcatccacggggtacccatgg tagatttgagaactttgggatattcacagcagcagcaggaaaagatcaagcccaaagtgaggtcgacagtcgcgcagcatcacgaagcgctggtgggtcatgggtttacacat gcccacatcgtagccttgtcgcagcacccctgcagcccttggcacggtcgccgtcaagtaccaggacatgattgcggcgttgccggaagccacacatgaggcgatcgtcggtgt ggggaaacagtggagcggagcccgagcgcttgaggccctgttgacggtcgcgggagagctgagagggcctccccttcagctggacacgggccagttgctgaagatcgcga agcggggaggagtcacggcggtcgaggcggtgcacgcgtggcgcaatgcgctcacgggagcacccctcaacCTGACCCCAGAGCAGGTCGTGGCCATT GCCTCGAATGGAGGGGGCAAACAGGCGTTGGAAACCGTACAACGATTGCTGCCGGTGCTTTGTCAGGCACACGGCCTC ACTCCGGAACAAGTGGTCGCCATTGCCTCGAATGGAGGGGGCAAACAGGCGTTGGAAACCGTACAACGATTGCTGCCG GTGCTGTGCCAAGCGCACGGTTTAACCCCAGAGCAGGTCGTGGCCATTGCCTCGAATGGAGGGGGCAAACAGGCGTTG GAAACCGTACAACGATTGCTGCCGGTGCTTTGTCAGGCACACGGCCTCACTCCGGAACAAGTGGTCGCAATCGCGAGC AATAACGGCGGAAAACAGGCTTTGGAAACGGTGCAGAGGCTCCTTCCAGTGCTGTGCCAAGCGCACGGCCTCACCCCA GAGCAGGTCGTGGCAATCGCCTCCAACATTGGCGGGAAACAGGCACTCGAGACTGTCCAGCGCCTGCTTCCCGTGCTT TGTCAGGCACACGGCCTCACTCCGGAACAAGTGGTCGCGATCGCAAGCCACGACGGAGGAAAGCAAGCCTTGGAAACA GTACAGAGGCTGTTGCCTGTGC TGTGCCAAGCGCACGGATTAACCCCAGAGCAGGTCGTGGCAATCGCCTCCAACATTGGCGGGAAACAGGCACTCGAGA CTGTCCAGCGCCTGCTTCCCGTGCTTTGTCAGGCACACGGCCTCACTCCGGAACAAGTGGTCGCAATCGCCTCCAACAT TGGCGGGAAACAGGCACTCGAGACTGTCCAGCGCCTGCTTCCCGTGCTGTGCCAAGCGCACGGACTTACCCCAGAGCA GGTCGTGGCCATTGCCTCGAATGGAGGGGGCAAACAGGCGTTGGAAACCGTACAACGATTGCTGCCGGTGCTGTGCCA AGCGCACGGCCTTACCCCAGAGCAGGTCGTGGCGATCGCAAGCCACGACGGAGGAAAGCAAGCCTTGGAAACAGTAC AGAGGCTGTTGCCTGTGCTGTGCCAAGCGCACGGACTAACCCCAGAGCAGGTCGTGGCGATCGCAAGCCACGACGGA GGAAAGCAAGCCTTGGAAACAGTACAGAGGCTGTTGCCTGTGCTGTGCCAAGCGCACGGGCTCACCCCAGAGCAGGTC GTGGCAATCGCCTCCAACATTGGCGGGAAACAGGCACTCGAGACTGTCCAGCGCCTGCTTCCCGTGCTTTGTCAGGCA CACGGCCTCACTCCGGAACAAGTGGTCGCAATCGCGAGCAATAACGGCGGAAAACAGGCTTTGGAAACGGTGCAGAGG CTCCTTCCAGTGCTGTGCCAAGCGCACGGGCTAACCCCAGAGCAGGTCGTGGCCATTGCCTCGAATGGAGGGGGCAAA CAGGCGTTGGAAACCGTACAACGATTGCTGCCGGTGCTTTGTCAGGCACACGGCCTCACTCCGGAACAAGTGGTCGCA ATCGCGAGCAATAACGGCGGAAAACAGGCTTTGGAAACGGTGCAGAGGCTCCTTCCAGTGCTGTGCCAAGCGCACGGA ctcacgcctgagcaggtagtggctattgcatccaatggaggggcagacccgcactggagtcaatcgtggcccagctttcgaggccggaccccgcgctggccgcactcacta atgatcatcttgtagcgctggcctgcctcggcggacgacccgccttggatgcggtgaagaaggggctcccgcacgcgcctgcattgattaagcggaccaacagaaggattccc gagaggacatcacatcgagtggcaggatcccagctggtgaagagcgagctggaggagaagaagtccgagctgcggcacaagctgaagtacgtgccccacgagtacatcg agctgatcgagatcgccaggaacagcacccaggaccgcatcctggagatgaaggtgatggagttcttcatgaaggtgtacggctacaggggaaagcacctgggcggaagc agaaagcctgacggcgccatctatacagtgggcagccccatcgattacggcgtgatcgtggacacaaaggcctacagcggcggctacaatctgcctatcggccaggccgac gagatggagagatacgtggaggagaaccagaccccggaataagcacctcaaccccaacgagtggtggaaggtgtaccctagcagcgtgaccgagttcaagttcctgttcgtg agcggccacttcaagggcaactacaaggcccagctgaccaggctgaaccacatcaccaactgcaatggcgccgtgctgagcgtggaggagctgctgatcggcggcgagat gatcaaagccggcaccctgacactggaggaggtgcggcgcaagttcaacaacggcgagatcaacttcagatcttgataactcgagtttaaacccgctgatcagcctcgactgt gccttctagttgccagccatctgttgtttgcccctccccgtgccttccttgaccctggaaggtgccactcccactgtcctttcctaataaaatgaggaaattgcatcgcattgtctgagta ggtgtcattctattctgggggggtggggtggggcaggacagcaagggggaggattgggaagacaatagcaggcatgctggggatgcggtgggctctatgggcggccgccttcta ctgggcggttttatggacagcaagcgaaccggaattgccagctggggcgccctctggtaaggttgggaagccctgcaaagtaaactggatggctttcttgccgccaaggatctg atggcgcaggggatcaagctctgatcaagagacaggatgaggatcgtttcgcatgattgaacaagatggattgcacgcaggttctccggccgcttgggtggagaggctattcgg ctatgactgggcacaacagacaatcggctgctctgatgccgccgtgttccggctgtcagcgcaggggcgcccggttctttttgtcaagaccgacctgtccggtgccctgaatgaac tgcaagacgaggcagcgcggctatcgtggctggccacgacgggcgttccttgcgcagctgtgctcgacgttgtcactgaagcgggaagggactggctgctattgggcgaagtg ccggggcaggatctcctgtcatctcaccttgctcctgccgagaaagtatccatcatggctgatgcaatgcggcggctgcatacgcttgatccggctacctgcccattcgaccacca agcgaaacatcgcatcgagcgagcacgtactcggatggaagccggtcttgtcgatcaggatgatctggacgaagagcatcacgggctcgcgccagccgaactgttcgccag gctcaaggcgagcatgcccgacggcgaggatctcgtcgtgacccatggcgatgcctgcttgccgaatatcatggtggaaaatggccgcttttctggattcatcgactgtggccggct ggtgtggcggaccgctatcaggacatagcgttggctaccgtgatattgctgaagagcttgcgcgaatggctgacgccttctcgtgcttacggtatcgccgcctcccgatgcatcagtg cacttttcggggaaatgtgcgcgaccctatttgtttatttttctaatacatcaaatatgtatccgctcatgagacaataacctgataaatgcttcaataatagcacgtgctaaaacttcattt ttaaatttaaaaggatctaggtgaagatccttttgataatctcatgaccaaaatcctttaacgtgagttttcgttccactgagcgtcagaccccgtagaaaagatcaaaggatcttcttg agatccttttttctgcgcgtaatctgctgcttgcaaacaaaaaaaccaccgctaccagcggtggtttgtttgccggatcaagagctaccaactcttttccgaaggtaactggcttcag cagagcgcagataccaaatactgttcttctagtgtagccgtagttaggccaccacttcaagaactctgtagcaccgcctacatacctcgctctgctaatcctgttaccagtggctgct gccagtggcgataagtcgtgtcttaccgggttggactcaagacgatagttaccggataaggcgcagcggtcgggctgaacgggggttcgt gcacacagcccagcttggagcgaacgacctacaccgaactgagatacctacagcgtgagctatgagaaagcgccacgcttcccgaagggag aaaggcggacaggtatccggtaagcggcagggtcggaacaggagagcgcacgagggagcttccagggggaaacgcctggtatctttatagt cctgtcgggtttcgccacctctgacttgagcgtcgatttttgtgatgctcgtcaggggggcggagcctatggaaaaacgccagcaacgcgg ccttttacggttcctggccttttgctggccttttgctcacatgttctt Protein: MDYKDHDGDYKDHDIDYKDDDDKMAPKKKRKVGIHGVPMVDLRTLGYSQQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIV ALSQHPAALGTVAVKYQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDTGQLLKIAKRGGVTAVEAVHA WRNALTGAPLNLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTP EQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQA LETVQRLLPVLCQAHGLTPEQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAH GLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDG GKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVL CQAHGLTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIA SNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGRPALESIVAQLSRPDPALAALTNDHLVALACLGGRPALDAVKKG LPHAPALIKRTNRRIPERTSHRVAGSQLVKSELEEKKSELRHKLKYVPHEYIELIEIARNSTQDRILEMKVMEFFMKVYGYRGKH LGGSRKPDGAIYTVGSPIDYGVIVDTKAYSGGYNLPIGQADEMERYVEENQTRNKHLNPNEWWKVYPSSVTEFKFLFVSGHF KGNYKAQLTRLNHITNCNGAVLSVEELLIGGEMIKAGTLTLEEVRRKFNNGEINFRS*

katnb1-R	Plasmid sequence: gctgcttcgcgatgtacgggccagatatacggctccggtgcccgtcagtgggcagagcgcacatcgcccacagtccccgagaagttgggggggaggggtcggcaattgaacc ggtgcctagagaaggtggcgcggggtaaactgggaaagtgatgtcgtgtactggctccgccttttcccgagggtgggggagaaccgtatataagtgcagtagtcgccgtgaac gttctttttcgcaacgggtttgccgccagaacacaatttaggtgacactatagaagctagcgtttaaacttaagctgatccactagtccagtgtggtggaattcgccatggactacaa agaccatgacggtgattataaagatcatgacatcgattacaaggatgacgatgacaagatggcccccaagaagaagaggaaggtgggcatccacggggtacccatggtag atttgagaactttgggatattcacagcagcagcaggaaaagatcaagcccaaagtgaggtcgacagtcgcgcagcatcacgaagcgctggtgggtcatgggtttacacatgcc cacatcgtagccttgtcgcagcaccctgcagcccttggcacggtcgccgtcaagtaccaggacatgattgcggcgttgccggaagccacacatgaggcgatcgtcggtgtggg gaaacagtggagcggagcccgagcgcttgaggccctgttgacggtcgcgggagagctgagagggcctccccttcagctggacacgggccagttgctgaagatcgcgaagc ggggaggagtcacggcggtcgaggcggtgcacgcgtggcgcaatgcgctcacgggagcacccctcaacCTGACCCCAGAGCAGGTCGTGGCGATCG CAAGCCACGACGGAGGAAAGCAAGCCTTGGAAACAGTACAGAGGCTGTTGCCTGTGCTTTGTCAGGCACACGGCCTCA CTCCGGAACAAGTGGTCGCAATCGCCTCCAACATTGGCGGGAAACAGGCACTCGAGACTGTCCAGCGCCTGCTTCCCG TGCTGTGCCAAGCGCACGGTTTAACCCCAGAGCAGGTCGTGGCAATCGCGAGCAATAACGGCGGAAAACAGGCTTTGG AAACGGTGCAGAGGCTCCTTCCAGTGCTTTGTCAGGCACACGGCCTCACTCCGGAACAAGTGGTCGCGATCGCAAGCC ACGACGGAGGAAAGCAAGCCTTGGAAACAGTACAGAGGCTGTTGCCTGTGCTGTGCCAAGCGCACGGCCTCACCCCAG AGCAGGTCGTGGCCATTGCCTCGAATGGAGGGGGCAAACAGGCGTTGGAAACCGTACAACGATTGCTGCCGGTGCTTT GTCAGGCACACGGCCTCACTCCGGAACAAGTGGTCGCCATTGCCTCGAATGGAGGGGGCAAACAGGCGTTGGAAACC GTACAACGATTGCTGCCGGTGCTGTGCCAAGCGCACGGATTAACCCCAGAGCAGGTCGTGGCAATCGCCTCCAACATT GGCGGGAAACAGGCACTCGAGACTGTCCAGCGCCTGCTTCCCGTGCTTTGTCAGGCACACGGCCTCACTCCGGAACAA GTGGTCGCAATCGCGAGCAATAACGGCGGAAAACAGGCTTTGGAAACGGTGCAGAGGCTCCTTCCAGTGCTGTGCCAA GCGCACGGACTTACCCCAGAGCAGGTCGTGGCCATTGCCTCGAATGGAGGGGGCAAACAGGCGTTGGAAACCGTACAA CGATTGCTGCCGGTGCTGTGCCAAGCGCACGGCCTTACCCCAGAGCAGGTCGTGGCAATCGCGAGCAATAACGGCGG AAAACAGGCTTTGGAAACGGTGCAGAGGCTCCTTCCAGTGCTGTGCCAAGCGCACGGACTAACCCCAGAGCAGGTCGT GGCGATCGCAAGCCACGACGGAGGAAAGCAAGCCTTGGAAACAGTACAGAGGCTGTTGCCTGTGCTGTGCCAAGCGCA CGGGCTCACCCCAGAGCAGGTCGTGGCCATTGCCTCGAATGGAGGGGGCAAACAGGCGTTGGAAACCGTACAACGATT GCTGCCGGTGCTTTGTCAGGCACACGGCCTCACTCCGGAACAAGTGGTCGCCATTGCCTCGAATGGAGGGGGCAAACA GGCGTTGGAAACCGTACAACGATTGCTGCCGGTGCTGTGCCAAGCGCACGGGCTAACCCCAGAGCAGGTCGTGGCAAT CGCCTCCAACATTGGCGGGAAACAGGCACTCGAGACTGTCCAGCGCCTGCTTCCCGTGCTTTGTCAGGCACACGGCCT CACTCCGGAACAAGTGGTCGCGATCGCAAGCCACGACGGAGGAAAGCAAGCCTTGGAAACAGTACAGAGGCTGTTGCC TGTGCTGTGCCAAGCGCACGGActcacgcctgagcaggtagtggctattgcatcccacgacgggggcagacccgcactggagtcaatcgtggcccagctttcg aggccggaccccgcgctggccgcactcactaatgatcatcttgtagcgctggcctgcctcggcggacgacccgccttggatgcggtgaagaaggggctcccgcacgcgcctg cattgattaagcggaccaacagaaggattcccgagaggacatcacatcgagtggcaggatcccagctggtgaagagcgagctggaggagaagaagtccgagctgcggca caagctgaagtacgtgccccacgagtacatcgagctgatcgagatcgccaggaacagcacccaggaccgcatcctggagatgaaggtgatggagttcttcatgaaggtgtac ggctacaggggaaagcacctgggcggaagcagaaagcctgacggcgccatctatacagtgggcagccccatcgattacggcgtgatcgtggacacaaaggcctacagcg gcggctacaatctgcctatcggccaggccgacgagatgcagagatacgtgaaggagaaccagacccggaataagcacatcaaccccaacgagtggtggaaggtgtaccct agcagcgtgaccgagttcaagttcctgttcgtgagcggccacttcaagggcaactacaaggcccagctgaccaggctgaaccacaaaaccaactgcaatggcgccgtgctga gcgtggaggagctgctgatcggcggcgagatgatcaaagccggcaccctgacactggaggaggtgcggcgcaagttcaacaacggcgagatcaacttctgataactcgagt ctagagggcccgtttaaacccgctgatcagcctcgactgtgccttctagttgccagccatctgttgtttgcccctccccgtgccttccttgaccctggaaggtgccactcccactgtcc tttcctaataaaatgaggaaattgcatcgcattgtctgagtaggtgtcattctattctgggggtggggtggggcaggacagcaagggggaggattgggaagacaatagcaggc atgctggggatgcggtgggctctatgggcggccgccttctactgggcggttttatggacagcaagcgaaccggaattgccagctggggcgccctctggtaaggttgggaagccct gcaaagtaaactggatggctttcttgccgccaaggatctgatggcgcaggggatcaagctctgatcaagagacaggatgaggatcgtttcgcatgattgaacaagatggattgc acgcaggttctccggccgcttgggtggagaggctattcggctatgactgggcacaacagacaatcggctgctctgatgccgccgtgttccggctgtcagcgcaggggcgcccgg ttctttttgtcaagaccgacctgtccggtgccctgaatgaactgcaagacgaggcagcgcggctatcgtggctggccacgacgggcgttccttgcgcagctgtgctcgacgttgtca ctgaagcgggaagggactggctgctattgggcgaagtgccggggcaggatctcctgtcatctcaccttgctcctgccgagaaagtatccatcatggctgatgcaatgcggcggct gcatacgcttgatccggctacctgcccattcgaccaccaagcgaaacatcgcatcgagcgagcacgtactcggatggaagccggtcttgtcgatcaggatgatctggacgaag agcatcacgggctcgcgccagccgaactgttcgccaggctcaaggcgagcatgcccgacggcgaggatctcgtcgtgacccatggcgatgcctgcttgccgaatatcatggtg gaaaatggccgcttttctggattcatcgactgtgccggctggtgtggcggaccgctatcaggacatagcgttggctaccgtgatattgctgaagagcttgcgcgaatggctgacgct tctcgtgcttacggtatcgccgcctcccgatgcatcagtgcacttttcggggaaatgtgcgcgaccctatttgtttatttttctaatacatcaaatatgtatccgctcatgagacaataacc tgataaatgcttcaataatagcacgtgctaaaacttcatttttaaatttaaaaggatctaggtgaagatccttttgataatctcatgaccaaaatcctttaacgtgagttttcgttccactg agcgtcagaccccgtagaaaagatcaaaggatcttcttgagatcctttttttctgcgcgtaatctgctgcttgcaaacaaaaaaaccaccgctaccagcggtggtttgtttgccggat caagagctaccaactctttttccgaaggtaactggcttcagcagagcgcagataccaaatactgttcttctagtgtagccgtagttaggccaccacttcaagaactctgtagcaccg cctacatacctcgctctgctaatcctgttaccagtggctgctgccagtggcgataagtcgtgtcttaccgggttggactcaagacgatagttaccggataaggcgcagcggtcggg ctgaacgggggttcgtgcacacagcccagcttggagcgaacgacctacaccgaactgagatacctacagcgtgagctatgagaaagcgccacgcttcccgaagggagaaa aggcggacaggtatccggtaagcggcagggtcggaacaggagagcgcacgagggagcttccagggggaaacgcctggtatctttatagtcctgtcgggtttcgccacctctg acttgagcgtcgatttttgtgatgctcgtcaggggggcggagcctatggaaaaacgccagcaacgcggccttttacggttcctggccttttgctggccttttgctcacatgttctt Protein: MDYKDHDGDYKDHDIDYKDDDDKMAPKKKRKVGIHGVPMVDLRTLGYSQQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIV ALSQHPAALGTVAVKYQDMIAALPEATHEAIVGVGKQWSGARALEALLTVAGELRGPPLQLDTGQLLKIAKRGGVTAVEAVHA WRNALTGAPLNLTPEQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPE QVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQAL ETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHG LTPEQVVAIASNNGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNNGG KQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNGGGKQALETVQRLLPVLC QAHGLTPEQVVAIASNGGGKQALETVQRLLPVLCQAHGLTPEQVVAIASNIGGKQALETVQRLLPVLCQAHGLTPEQVVAIAS HDGGKQALETVQRLLPVLCQAHGLTPEQVVAIASHDGGRPALESIVAQLSRPDPALAALTNDHLVALACLGGRPALDAVKKGL PHAPALIKRTNRRIPERTSHRVAGSQLVKSELEEKKSELRHKLKYVPHEYIELIEIARNSTQDRILEMKVMEFFMKVYGYRGKHL GGSRKPDGAIYTVGSPIDYGVIVDTKAYSGGYNLPIGQADEMQRYVKENQTRNKHINPNEWWKVYPSSVTEFKFLFVSGHFK GNYKAQLTRLNHKTNCNGAVLSVEELLIGGEMIKAGTLTLEEVRRKFNNGEINF*

Extended Experimental Procedures

Human Subjects

We collected peripheral blood samples from the affected individuals and one or both of their parents after obtaining written informed consent according to the protocols approved by the participating institutions. All human studies were reviewed and approved by the institutional review board of Boston Children's Hospital, Harvard Medical School and all local institutions.

Homozygosity Mapping, Exome Sequencing and Filtering in Family 1

Genomic DNA was extracted from saliva samples using the Oragen Purifier Kit. Proband 1, the affected cousin, and their unaffected parents and siblings were genotyped using Illumina Human OmniExpress 12v1 BeadChips following manufacturer's instructions. Call rates were above 99%. Gender and relationship were verified using Illumina BeadStudio software. Mapping was performed by searching for shared regions that are homozygous and identical-by-descent (IBD) in the 2 affected individuals using custom programs written in the Mathematica (Wofram Research, Inc.) data analysis software. Candidate regions were further refined by exclusion of common homozygous segments with any unaffected family members. The confidence criteria to identify IBD blocks were a minimum of 3 cM and a SNP density over the 5th percentile. Centimorgan distances between SNPs were determined from the HapMap Phase II recombination rate map. We identified a single shared identical-by-descent candidate locus totaling 9 Mb on Chromosome 16 (hg19; chr16:55,216,971-64,204,977).

Whole exome sequencing of Proband 1 was performed using the Illumina TruSeq Exome Enrichment Kit for exome capture using 1 ug of genomic DNA. Illumina

HiSeq2000 High-output mode was used for sequencing as 100 bp paired-end runs at the UCLA Clinical Genomics Centre and at the UCLA Broad Stem Cell Research Centre. 6 samples were pooled before the exome capture and sequenced across 2 flowcell lanes. Sequence reads were aligned to the human reference genome (Human GRCh37 (hg19) build) using Novoalign (v2.07, http://www.novocraft.com/main/index.php). PCR duplicates were identified by Picard (v1.42, http://picard.sourceforge.net/) and Genome Analysis Toolkit (v1.1, http://www.broadinstitute.org/gatk/)(McKenna et al., 2010) was used to re-align indels, recalibrate the quality scores, call, filter, recalibrate and evaluate the variants. SNVs and INDELs across the sequenced protein-coding regions and flanking junctions were annotated using Variant Annotator X (VAX), a customized Ensembl Variant Effect Predictor(Yourshaw et al., 2014). Each variant was annotated with information including gene names and accession numbers, reference variant, variant consequences, protein positions and amino-acid changes, conservation scores, population minor allele frequencies, expression pattern, and PolyPhen2(Adzhubei et al., 2010), SIFT(Kumar et al., 2009), Condel(González-Pérez and López-Bigas, 2011) predictions. An average coverage of 62X was achieved across the exome with 86.2% of these bases covered at ≥10X. A total of 19,247 variants were identified across the RefSeq protein coding exons and flanking introns (+/-2bp). Of these, 1515 variants (76 homozygous and 1439 heterozygous variants) were protein-changing variants with population minor allele frequencies <1%.

A single, unique homozygous start codon mutation in *KATNB1* (c.1A>G, hg19 chr16: 57,771,156) was present in the region of IBD. Sanger sequencing confirmed and

validated mutation segregation in all tested family members, using forward primer 5'-TTCTGTCTGCCTGGATGTG and reverse primer 5'-GAAACAAAGCTGTGGGTTCAG.

Exome Sequencing and Filtering in Proband 2

Exome capture was performed using TruSeq Exome Enrichment kit (Illumina, San Diego, CA, USA) following the manufacturer's protocol. Samples were prepared as an Illumina sequencing library, and in the second step, the sequencing libraries were enriched for the desired target using the Illumina Exome Enrichment protocol. The captured libraries were sequenced using the Illumina HiSeq 2000 Sequencer. The reads were mapped against UCSC hg19 (http://genome.ucsc.edu/) by BWA (Burrows-Wheeler Aligner) (http://bio-bwa.sourceforge.net/). SNPs and Indels were detected by SAMTOOLS (http://samtools.sourceforge.net/).

Exome-derived data were filtered for variants as follows. A total of 82073 exome variants were identified, of which 37,114 were homozygous, consistent with inheritance from consanguineous parents. We excluded all variants that did not reside in a coding or splice site region, or were present in dbSNP or the 1000Genomes database, leaving 175 variants. Subsequent steps involved removal of all variants which resided outside the autozygome, or which were present in our Saudi population specific database (167 exomes) at an allele frequency greater than 0.005, or listed within the Exome Variant Server, leaving only 4 variants. Finally the pathogenicity of all surviving missense variants was established using Polyphen-2 and SIFT, and any variant predicted benign by both algorithms was eliminated, with 3 variants remaining.

Gene	Exonic Function	Change	Present in control Saudi database (167 exomes)
GLYATL3	frameshift deletion	NM_001010904:exon6:c.577_578del: p.193_193del	Yes
GJA10	nonsynonymous SNV	NM_032602:exon1:c.669G>T: p.L223F	Yes
KATNB1	nonsynonymous SNV	NM_005886:exon3:c.97G>T: p.G33W	No

KATNB1 was the only gene with variants that intersected with candidate variants in Family 1 and Proband 3, and only the *KATNB1* missense variant was completely absent from our Saudi population specific database (c.97G>T, hg19 chr16:57,775,655). This variant was confirmed by Sanger sequencing using the forward primer 5'-GGTCACCCCTCCTTCACAAG and reverse primer 5'- TCACTCTTCAGCCTCCAACC.

Genome-wide SNP genotyping and homozygosity mapping in Proband 3

Proband 3 and his father were genotyped by the Illumina Human610-Quad BeadChip array according to manufacturers' instructions. Though the parents of the affected individual were not reported to be consanguineous, there were 11 blocks of homozygosity over 2.0 cM in Proband 3 that were not shared by his father. The largest block was 13.4 cM in size, suggesting that the parents are likely to be distantly related.

Chromosome	Interval	Size (cM)	Size (Mb)
16	rs8046668; rs12373064	13.4	14.1
14	rs11624171; rs4562969	9.5	5.8
16	rs2966097; rs7195057	6	1.7
5	rs1549221; rs2913861	5.5	2.9
5	rs3765045; rs13158449	2.9	1.9
20	rs242435; rs6129021	2.6	0.8
21	p terminus; rs2207843	2.6	4.3
1	rs7541950; rs11205321	2.6	2.4
9	rs10116618; rs28651528	2.4	4.8
9	rs28651528; rs10869153	2.2	26.8
14	rs2401756; rs9323838	2.1	0.9

Array capture and exon sequencing in Proband 3

We performed array capture followed by high-throughput Illumina sequencing of all coding exons within the blocks of homozygosity over 2.0 cM in Proband 3 that were not shared by his father. Array capture and sequencing were performed as previously described(Yu et al., 2010) in Proband 3. Briefly, a custom microarray was designed with oligonucleotide probes targeting all exons within the candidate interval. DNA bound to the array was amplified and used to generate sequencing libraries. The average depth of target coverage was 138X, with 86.6% of bases covered by ≥10 reads. After filtering for standard quality and mapping confidence, a total of 1297 potential homozygous sequence variants were identified.

Variant analysis and sequencing confirmation in Proband 3

Variants were annotated with respect to predicted molecular effect and presence in known population datasets using ANNOVAR(Wang et al., 2010) as well as the genomic mutation consequence calculator(Major, 2007). To identify candidate pathogenic variants, variants were filtered to exclude those found in known population datasets with a minor allele frequency (MAF) of 1% or greater in the 1000 Genomes Project (http://www.1000genomes.org) or the NHLBI Exome Sequencing Project (http://evs.gs.washington.edu/EVS), and to exclude those predicted to be molecularly nonimpactful (intergenic, synonymous, or intronic and greater than 2bp from a splice junction). Of the 1297 homozygous variants identified in the individual, 74 were protein-altering, and only 7 variants were rare (MAF 0.01 as above) and protein-altering. Of the 7 variants, 3 are present in dbSNP132 and 2 others are present at very low frequency in EVS; these 5 were all predicted to be tolerated by Polyphen-2 and SIFT, leaving two remaining candidate variants.

Gene	Exonic Function	Change	Exome Variant Server 6500 allele frequency	dbSNP132	Polyphen -2	SIFT
TPTE	nonsynonymous SNV	NM_199260:c.A1042G: p.K348E		rs212146	NA	Tolerated
LOC100288966,POTED	nonsynonymous SNV	NM_001257362:c.G337A: p.G113S		rs6517869	Benign	Tolerated
LOC100288966,POTED	nonsynonymous SNV	NM_001257362:c.A1603G: p.M535V		rs78643169	Benign	Tolerated
CDC42BPB	nonsynonymous SNV	NM_006035:c.C2948T: p.A983V	0.000923		Benign	Tolerated
PROP1	nonsynonymous SNV	NM_006261:c.C425T: p.A142V	0.00246		Benign	Tolerated
NOP16	frameshift insertion	NM_001256539:c.586_587insAC: p.R196fs				
KATNB1	splicing	NM_005886:exon6:c.432+1G>A				

Of these 2 variants, whole-exome sequencing of Probands 1 and 2 led to the elimination the NOP16 variant, leaving a single remaining candidate. This was a homozygous single base pair change in *KATNB1* in the first base of the intron (c.432+1G>A, hg19 chr16:57,785,215). This variant was verified by Sanger sequencing with forward primer 5'-TTGTCTCCGTGGGGAGTAAC and reverse primer 5'-CTAGGGGATTCATGCAGAGC. The primers for Sanger confirmation of this variant were designed with Primer3 (http://frodo.wi.mit.edu/primer3/)(Rozen and Skaletsky, 2000).

Sequencing of *KATNB1* in Controls

The missense mutation was screened against 178 ethnically matched Saudi Arabian controls and 82 tribe-specific controls. The splice site mutation is absent from >800 exomes sequenced in the Walsh lab (many of whom are from individuals from the Middle East with brain malformations or autism) and from the Exome Variant Server (EVS, http://evs.gs.washington.edu/EVS/). No allele carrying the mutation was found in these controls.

Animals

All animal experimentation was carried out under protocols approved by the IACUCs of Harvard Medical School and Boston Children's Hospital. Targeted gene-trap ES cells (EUCOMM) were injected into C57BL/6 blastocysts to generate chimeric mice. Male chimeras were be bred to WT C57BL/6 background females (Charles River laboratories) to transmit the gene-trap allele. The mice are maintained on a C57/Bl6 background. For BrdU pulse-labeling experiments, E12.5 timed pregnant mice were injected with 75 mg/kg BrdU intraperitoneally and euthanized 30 minutes later. For histological analysis, brains were perfused with PBS and then followed by 4% PFA before paraffin embedding and sectioning, or cryoprotected in 30% sucrose before embedding in OCT (Tissue-Tek) and cryosectioning. Genotyping primers are available upon request.

Transcription activator-like effector nucleases (TALENs) targeting the exon 6-intron 6 boundary of zebrafish *katnb1* (ENSDARG00000005456) were designed by and ordered from ZGENEBIO Biotech Co. LTD (Supplementary Table 2). Plasmids were linearized with NotI and RNA was subsequently synthesized using the mMESSAGE

mMachine SP6 Transcription kit (Life Technologies). To generate *katnb1* mutants, 150-200pg of RNA was injected into eggs (*albino/slc45a2*) at the 1-cell stage. Fish in the F1 generation, stemming from brother-sister matings of injected fish, were prescreened for induced mutations by heteroduplex mobility assay(Ota et al., 2013). Genotyping was subsequently confirmed by colony PCR and Sanger sequencing of individual fish to identify the specific alleles. Genotyping primers are available upon request.

Culture Systems

HEK293T cells were grown in Dulbecco's Modified Eagle Medium, High Glucose, High Glucose (HyClone) with 10% fetal bovine serum (FBS, Gibco) and 1mM penicillin, streptomycin and L-glutamine. Patient point mutations were generated using QuikChange kit (Stratagene). For tagged *KATNB1* mini-gene and cDNA overexpression, we transfected 293T cells with specified plasmids using Fugene6 (Roche). Constructs were generated using cDNA for Refseq gene ID 10300.

Primary fibroblasts were established from Proband 1 and were maintained in RPMI-1640 (Gibco) with 10% FBS, and 1mM Penicillin and Streptomycin. Human iPSCs were generated according to published protocols(Takahashi et al., 2007) using retroviral vectors encoding the human cDNAs of *KLF4*, *SOX2*, *OCT4*, and *C-MYC* (Addgene) and grown in Knockout DMEM containing 20% knockout serum replacement, 2 mM L-glutamine, 1% NEAA, 0.1 mM b-mercaptoethanol, 1% NEAA, 0.1% penicillin and streptomycin, and 4 ng/ml bFGF on irradiated fibroblast feeders. iPSC colonies were manually picked after 3–4 weeks and cultured on irradiated MEFs. Control cell lines are from an unrelated, unaffected ethnically-matched individual.

For neural stem cell (NSC) generation, iPSCs were dissociated into single cells and 10,000 cells/well were seeded into a low attachment U-bottom 96-well plate to form cell aggregates of embryoid bodies (EB) in neural induction medium (DMEM/F12 supplemented with 20% KOSR, 2mM L-glutamine, 0.2mM NEAA, 0.1mM 2-mercaptoethanol and 1mM Sodium Pyruvate) for 7 days. EBs were then adhered onto laminin-coated plates in neural precursor medium (Neurobasal medium supplemented with 2mM L-glutamine, 1% B27, 1% N2 and 20ng/ml bFGF) for the formation of rosette-like structures for additional 7-14 days. The neural rosettes were manually isolated and further expanded as NSCs in neural induction medium. For formation of neurospheres, 8000 cells/well of NSCs were seeded into a low attachment U-bottom 96-well plate in neural precursor medium to form neurospheres of equal sizes. Sphere size was measured every day for 6 days.

Lymphoblastoid cell lines were established from Proband 2 following Epstein-Barr virus (EBV) immortalization and were maintained in RPMI-1640 (Gibco) with 10% FBS, and 1mM Penicillin and Streptomycin. For imaging, they were grown on Poly-L-Ornithine (Sigma) coated coverslips. Control cell lines are from an unrelated, unaffected individual.

Gene-trap MEFs were isolated from E13.5 embryos and dissociated by trypsinization. Cells were grown in Dulbecco's Modified Eagle Medium, High Glucose (HyClone) with 15% fetal bovine serum (FBS, Gibco) and 1mM penicillin, streptomycin and L-glutamine. Cilia were induced by serum starvation for 48 hrs in Optimem (Gibco) media, then treated with DMSO or 100nM SAG for an additional 24 hours.

RT-PCR and Quantitative RT-PCR

For RT-PCR, RNA was harvested using RNAeasy Plus kit (Qiagen). 1 ug of RNA was used to synthesize cDNA using Superscript First Strand Synthesis kit (Invitrogen).

For qRT-PCR, total RNA was extracted from Proband 1-derived fibroblasts and iPSCs using Nucleospine RNAII Kit (Macherey-Naga) with DNAse digestion. RNA was quantified using ND-1000 spectrophotometer (Biofrontier Technology) and first strand cDNA was produced with M-MulV reverse transcriptase (Biolabs) using 1 µg of total RNA input. Quantitative RT-PCR, normalized to *GAPDH* expression levels, was performed with the SYBER Green Mix (Roche). P**rimers** were as follows: *KATNB1* forward 5'-ACTGGTGCTGGGCAAAG; *KATNB1* reverse 5'-TGCAGTTGGGCTTGTTGA; *KATNA1* forward 5'-GGCGGGTTTGTCTTGAAATTC; *KATNA1* reverse 5'-GTACTGCCCAGTTCAATCGT.

For Shh qRT-PCR, MEFs were plated in 6-well plates, starved overnight in Opti-MEM, and then treated with DMSO or 100nM SAG for an additional 24 hours. RNA was harvested using an RNAeasy Plus kit (Qiagen). 1 ug of RNA was used to synthesize cDNA using a First Strand cDNA synthesis kit (Maxima). The resulting cDNA was probed with primers to Gli1, Patched and beta-Actin and amplified using Sybr GreenER (Invitrogen). Primer and probe sequences as previously described(Garcia-Gonzalo et al., 2011).

Western Blot

Lysate was prepared using RIPA buffer (50 mM Tris-HCl 7.4, 150 mM NaCl, 1% NP-40, 0.5% sodium deoxycholate, and 0.1% SDS) containing a protease inhibitor cocktail mix (Roche). Protein concentrations were normalized using the bicinchoninic

acid assay (Thermo Scientific). Extracted proteins were separated by electrophoresis on Bis-Tris polyacrylamide gels (Life Technologies) or SDS polyacrylamide gels with DTT, followed by transfer on polyvinylidene difluoride membrane (Millipore).

For quantitative Western blotting, membranes were blocked for 1 hr in Odyssey Blocking Buffer (LICOR Biosciences) at room temperature, incubated them with primary antibodies according to the antibody manufacturer's instructions, and then incubated them with fluorescent-dye-conjugated secondary antibodies (LICOR Biosciences). Bands were detected, normalized and quantified with the Odyssey Infrared Imaging System (Li-Cor Biosciences).

Antibodies: Flag (Cell Signaling, 1:1000); GAPDH (Abcam, 1:2000); Katanin p60 (Proteintech, 1:200; Atlas, 1:200); Katanin p80 (Proteintech, 1:500; Santa Cruz, 1:200; LSBio, 1:2000; Sigma, 1:100)

Immunostaining

Cells were grown on glass coverslips and fixed in either 4% PFA for 20 minutes at room temperature, or 100% Methanol for 3 minutes at -20C. Cells were blocked (5% serum in PBS-0.1% Triton-X), incubated with antibodies and mounted using Gelvatol or Fluoromount-G (Southern Biotech) mounting media.

Embryos were drop fixed in 4% PFA overnight, followed by PBS wash. Paraffin sections (5 um) of brains were rehydrated and subjected to antigen retrieval with Retrievagen A (BD Pharmingen), followed by blocking (5% serum in PBS), permeabilization (0.1% Triton X-100), and antibody incubation. Alternatively, cryosections (50 um) were permeabilized (0.1% Triton X-100), blocked, and incubated with antibodies (24h-48h).

All samples were counterstained with DAPI or Hoechst 33258 (Sigma). Images were acquired with a Zeiss LSM 700 on an upright Axio Imager.M2 equipped with 405, 488, 555 and 639 nm solid state lasers, motorized stage and a CCD AxioCam MR; an inverted Zeiss Axio Observer D1; or a DeltaVision RT GE microscope.

For Gli3 immunocytochemistry quantification, Gli3 intensity was measured at the ciliary tip and background intensity was subtracted.

Antibodies: Acetylated tubulin (Abcam, 1:500; MeOH); Arl13b (kind gift of Tamara Caspary, 1:500; PFA); Beta catenin (BD Biosciences, 1:100; PFA); Beta tubulin (Abcam, 1:400; PFA); BrdU (AbD Serotec, 1:500; PFA); CAMSAP2 (Proteintech, 1:1000; MeOH); Centrin (Millipore, 1:1000; MeOH); Centromere (Antibodies Inc, 1:100; PFA); Cep164 (kind gift of Erich Nigg, 1:1000; MeOH); Cep170 (Abcam, 1:500; PFA); Cep63 (Millipore, 1:500; MeOH); Cleaved Caspase 3 (Abcam, 1:500; PFA); Cyclin A (Santa Cruz, 1:200; MeOH); DCX (Santa Cruz, 1:250; PFA); EB1 (BD Biosciences, 1:250; MeOH); Gamma tubulin (Sigma, 1:200, MeOH; Abcam, 1:300, PFA); GFP (Abcam, 1:1000; PFA); Gli3 (kind gift of Suzie Scales, 1:1000; MeOH); GT335 poly glutamylated tubulin (Enzo Life Sciences, 1:500; PFA); IFT88 (Proteintech, 1:500; MeOH); Katanin p60 (Proteintech, 1:500; Atlas, 1:250; PFA); Katanin p80 (Proteintech, 1:500; Santa Cruz, 1:500; Sigma, 1:2000; PFA); LacZ (MP biomedicals, 1:250; PFA); Ninein (kind gift of James Sillibourne, 1:2500; MeOH); OCT4 (Santa Cruz, 1:100; PFA); Pericentrin (Abcam, 1:1000; PFA); Phosphorylated Histone H3 (Cell Signaling, 1:1000, MeOH); Smoothened (Abcam, 1:500; PFA); Sox2 (Santa Cruz, 1:500; R&D, 1:200; PFA); SSEA4 (Millipore, 1:100; PFA); STIL (Bethyl Labs, 1:200; MeOH); Tbr2 (Abcam,

1:500; PFA); TRA-1-60 (Millipore, 1:100; PFA); TRA-1-81 (Millipore, 1:100; PFA); Yo Pro (Invitrogen, 1:500; PFA)

Electron Microscopy

Cells were grown on aclar coverslips and fixed for 20 minutes in a 1:1 dilution of culture media to 5% Glutaraldehyde, 2.5% Paraformaldehyde, 0.06 % picric acid in 0.2 M sodium cacodylate buffer (pH 7.4), washed 3x in 0.1M Cacodylate buffer, then post-fixed in 1% Osmium tetroxide (OsO_4)/1.5% Potassiumferrocyanide ($KFeCN_6$) for 30 min, washed in water 3x and incubated in 1% aqueous uranyl acetate for 30min followed by 2 washes in water and subsequent dehydration in grades of alcohol (5min each; 50%, 70%, 95%, 2x 100%). Cells were embedded in TAAB Epon (Marivac Canada Inc. St. Laurent, Canada) and polymerized at 60 degrees C for 48 hrs. After polymerization the aclar coverslip was peeled off and 1mm squares of the embedded monolayer was glued onto an empty Epon block for sectioning.

Ultrathin serial sections (80nm) were cut on a Reichert Ultracut-S microtome, picked up on to formvar/carbon coated copper grids, stained with lead citrate and examined in a JEOL 1200EX Transmission electron microscope. Images were recorded with an AMT 2k CCD camera.

Statistical Analyses

All data are represented as Mean ± SEM. Analyses were performed using at least 3 samples each from 3 different animals of each genotype, or using at least 2 replicates for Proband-derived cell lines. ANOVAs were analyzed using Bonferroni test correction for multiple comparisons. No F-tests were significant for unequal variances; therefore, unpaired t-tests were analyzed assuming equal variances

References

Adzhubei, I.A., Schmidt, S., Peshkin, L., Ramensky, V.E., Gerasimova, A., Bork, P., Kondrashov, A.S., and Sunyaev, S.R. (2010). A method and server for predicting damaging missense mutations. Nat Meth *7*, 248–249.

Garcia-Gonzalo, F.R., Corbit, K.C., Sirerol-Piquer, M.S., Ramaswami, G., Otto, E.A., Noriega, T.R., Seol, A.D., Robinson, J.F., Bennett, C.L., Josifova, D.J., et al. (2011). A transition zone complex regulates mammalian ciliogenesis and ciliary membrane composition. Nat Genet *43*, 776–784.

González-Pérez, A., and López-Bigas, N. (2011). Improving the assessment of the outcome of nonsynonymous SNVs with a consensus deleteriousness score, Condel. Am J Hum Genet *88*, 440–449.

Kumar, P., Henikoff, S., and Ng, P.C. (2009). Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nat Protoc *4*, 1073–1081.

Major, J.E. (2007). Genomic mutation consequence calculator. Bioinformatics *23*, 3091–3092.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., et al. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res *20*, 1297–1303.

Ota, S., Hisano, Y., Muraki, M., Hoshijima, K., Dahlem, T.J., Grunwald, D.J., Okada, Y., and Kawahara, A. (2013). Efficient identification of TALEN-mediated genome modifications using heteroduplex mobility assays. Genes Cells *18*, 450–458.

Rozen, S., and Skaletsky, H. (2000). Primer3 on the WWW for general users and for biologist programmers. Methods Mol Biol *132*, 365–386.

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., and Yamanaka, S. (2007). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell *131*, 861–872.

Wang, K., Li, M., and Hakonarson, H. (2010). ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res *38*, e164.

WHO Multicentre Growth Reference Study Group (2006). WHO Child Growth Standards based on length/height, weight and age. Acta Paediatr Suppl *450*, 76–85.

Yourshaw, M., Taylor, S.P., Rao, A.R., Martín, M.G., and Nelson, S.F. (2014). Rich annotation of DNA sequencing variants by leveraging the Ensembl Variant Effect Predictor with plugins. Brief. Bioinformatics.

Yu, T.W., Mochida, G.H., Tischfield, D.J., Sgaier, S.K., Flores-Sarnat, L., Sergi, C.M., Topçu, M., McDonald, M.T., Barry, B.J., Felie, J.M., et al. (2010). Mutations in WDR62, encoding a centrosome-associated protein, cause microcephaly with simplified gyri and abnormal cortical architecture. Nat Genet *42*, 1015–1020.