The American Journal of Human Genetics, Volume *105*

Supplemental Data

Bi-allelic Loss of Human *APC2*, Encoding Adenomatous Polyposis Coli Protein 2, Leads to Lissencephaly, Subcortical Heterotopia, and Global Developmental Delay

Sangmoon Lee, Dillon Y. Chen, Maha S. Zaki, Reza Maroofian, Henry Houlden, Nataliya Di Donato, Dalia Abdin, Heba Morsy, Ghayda M. Mirzaa, William B. Dobyns, Jennifer McEvoy-Venneri, Valentina Stanley, Kiely N. James, Grazia M.S. Mancini, Rachel Schot, Tugba Kalayci, Umut Altunoglu, Ehsan Ghayoor Karimiani, Lauren Brick, Mariya Kozenko, Yalda Jamshidi, M. Chiara Manzini, Mehran Beiraghi Toosi, and Joseph G. Gleeson

Supplemental Figures

Figure S1

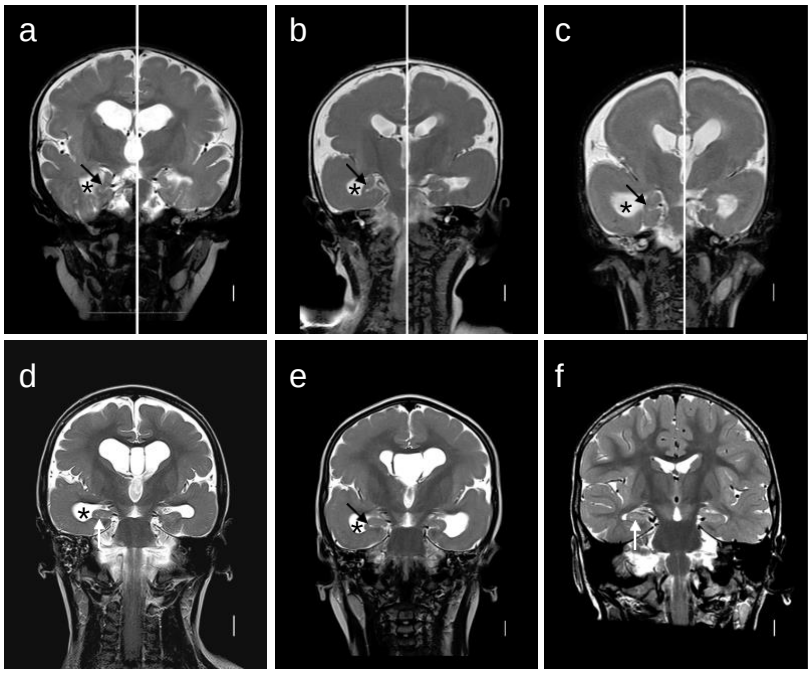

Figure S1. Brain imaging in *APC2*-lissencephaly in 5 children highlighting hippocampal malformations. Subjects 5-II-1 (**a**), 6-III-1 (**b**), 6-III-2 (**c**), 7-III-3 (**d**), 7-III-4 (**e**), and a normal control (**f**). The T2-weighted coronal images through the posterior frontal lobes and hippocampi showed globular and open hippocampi (**a-c**, **e**) that were usually under developed (**a-c**). The hippocampi in one child appear normal on these images (**d**), but all five subjects have moderately enlarged temporal horns (asterisks in **a-e**), which is commonly seen with hippocampal malformations associated with lissencephaly. All images are T2-weighted.

Supplemental Table

Table S1

	Gene	Zygosity	Chr	Pos	Ref	Alt	cDNA	AAChange	Segregated	OMIM	Effect	Impact	Transcript	dbSNP	gnomAD_AF	SIFT	PolyPhen	CADD_PHRED	Associated recessive disease
Family 1	VSTM4	hom	10	50272764	T	C	c.652A>G	p.(Lys218Glu)			missense_variant	MODERATE	NM_001031746.4		0.000012	deleterious(0)	probably_damagi	24.1	
	KCNMA1	hom	10	78651483	G	A	c.2986-6C>T	p.?		600150	splice_region_var	MODERATE	NM_001014797.2	201087232	0.000008	-	-	-	
	TGM7	hom	15	43584295	T	A	c.440A>T	p.(Glu147Val)		606776	missense_variant	MODERATE	NM_052955.2	181416302	0.000060	deleterious(0.01)	probably_damagi	26.5	
	FBN1	hom	15	48758054	G	A	c.4749C>T	p.(Ser1583=)		134797	splice_region_var	MODERATE	NM_000138.4		-	-	-	12.02	
	WDR18	hom	19	989769	C	T	c.329C>T	p.(Thr110Ile)	Y		missense_variant	MODERATE	NM_024100.3		-	deleterious(0.04)	probably_damagi	25.4	
	APC2	hom	19	1457116	C	T	c.1081C>T	p.(Gln361*)	Y	612034	stop_gained	HIGH	NM_005883.2		0.000000	-	-	36	
Family 2	AMY2A	hom	1	104160661	G	A	c.254G>A	p.(Cys85Tyr)		104650	missense_variant	MODERATE	NM_000699.3		-	tolerated(0.2)	probably_damagi	17.93	
	EPHA5	hom	4	66231748	A	C	c.1952T>G	p.(Ile651Ser)		600004	missense_variant	MODERATE	NM_004439.7		-	deleterious(0.02)	possibly_damagin	24.7	
	HOXC10	hom	12	54379276	C	T	c.233C>T	p.(Ser78Phe)		605560	missense_variant	MODERATE	NM_017409.3		-	deleterious(0.01)	benign(0.276)	25.9	
	BRCA2	hom	13	32905126	C	G	c.752C>G	p.(Thr251Arg)		600185	missense_variant	MODERATE	NM_000059.3	587781513	0.000004	tolerated(0.18)	probably_damagi	7.681	Fanconi anemia (MIM 605724)
	CDC42BPB	hom	14	103410287	G	A	c.4349C>T	p.(Pro1450Leu)		614062	missense_variant	MODERATE	NM_006035.3		0.000016	tolerated(0.76)	benign(0.003)	17.13	
	APC2	hom	19	1465182	C	T	c.1882C>T	p.(Gln628*)	Y	612034	stop_gained	HIGH	NM_005883.2		-	-	-	40	
	DEPDC5	hom	22	32188752	G	A	c.716G>A	p.(Arg239Gln)		614191	missense_variant	MODERATE	NM_001007188		0.000010	deleterious(0.03)	possibly_damagin	25.3	
Family 3	MIDN	hom	19	1255457	G	A	c.893G>A	p.(Arg298Gln)		606700	missense_variant	MODERATE	NM_177401.4	143719550	0.000295	deleterious(0)	possibly_damagin	23.5	
	APC2	hom	19	1469940	C	-	c.6645delC	p.(Ala2217Profs*118)	Y	612034	frameshift_variar	HIGH	NM_005883.2		-	-	-	-	
Family 4	SPEG	hom	2	220348409	T	G	c.6224T>G	p.(Leu2075Arg)		615950	missense_variant	MODERATE	NM_005876.4	892492321	0.000016	deleterious(0)	probably_damagi	24.3	Centronuclear myopathy 5 (MIM 615959)
	UGT2B17	hom	4	69416566	A	T	c.1142T>A	p.(Ile381Asn)		601903	missense_variant	MODERATE	NM_001077.3	749202776	0.000020	deleterious(0)	probably_damagi	26.3	
	ZFHX4	hom	8	77618200	G	A	c.1877G>A	p.(p.Arg626Lys)		606940	missense_variant	MODERATE	NM_024721.4	751323188	0.000032	deleterious(0.03)	possibly_damagin	14.46	
	FGF6	hom	12	4554487	G	A	c.250C>T	p.(Arg84Trp)		134921	missense_variant	MODERATE	NM_020996.1	373061794	0.000040	deleterious(0.02)	possibly_damagin	25	
	SERPINF1	hom	17	1675327	G	C	c.601G>C	p.(Asp201His)		172860	missense_variant	MODERATE	NM_001329903.1	137997656	0.000008	tolerated(0.17)	benign(0.404)	6.925	Osteogenesis imperfecta, type VI (MIM 613982)
	SMYD4	hom	17	1703330	T	-	c.1358delA	p.(Gln453Argfs*2)			frameshift deletic	HIGH	NM_052928.2	755710047	0.000096	-	-	32	
	APC2	hom	19	1462016	CA	-	c.1694_1695delCA	p.(Thr565Argfs*50)	Y	612034	frameshift deletic	HIGH	NM_005883.2	.	-	-	-	-	
	ATP4A	hom	19	36046673	G	-	c.1911delC	p.(Ile638Leufs*26)		137216	frameshift deletic	HIGH	NM_000704.2	.	-	-	-	-	
	AX746638	hom	19	36806475	A	T	c.143T>A	p.(Leu48Gln)			missense_variant	MODERATE	NR_029389.1	2967481	-	-	-	-	
	SRRM5;ZNF5	hom	19	44116719	G	A	c.491G>A	p.(Gly164Asp)			missense_variant	MODERATE	NM_001145641.1	.	-	deleterious(0.04)	possibly_damagin	22.8	
	MARK4	hom	19	45790731	C	T	c.1303C>T	p.(Pro435Ser)		606495	missense_variant	MODERATE	NM_001199867.1	.	-	tolerated(0.16)	benign(0.3)	24.6	
	RPL18	hom	19	49121116	T	C	c.22A>G	p.(Asn8Asp)		604179	missense_variant	MODERATE	NM_000979.3	.	-	tolerated(0.25)	benign(0.014)	19.96	
	NTN5	hom	19	49165133	T	C	c.1271A>G	p.(Gln424Arg)			missense_variant	MODERATE	NM_145807.1	760927020	0.000133	tolerated(1)	benign(0.003)	0.001	
	LRRC4B	hom	19	51051969	C	T	c.127G>A	p.(Val43Met)			missense_variant	MODERATE	NM_001080457.1	753942999	0.000236	tolerated(0.07)	benign(0.005)	23.2	
	RIMBP3C	hom	22	21900617	A	G	c.4649T>C	p.(Ile1550Thr)		612701	missense_variant	MODERATE	NM_001128633.1	484252	-	tolerated(1)	benign(0)	0.001	
Family 8	IGSF9	hom	1	159899764	T	A	c.2066A>T	p.Asp689Val		609738	missense_variant	MODERATE	NM_001135050.1	201810988	0.000300	deleterious	probably_damagi	26.5	
	MEGF6	hom	1	3418428	G	A	c.2246C>T	p.Ala749Val		604266	missense_variant	MODERATE	NM_001409.3	200472001	0.001700	tolerated	probably_damagi	22.5	
	HS1BP3	hom	2	20840790	G	A	c.349C>T	p.Arg117Cys		609359	missense_variant	MODERATE	NM_022460.3	377728516	0.000050	deleterious	probably_damagi	29.7	
	CDHR4	hom	3	49834383	G	A	c.578C>T	p.Ser193Phe			missense_variant	MODERATE	NM_001007540.3	560357842	0.00009	-	probably_damagi	24.7	
	LAMB2	hom	3	49160696	G	A	c.4093C>T	p.Arg1365Trp		150325	missense_variant	MODERATE	NM_002292.3	751854328	0.0002	deleterious	probably_damagi	26.3	Nephrotic syndrome, type 5, with or without ocular abnormalities (MIM 614199), Pierson syndrome (MIM 609049)
	SEMA3B	hom	3	50311438	C	A	c.1086C>A	p.His363Asn		601281	missense_variant	MODERATE	NM_004636.3	782238556	0.000060	deleterious	-	20.5	
	TREX1	hom	3	48508733	G	A	c.679G>A	p.Gly227Ser		606609	missense_variant	MODERATE	NM_033629.5	113107733	0.0002	tolerated	-	15.7	Aicardi-Goutieres syndrome 1, dominant and recessive (MIM 225750)
	ULK4	hom	3	41291010	C	T	c.3734G>A	p.Arg1245Gln		617010	missense_variant	MODERATE	NM_017886.3	756001134	0.00002	tolerated	benign	<10	
	FILIP1	hom	6	76063397	G	A	c.487C>T	p.Arg163Trp		607307	missense_variant	MODERATE	NM_015687.4	759270192	0.0001	deleterious	probably_damagi	28.4	
	GABRR2	hom	6	89974256	C	T	c.961G>A	p.Val321Ile		137162	missense_variant	MODERATE	NM_002043.4	2228644	-	tolerated	-	15	
	TAS2R60	hom	7	143141342	G	C	c.797G>C	p.Ser266Thr		613968	missense_variant	MODERATE	NM_177437.1	-	-	tolerated	benign	<10	
	CSMD1	hom	8	3611478	C	T	c.905G>A	p.Arg302His		608397	missense_variant	MODERATE	NM_033225.5	754405745	0.00002	-	probably_damagi	23.9	
	FAM189A1	hom	15	29415846	C	T	c.1316G>A	p.Arg439His			missense_variant	MODERATE	NM_015307.1	61736883	0.000500	tolerated	benign	14.6	
	HERC2	hom	15	28467280	T	C	c.5546A>G	p.Lys1849Arg		605837	missense_variant	MODERATE	NM_004667.5	201821203	0.00057233	-	benign	17.1	Mental retardation, autosomal recessive 38 (MIM 615516)
	FSCN2	hom	17	79503213	G	A	c.1025G>A	p.Arg342Gln		607643	missense_variant	MODERATE	NM_001077182.2	374441539	0.000090	tolerated	benign	23.6	
	MAPRE2	hom	18	32558480..32558483	GAA	-	c.-2_2delGAAT	p.Met1		605789	initiation_codon	HIGH	NM_001143827.2	764635254	0.0015	-	-	22.1	
	APC2	hom	19	1462177	G	C	c.1853+1G>C		Y	612034	splice site disrupt	HIGH	NM_005883.2	-	-	-	-	32	

Table S1. High impact homozygous variants returned from whole exome sequencing of Families 1-4. In each family, a homozygous damaging mutation in *APC2* was determined to be most likely causative based upon objective filtering criteria (yellow).

Supplemental Methods

Study samples

We performed whole exome sequencing (WES) in 8 families with affected(s) displaying features consistent with lissencephaly, where prior gene panels and microarray studies proved negative at identifying a cause of disease. Subjects were enrolled in IRB-approved research studies at the University of California, San Diego or their home institution (Institute for Clinical Genetics, TU Dresden, Germany, University of Washington, National Research Center Egypt, St. George's University of London, Erasmus University, Istanbul University, The George Washington University and Mashhad University).

Exome sequencing and variant calling

Blood was acquired from informed, consenting individuals or their surrogates, according to institutional guidelines, and DNA extracted using established protocols. In solution exome capture was preformed using the SureSelect Human All Exome 50 Mb Kit (Agilent Technologies, USA) or xGen exome research panel (Integrated DNA Technologies, USA) with 100- or 150-bp paired-end read sequences generated on a HiSeq4000 or NextSeq500 instruments (Illumina, Inc. USA). Sequences were aligned to hg19 and variants identified through the GATK pipeline or CLC Biomedical Genomics Workbench (Qiagen, Hilden, Germany). Variations were annotated with in-house software, Annovar, Variant Effect Prediction software or CLC Biomedical Genomics Workbench to define population-specific allele frequencies from 1000 Genomes, the Greater Middle East Variome, dbSNP, and gnomAD, along with the transcript-specific predicted effect on the protein. All variants were prioritized by allele frequency, conservation, and predicted effect on protein function.

Variant prioritization

Variants were prioritized for each family using the following criteria:

1. The variant was predicted to perturb protein function. All synonymous and intronic variants were excluded unless the variant was within a predicted splice site (+ or -2 bp from splice junction). Any variation that was predicted to alter gene expression or protein function was included. These included nonsynonymous variations in coding regions (i.e. missense) or

alterations resulting in frameshifts, premature stop codons, loss of stop codons, coding INDELS, and splice sites (i.e. ±2 nucleotides from an exon junction).

2. The variant was rare as defined by allele frequency of less than 0.1% in either gnomAD or GME variomes.

3. The variant was present in a region of homozygosity as defined by HomozygosityMapper or parametric linkage analysis for consanguineous families.

4. The variant was conserved evolutionary as determined by a number of conservation scores including GERP, PhastCons, and PolyPhen2. Variations with negative GERP scores or vertebrate PhastCons scores less than 0.8 were excluded. Typical conservation criteria for the candidate genes provided in this study were GERP > 4 and vertebrate PhastCons > 0.9.

5. The variant was confirmed using Sanger sequencing and segregated with the disease in the family pedigree according to a strictly recessive mode of inheritance with full expressivity and absent phenotype in heterozygous carriers.

All variants following the above criteria were considered for each family independent of its predicted severity (i.e. no variants were excluded based upon type of mutation).

Sanger sequencing

Primers for Sanger sequencing were designed using the Primer3 program (U. Massachusetts) and tested for specificity using the Alamut Visual 2.7.1 software. PCR products were treated with Exonuclease I (Fermentas) and Shrimp Alkaline Phosphatase (USB Corp) and sequenced using the Big Dye terminator cycle sequencing kit v.3.1 (Applied Biosystems) on an ABI DNA analyzer (Applied Biosystems). Sequence data were analyzed using ApE1® software.