

Supplemental Information

A Primate-Specific Isoform of *PLEKHG6*

Regulates Neurogenesis and Neuronal Migration

Adam C. O'Neill, Christina Kyrousi, Johannes Klaus, Richard J. Leventer, Edwin P. Kirk, Andrew Fry, Daniela T. Pilz, Tim Morgan, Zandra A. Jenkins, Micha Drukker, Samuel F. Berkovic, Ingrid E. Scheffer, Renzo Guerrini, David M. Markie, Magdalena Götz, Silvia Cappello, and Stephen P. Robertson

Table S1: List of candidate *de novo* mutations identified in 65 PH trios analysed in this study. Related to Figure 1

Individual	Candidate gene(s)	Chr	Position	Nucleotide	Protein
769	<i>CALM3</i>	19	47111760	c.200C>T	p.Pro67Leu
780	<i>PIK3R1</i>	5	67522777	c.274C>T	p.Pro92Ser
1226	<i>CLEC16A</i>	16	11136118	c.1668G>C	p.K556N
1291	<i>GOLGB1</i>	3	121414605	c.4765CCTTdel	p.Glu1589fs*
1337	<i>C17orf104</i>	17	42744834	c.1555C>G	p.Gln519Glu
1490	<i>SCAF8</i>	6	155124724	c.830A>C	p.Glu277Ala
1492	<i>GATAD2A</i>	19	19576161	c.7G>A	p.Glu3Lys
	<i>GOLGA4</i>	3	37330777	c.643A>G	p.Arg215Gly
	<i>EFCAB1</i>	8	49637324	c.614G>A	p.Pro205Ser
1493	<i>RYR2</i>	1	237881777	c.10510G>A	p.Glu3504Lys
2317	<i>CNTROB</i>	17	7849070	c.1759C>T	p.Pro587Ser
2506	<i>JAKMIP3</i>	10	133967302	c.2107C>T	p.Arg703Trp
2808	<i>CNTN2</i>	1	205027694	IVS4+2A>G	Splice defect
	<i>TNMD</i>	X	99854069	c.642insA	p.G215Rfs*3
	<i>ACIN1</i>	14	23549258	c.1286A>G	p.K429R
3037	<i>LMX1A</i>	1	165324772	c.24_25insA	p.Glu9Argfs*49
3040	<i>LRRK2</i>	12	40631906	IVS5-1G>T	Splice defect
	<i>KLHL10</i>	17	40004255	c.1523C>T	p.Arg508His
	<i>KIFAP3</i>	1	170003611	c.644C>G	p.Thr215Ser
3051	<i>ARHGEF10</i>	8	1900958	c.3485C>T	p.Pro1162Leu
3054	<i>GFRA1</i>	10	117853290	c.923C>G	p.Cys308Ser
3107	<i>ATAD3C</i>	1	1389837	c.335C>T	p.Thr112Met
3160	<i>LRIG3</i>	12	59280626	c.958C>T	p.Gly320Ser
3171	<i>SPRY2</i>	13	80911693	c.148G>A	p.Arg50*
3176	<i>FAM46D</i>	X	79699198	c.1160G>A	p.Gly387Asp
3183	<i>GLUL</i>	1	182357871	c.1A>C	p.Met1Arg
3242	<i>LGALS3BP</i>	17	76968308	c.1108G>A	p.Glu370Lys
3288	<i>CDK13</i>	7	40085606	c.2525A>G	p.Asn842Ser
3296	<i>SAMD11</i>	1	878375	c.1501C>T	p.Leu501Phe
3338	<i>OSBPL1A</i>	18	21921567	c.338G>C	p.Ser113Thr
	<i>PTPRJ</i>	11	48145255	c.708*>+C	p.Cys236Cysfs*5
3341	<i>GPR98</i>	5	90079771	c.13550T>C	p.Ile4517Thr
	<i>NOL10</i>	2	10811767	c.299G>T	p.Gly100Val

Table S1: continued

Individual	Candidate gene(s)	Chr	Position	Nucleotide	Protein
3344	<i>PRPF6</i>	20	62612665	c.67C>T	p.Arg23Trp
3350	<i>ZNF790</i>	19	37310002	c.1244T>C	p.Ile415Thr
3390	<i>ZNF12</i>	7	6731756	c.817G>C	p.Glu273Gln
3412	<i>HNRNPA0</i>	5	137089208	c.548G>T	p.Gly183Val
3433	<i>DPP9</i>	19	4703897	IVS5-C>T	Splice defect
3480	<i>PTPRD</i> *	9	8528586	c.546G>C	p.Arg182Ser
	<i>MYH2</i>	17	10440684	c.1763T>G	p.Val588Gly
3483	<i>PI4KA</i>	22	21073066	c.5161G>A	p.Gly1721Ser
3486	<i>NEB</i>	2	152521300	c.5314_5315delTC	p.Ser1772fs
	<i>MEX3B</i>	15	82336718	c.493G>A	p.Val165Met
3489	<i>SUSD2</i>	22	24581633	c.1075C>T	p.Arg359Trp
3492	<i>RAB11FIP5</i>	2	73303277	c.1602delT	p.Pro534fs
3498	<i>TTN</i>	2	179449471	c.64897C>T	p.Arg21633Trp
3504	<i>ZNF117</i>	7	64438651	c.1298T>C	p.Ile433Thr
	<i>ABAT</i> *	16	8875151	c.1426T>G	p.Ser476Ala
	<i>ZC3H4</i>	19	47570329	c.3196G>A	p.Asp1066Asn
	<i>DCBLD2</i>	3	98538048	c.1085C>A	p.Thr362Lys

* Transcript ENST00000537002 only. ‡ Transcript ENST00000569156 only. Both transcripts are also only present within *Homo sapiens*.

Table S2: List of candidate biallelic mutations identified in 65 PH trios analysed in this study. Related to Figure 1

Individual	Candidate gene(s)	Chr	Position	Nucleotide	Protein
533	<i>F5</i>	1	169509859	c.4449T>G	p.Met1490Arg
	<i>F5</i>	1	169515818	c.1624A>G	p.Ile542Val
1203	<i>MOB2</i>	11	1502019	c.207delC	p.Phe69Phefs*127
	<i>MOB2</i>	11	1491530	c.679C>T	p.Glu227Lys
1291	<i>SVEP1</i>	9	113208279	c.4301C>T	p.Gly1434Asp
	<i>SVEP1</i>	9	113192266	c.5549G>A	p.Pro1850Leu
1493	<i>COL5A1</i>	2	189929302	c.1697G>A	p.Pro566Leu
	<i>COL5A1</i>	2	189969003	c.323C>T	p.Gly108Asp
2758	<i>PLEKHG6</i>	12	6422338	c.28delG	p.Glu10Argfs*40
2859	<i>HKDC1</i>	10	70987024	c.125G>A	p.Arg42Gln
	<i>HKDC1</i>	10	71000508	IVS6-1G>A	Splice defect
3046	<i>CCDC88B</i>	11	64118952	c.2963C>A	p.Ala988Glu
	<i>CCDC88B</i>	11	64119728	c.3126C>T	p.Arg1076Trp
3096	<i>ARHGAP39</i>	8	145773344	c.24_26delTCG	p.Gln376del
3168	<i>LCNL1</i>	9	139879399	c.431C>G	p.Pro144Arg
	<i>LCNL1</i>	9	139879435	c.467T>C	p.Leu156Pro
3176	<i>ECE2</i>	3	183995077	c.655C>G	p.Arg219Gly
	<i>ECE2</i>	3	184008893	c.2253C>T	p.Arg752Trp
3242	<i>SERPINA9</i>	14	94936088	c.90C>A	p.Tyr30*
	<i>SERPINA9</i>	14	94929482	c.1256C>T	p.Ala419Val
3412	<i>TTN</i>	2	179621363	c.10327G>T	p.Glu3443*
	<i>TTN</i>	2	179410304	c.68914G>A	p.Asp22972Asn
3415	<i>TPGS1</i>	19	507765	c.259C>T	p.Pro87Ser
	<i>TPGS1</i>	19	519341	c.7919G>T	p.Arg264Leu
3483	<i>BHLHE22</i>	8	65493448	c.101C>G	p.Ala34Gly
	<i>BHLHE22</i>	8	65494020	c.695_703delGCAGCAGCA	p.Ser232_Ser234del
3495	<i>CABP1</i>	12	121093653	c.43_44delGC	p.Ala15fs
	<i>CABP1</i>	12	121093759	c.146G>A	p.Arg49His
3498	<i>EYS</i>	6	65532559	c.3149C>G	p.Pro1050Arg
	<i>EYS</i>	6	64488046	c.7751C>G	p.Thr2584Ser
3501	<i>ITIH5</i>	10	7621958	c.1178T>C	p.Ile393Thr
	<i>ITIH5</i>	10	7621932	c.1204G>A	p.Val402Ile

Table S3: Observed and expected de novo variants in six gene-sets of patients with PH analysed in this study. Related to Figure 1

Gene set [‡]	No. of genes	Obs	Nonsense		Obs	missense	
			Exp	P-value		Exp	P-value
FMRP	842	1	0.843	0.569	7	3.579	0.099
Chromatin	428	0	0.292	1.000	1	1.236	1.000
Embryonic	1,912	0	1.112	0.634	6	4.732	0.485
PSD	1,445	2	0.803	0.192	5	3.414	0.403
Essential	1,750	3	1.048	0.089	4	4.457	1.000
Mendelian	256	1	0.198	0.179	1	0.841	0.569

[‡]The six different gene sets analysed in this test, for definitions of names see Supplemental Experimental Procedures. Exact binomial test (two-tailed).

Table S4: Phenotypic description of patient (i.d 2758) with loss-of-function variant in isoform four of *PLEKHG6*. Related to Figure 1

Individual	Clinical description
2758	<p>The individual who was homozygous for the c.26delG variant was the first born male child of unrelated healthy parents. The pregnancy was unremarkable, there was no exposure to known teratogens and his delivery was at term. His birth weight was on the 50th centile and head circumference on the 90th centile. He was hypotonic in the neonatal period with some episodes of hypoglycemia which responded to treatment satisfactorily. He had bilateral undescended testes and was noted to be dysmorphic. A MRI scan of the head revealed bilateral periventricular nodular heterotopia, predominating in the occipital horns and trigone. There was a megacisterna magna, thinning of the corpus callosum and an ectopic posterior pituitary. He developed localized bronchiectatic changes in his right middle lobe in the first three years of life but a CT chest performed at age 7 demonstrated resolution of these changes. Now aged 7 years he has mild-mod developmental delay but has never had a seizure. Evaluation of his pituitary function has demonstrated no abnormality. His head circumference has tracked the 75th centile for age.</p>

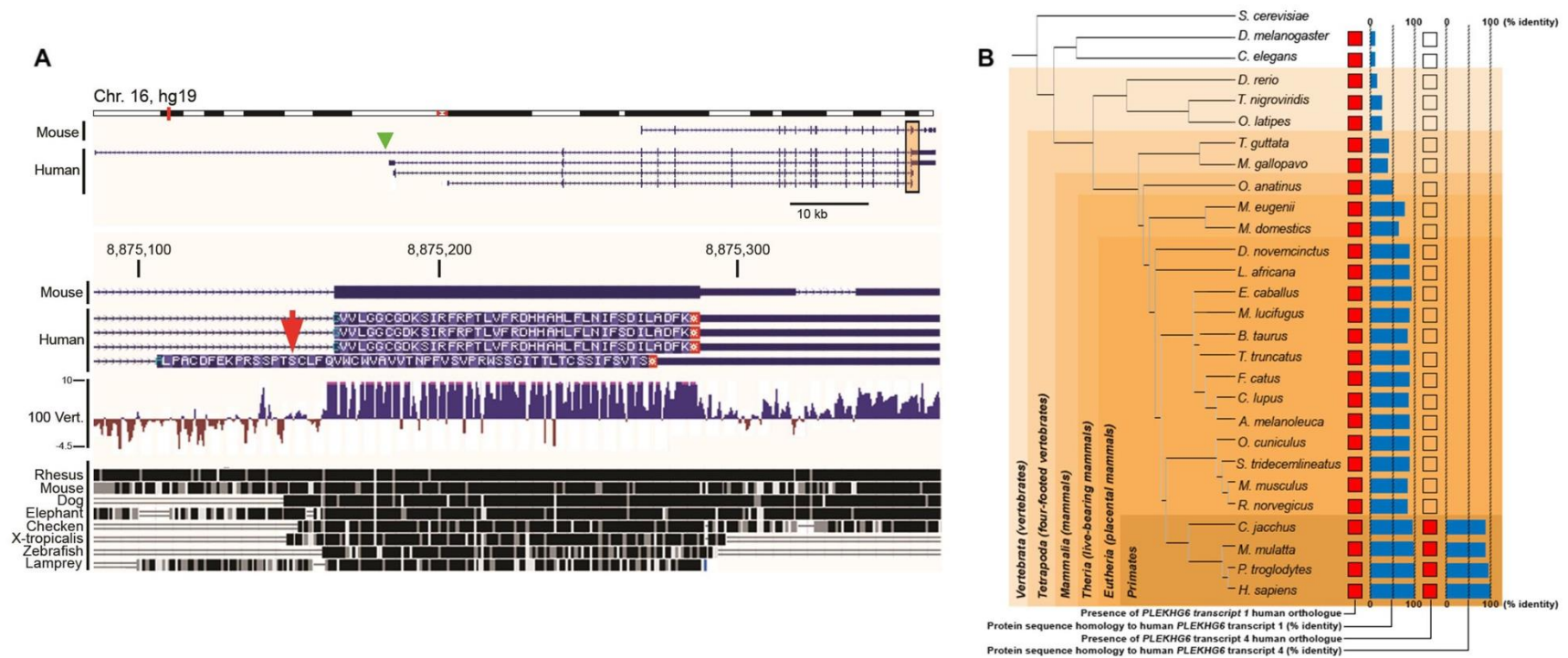


Figure S1. Additional evolutionarily dynamic isoforms identified through whole-exome sequencing of patients with PH. Related to Figure 1. (A) UCSC Genome Browser tracks illustrating the *ABAT* loci variant in a patient with PH in an isoform that is not present in mice, but has some experimental support to be present in humans. Top panel outlines the entire locus and the various isoforms annotated experimentally in mice and humans. Orange shade highlights the region shown at higher resolution in the bottom panel. Red arrow identifies the site of the variant in the patient (specific nucleotide residue is outlined in tables S1). 100 Vert. track outlines multiple alignment data for 100 vertebrate species and measurement of evolutionary conservation. Conservation track is also outlined for eight species of varying phylogenetic distances from humans. Areas of black and grey indicate ‘well’ and ‘less’ conserved regions, respectively (as defined by UCSC Genome Browser). Double line represents areas of excessive evolutionary distance between species generating long stretches of reduced homology; same for areas with no double lines or shading. Green arrow head indicates the location of a ‘newly enriched primate specific’ *cis*-regulatory site within the locus, as outlined in (Vermunt et al., 2016). **(B)** Phylogenetic tree outlining the evolutionary relationship among the various species. Red squares indicate orthologue annotations supported by Ensembl, NCBI, and UCSC genome bioinformatics data. Blue shades outline protein sequence homology (% identity) for isoform one and four of humans, first and second column, respectively.

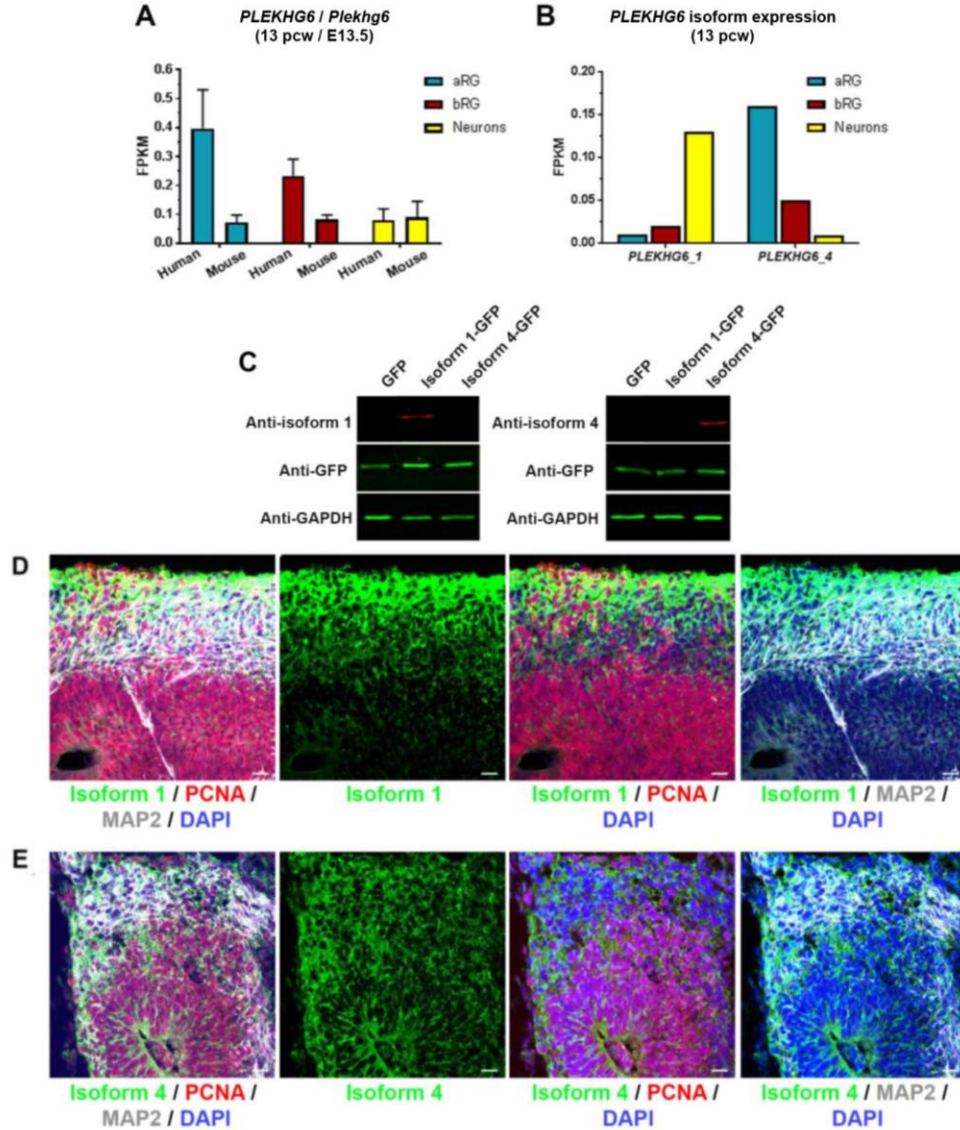


Figure S2. *PLEKHG6_1* and *PLEKHG6_4* are differentially regulated in developing human brain tissue. Related to Figure 1. (A) Normalised *PLEKHG6* and *Plekhg6* expression levels in 13 post-conceptional week (pcw) human or embryonic day (E) 13.5 mouse apical radial glia (aRG), basal radial glia (bRG), and post-mitotic neurons (N). (B) Normalised expression of *PLEKHG6* isoforms 1 and 4 (*PLEKHG6_1* and *PLEKHG6_4*, respectively) of human *PLEKHG6* in the three cell types at 13 pcw. Data in (A) and (B) adapted from (Florio et al., 2015); Error bars, SD. Fragments Per Kilobase of transcript per Million mapped reads (FPKM). n =2 biological replicates. (C) HEK293 cells were transiently transfected with a bi-cistronic vector expressing both green fluorescent protein (GFP) and isoform one or four of *PLEKHG6*. At 24 h after transfection, total cell lysates were prepared and subjected to western blot analysis using anti-isoform one or four antibodies, anti-GFP (transfection control) and anti-GAPDH as a loading control. (D, E) Micrographs of sections of cerebral organoids derived from human iPSCs (day 57 in culture) showing isoform one localisation with MAP2+ cells and isoform four localisation with PCNA+ and MAP2+ cells. Scale bar represents 30 μ m.

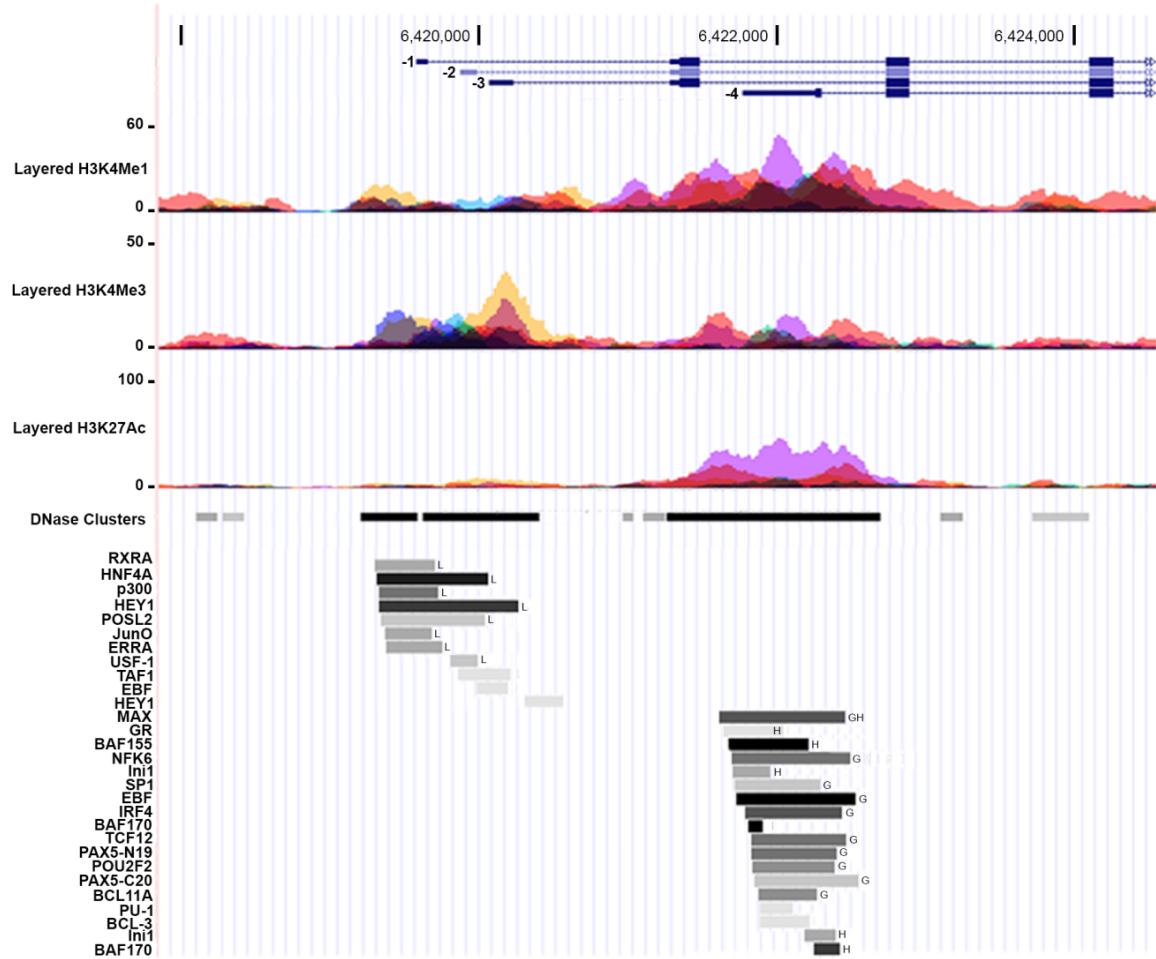


Figure S3. *PLEKHG6* has alternate isoforms that are potentially differentially regulated in distinct cell types. Related to Figure 1. Top panel indicates the relative genomic organization of the 5' non-coding first exons of isoforms one to four. Middle panel, reads per million mapped (RPM) normalized ChIP-seq reads for H3K4Me1, H3K4Me3 and H3K27Ac, layered across seven cell lines as outlined by ENCODE (axis limit 100 RPM). Sites of DNase hypersensitivity overlap defined *cis*-regulatory sites. Bottom panel, outlines transcription factor binding analysis at the *PLEKHG6* locus as defined by ChIP-seq data from ENCODE. In total 91 cell lines were assayed for a total of 161 transcription factors. The name of the transcription factors is identified to the left, with the line detected also labeled: L, hepatocellular carcinoma; G, lymphoblastoid; H, HeLa. Note transcription factors overlapping isoform one locus were all identified in hepatocellular carcinoma. In contrast, transcription factors overlapping isoform four were detected in either lymphoblastoid or HeLa cells.

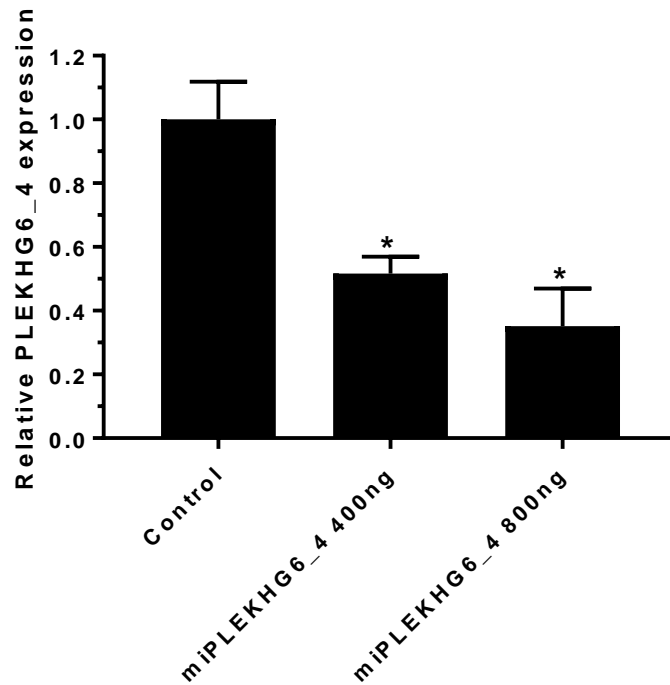


Figure S4. Validation of the miRNA targeting the primate specific isoform of PLEKHG6 (PLEKHG6_4). Related to Figure 2. Efficiency of miRNA knockdown targeted by PLEKHG6_4 (miPLEKHG6_4) at 400ng and 800ng was determined by transient co-transfection with a PLEKHG6_4 expression plasmid (pcDNA3.1V5/His-PLEKHG6_4), and data expressed as relative PLEKHG6_4 expression. Relative expression of PLEKHG6_4 was calculated for each miPLEKHG6_4 sample as the ratio of PLEKHG6_4 expression to mock transfected control via Western blot, using anti-V5 antibody to detect PLEKHG6 expression. Data presented as mean (\pm S.E.M). Mann-Whitney U test; * $p < 0.05$. $n = 3$.

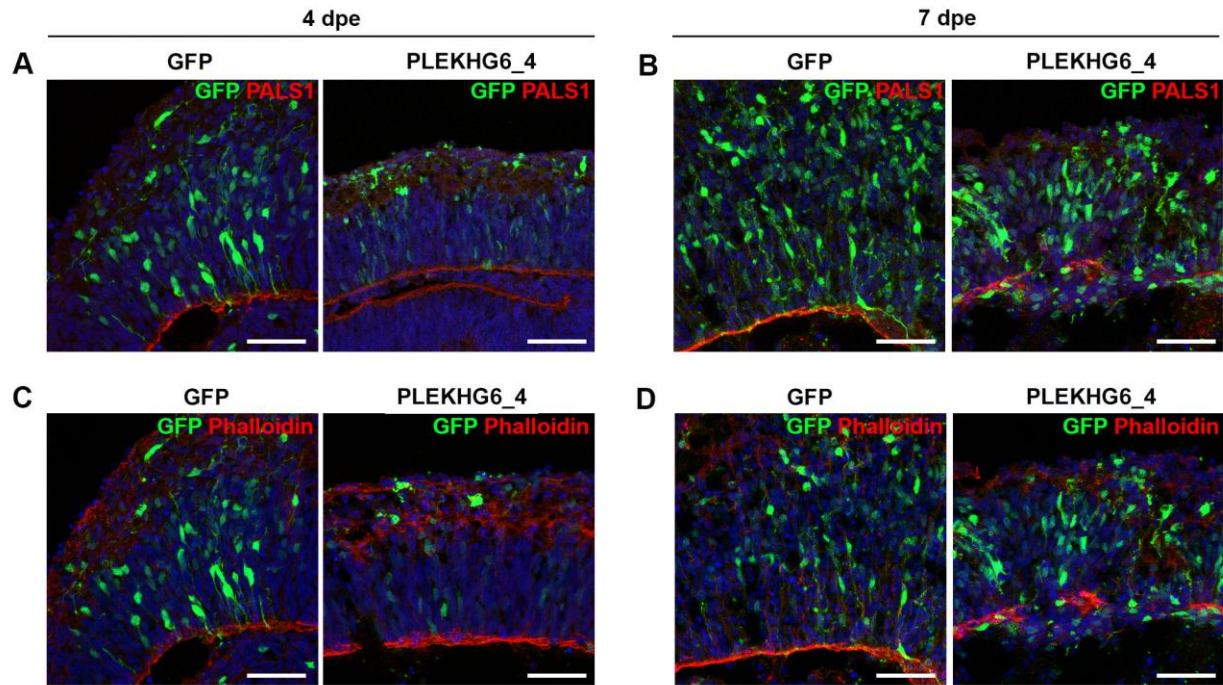


Figure S5. Progressive disruption of ventricular integrity in organoids overexpressing PLEKHG6_4. Related to Figure 3. (A, D) Micrographs sections of day 42 human cerebral organoids electroporated with GFP/empty vector control or human *PLEKHG6* isoform 4 (*PLEKHG6_4*) and analysed four (A, C) or seven (B, D) days post electroporation, 4 dpe and 7 dpe, respectively. Sections were then immunostained for PALS1 or Phalloidin as indicated. Scale bar represents 30μm.

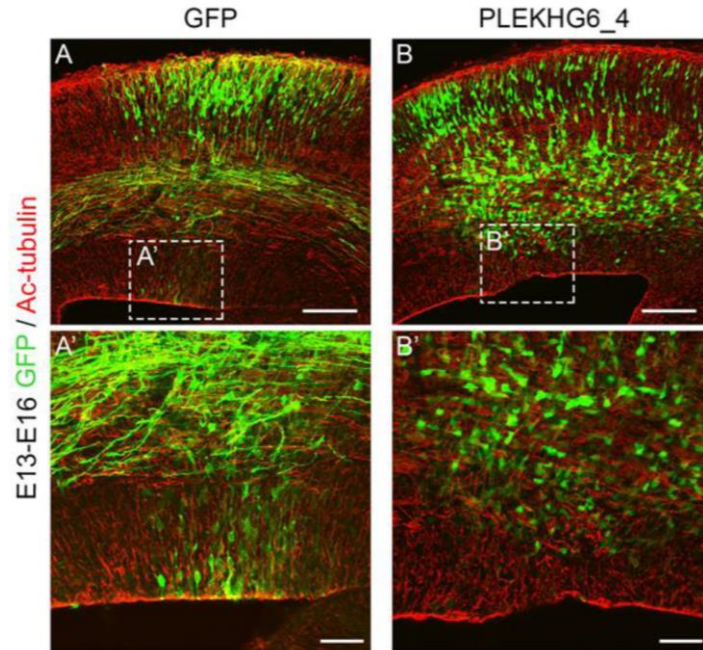


Figure S6. PLEKHG6_4 expression in the developing mouse cortex disrupts radial glia processes. Related to Figure 4. Coronal micrograph sections of E16 mouse cerebral cortices electroporated at E13 with (A) GFP/empty vector control or (B) human PLEKHG6 isoform 4 (PLEKHG6_4) and stained for GFP and acetylated tubulin (Ac-tubulin). (A', B') Higher magnifications of (A) and (B) showing disrupted radial glia processes upon PLEKHG6_4 overexpression. Scale bar represents (A, B) 100 μ m and (A', B') 50 μ m.

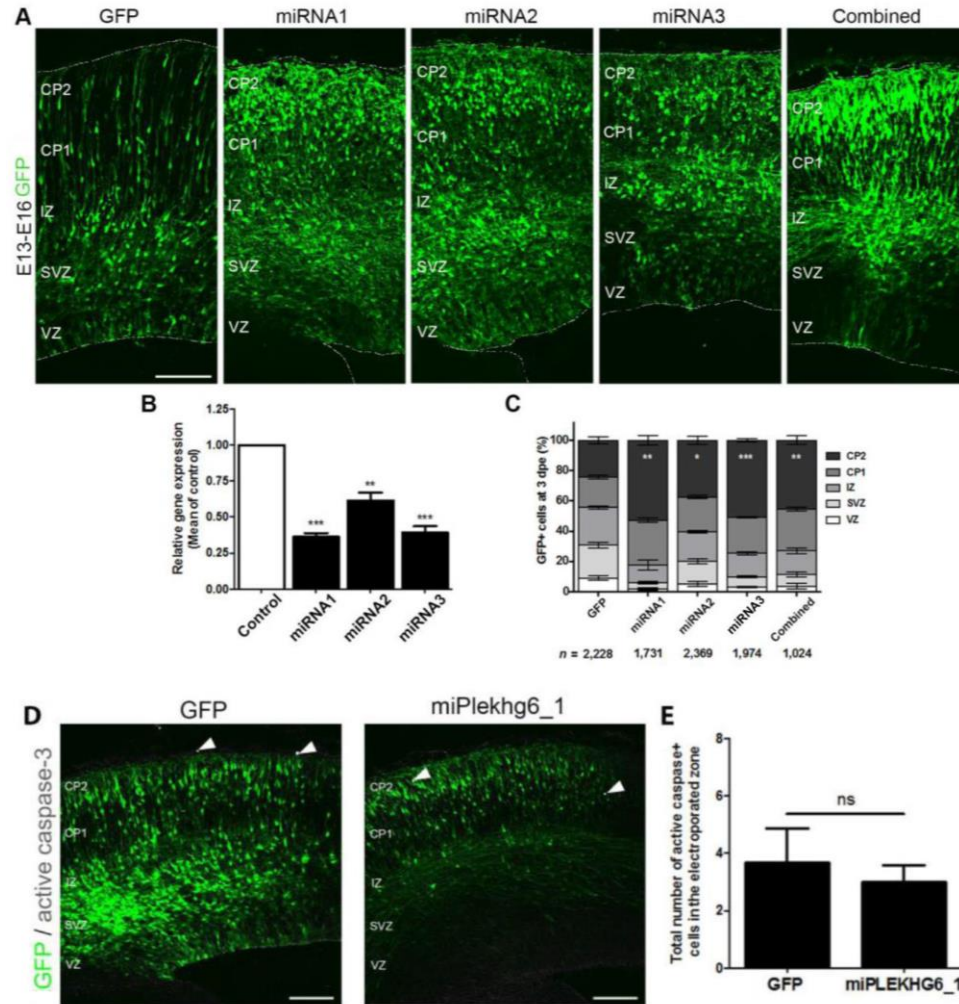


Figure S7. microRNA (miRNA) validation studies. Related to Figure 5. (A) Coronal micrograph sections of E16 mouse cerebral cortices electroporated at E13 with GFP/empty vector control, *Plekhg6* microRNAs (miRNA1,2,3 or combined). (B) Mean (\pm s.e.m) *Plekhg6* expression in P19 cells (n=3) 24 hours post transfection (1.5 μ g) with one of three miRNA targeting *Plekhg6* relative to control treated sample (empty vector) expression using the delta-delta Ct method, as determined by real-time PCR. Expression was normalised against *GAPDH* and *DIMT1* housekeeping genes in the same sample using the relative standard curve method. (C) Quantification of the distribution of GFP-expressing (GFP+) cells transfected with GFP/empty vector alone or *Plekhg6* miRNAs (miRNA1, 2 and 3) 3 days after electroporation (mean \pm s.e.m). (D) Coronal micrograph sections of E16 mouse cerebral cortices electroporated at E13 with GFP/empty vector control or *Plekhg6* targeting miRNAs (miPlekhg6). White arrows outline cells staining for active-caspase within the electroporated zone. Quantification of the total number of active caspase cells in the electroporated region within the treatment groups are presented as mean \pm s.e.m (E). At least three embryos analysed for each condition. n, total number of GFP+ cells counted per condition. Mann-Whitney U test; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$. Scale bar represents 100 μ m.