Figure S1: Study Overview, Related to Table 1

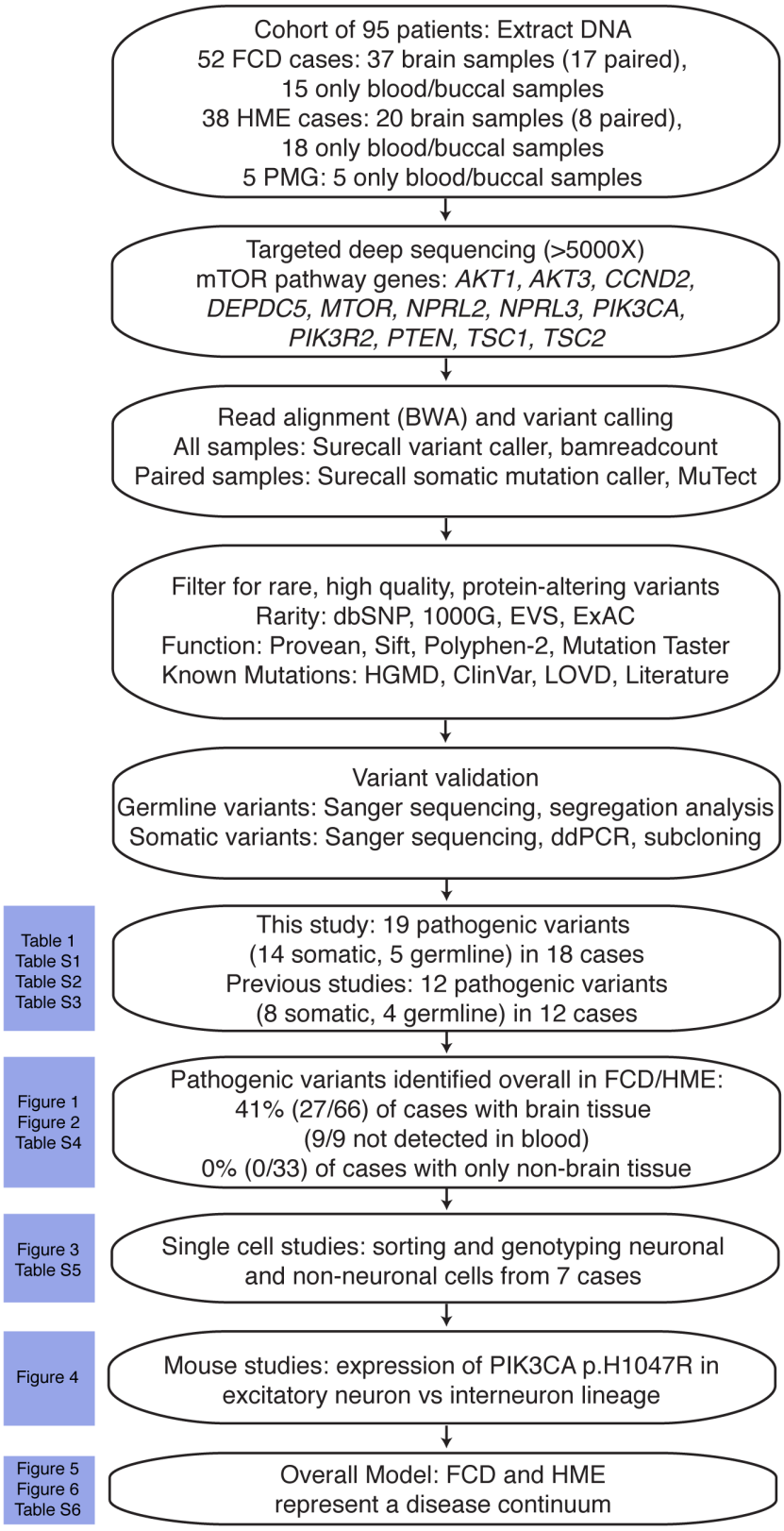

Figure S1: Study Overview, Related to Table 1: DNA extracted from brain, blood, and buccal samples from 95 FCD, HME, and PMG cases was sequenced at high coverage (>5000x) using targeted panels of mTOR pathway genes. After variant calling using four different methods and filtering for high quality, rare, protein-altering variants, validation was performed using Sanger sequencing, ddPCR, and/or subcloning. To investigate the role of cell type-specificity of abnormal mTOR activation in dysplasia pathogenesis, we then performed single cell sequencing for somatic mutation positive FCD and HME cases and studied mouse models with conditional expression of constitutively active *PIK3CA* p.H1047R in either the dorsal telencephalic or the interneuron lineage.

Table S3: Pathogenic Variants Identified in Our Previous Studies, Related to Figure 2

Subject	Diagnosis	Gene	Mutation	HGVS	Type	AAF	Comments
HME-16	HME	*AKT3*	Ms	p.E17K	Somatic (Not in blood)	17.4% Brain	Identified in Poduri et al., 2012
HME-17	HME	*AKT3*	CNV	Chromosome 1q tetrasomy	Somatic (Not in blood)	N/A	Identified in Poduri et al., 2012
HME-18	HME	*AKT3*	CNV	Chromosome 1q CN increase	Somatic	N/A	Identified in Poduri et al., 2012
FCD-1	FCD IIb	*DEPDC5*	Fs	p.N261Kfs*11	Germline	52.2%	Loss of function, Identified in D'Gama et al., 2015
FCD-2	FCD IIb	*DEPDC5*	Sp	c.624+1G>A	Germline	50.0%	Loss of function, Identified in D'Gama et al., 2015
HME-1	HME	*DEPDC5*	Fs	p.N45Qfs*3	Germline	61.2%	Loss of function, Identified in D'Gama et al., 2015
HME-2	HME	*MTOR*	Ms	p.C1483Y	Somatic	5.9-18.4%	Previously identified in HME (Lee et al., 2012), Functional studies suggest pathogenic (Grabiner et al., 2014)
HME-6	HME	*PIK3CA*	Ms	p.H1047R	Somatic	3.7-12.5%	Identified in D'Gama et al., 2015, Previously identified in CLOVES (Kurek et al., 2012) and FCD (Jansen et al., 2015)
HME-5	HME	*PIK3CA*	Ms	p.E545K	Somatic	8.1-16.7%	Previously identified in HME (Lee et al., 2012) and MCAP (Riviere et al., 2012)
HME-3	HME	*PIK3CA*	Ms	p.E542K	Somatic	8.8-27.5%	Identified in D'Gama et al., 2015, Previously identified in CLOVES (Kurek et al., 2012) and HME (Jansen et al., 2015)
HME-4	HME	*PIK3CA*	Ms	p.E545K	Somatic	18.1-44.4%	Previously identified in HME (Lee et al., 2012) and MCAP (Riviere et al., 2012)
HME-11	HME	*TSC2*	Ms	p.R1713H	Germline	51.0%	Identified in D'Gama et al., 2015, Previously identified in TSC (Hirfanoglu et al., 2010), Functional studies suggest pathogenic (Hoogeveen-Westerveld et al., 2011)

Table S5: Single Cell Genotyping Details, Related to Figure 3

Case	Cell Population	% Cells with mutation	SE	Range		Cells with mutation/ Total cells	p value
FCD-6	NeuN+	7.1%	1.8%	5.3%	8.9%	15/211	0.0005
	NeuN-	0.5%	0.5%	0.0%	1.1%	1/188	
FCD-14	NeuN+	24.0%	4.2%	19.8%	28.2%	25/104	0.00003
	NeuN-	0.0%	0.0%	0.0%	0.0%	0/50 (<1/50)	
FCD-12	NeuN+	6.2%	2.1%	4.0%	8.3%	8/130	0.5
	NeuN-	8.7%	4.2%	4.5%	12.9%	4/46	
FCD-8	NeuN+	20.0%	4.0%	16.0%	24.0%	20/100	0.008 (does not survive multiple testing correction)
	NeuN-	36.3%	4.3%	32.0%	40.6%	45/124	
HME-22	NeuN+	9.0%	2.5%	6.5%	11.4%	12/134	0.00002
	NeuN-	40.0%	7.7%	32.3%	47.7%	16/40	
HME-16	NeuN+	60.0%	5.2%	54.8%	65.2%	54/90	0.00001
	NeuN-	26.8%	4.9%	21.9%	31.7%	22/82	
HME-23	NeuN+	77.0%	4.5%	72.5%	81.5%	67/87	7E-09
	NeuN-	24.4%	6.4%	18.0%	30.9%	11/45	

Supplemental Experimental Procedures

Panel design, DNA library preparation, and DNA sequencing

We designed two custom Haloplex panels using Agilent's Suredesign software. The first panel targeted 10 genes: *AKT1*, *AKT3*, *CCND2*, *DEPDC5*, *MTOR*, *PIK3CA*, *PIK3R2*, *PTEN*, *TSC1*, and *TSC2*. The panel generated a target region of 36.6kbp that includes all exons, exon-intron boundaries, and 10bp of flanking sequence for each gene, and the design is predicted to cover 99.5% of the target region. The second panel targeted 12 genes, with the addition of *NPRL2* and *NPRL3*. The panel generated a target region of 40.41kbp that similarly includes all exons, exon-intron boundaries, and 10bp of flanking sequence for each gene, and the design is predicted to cover 99.5% of the target region. Samples were run on one or both panels.

DNA for library preparation was extracted from available tissues of 95 patients, which included surgically resected brain samples from 57 patients (37 FCD, 17 also had blood samples, and 20 HME, 8 also had blood samples; blood samples were sequenced and/or used for validation) and only blood or buccal samples from the remaining 38 patients (15 FCD, 18 HME, and 5 PMG). Paired-end, barcoded libraries were prepared per the manufacturer's protocol using 225ng of DNA from each sample and custom Haloplex Target Enrichment Kits (Agilent). Paired-end sequencing was performed on HiSeq sequencers (Illumina) at the HMS Biopolymers Core and we achieved coverage >5000X (average 5069X for Panel 1 and 6494X for Panel 2).

DNA sequencing analysis

To identify variants, we performed four different analyses. First, we performed germline and somatic variant calling for each sample using Agilent's software Surecall, which utilizes BWA for alignment and SAMtools or their proprietary caller for variant calling (Li and Durbin, 2009; Li et al., 2009). Importantly, Surecall allow users to optimize the settings and call variants with as low as 1% alternate allele frequency (while default settings are optimized for detecting germline variants). Second, we performed paired analysis for somatic variants using MuTect for cases where paired samples (brain and non-brain) were available for sequencing (Cibulskis et al., 2013). As MuTect was designed to detect somatic variants when comparing "tumor" to "normal" samples, it can detect very low frequency somatic variants. Third, we performed an additional paired analysis for somatic variants using Surecall's somatic variant caller for cases where paired samples (brain and non-brain) were available for sequencing. This caller is a relatively new feature of the Surecall software that similarly can detect very low frequency somatic variants. Finally, we performed a literature review and created a list of mutant alleles reported for FCD, HME, and related phenotypes, and performed a BAM read count analysis for each of those alleles for each sample to identify potential somatic variants missed by the other approaches. This comprehensive approach provided information about read counts at every assayed genomic location regardless of whether any alternate allele reads were present, and thus allowed detection of all alternate allele frequencies.

For each analysis, manual review of every variant was performed to filter for high quality, rare, and protein-altering variants and to distinguish potential somatic variants and false positives. Interestingly, every variant we validated as pathogenic was detected using the Surecall analysis approach; several were also detected using at least one additional approach. Surecall was likely highly successful due to the very high depth of coverage achieved, which made it easier to identify even very low frequency somatic variants and discriminate such variants from false positives.

Rare variants (minor allele frequency ≤ 1%) in genes in the mTOR pathway were filtered using dbSNP 137 (http://www.ncbi.nlm.nih.gov/SNP/) (Sherry et al., 2001), the 1000 Genomes Project (http://browser.1000genomes.org/index.html) (Abecasis et al., 2012), the Exome Variant Server (http://evs.gs.washington.edu/EVS/), and the Exome Aggregation Consortium (ExAC) (Lek et al., 2016). Previously reported mutations were identified using the Human Gene Mutation Database (http://www.hgmd.cf.ac.uk/ac/index.php) (Stenson et al., 2014), ClinVar (http://www.ncbi.nlm.nih.gov/clinvar/) (Landrum et al., 2014), and the Leiden Open Variation Database (http://chromium.liacs.nl/LOVD2/home.php) for *TSC1/2*. We used Provean (http://provean.jcvi.org/index.php) (Choi et al., 2012), Sift (http://sift.jcvi.org/) (Ng and Henikoff, 2003), Polyphen 2 (http://genetics.bwh.harvard.edu/pph2/) (Adzhubei et al., 2010), and Mutation Taster (http://www.mutationtaster.org/) (Schwarz et al., 2014) to assess for pathogenicity. Variants were considered somatic if (1) NGS showed an alternate allele frequency (AAF) <50% and we validated the AAF using ddPCR or subcloning for cases where only one tissue was available, and/or (2) the variant was present in brain tissue but not in non-brain tissue for cases where multiple tissues were available.

Variant Validation

For ddPCR, we designed primers flanking the mutation site and probes specific to the reference or mutant allele. A mix of ddPCR Super Mix (Bio-Rad, Hercules, CA), mutant and reference probes (0.25 μM each), forward

and reverse primers (0.9 μM each), and 30 ng of sample DNA was emulsified into 20,000 droplets using a QX100 Droplet Generator (Bio-Rad, Hercules, CA). PCR was performed using the following cycles: 10 min at 95°C, 40 cycles of 30 sec at 94°C and 60 sec at 52-60°C (see table below and (Luks et al., 2015) for specific annealing temperatures), 10 min at 98°C. Samples were analyzed using a QX100 Droplet Reader and QuantaSoft software (Bio-Rad, Hercules, CA).

ddPCR Primers and Probes

Mutation	**Primers (Forward, Reverse)**	**Probes (Reference, Mutant)**	**Annealing Temperature**
AKT3 p.E17K	GTCTTCAAAAGGAAGTATCTTG	ATATATTCTCCTACATGAGGAAAGC	52
	CTGTTCATTTTTTTCAGCA	TATATTTTCCTACATGAGGAAAGCAC	
MTOR p.L1460P	TCTGGGTCGTCCTTGTTG	AGGCCACAAGGGCATCCTCC	57
	GGCTACCTGGTATGAGAAACTG	AGGCCACAGGGGCATCCTC	
MTOR p.C1483R	TGCCATCATTCTAGGAAGCTC	AGGCCTCGAGGCAGCGCATG	57
	GACACCAACAAGGACGACC	AGGCCTCGAGGCGGCGCAT	
MTOR p.S2215Y	TGCCATTTCAGGGTTTCTG	CGAAGAGATGTTGGGTCATTGGCC	57
	GCTCTTCGGCCTGGTTAAC	CGAAGATATGTTGGGTCATTGGCCAG	
MTOR p.S2215F	TGCCATTTCAGGGTTTCTG	CGAAGAGATGTTGGGTCATTGGCC	57
	GCTCTTCGGCCTGGTTAAC	CCGAAGAAATGTTGGGTCATTGGCC	
MTOR p.T1977R	GTTGGCTGCATTGTGCC	TGGTAGACTTAGAAGCCACTGTCAG	57
	TCCCTCAGGCCCTCATC	TGGTAGACTTAGAAGCCACTCTCAG	
PIK3CA p.E542K	CAGCTCAAAGCAATTTCTAC	CTGAAATCACTGAGCAGGAGA	52
	CACTTACCTGTGACTCCAT	CTAAAATCACTGAGCAGGAGA	
TSC1 p.Q55*	CTTGTCATGTGGCTCTTGC	CAATGCCGGCTGAGAGCTGGTTT	57
	TATGCTTGTAAACACCTTGGTG	ATGCCGGCTAAGAGCTGGTTTCC	
TSC2 p.R751*	GGATTTGGTCATCAGCTTTCA	TGGAGCGGCTCCGAGGCG	57
	AGGTGCAAGTCAGTTCTGGAG	CTGGAGCGGCTCTGAGGCG	
TSC2 p.E1558K	TGATGCCACCCTGCCT	AGAGCAACAGCGAGCTCGCCATC	57
	TCCGTGTACCTGTAGGAGCC	AGAGCAACAGCAAGCTCGCCATCC	
TSC2 p.Y587*	TCTCTGCAGACCAAGCTGTA	CACGCCACGCGTGTGTATGAG	57
	GTGCTTGTAGTGGAGCTGAAT	CCACGCCACGCGTGTATGAG	

For subcloning, the original DNA was amplified using polymerase chain reaction (PCR), subcloned into a TOPO TA vector (Invitrogen, Carlsbad, CA), and transformed into TOP10 chemically competent *Escherichia coli* cells (Invitrogen, Carlsbad, CA); multiple clones were then isolated and sequenced.

Supplemental References

Abecasis, G.R., Auton, A., Brooks, L.D., DePristo, M.A., Durbin, R.M., Handsaker, R.E., Kang, H.M., Marth, G.T., and McVean, G.A. (2012). An integrated map of genetic variation from 1,092 human genomes. Nature 491, 56-65.

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. Nature methods 7, 248-249.

Choi, Y., Sims, G.E., Murphy, S., Miller, J.R., and Chan, A.P. (2012). Predicting the functional effect of amino acid substitutions and indels. PloS one 7, e46688.

Cibulskis, K., Lawrence, M.S., Carter, S.L., Sivachenko, A., Jaffe, D., Sougnez, C., Gabriel, S., Meyerson, M., Lander, E.S., and Getz, G. (2013). Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. Nature biotechnology 31, 213-219.

Landrum, M.J., Lee, J.M., Riley, G.R., Jang, W., Rubinstein, W.S., Church, D.M., and Maglott, D.R. (2014). ClinVar: public archive of relationships among sequence variation and human phenotype. Nucleic Acids Res 42, D980-985.

Lek, M., Karczewski, K.J., Minikel, E.V., Samocha, K.E., Banks, E., Fennell, T., O'Donnell-Luria, A.H., Ware, J.S., Hill, A.J., Cummings, B.B.*, et al.* (2016). Analysis of protein-coding genetic variation in 60,706 humans. Nature 536, 285-291.

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25, 1754-1760.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics 25, 2078-2079.

Luks, V.L., Kamitaki, N., Vivero, M.P., Uller, W., Rab, R., Bovee, J.V., Rialon, K.L., Guevara, C.J., Alomari, A.I., Greene, A.K.*, et al.* (2015). Lymphatic and other vascular malformative/overgrowth disorders are caused by somatic mutations in PIK3CA. J Pediatr 166, 1048-1054 e1041-1045.

Ng, P.C., and Henikoff, S. (2003). SIFT: Predicting amino acid changes that affect protein function. Nucleic Acids Res 31, 3812-3814.

Schwarz, J.M., Cooper, D.N., Schuelke, M., and Seelow, D. (2014). MutationTaster2: mutation prediction for the deep-sequencing age. Nature methods 11, 361-362.

Sherry, S.T., Ward, M.H., Kholodov, M., Baker, J., Phan, L., Smigielski, E.M., and Sirotkin, K. (2001). dbSNP: the NCBI database of genetic variation. Nucleic Acids Res 29, 308-311.

Stenson, P.D., Mort, M., Ball, E.V., Shaw, K., Phillips, A., and Cooper, D.N. (2014). The Human Gene Mutation Database: building a comprehensive mutation repository for clinical and molecular genetics, diagnostic testing and personalized genomic medicine. Human genetics 133, 1-9.