De novo germline and postzygotic mutations in AKT3, PIK3R2 and PIK3CA cause a spectrum of related megalencephaly syndromes

Jean-Baptiste Rivière, Ghayda M Mirzaa, Brian J O’Roak, Margaret Beddaoui, Diana Alcantara, Robert L Conway, Judith St-Onge, Jeremy A Schwartzentruber, Karen W Gripp, Sarah M Nickel, Thea Worthylake, Christopher T Sullivan, Thomas R Ward, Hailly E Butler, Nancy A Kramer, Beate Albrecht, Christine M Armour, Linlea Armstrong, Oana Calusereu, Cheryl Cytrynbaum, Beth A Drolet, A Michel Innes, Julie L Lauzon, Angela E Lin, Grazia M S Mancini, Wendy S Meschino, James D Reggin, Anand K Saggar, Tally Lerman-Sagie, Gökhan Uyanik, Rosanna Weksberg, Birgit Zirn, Chandree L Beaulieu, Finding of Rare Disease Genes (FORGE) Canada Consortium, Jacek Majewski, Dennis E Bulman, Mark O’Driscoll, Jay Shendure, John M Graham Jr, Kym M Boycott & William B Dobyns

Megalencephaly-capillary malformation (MCAP) and megalencephaly-polymicrogyria-polydactyly-hydrocephalus (MPPH) syndromes are sporadic overgrowth disorders associated with markedly enlarged brain size and other recognizable features. We performed exome sequencing in 3 families with MCAP or MPPH, and our initial observations were confirmed in exomes from 7 individuals with MCAP and 174 control individuals, as well as in 40 additional subjects with megalencephaly, using a combination of Sanger sequencing, restriction enzyme assays and targeted deep sequencing. We identified de novo germline or postzygotic mutations in three core components of the phosphatidylinositol 3-kinase (PI3K)-AKT pathway. These include 2 mutations in AKT3, 1 recurrent mutation in PIK3R2 in 11 unrelated families with MPPH and 15 mostly postzygotic mutations in PIK3CA in 23 individuals with MCAP and 1 with MPPH. Our data highlight the central role of PI3K-AKT signaling in vascular, limb and brain development and emphasize the power of massively parallel sequencing in a challenging context of phenotypic and genetic heterogeneity combined with postzygotic mosaicism.

As described in our recent clinical analysis in 42 children with MCAP or MPPH, the former consists of megalencephaly (or sometimes hemimegalencephaly), associated growth dysregulation with variable asymmetry, developmental vascular anomalies, distal limb malformations (syndactyly and polydactyly), variable cortical malformation and a mild connective tissue dysplasia (Fig. 1 and Supplementary Figs. 1 and 2). The patchy skin vascular malformations and asymmetric overgrowth seen in MCAP meet the criteria for the Klippel-Trenaunay subtype of vascular malformations and deregulated growth, which suggests that postzygotic mosaicism may be present in a subset of cases. The phenotype of MPPH resembles that of MCAP but lacks vascular malformations and syndactyly. We hypothesized that MCAP and MPPH result from mutations in the same pathway, and we therefore studied them together. Given the absence of recurrence of both syndromes in all reported families, we conducted exome sequencing in two parent-proband trios—one with MCAP (subject LR08-018; Fig. 1) and one with clinical features overlapping both MCAP and MPPH (subject LR08-018; Fig. 1)—and searched for de novo mutations. We also performed exome sequencing in the oldest of three affected siblings with MPPH (subject LR00-016a; Fig. 1) in the first known instance of familial recurrence of this syndrome, assuming either autosomal recessive inheritance or germline mosaicism in one parent.

We identified 247–254 rare protein-altering variants per proband that were not found in public variant databases or 112 other exomes (Online Methods and Supplementary Table 1). Analysis of the trio with proband LR08-018 identified a de novo mutation in AKT3 (c.1393C>T; p.Arg465Trp; Supplementary Table 2). Sanger sequencing of AKT3 in another 40 individuals with megalencephaly identified a second de novo mutation in this gene in subject LR11-354 (c.686A>G; p.Asn229Ser; Table 1), which supports mutations in AKT3 as a rare cause of megalencephaly (P = 0.002, calculated as the likelihood of observing a second de novo mutation in AKT3; Online Methods). AKT3 encodes the brain-predominant isoform of the AKT serine/threonine kinase, which is a major downstream mediator of PI3K signaling, leading us to focus on genes in this pathway in other individuals with megalencephaly.

A full list of affiliations appears at the end of the paper.

Received 30 March; accepted 29 May; published online 24 June 2012; doi:10.1038/ng.2331
Analysis of LR00-016a1 revealed four genes with one homozygous or two compound heterozygous mutations, consistent with autosomal recessive inheritance. However, filtering of variants using genome-wide SNP data from all three affected siblings excluded all four chromosomal loci (Online Methods and Supplementary Table 3). We next manually examined the list of 247 rare variants and identified a heterozygous mutation in PIK3R2 (c.1117G>A; p.Gly373Arg; Table 1), which encodes the p85β regulatory subunit of class IA PI3K13. Sanger sequencing confirmed the presence of the mutation in all three affected siblings and its absence in saliva and blood in both parents and the unaffected sister, showing germline mosaicism in one parent. Sequencing of PIK3R2 in 40 individuals with megalencephaly identified the same nucleotide change in 10 additional subjects with MPPH, and this mutation was shown to be de novo in all 7 subjects for whom parental DNA was available (Table 1). The mutation occurred at a CpG dinucleotide, which might explain its recurrence12.

We identified four candidate de novo variants in the trio with proband LR09-006 and confirmed a de novo substitution in the exon 3 splice-donor consensus sequence of BPDL1 (Supplementary Table 2). Screening of BPDL1 in 12 additional subjects detected no other de novo mutations, thus failing to support a causative role for this sequence change. Given the clinical presentation of MCAP and the recent report of postzygotic mutations of AKT1 in Proteus syndrome13, we speculated that MCAP might result from low-level mosaic mutations missed by our standard pipeline. In parallel, we performed a second analysis of de novo mutations in this trio by including the raw variants that did not meet our initial hard-filtering criteria (Online Methods). The second approach identified a missense change in PIK3CA (c.2740G>A; p.Gly914Arg; Supplementary Table 2), which encodes the p110α catalytic subunit of class IA PI3K14. This mutation was supported by 20 of 177 reads (11%) in the exome sequencing data, and it was confirmed to be de novo and mosaic by Sanger sequencing and a custom restriction enzyme assay (Online Methods and Supplementary Fig. 3a). We then sequenced the coding exons of PIK3CA in 29 individuals with megalencephaly with no mutations in AKT3 or PIK3R2 and identified 14 additional PIK3CA mutations, with mutant allele frequencies ranging from 10–50% (Table 1 and Supplementary Table 4).

Standard variant calling in exomes from seven additional subjects with MCAP identified a mutation of PIK3CA (c.1133G>A; p.Cys378Tyr) that was supported by 68 of 250 reads (27%) in subject 44735 (Table 1 and Online Methods). This mutation showed variable levels of mosaicism depending on the tissue tested (Supplementary Fig. 3b). Manual inspection of the Sequence Alignment/Map (SAM) files of the remaining six subjects with MCAP of unknown cause using the Integrative Genomics Viewer (IGV)15 revealed other candidate mosaic mutations in PIK3CA, with mutations represented by 2–15% of the total reads. To differentiate between sequencing errors and putative mutations, we systematically searched for low-level mosaicism of PIK3CA in the exomes from these 6 individuals and 174 control subjects (Online Methods and Supplementary Table 5). Although data from both cohorts contained many sites with 1 or 2 variant reads (most expected to be sequencing errors), we found only 12 variant sites supported by 3 or more reads. A significantly higher frequency of such sites was found in the MCAP cohort. Using a threshold of four variant reads, we identified candidate mosaic mutations in five of six affected individuals and none in control individuals (Supplementary Table 6). Sanger sequencing, a custom restriction enzyme assay or both confirmed all five mutations (Supplementary Table 4). The only remaining subject with MCAP of unknown cause (115422) had a mutation encoding a p.Ala1035Val variation that was supported by 3 of 185 reads (Supplementary Table 6), which was confirmed by Sanger sequencing and restriction enzyme assay in DNA from saliva and buccal swab. We previously found and validated the same nucleotide substitution in four tissues from another subject with MCAP (LR11-270; Supplementary Fig. 3c). Six other variant sites supported by three reads were identified in five individuals, including two sites in subject 162-001P with MCAP in addition to the pathogenic mutation, neither of which was confirmed by Sanger sequencing in blood or saliva, and four sites in control individuals. This suggests that variant sites supported by one to three reads are mostly sequencing artifacts (Supplementary Table 6).

Given the limitations of Sanger sequencing for detecting low-level mosaic mutations, we performed targeted ultra-deep sequencing (coverage of >10,000 reads) of 5 mutation sites in 15 mutation-negative affected individuals, as well as in known mutation carriers and control individuals (Online Methods and Supplementary Table 7). This experiment confirmed all previously identified mutations and detected two additional low-level mosaic mutations missed by Sanger sequencing (Table 2). Both were confirmed by a second deep-sequencing experiment and showed mutant allele frequencies ranging from 1–8%.

Class IA PI3Ks are heterodimeric enzymes that convert phosphatidylinositol 4,5-bisphosphate (PIP2) to phosphatidylinositol 4,5-trisphosphate (PIP3), a reaction that is reversed by the PTEN phosphatase16. This reaction leads to activation of the PI3K-AKT-mTOR network, which can be monitored by examining the extent of...
phosphorylation of downstream targets, such as S6 ribosomal protein and eukaryotic translation initiation factor 4E-binding protein 1 (4E-BP1)\(^\text{17}\). To assess the impact of AKT3, PIK3R2 and PIK3CA mutations on PI3K activity, we used immunostaining to compare PIP\(_3\) amounts in lymphoblastoid cell lines derived from four mutation carriers with megalencephaly to those in control and P1TEN-mutant cells (Fig. 2). Consistent with elevated PI3K activity—and similar to what is seen with P1TEN loss—all three lines with PIK3R2 or PIK3CA mutations showed significantly more PIP\(_3\) staining than control cells (Fig. 2b), as well as greater localization of active phosphoinositide-dependent kinase 1 (PDK1) to the cell membrane\(^\text{18}\) (Supplementary Fig. 4). Treatment with the PI3K inhibitor PI-103 resulted in less PIP\(_3\) in the PIK3R2 p.Gly373Arg and PIK3CA p.Glu453del mutant lines, confirming that these observations are PI3K-dependent (Fig. 2c). We found no evidence for increased PI3K activity in the AKT3-mutant line, consistent with a mutation affecting a downstream effector of PI3K. Protein blot analysis showed higher amounts of phosphorylated S6 protein and 4E-BP1 in all mutant cell lines compared to controls (Supplementary Fig. 5). Although PI-103 treatment reduced S6 phosphorylation in control and mutant lines, the latter showed relative resistance to PI3K inhibition, consistent with elevated signaling through the pathway. Taken together, these observations support the conclusion that the megalencephaly-associated mutations result in higher PI3K activity and PI3K-mTOR signaling.

The PI3Ks are a highly conserved family of signaling enzymes that regulate a wide range of processes, including cell growth, proliferation, survival, migration, metabolism, angiogenesis, apoptosis, tumorigenesis and brain development\(^\text{16,17,19–21}\). As the predominant downstream effector of PI3K signaling, AKT kinases are involved in a wide range of human diseases and have a critical role in growth regulation\(^\text{15}\). Extensive mouse and human data have shown that loss- and gain-of-function mutations in the genes encoding the AKT isoforms lead to opposite phenotypes\(^\text{13,22–25}\). Several reports have shown that heterozygous loss of AKT3 in humans and homzygous loss in mice cause microcephaly,\(^\text{26–28}\), whereas a heterozygous missense mutation in AKT3 (encoding p.Asp219Val) resulting in increased kinase activity was shown to cause seizures and brain overgrowth in mice\(^\text{29}\).

This phenotype resembles the clinical presentation of the two AKT3 mutations identified in the patients, raising the possibility that PI3K-mTOR signaling is uniquely sensitive to the loss of its regulatory subunit, AKT3.
mutation carriers in our study (Supplementary Table 8). Further, another group recently reported that somatic activating mutations in AKT3 in brain cause hemimegalencephaly.\(^{30}\) Many subjects in our series had asymmetric brain enlargement, and several were diagnosed with hemimegalencephaly (Supplementary Fig. 2). These data, combined with our observations of higher S6 and 4E-BP1 phosphorylation in a lymphoblastoid cell line harboring the AKT3 p.Arg465Trp alteration, strongly suggest a gain-of-function mechanism resulting in enhanced AKT activity for the two AKT3 mutations reported here.

Class IA PI3K dimers are composed of a p110 catalytic subunit and a p85 regulatory subunit, each of which has three isoforms encoded by three genes.\(^{17}\) Mutations in five of these genes (PIK3CA, PIK3CB, PIK3CD, PIK3R1 and PIK3R2) have been observed in many human cancers.\(^{31—34}\) Our data show that mutations affecting the p85β (PIK3R2) regulatory and p110α (PIK3CA) catalytic subunits are a common cause of megalencephaly syndromes, albeit with a clear genotype-phenotype correlation, as PIK3R2 and PIK3CA mutations are associated with MPPH (\(P = 3.3 \times 10^{-6}\)) and MCAP (\(P = 1.0 \times 10^{-6}\)), respectively (Online Methods and Supplementary Table 9). Both PIK3R1 and PIK3R2 have oncogenic potential, and alterations—including the glycine-to-arginine substitution in PIK3R2 found in MPPH (p.Gly373Arg) and substitution of the homologous amino-acid residue in PIK3R1 (p.Gly376Arg)—have been found in cancer.\(^{32}\) Available functional studies showed that several of these alterations disrupt the inactive conformation of the PI3K dimer and maintain the catalytic subunit in a

**Figure 2** PIP\(_3\) levels in lymphoblastoid cell lines derived from an unaffected control (WT), an individual with Cowden disease (GM10080) and four individuals with megalencephaly. The individual with Cowden disease (PTEN p.Glu261Ter) served as a positive control. (a) Indirect immunofluorescence staining of PIP\(_3\) in exponentially growing lymphoblastoid cell lines. DAPI, 4',6-diamidino-2-phenylindol. Scale bar, 10 μm. (b) Per-cell quantification of PIP3 using the signal intensity of staining with antibody to PIP3 (a.u., arbitrary units). Statistically significant differences compared to control cells are indicated: \(* P < 0.05\), two-tailed t test, assuming unequal variance, \(n = 30–50\) cells per cell line. Error bars, s.d. (c) Staining for PIP3 (green) and DAPI (blue) in cell lines from LR00-016a1 and LR05-204 following treatment with the PI3K inhibitor PI-103 (5 μM for 16 h). Scale bar, 10 μm.

### Table 2 PIK3CA mutations identified or confirmed by deep sequencing of five mutation sites

| Subject ID | Source of DNA | Mutation coordinates (hg19) | Amino-acid change | Total reads\(^{a}\) | Mutant reads\(^{a}\) | Percent mutant allele | Mean mutant allele freq.\(^{b}\) (× 10\(^{-4}\)) | Max mutant allele freq.\(^{b}\) (× 10\(^{-4}\)) |
|------------|---------------|-----------------------------|-------------------|-----------------|-----------------|----------------------|---------------------------------|---------------------------------|
| **Confirmed PIK3CA mutations** | | | | | | | | |
| LR06-220 | Blood | Chr. 3: g.178916876G>A | p.Arg88Gln | 149,513 | 64,260 | 43 | 1.7 | 2.9 |
| 44735 | Blood | Chr. 3: g.178922364G>A | p.Cys78Tyr | 285,251 | 84,600 | 30 | 0.4 | 1.1 |
| LR08-261 | Blood | Chr. 3: g.178938934G>A | p.Glu726Lys | 293,540 | 35,865 | 12 | 0.5 | 0.7 |
| Buccal swab | | | | | | | | |
| LR06-333 | LCL | Chr. 3: g.178938934G>A | p.Glu726Lys | 7,087 | 2,110 | 41 | 0.5 | 0.7 |
| LR09-006 | Blood | Chr. 3: g.178947865G>A | p.Gly914Arg | 392,036 | 61,427 | 16 | 19 | 23 |
| LR11-070 | LCL | Chr. 3: g.178947865G>A | p.Gly914Arg | 663,398 | 102,811 | 15 | 19 | 23 |
| LR06-341 | LCL | Chr. 3: g.178947865G>A | p.Gly914Arg | 327,763 | 56,313 | 17 | 19 | 23 |
| LR11-069 | LCL | Chr. 3: g.178952018A>G | p.Thr1025Ala | 3,586,333 | 103,171 | 19 | 19 | 23 |
| LR11-068 | Blood | Chr. 3: g.178952018A>G | p.Thr1025Ala | 3,586,333 | 103,171 | 19 | 19 | 23 |
| **New PIK3CA mutations** | | | | | | | | |
| LR06-342 | Saliva | Chr. 3: g.178916854G>A | p.Glu81Lys | 51,268 | 4,227 | 8 | 0.4 | 0.5 |
| LR11-068 | Blood | Chr. 3: g.178916876G>A | p.Arg88Gln | 117,487 | 2,921 | 2 | 1.7 | 2.9 |

**Validation of new PIK3CA mutations**

| Subject ID | Source of DNA | Mutation coordinates (hg19) | Amino-acid change | Total reads\(^{a}\) | Mutant reads\(^{a}\) | Percent mutant allele | Mean mutant allele freq.\(^{b}\) (× 10\(^{-4}\)) | Max mutant allele freq.\(^{b}\) (× 10\(^{-4}\)) |
|------------|---------------|-----------------------------|-------------------|-----------------|-----------------|----------------------|---------------------------------|---------------------------------|
| LR06-342 | Saliva | Chr. 3: g.178916854G>A | p.Glu81Lys | 135,801 | 993 | 1 | 0.2 | 0.2 |
| LR11-068 | Blood | Chr. 3: g.178916876G>A | p.Arg88Gln | 82,864 | 2,783 | 3 | 0.2 | 0.2 |
| LR11-068 | Saliva | Chr. 3: g.178916876G>A | p.Arg88Gln | 282,771 | 3,378 | 1 | 0.2 | 0.3 |
| LR11-068 | Saliva | Chr. 3: g.178916876G>A | p.Arg88Gln | 45,578 | 1,714 | 4 | 0.2 | 0.3 |

LCL, lymphoblastoid cell line; chr., chromosome; freq., frequency.

\(^{a}\)Only bases with base quality of ≥20 were considered.

\(^{b}\)Average frequency of the mutant allele in control individuals.

\(^{c}\)Highest mutant allele frequency in control individuals.
high-activity state. Our observations in lymphoblastoid cells derived from subject LR00-016a1 show that the p.Gly373Arg alteration results in increased PI3K activity and elevated PI3K-mTOR signaling, further supporting this mechanism.

We identified 24 affected individuals with PIK3CA mutations, and all but three (LR06-220, LR11-153 and LR11-230) showed evidence of postzygotic mosaic. These mutations were discovered and confirmed by independent experiments using four different methods and multiple, mostly uncultured tissues, thus ruling out the possibility of cell culture or technology artifacts. As in the study on Proteus syndrome, we observed lower levels of mosaicism in blood compared to other tissues (Supplementary Table 4).

Somatic activating mutations in PIK3CA are frequently observed in several common human tumor types. This may partly explain the mildly higher (~3%) incidence of cancer in individuals with MCAP and MPPH (Supplementary Table 10 and Supplementary Note). Three mutation hotspots (encoding p.Glu545Lys, p.Glu545Lys and His1047Arg) account for 80% of all tumor-associated PIK3CA mutations and show the highest oncogenic activity. Although 13 of 15 PIK3CA mutations found in our megalencephaly cohort are reported in the Catalogue of Somatic Mutations in Cancer (COSMIC; Table 1), only 1 subject with megalencephaly carried 1 of these severe PIK3CA cancer mutation hotspots (LR12-033; Table 1). Notably, the brain phenotype of this individual appeared more severe than that of other affected subjects (Supplementary Fig. 2). Among the MCAP-causing mutations previously associated with cancer, functional data reported for the mutations encoding p.Arg88Gln, p.Glu365Lys, p.Glu365Lys, p.Met1043Ile and p.His1047Tyr show increased lipid kinase activity resulting in constitutive PI3K signaling, consistent with our observations of increased PI3K activity and signaling in lymphoblastoid cell lines from subjects with MCAP carrying PIK3CA mutations.

Altogether, we identified germline or postzygotic mutations of AKT3, PIK3R2 or PIK3CA in 37 of 50 (74%) unrelated probands (Fig. 3 and Supplementary Table 9). Mutations of all three genes led to pre- and postnatal overgrowth of brain and variably body. The brain overgrowth phenotype universally included significant megalencephaly and reduced extra-axial spaces and was frequently accompanied by polymicrogyria, hydrocephalus and cerebellar tonsillar ectopia, including Chiari malformation (Supplementary Table 8). Individuals with MCAP also presented striking—often asymmetric—overgrowth of many other tissues, especially vascular, subcutaneous, connective and lymphatic tissues. Considering the clinical complexity of these syndromes and the potential importance of our findings for clinical care, we have summarized our provisional recommendations for clinical management in the Supplementary Note.

Although we cannot rule out the possibility of low-level mosaic mutations that are undetectable by Sanger sequencing, none of the 13 individuals with MCAP of unknown cause had mutations in PTEN or the four other genes encoding subunits of class IA PI3K (Online Methods). Further, some of these subjects might carry mosaic PIK3CA mutations undetectable by Sanger sequencing in the tissues tested. Our data emphasize the challenges of detecting low-level mosaic mutations but, nonetheless, highlight the power of massively parallel sequencing for discovering postzygotic mutations across the genome.

MPPH and MCAP share clinical features with Proteus syndrome and a spectrum of overgrowth disorders resulting from loss-of-function mutations in PTEN, including Bannayan-Riley-Ruvalcaba syndrome, Cowden disease and autism with severe megalencephaly. Our results extend the list of overgrowth syndromes associated with aberrant PI3K-akt signaling and show that dysfunction in this pathway causes a constellation of brain, vascular and limb malformations. Finally, our findings, in combination with the development of PI3K inhibitors for the treatment of human disease, may open the door to new therapeutic opportunities for megalencephaly and other developmental disorders.

METHODS

Methods and any associated references are available in the online version of the paper.

Accession codes. Exome sequencing data from individuals in the three index families have been deposited in the database of Genotypes and Phenotypes (dbGaP) under accession phs000455.v1.p1.

Note: Supplementary information is available in the online version of the paper.

ACKNOWLEDGMENTS

We wish to thank all of the children and families in this study, their referring physicians and the M-CM Network (see URLs) for their help with this project over many years. We thank the members of the Northwest Genomics Center and the McGill University and Genome Quebec Innovation Centre for their excellent technical assistance. We also thank the Finding of Rare Disease Genes (FORGE) Canada Consortium, especially J. Marcadier for her contribution to the infrastructure.

Figure 3 Distribution of alterations in AKT3, PIK3R2 and PIK3CA. The activating Akt3 variation in mouse (p.Asp219Val) is indicated in gray. For recurrent variations, the number of occurrences is indicated in parentheses. PH, pleckstrin homology domain; SH2 and SH3, Src-homology 2 and 3 domains; RhoGAP, Rho GT-Pase–activating protein domain; p85 BD and Ras BD, p86- and Ras-binding domains; C2, protein kinase C homology–2 domain. The PIK3CA mutations affected a total of 15 residues, mainly localized in the p85-binding, C2, protein kinase C homology–2 domain.

Supplementary Note

Table 1

Distribution of alterations in AKT3, PIK3R2 and PIK3CA. The activating Akt3 variation in mouse (p.Asp219Val) is indicated in gray. For recurrent variations, the number of occurrences is indicated in parentheses. PH, pleckstrin homology domain; SH2 and SH3, Src-homology 2 and 3 domains; RhoGAP, Rho GT-Pase–activating protein domain; p85 BD and Ras BD, p86- and Ras-binding domains; C2, protein kinase C homology–2 domain. The PIK3CA mutations affected a total of 15 residues, mainly localized in the p85-binding, C2 and catalytic lipid kinase domains.

Supplementary Note

Table 2

Distribution of alterations in AKT3, PIK3R2 and PIK3CA. The activating Akt3 variation in mouse (p.Asp219Val) is indicated in gray. For recurrent variations, the number of occurrences is indicated in parentheses. PH, pleckstrin homology domain; SH2 and SH3, Src-homology 2 and 3 domains; RhoGAP, Rho GT-Pase–activating protein domain; p85 BD and Ras BD, p86- and Ras-binding domains; C2, protein kinase C homology–2 domain. The PIK3CA mutations affected a total of 15 residues, mainly localized in the p85-binding, C2 and catalytic lipid kinase domains.

Supplementary Note

Table 3

Distribution of alterations in AKT3, PIK3R2 and PIK3CA. The activating Akt3 variation in mouse (p.Asp219Val) is indicated in gray. For recurrent variations, the number of occurrences is indicated in parentheses. PH, pleckstrin homology domain; SH2 and SH3, Src-homology 2 and 3 domains; RhoGAP, Rho GT-Pase–activating protein domain; p85 BD and Ras BD, p86- and Ras-binding domains; C2, protein kinase C homology–2 domain. The PIK3CA mutations affected a total of 15 residues, mainly localized in the p85-binding, C2 and catalytic lipid kinase domains.

Supplementary Note

Table 4

Distribution of alterations in AKT3, PIK3R2 and PIK3CA. The activating Akt3 variation in mouse (p.Asp219Val) is indicated in gray. For recurrent variations, the number of occurrences is indicated in parentheses. PH, pleckstrin homology domain; SH2 and SH3, Src-homology 2 and 3 domains; RhoGAP, Rho GT-Pase–activating protein domain; p85 BD and Ras BD, p86- and Ras-binding domains; C2, protein kinase C homology–2 domain. The PIK3CA mutations affected a total of 15 residues, mainly localized in the p85-binding, C2 and catalytic lipid kinase domains.

Supplementary Note

Table 5

Distribution of alterations in AKT3, PIK3R2 and PIK3CA. The activating Akt3 variation in mouse (p.Asp219Val) is indicated in gray. For recurrent variations, the number of occurrences is indicated in parentheses. PH, pleckstrin homology domain; SH2 and SH3, Src-homology 2 and 3 domains; RhoGAP, Rho GT-Pase–activating protein domain; p85 BD and Ras BD, p86- and Ras-binding domains; C2, protein kinase C homology–2 domain. The PIK3CA mutations affected a total of 15 residues, mainly localized in the p85-binding, C2 and catalytic lipid kinase domains.

Supplementary Note

Table 6

Distribution of alterations in AKT3, PIK3R2 and PIK3CA. The activating Akt3 variation in mouse (p.Asp219Val) is indicated in gray. For recurrent variations, the number of occurrences is indicated in parentheses. PH, pleckstrin homology domain; SH2 and SH3, Src-homology 2 and 3 domains; RhoGAP, Rho GT-Pase–activating protein domain; p85 BD and Ras BD, p86- and Ras-binding domains; C2, protein kinase C homology–2 domain. The PIK3CA mutations affected a total of 15 residues, mainly localized in the p85-binding, C2 and catalytic lipid kinase domains.

Supplementary Note

Table 7

Distribution of alterations in AKT3, PIK3R2 and PIK3CA. The activating Akt3 variation in mouse (p.Asp219Val) is indicated in gray. For recurrent variations, the number of occurrences is indicated in parentheses. PH, pleckstrin homology domain; SH2 and SH3, Src-homology 2 and 3 domains; RhoGAP, Rho GT-Pase–activating protein domain; p85 BD and Ras BD, p86- and Ras-binding domains; C2, protein kinase C homology–2 domain. The PIK3CA mutations affected a total of 15 residues, mainly localized in the p85-binding, C2 and catalytic lipid kinase domains.

Supplementary Note

Table 8

Distribution of alterations in AKT3, PIK3R2 and PIK3CA. The activating Akt3 variation in mouse (p.Asp219Val) is indicated in gray. For recurrent variations, the number of occurrences is indicated in parentheses. PH, pleckstrin homology domain; SH2 and SH3, Src-homology 2 and 3 domains; RhoGAP, Rho GT-Pase–activating protein domain; p85 BD and Ras BD, p86- and Ras-binding domains; C2, protein kinase C homology–2 domain. The PIK3CA mutations affected a total of 15 residues, mainly localized in the p85-binding, C2 and catalytic lipid kinase domains.

Supplementary Note

Table 9

Distribution of alterations in AKT3, PIK3R2 and PIK3CA. The activating Akt3 variation in mouse (p.Asp219Val) is indicated in gray. For recurrent variations, the number of occurrences is indicated in parentheses. PH, pleckstrin homology domain; SH2 and SH3, Src-homology 2 and 3 domains; RhoGAP, Rho GT-Pase–activating protein domain; p85 BD and Ras BD, p86- and Ras-binding domains; C2, protein kinase C homology–2 domain. The PIK3CA mutations affected a total of 15 residues, mainly localized in the p85-binding, C2 and catalytic lipid kinase domains.
Megalencephaly and perisylvian polymicrogyria with postaxial α Insulin resistance and a diabetes mellitus–like syndrome in Cellular function of phosphoinositide 3-kinases: implications for 43 x 22 Nature Ge 43 x 73 12. 8. 3. 2. Reprints and permissions information is available online at http://www.nature.com/ 450. N.A.K. and C.L.B. provided administrative support and recruited the study 459. J.M.G., K.M.B. and W.B.D. recruited and evaluated the study subjects. H.E.B., 477. contributed to the genetics experiments. J.-B.R., J.A.S. and B.J.O. performed the 495. designed and performed the genetics experiments. M.B., T.W, C.T.S. and T.R.W. 513. J.-B.R., G.M.M., K.M.B. and W.B.D. designed the study. J.-B.R., B.J.O. and J.S.-O. 531. GO (HL-102923), Broad GO (HL-10925), Seattle GO (HL-102926), Heart GO 550. the NIEHS Environmental Genome Project (contract HHSN273200800010C) 568. contributed to the genetics experiments. G.M.M., R.L.C., K.W.G., S.M.N., B.A., C.M.A., L.A., O.C., C.C., B.A.D., 586. J.A. and K.W.B. designed the study. J.-B.R., G.M.M. and R.L.C. wrote the manuscript. 598. COMPETING FINANCIAL INTERESTS 610. The authors declare no competing financial interests. 623. Published online at http://www.nature.com/doifinder/10.1038/ng.2331. 630. Reprints and permissions information is available online at http://www.nature.com/ 640. reprints/index.html.

1. Clayton-Smith, J. et al. Macrocephaly with cutis marmorata, haemangioma and syndactyly—a distinctive overgrowth syndrome. Clin. Dysmorphol. 6, 291–302 (1997).

2. Mirzaa, G. et al. Megalencephaly and perisylvian polymicrogyria with postaxial polydactyly and hydrocephalus: a rare brain malformation syndrome associated with mental retardation and seizures. Neuropediatrics 35, 353–359 (2004).

3. Moore, C.A. et al. Macrocephaly-cutis marmorata telangiectasia congenita: a distinct disorder with developmental delay and connective tissue abnormalities. Am. J. Med. Genet. A. 151A, 269–291 (2012).

4. Conway, R.L. et al. Neuroimaging findings in macrocephaly-capillary malformation: a longitudinal study of 17 patients. Am. J. Med. Genet. A. 143A, 2981–3008 (2007).

5. Oudkerk, C.E. et al. A proposal for classification of entities combining vascular malformations and deregulated growth. Eur. J. Med. Genet. 54, 262–271 (2011).

6. Happle, R. Lethal genes surviving by mosaicism: a possible explanation for sporadic birth defects involving the skin. J. Am. Acad. Dermatol. 16, 899–906 (1997).

7. Gripp, K.W. et al. Significant overlap and possible identity of macrocephaly capillary malformation and megalencephaly polymicrogyria-polydactyly hydrocephalus syndrome. Am. J. Med. Genet. A. 149A, 688–687 (2009).

8. Brodbbeck, D., Cron, P. & Hemmings, B.A. Human protein kinase B...
1Center for Integrative Brain Research, Seattle Children’s Hospital, Seattle, Washington, USA. 2Department of Human Genetics, University of Chicago, Chicago, Illinois, USA. 3Department of Genome Sciences, University of Washington, Seattle, Washington, USA. 4Children’s Hospital of Eastern Ontario Research Institute, University of Ottawa, Ottawa, Ontario, Canada. 5Genome Damage & Stability Centre, University of Sussex, Falmer, Brighton, UK. 6Department of Pediatrics and Human Development, Michigan State University, East Lansing, Michigan, USA. 7Genome Quebec Innovation Centre, McGill University, Montreal, Quebec, Canada. 8Division of Medical Genetics, A.I. duPont Hospital for Children, Wilmington, Delaware, USA. 9Department of Genetics, Children’s Hospital of Eastern Ontario, Ottawa, Ontario, Canada. 10Cedars-Sinai Medical Center, Medical Genetics Institute, Los Angeles, California, USA. 11Department of Human Genetics, University Hospital Essen, Essen, Germany. 12Department of Paediatrics, Queen’s University, Kingston, Ontario, Canada. 13Department of Medical Genetics, University of British Columbia, Vancouver, British Columbia, Canada. 14Department of Medical Genetics, University of Calgary, Calgary, Alberta, Canada. 15Division of Clinical and Metabolic Genetics, Hospital for Sick Children, Toronto, Ontario, Canada. 16Department of Dermatology, Medical College of Wisconsin, Milwaukee, Wisconsin, USA. 17Department of Pediatrics, Medical College of Wisconsin, Milwaukee, Wisconsin, USA. 18Department of Medical Genetics, MassGeneral Hospital for Children, Boston, Massachusetts, USA. 19Department of Clinical Genetics, Erasmus Medical Center, Rotterdam, The Netherlands. 20Department of Genetics, North York General Hospital, Toronto, Ontario, Canada. 21Providence Child Neurology, Providence Sacred Heart Medical Center and Children’s Hospital, Spokane, Washington, USA. 22Clinical Genetics Department, St George’s Hospital, University of London, London, UK. 23Pediatric Neurology Unit, Wolfson Medical Center, Holon, Israel. 24Institute of Human Genetics, University Medical Center Hamburg-Eppendorf, Hamburg, Germany. 25Department of Neuropediatrics, University of Goettingen, Goettingen, Germany. 26Membership of the Steering Committee for the Consortium is provided in the Supplementary Note. 27Department of Human Genetics, McGill University, Montreal, Quebec, Canada. 28Ottawa Hospital Research Institute, University of Ottawa, Ottawa, Ontario, Canada. 29Department of Pediatrics, University of Washington, Seattle, Washington, USA. 30Department of Neurology, University of Washington, Seattle, Washington, USA. Correspondence should be addressed to W.B.D. (wbd@uw.edu).
ONLINE METHODS

Study subjects. The study included 52 affected individuals from 50 unrelated families. Informed consent was obtained from all subjects, and the study was approved by the institutional review boards at participating institutions: the University of Chicago, Seattle Children’s Hospital, Cedars-Sinai Medical Center in Los Angeles and Children’s Hospital of Eastern Ontario Research Institute in Ottawa. Individuals were diagnosed as having MCAP or MPPH using our published diagnostic criteria. Genomic DNA was extracted from different tissues using standard procedures. When parental DNA samples were available, paternity-maternity testing was performed by genotyping six highly polymorphic short tandem repeats. All exome capture and sequencing experiments were performed using genomic DNA from whole blood.

Exome sequencing of the three index families. Exome capture and sequencing of trio LR08-018 was performed at the PerkinElmer DNA Sequencing Service. Libraries were generated using the 50 Mb SureSelect human exome kit (Agilent) and sequenced on a HiSeq 2000 (Illumina) according to the manufacturer’s recommendations for paired-end 100-bp reads. Exome sequencing of trio LR09-006 and subject LR00-016a1 were carried out at the University of Washington Genome Sciences Genomic Resource Center using the SeqCap EZ Exome Library v2.0 capture kit (Roche) combined with sequencing of paired-end 76-bp reads on a Genome Analyzer Ix (Illumina). Reads were aligned to the human reference genome (hg19) as previously described. Single-nucleotide variants and small indels were identified using the GATK Unified Genotyper and a variant quality score of ≥10 and were annotated using using SeqExpress SNP annotation (see URLs). Variants were then filtered using standard hard-filtering parameters. Specifically, only variants with a quality score of ≥30, sequencing depth of ≥4, quality/depth ratio of ≥2,5 length of homopolymer run of ≤5.0 and allele balance of ≤0.80 were considered for downstream analysis. We also used exome sequencing data from 112 individuals sequenced at the University of Washington Genome Sciences or PerkinElmer DNA Sequencing Service. These included 95 unrelated control individuals from the University of Washington National Institute of Environmental Health Sciences (NIEHS) Exome Project (see URLs) and 17 in-house exomes from healthy individuals and individuals with unrelated disease phenotypes.

Family-based analysis. For each index subject (LR08-018, LR09-006 and LR00-016a1), we generated a list of rare variants by focusing on protein-altering and splice-site DNA changes that were absent from dbSNP build 132, the 1000 Genomes Project data and 112 control exomes. As described previously, we systematically identified potential de novo mutational events from the list of rare variants generated for subjects LR08-018 and LR09-006 by focusing on variants supported by at least two reads in the proband and absent in both parents (less than two variant reads) at base-pair positions covered by at least four reads in the entire trio. Candidate de novo mutations were manually inspected using IGV. This approach was unsuccessful for MCAP trio LR09-006, so we relaxed our criteria to include all variants that did not meet the initial hard-filtering criteria used for variant calling and generated a second list of rare variants and candidate de novo events using the aforementioned criteria.

Genomic DNA samples from LR00-016a1 and both affected siblings were genotyped using the InfiniumII HumanHap610 Quad BeadChip array (Illumina) at the Center for Applied Genomics at the Children’s Hospital of Philadelphia before the current study. We assumed autosomal recessive inheritance and selected genes with either homozygous or two heterozygous rare variants. We then used the genome-wide SNP data obtained for all three affected siblings to further reduce the number of candidate genes. Assuming a dominant mode of inheritance, we manually examined candidate homozygous variants from the list of rare variants generated for LR00-016a1.

Additional subjects with MCAP and control exomes. Exome capture and sequencing from seven unrelated individuals with MCAP was performed at the McGill University and Genome Quebec Innovation Centre using the 50 Mb SureSelect human exome kit and sequencing of 100-bp paired-end reads on the Illumina HiSeq 2000. We generated over 15 Gb of sequence per individual, such that approximately 90% of the coding bases of the genome as defined by the consensus coding sequence project were covered by at least 20 reads. Reads were aligned to hg19 with the Burrows-Wheeler Aligner (BWA) and duplicate reads were marked using Picard (see URLs) and excluded from downstream analysis. For each sample, single-nucleotide variants and indels were called using SAMtools pileup and varFilter with the base alignment quality adjustment disabled and were then filtered for quality with the requirement that at least 20% of reads supported the variant call. Due to low-level mosaicism in most subjects with MCAP, this initial analysis missed all but one mutation of PIK3CA, which were later found using other methods. The study also included exome sequencing data from 174 individuals sequenced at the University of Washington Genome Sciences using the SeqCap EZ Exome Library v2.0 and an Illumina sequencing platform. These 174 individuals comprised unrelated healthy individuals of European descent who are parents of children with sporadic autism.

To detect low-level mosaicism, we systematically analyzed the 3,287 coding and splice-site bases of PIK3CA in exomes from the remaining six unsolved MCAP cases and 174 unrelated control individuals. Standard variant calling in the subject and control data sets identified 36 heterozygous variants supported by 41–66% of reads and one homozygous change found in 99% of reads. These sites were considered to be germline variants and were excluded from downstream analysis. To detect candidate mosaic changes, we first selected all sites with at least one read not matching the reference sequence using SAMtools (v0.1.7) pileup. Raw variant sites were then run through a custom pipeline to identify candidate mosaic changes supported by at least one, two, three or four variant reads using a base-quality threshold of 20 and requiring at least 1% of reads to support the variant.

Sanger sequencing. We amplified the coding exons of AKT3 (NM_005465.4 and NM_181690), BDP1 (NM_018429.2), PIK3R1 (NM_181523, NM_181524.1 and NM_181504.3), PIK3R2 (NM_005027.2), PIK3R3 (AK302049.1 and EU832531.1), PIK3CA (NM_006218.2), PIK3CB (NM_006219.2), PIK3CD (NM_005026.3 and U57843.1) and PTEN (NM_000314.4) using custom intronic primers and standard PCR protocols (primer sequences available on request) combined with Sanger sequencing. Amplicons were sequenced at the University of Washington High-Throughput Genomics Unit and the Seattle BioMed Sequencing Core Facility. All mutations were tested in at least two independent amplification and sequencing reactions in the proband and available relatives. Sequence traces were analyzed using Mutation Surveyor v3.97 (SoftGenetics). For all mutations, nucleotide-level conservation and impact of amino-acid substitutions were assessed using genomic evolutionary rate profiling (GERP) and Grantham matrix scores, respectively.

Restriction enzyme and genotyping assays. Six mutation sites (encoding p.Glu453del, p.Glu726Lys, p.Gly914Arg, p.Ala1035Val, p.Met1043Ile and p.His1047Tyr) were confirmed by PCR using HotStarTaq Plus DNA polymerase (Qiagen) and primer pairs including a carboxyfluorescein (FAM)-labeled primer and an unlabeled primer with a GTTTCTT sequence at the 5′ end. Amplification products were run on a 7% non-denaturing polyacrylamide gel (253:1) and visualized using the FOS-350 laser image analyzer (Applied Biosystems), and results were analyzed using GenMarker software (SoftGenetics). The frequency of the mutant allele was calculated by mutant peak area / (wild-type peak area + mutant peak area).

The PIK3CA mutation encoding the p.Glu4726Lys alteration is located in a 7.2-kb segmental duplication (chr. 3: 178,932,477–178,939,690) sharing 98% identity with a region on chromosome 22 (chr 22: 17,049,390–17,056,254). To rule out the possibility of non-specific amplification, we amplified a 12.2-kb fragment using TaKaRa LA Taq and primers flanking the segmental duplication (primer sequences are listed in Supplementary Table 1). We then used the long-range PCR product as a template for the restriction enzyme assay.

The PIK3CA mutation encoding the p.Glu453del alteration was confirmed by standard PCR using a forward FAM-labeled primer and an unlabeled reverse primer to amplify the 238-bp and 235-bp fragments corresponding to the wild-type and mutant alleles, respectively.

Deep sequencing of mutation sites. We performed targeted ultra-deep sequencing of five mutation sites in PIK3CA in a series of subjects with MCAP and...
MPPH and their unaffected parents, using chimeric oligonucleotides containing Illumina adaptor sequences combined with locus-specific primers. Reverse oligonucleotides also contained 12 unique octamer barcodes for multiplexing of up to 12 samples per lane (Supplementary Table 12). Amplification was carried out in a MiniOpticon Real-time PCR system (Bio-Rad) using the iProof High-Fidelity Master Mix (Bio-Rad), 50 ng of genomic DNA and SYBR Green. Samples were removed from the PCR machine before fluorescence began to plateau. Amplicons were purified using Ampure beads (Agencourt) and were directly sequenced on an Illumina HiSeq 2000 using 100-bp paired-end reads. Sequencing reads were aligned to hg19 using BWA47. We used SAMtools (v0.1.7) pileup48 combined with custom scripts to select variant sites supported by at least 1% of reads with a base-quality score threshold of 20.

The two newly identified mutations (encoding p.Glu81Lys and p.Arg88Gln in subjects LR06-342 and LR11-068, respectively) were validated by an independent ultra-deep sequencing experiment, including two DNA sources per affected individual and DNA samples from the parents of LR11-068. All six DNA samples were amplified and sequenced twice using octamer barcodes.

**Lymphoblastoid cell lines.** Epstein-Barr virus (EBV)-immortalized lymphoblastoid cell lines were established from peripheral blood samples of four subjects with megalencephaly (LR00-016A, LR08-018, LR05-204 and LR09-006), one unaffected individual and one individual with Cowden disease (GM10080). Lymphoblastoid cell lines were cultured in RPMI medium (Fisher Scientific) with 15% FCS, 2-mercaptoethanol) and sonicating samples for 15 s at 30% amplitude with a micro-tip (Sigma-Aldrich). The following antibodies were obtained from Cell Signaling: Monoclonal antibody to phosphatidylinositol 3,4,5-trisphosphate (PIP3) was from Caltag-Medsystems (D145-3, clone NN111.1.1). The following antibodies were obtained from Cell Signaling Technology: PDK1-pSer241 (3061), S6-pSer240/244 (2215), 4E-BP-pThr37/46 (2855) and β-tubulin (2128). Conjugated secondary antibodies, anti-mouse IgG–FITC (whole molecule; F0257) and anti-rabbit IgG–Cy3 F(ab)2 fragment (2855) and anti-fade mounting medium (Vectashield). Slides were analyzed using the Zeiss AxioPlan platform. Images were captured using SimplePCI software (see URLs) and the same exposure time for each sample. Images (TIFF format) were then used to quantify signal (RGB) using ImageJ software (see URLs).

**SDS-PAGE and protein blots.** Whole-cell extracts were prepared by lysing the cells in urea buffer (9 M urea, 50 mM Tris-HCl at pH 7.5 and 10 mM 2-mercaptoethanol) and sonicating samples for 15 s at 30% amplitude with a micro-tip (Sigma-Aldrich). Protein in the supernatant was quantified by Bradford assay. Extracts were resolved by SDS-PAGE and were semi-dry blotted onto polyvinylidene fluoride (PVDF) membranes for blocking and antibody incubation.

**Statistical analysis.** The likelihood of observing a second de novo mutation in AKT3 was calculated by $P = 1.28 \times 10^{-8} \times 1.540 \times 98$, where $1.28 \times 10^{-8}$ is the estimated haploid substitution rate49, 1.540 is the number of coding and splice-site bases of AKT3 (NM_005465.4 and NM_181690) and 98 is the number of chromosomes tested (49 unrelated affected individuals). We used two-tailed Fisher’s exact tests for genotype-phenotype correlations.

44. Riviere, J.B. et al. De novo mutations in the actin genes ACTB and ACTG1 cause Baraitser-Winter syndrome. Nat. Genet. 44, 440–444, S1–S2 (2012).
45. DePristo, M.A. et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat. Genet. 43, 491–498 (2011).
46. 1000 Genomes Project Consortium. A map of human genome variation from population-scale sequencing. Nature 467, 1061–1073 (2010).
47. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25, 1754–1760 (2009).
48. Li, H. et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25, 2078–2079 (2009).
49. O’Roak, B.I. et al. Exome sequencing in sporadic autism spectrum disorders identifies severe de novo mutations. Nat. Genet. 43, 585–589 (2011).
50. Cooper, G.M. et al. Single-nucleotide evolutionary constraint scores highlight disease-causing mutations. Nat. Methods 7, 250–251 (2010).
51. Grantham, R. Amino acid difference formula to help explain protein evolution. Science 185, 862–864 (1974).
52. Brownstein, M.J., Carpten, J.D. & Smith, J.R. Modulation of non-templated nucleotide addition by Taq DNA polymerase: primer modifications that facilitate genotyping. Biotechniques 20, 1004–1006, 1008–1010 (1996).
53. Lynch, M. Rate, molecular spectrum, and consequences of human mutation. Proc. Natl. Acad. Sci. USA 107, 961–968 (2010).