Mosaicism is increasingly recognized as a cause of developmental disorders with the advent of next-generation sequencing (NGS). Mosaic mutations of PIK3CA have been associated with the widest spectrum of phenotypes associated with overgrowth and vascular malformations. We performed targeted NGS using 2 independent deep-coverage methods that utilize molecular inversion probes and amplicon sequencing in a cohort of 241 samples from 181 individuals with brain and/or body overgrowth. We identified PIK3CA mutations in 60 individuals. Several other individuals (n = 12) were identified separately to have mutations in PIK3CA by clinical targeted-panel testing (n = 6), whole-exome sequencing (n = 5), or Sanger sequencing (n = 1). Based on the clinical and molecular features, this cohort segregated into three distinct groups: (a) severe focal overgrowth due to low-level but highly activating (hotspot) mutations, (b) predominantly brain overgrowth and less severe somatic overgrowth due to less-activating mutations, and (c) intermediate phenotypes (capillary malformations with overgrowth) with intermediately activating mutations. Sixteen of 29 PIK3CA mutations were novel. We also identified constitutional PIK3CA mutations in 10 patients. Our molecular data, combined with review of the literature, show that PIK3CA-related overgrowth disorders comprise a discontinuous spectrum of disorders that correlate with the severity and distribution of mutations.
PIK3CA-associated developmental disorders exhibit distinct classes of mutations with variable expression and tissue distribution


1Division of Genetic Medicine, Department of Pediatrics, University of Washington, Seattle, Washington, USA. 2Center for Integrative Brain Research and 3Center for Developmental Biology and Regenerative Medicine, Seattle Children’s Research Institute, Seattle, Washington, USA. 4Pediatric Neurology, Neurogenetics and Neurobiology Unit and Laboratories, Neurosciences Department, A. Meyer Children’s Hospital, University of Florence, Florence, Italy. 5Department of Genetics, Stanford University School of Medicine, Stanford, California, USA. 6Department of Medical Genetics, Kasturba Medical College, Manipal University, Manipal, Karnataka, India. 7Department of Genome Sciences, University of Washington, Seattle, Washington, USA. 8Whole Exome Sequencing Program, GeneDx, Gaithersburg, Maryland, USA. 9Regional Genetics Program, The Children’s Hospital of Eastern Ontario, Ottawa, Ontario, Canada. 10Department of Pediatrics, and Pediatric Neurology and Developmental Medicine, University Children’s Hospital, Tübingen, Germany. 11Mount Sinai Hospital, The Prenatal Diagnosis and Medical Genetics Division, Department of Obstetrics and Gynecology, and 12Department of Pediatrics, Division of Clinical and Metabolic Genetics, University of Toronto, Toronto, Ontario, Canada. 13Center for Human Genetics, University Hospitals Case Medical Center, Cleveland, Ohio, USA. 14Department of Pediatrics, Division of Medical Genetics, Saint Louis University, St. Louis, Missouri, USA. 15Department of Pediatrics, University of Vermont College of Medicine, Burlington, Vermont, USA. 16Department of Clinical Genetics, Royal North Shore Hospital, St Leonards, New South Wales, Australia. 17Belfast Health and Social Care Trust, Belfast, United Kingdom. 18Genetics, Metabolism and Pediatrics, London, Ontario, Canada. 19Clinical Genetics, Center for Personalized Medicine, Children’s Hospital Los Angeles, Keck School of Medicine at University of Southern California, Los Angeles, California, USA. 20Neurology Unit, Schneider Children’s Medical Center of Israel, Petach Tikva, and Sackler School of Medicine, Tel Aviv University, Tel Aviv, Israel. 21School of Medicine, Griffith University, Brisbane, Queensland, Australia. 22Division of Genetics, Department of Pediatrics, and 23Department of Human Genetics, University of Michigan, Ann Arbor, Michigan, USA. 24Genetics Service, Department of Pediatric Medicine, KK Women’s and Children’s Hospital, Singapore. 25British Columbia Medical Genetics Provincial Program, University of British Columbia, Vancouver, British Columbia, Canada. 26Service de génétique médicale, Centre Hospitalier Universitaire Vaudois CHUV, Switzerland. 27Greenwood Genetic Center, Greenwood, South Carolina, USA. 28Department of Dermatology, Children Hospital Prof. Dr. J. P. Garrahan, Buenos Aires, Argentina. 29Children’s Hospital of Michigan, Wayne State University, Detroit, Michigan, USA. 30Center for Human Genetics, University Hospitals Leuven and KU Leuven, Leuven, Belgium. 31South West Thames Regional Genetics Service, St George’s University NHS Foundation Trust, London, and Section of Cancer Genetics, Institute of Cancer Research, Sutton, United Kingdom. 32Department of Pediatrics and Ophthalmology, University of Minnesota, Minneapolis, Minnesota, USA. 33Department of Child Neurology, University of Florida, Jacksonville, Florida, USA. 34Department of Genetics, Liverpool Hospital, Liverpool, New South Wales, Australia. 35Department of Neurology, University of Washington, Seattle, Washington, USA. 36Providence Child Neurology, Providence Sacred Heart Medical Center and Children’s Hospital, Spokane, Washington, USA. 37Department of Pediatrics
Mosaicism is increasingly recognized as a cause of developmental disorders with the advent of next-generation sequencing (NGS). Mosaic mutations of PIK3CA have been associated with the widest spectrum of phenotypes associated with overgrowth and vascular malformations. We performed targeted NGS using 2 independent deep-coverage methods that utilize molecular inversion probes and amplicon sequencing in a cohort of 241 samples from 181 individuals with brain and/or body overgrowth. We identified PIK3CA mutations in 60 individuals. Several other individuals (n = 12) were identified separately to have mutations in PIK3CA by clinical targeted-panel testing (n = 6), whole-exome sequencing (n = 5), or Sanger sequencing (n = 1). Based on the clinical and molecular features, this cohort segregated into three distinct groups: (a) severe focal overgrowth due to low-level but highly activating (hotspot) mutations, (b) predominantly brain overgrowth and less severe somatic overgrowth due to less-activating mutations, and (c) intermediate phenotypes (capillary malformations with overgrowth) with intermediately activating mutations. Sixteen of 29 PIK3CA mutations were novel. We also identified constitutional PIK3CA mutations in 10 patients. Our molecular data, combined with review of the literature, show that PIK3CA-related overgrowth disorders comprise a discontinuous spectrum of disorders that correlate with the severity and distribution of mutations.

Introduction

Mosaicism refers to a biological phenomenon in which an individual derived from a single fertilized egg has two or more populations of cells with different genotypes. This process can be the result of spontaneous mutations occurring at different times during the life course of a multicellular organism, and is a widespread phenomenon during the normal aging process (1). Here, we will refer to somatic mosaicism as a process occurring strictly during development (postzygotically), excluding mosaicism confined to germ cells (germ-line mosaicism) and somatic mosaicism associated with cancer (2). While this phenomenon has been known for many years, the association of mosaic mutations with human disease began with discovery of mosaic chromosomal disorders, and patchy (or segmental) manifestations of Mendelian disorders (3, 4). The first sequencing-based proof of mosaic aberrations came in 1991 with discovery of activating mutations of the GNAS gene in McCune-Albright syndrome (5, 6). Subsequently, an increasing number of other conditions, largely congenital or childhood-onset developmental disorders, have been associated with mosaic mutations. Examples include Proteus syndrome (AKT1), Sturge-Weber syndrome (GNAQ), neurocutaneous melanosis (NRAS), epidermal nevi and linear nevus sebaceous syndrome (HRAS, KRAS, PIK3CA), Maffucci and Ollier syndromes (IDH1, IDH2), and encephalocraniocutaneous lipomatosis (FGFR1); disorders first proposed to be mosaic based on their asymmetric, patchy presentation, and lack of familial recurrence (7, 8). Mosaic mutations of these genes were primarily identified with the advent of massively parallel or next-generation sequencing (NGS) methods that facilitate detection of low-frequency variation, especially in affected tissues.

However, there are several limitations in our knowledge of mosaic developmental disorders, particularly which molecular diagnostic methods are best, which tissues should be assayed, and the nature and rela-
The relationship between phenotypes and genotypes. The low-level mosaicism found in these disorders poses significant challenges to conventional clinical diagnostic approaches. Current molecular methods in standard clinical laboratories rely on Sanger sequencing for single-gene analysis, and standard-depth NGS methods including targeted panels or whole-exome sequencing (with many clinical diagnostic labs using 20× to 30× as the minimum depth of coverage for putatively constitutional or germline disorders) (9). However, these methods are not sensitive enough to detect low-level mutations, particularly when the level of mosaicism is very low (<5%–10%) in the tissue from which DNA is isolated (typically peripheral blood lymphocytes).

Of the genes associated with mosaicism in developmental disorders, mutations of \( \text{PIK3CA} \) have been associated with the widest spectrum of developmental phenotypes to date, with most described as distinct clinical entities long before the link to \( \text{PIK3CA} \) was discovered (10). \( \text{PIK3CA} \) encodes the alpha catalytic subunit of phosphatidylinositol-4,5-bisphosphate 3-kinase, a central member of the phosphatidylinositol 3-kinase (PI3K) enzyme family (11). \( \text{PIK3CA} \) functions as an oncogene, and activating (or gain-of-function) mutations of \( \text{PIK3CA} \) are widely seen in human cancers (12, 13). The most common \( \text{PIK3CA} \) mutations are p.Glu542Lys, p.Glu545Lys, and p.His1047Arg, which are seen in ~80% of somatic tissues in cancer (and therefore termed “hotspot” mutations). Based on their degree of activity, several other classes of \( \text{PIK3CA} \) mutations have been described including strong, intermediate, and weak (13). Some of these same gain-of-function \( \text{PIK3CA} \) mutations have been recently reported in a range of pediatric developmental phenotypes. These disorders are broadly characterized by cutaneous vascular malformations with segmental overgrowth and involve multiple tissues or body regions. These conditions have been variably classified as Klippel-Trenaunay (KTS), congenital lipomatosis with overgrowth, vascular malformations, epidermal nevi, and skeletal/scoliosis/spinal abnormalities (CLOVES), megalencephaly-capillary malformation syndrome (MCAP), and dysplastic megalencephaly (DMEG), but the spectrum and differences between these disorders have not been defined.
Here, we report the largest series of patients to date with developmental disorders associated with PIK3CA mutations, identified using 2 orthogonal deep-targeted NGS methods that utilize molecular inversion probes (MIPs) and amplicon sequencing, to determine the mutational spectrum and quantify mutation levels in multiple tissues from affected individuals. By combining our data with previous reports, we assess available clinical molecular diagnostic methods, the most ideal tissue samples to be assayed, and derive genotype-phenotype correlations in this spectrum of disorders. We show that the phenotypes are not simply related to the level and distribution of PIK3CA mutations, but also to the class of mutation. Specifically, we show that CLOVES and KTS, as well as most localized lesions, are caused by one of the three most common oncogenic (“hotspot”) mutations or occasionally by strong or intermediate mutations, while MCAP is caused largely by a different set of less-activating mutations, with a much wider mutational spectrum that overlaps at only a few strong mutations (13). We also show that mutations are not equally detectable from apparently unaffected but easily available “surrogate” tissues, with levels of mosaicism on average lower in peripheral blood lymphocytes than saliva, and lower in saliva than skin fibroblasts.

Results
We screened a cohort of 241 DNA samples from 181 individuals with brain and body overgrowth for mutations in the PIK3CA gene using two deep NGS methods that utilize MIPs (14, 15) and amplicon sequencing. This cohort included 131 individuals with features of MCAP, 19 with diffuse brain overgrowth or megalencephaly, and 31 with various forms of somatic overgrowth and vascular malformations including CLOVES, KTS, and capillary malformations with overgrowth (Table 1 and Supplemental Figure 1; supplemental material available online with this article; doi:10.1172/jci.insight.87623DS1). We identified mutations of PIK3CA in 60 individuals (33%). Several other patients (n = 12) were identified to have mutations separately by clinical targeted-panel testing (n = 6), clinical whole-exome sequencing (n = 5), or Sanger sequencing (n = 1), and were included to highlight the clinical and molecular variability of this spectrum of disorders. This cohort includes one of our previously published MCAP patients, as we have analyzed several newly acquired tissue samples from this patient in this study (16). The remaining 23 patients with MCAP who were previously published were excluded from this study (16). The mutations identified, levels of mosaicism, and tissue distribution in this cohort are listed in Tables 2, 3, and 4. The levels of mosaicism in all tissues of all mutation-positive patients and their parents using all molecular methods are listed in Supplemental Table 1.

Molecular results
Levels of mosaicism and tissue distribution. We identified 29 mutations of PIK3CA, including 16 novel mutations that have not been previously identified in developmental disorders, to our knowledge. Oncogenic mutations at all of these amino acid residues were present in the Catalogue of Somatic Mutations in Cancer (COSMIC; http://cancer.sanger.ac.uk/cosmic). None of the mutations were present in public databases including the 1000 Genomes project (http://www.1000genomes.org), the NHLBI Exome Variant Server (http://evs.gs.washington.edu/EVS), or the Exome Aggregation Consortium (ExAC; http://exac.broadinstitute.org), with the exception of p.Glu545Lys and p.His1047Arg, which were present at very low frequencies in ExAC (<0.0001). Alternative allele percentages (the percentage of alternate or mutant reads to total reads; AAP) ranged from 1% to 59.5% across the entire cohort (Tables 2, 3, and 4). We observed substantial variation in AAPs in multiple samples from the same individual, with levels of mosaicism in blood (peripheral lymphocytes) ranging from undetectable to 59.5% (mean AAP 16.5%; 43 samples), saliva (lymphocytes and epithelial cells) from undetectable to 54% (mean AAP 24%; 37 samples), and skin-derived fibroblasts from 4% to 60% (mean AAP 26.4%; 20 samples). The AAPs grouped by sample type (blood, saliva, skin fibroblasts) are shown in Figure 1, whereas the AAPs for each sample type in all mutation-positive individuals are shown in Supplemental Figure 2. For a few individuals, we tested additional sample types including buccal swabs (n = 4 samples), surgically resected tonsils (lymphoid tissue; n = 2), and occipital bone dura mater removed during posterior fossa decompression surgery (n = 1).

Diagnostic testing methods. The average coverage of single-molecule MIPs (smMIPs) was 450.74×, with an average of 71.27% of coding regions covered by 50 or more capture events, and with 2% being the lower threshold of mutant allele detection. All amino acid residues in which mutations were previously identified were well covered by smMIPs in the entire cohort (Supplemental Figure 3). All mutations identified via smMIPs were confirmed using a second orthogonal method (amplicon sequencing), when sufficient DNA
Amplicon sequencing yielded high coverage data for all mutations, with a mean coverage of 1975×. None of these mutations were present in the parents of affected individuals by amplicon sequencing. Overall, AAPs detected by amplicon sequencing were similar to smMIP results, although the latter were usually slightly higher for the same samples (Supplemental Figure 4). This may be partly due to allelic bias/preferential amplification of amplicon sequencing when compared to smMIPs.

Clinical testing identified several more patients with PIK3CA mutations. Mutations in 6 individuals were detected by targeted NGS using the Agilent SureSelect Capture technology. Another 5 were identified using standard-depth whole-exome sequencing (WES) on blood-derived DNA, and 1 patient was identified to have a mutation by standard Sanger sequencing. AAPs in these samples ranged from 41%–59.5%, suggesting constitutional mutations, with the exception of 1 patient (LR15-238) found to have the p.Arg93Gln mutation...

Table 2. PIK3CA mutations and levels of mosaicism — part 1 (n = 22 patients) [PIK3CA: NM_006218.2]

| Sample # | Diagnosis | Position | cDNA change | Amino acid change | Alternative/total alleles (% alt alleles) | Blood | Saliva | FB | Other tissues | N COSMIC | INH | Type
|-----------|-----------|----------|-------------|------------------|----------------------------------------|-------|-------|---|-------------|----------|-----|------
| LR14-323<sup>a</sup> | MCAP | 3:178916891 | c.278G>A | p.Arg93Gln<sup>a</sup> | 47%<sup>SS</sup> -50%<sup>inher</sup> | – | – | – | 14 | De novo | Const.
| LR15-238 | OVG | 3:178916891 | c.278G>A | p.Arg93Gln<sup>a</sup> | 19%<sup>WES</sup> | – | – | – | 14 | De novo | Mosaic
| LR01-060<sup>b,c</sup> | MCAP | 3:178916924 | c.311C>T | p.Pro104Leu<sup>a</sup> | – | 41%<sup>MIP</sup> 49%<sup>MIP</sup> | – | 5 | De novo | Presumed Const.
| LR13-359 | MCAP | 3:178916924 | c.311C>T | p.Pro104Leu<sup>a</sup> | 11%<sup>MIP</sup> 42%<sup>MIP</sup> | – | – | – | 5 | De novo | Mosaic
| LR15-076<sup>d</sup> | MEG | 3:178921548 | c.1030G>A | p.Val344Met<sup>a</sup> | 41.5%<sup>WES</sup> | – | – | – | 13 | De novo | Const.
| LR12-365 | MCAP | 3:178921552 | c.1034A>C | p.Asn345Thr<sup>a</sup> | – | 39%<sup>MIP</sup> | – | – | 3 | Mother negative | Presumed Const.
| LR11-076 | MCAP | 3:178921566 | c.1048G>A | p.Pro104Leu<sup>a</sup> | 12%<sup>Amp</sup> 32%<sup>MIP</sup> | – | 10%<sup>LB,Amp</sup> | 3 | De novo | Mosaic
| LR13-036<sup>e</sup> | MEG-macrosmia | 3:178922324 | c.1039G>A | p.Glu365Lys<sup>a</sup> | 52%<sup>WES</sup> | – | – | – | 10 | De novo | Const.
| LR13-264 | MCAP | 3:178922324 | c.1093G>A | p.Glu365Lys<sup>a</sup> | 2%<sup>SS</sup> | – | 23%<sup>SS</sup> | – | 10 | NA | Mosaic
| LR11-374 | MCAP | 3:178922364 | c.1133G>A | p.Cys378Tyr | 3%<sup>MIP</sup> 40%<sup>MIP</sup> | – | – | – | 2 | De novo | Mosaic
| LR11-418 | MCAP | 3:178922364 | c.1133G>A | p.Cys378Tyr | 2%<sup>MIP</sup> | – | – | – | 2 | Mother negative | Mosaic
| LR12-382<sup>a</sup> | MCAP | 3:178922364 | c.1133G>A | p.Cys378Tyr | – | 7%<sup>MIP</sup> | – | – | 2 | De novo | Mosaic
| LR14-278<sup>c</sup> | MCAP | 3:178922364 | c.1133G>A | p.Cys378Tyr | – | 50%<sup>WES</sup> | – | – | 2 | De novo | Const.
| LR11-200 | MCAP | 3:178928067 | c.1345C>A | p.Pro449Thr<sup>a</sup> | – | – | 37%<sup>MIP</sup> | – | 4 | NA | Mosaic
| LR14-358<sup>c</sup> | MCAP | 3:178928067 | c.1345C>A | p.Pro449Thr<sup>a</sup> | – | – | – | – | 4 | De novo | Const.
| LR11-392 | MCAP | 3:178928079 | c.1357G>A | p.Glu453Lys<sup>a</sup> | 2%<sup>MIP</sup> | – | – | – | 25 | NA | Mosaic
| LR12-070 | OVG-MD | 3:178928079 | c.1357G>A | p.Glu453Lys<sup>a</sup> | – | 0%<sup>MIP</sup> 25%<sup>MIP</sup> | – | 25 | NA | Mosaic
| LR12-184 | OVG | 3:178928079 | c.1357G>A | p.Glu453Lys<sup>a</sup> | – | 1%<sup>Amp</sup> 1%<sup>SS</sup> 24%<sup>MIP</sup> 5%<sup>-17%</sup> several skin samples, Amp | 25 | De novo | Mosaic
| LR12-329 | OVG-MD | 3:178928079 | c.1357G>A | p.Glu453Lys<sup>a</sup> | – | 6%<sup>MIP</sup> | – | – | 25 | Mosaic
| LR13-048<sup>b,c</sup> | MCAP | 3:178928078 | c.1359_1361delAGA | p.Glu453Lys<sup>a</sup> | – | 19%<sup>SS</sup> | – | 1 | Mother negative | Const.

Abbreviations: ABD, adaptor-binding domain; Amp, amplicon sequencing; BS, buccal swab; FB, skin-derived fibroblasts; INH, inheritance; LB, lymphoblastoid cell line; MEG, megalencephaly; MCAP, megalencephaly-capillary malformation syndrome; MD, macrodactyly; MIPs, molecular inversion probes; NA, unavailable; OVG, overgrowth; Sanger, Sanger sequencing; SS, Agilent SureSelect; WES, whole-exome sequencing. Protein domains are adapted from: http://www.uniprot.org/uniprot/P42336#family_and_domains. Notes: *Novel PIK3CA mutations. *Expanded or newly-published molecular results on previously published patients: LR01-060 (patient 3) (57), LR12-382 (patient 7) (59), and LR13-048 (patient 1) (60). Constitutional or apparently constitutional mutations. Mutations were assigned conservatively constitutional when AAP was greater than 35% in blood (lymphocyte-derived) DNA. Mutations with AAPs less than 35%, particularly in several tissue samples, were assigned as mosaic. Mutations were presumed to be constitutional when the AAP was greater than 35% in nonblood-derived DNA (particularly saliva). The numbers of COSMIC mutations are based on those present when the COSMIC database was last accessed in May 2016.

was available. Amplicon sequencing yielded high coverage data for all mutations, with a mean coverage of 1975×. None of these mutations were present in the parents of affected individuals by amplicon sequencing. Overall, AAPs detected by amplicon sequencing were similar to smMIP results, although the latter were usually slightly higher for the same samples (Supplemental Figure 4). This may be partly due to allelic bias/preferential amplification of amplicon sequencing when compared to smMIPs.

Clinical testing identified several more patients with PIK3CA mutations. Mutations in 6 individuals were detected by targeted NGS using the Agilent SureSelect Capture technology. Another 5 were identified using standard-depth whole-exome sequencing (WES) on blood-derived DNA, and 1 patient was identified to have a mutation by standard Sanger sequencing. AAPs in these samples ranged from 41%–59.5%, suggesting constitutional mutations, with the exception of 1 patient (LR15-238) found to have the p.Arg93Gln mutation...
## Table 3. PIK3CA mutations and levels of mosaicism — part 2 (*n* = 19 patients) [PIK3CA: NM_006218.2]

<table>
<thead>
<tr>
<th>Sample #</th>
<th>Diagnosis</th>
<th>Position</th>
<th>cDNA change</th>
<th>Amino acid change</th>
<th>Alternative/total alleles (% alt alleles)</th>
<th>N COSMIC</th>
<th>INH</th>
<th>Typea</th>
</tr>
</thead>
<tbody>
<tr>
<td>LR04-078</td>
<td>MCAP</td>
<td>3:178916957</td>
<td>c.344G&gt;C</td>
<td>p.Arg115Pro</td>
<td>6%MIPs -</td>
<td>1</td>
<td>NA</td>
<td>Mosaic</td>
</tr>
<tr>
<td>LR11-397</td>
<td>MCAP</td>
<td>3:178916957</td>
<td>c.344G&gt;C</td>
<td>p.Arg115Pro</td>
<td>11%MIPs - 27%MIPs -</td>
<td>1</td>
<td>De novo</td>
<td>Mosaic</td>
</tr>
<tr>
<td>LR12-001b</td>
<td>MCAP</td>
<td>3:178916957</td>
<td>c.344G&gt;C</td>
<td>p.Arg115Pro</td>
<td>- 11%MIPs 16%MIPs -</td>
<td>1</td>
<td>NA</td>
<td>Mosaic</td>
</tr>
<tr>
<td>LR12-080</td>
<td>MCAP</td>
<td>3:178916957</td>
<td>c.344G&gt;C</td>
<td>p.Arg115Pro</td>
<td>- 25%MIPs -</td>
<td>1</td>
<td>De novo</td>
<td>Mosaic</td>
</tr>
</tbody>
</table>

### Linker region (AA 106–186) *n* = 5

<table>
<thead>
<tr>
<th>Sample #</th>
<th>Diagnosis</th>
<th>Position</th>
<th>cDNA change</th>
<th>Amino acid change</th>
<th>Alternative/total alleles (% alt alleles)</th>
<th>N COSMIC</th>
<th>INH</th>
<th>Typea</th>
</tr>
</thead>
<tbody>
<tr>
<td>LR13-197</td>
<td>MEG-MDEG</td>
<td>3:178936091</td>
<td>c.1633G&gt;A</td>
<td>p.Glu545Lys</td>
<td>0%Amp 26%MIPs -</td>
<td>1371</td>
<td>De novo</td>
<td>Mosaic</td>
</tr>
<tr>
<td>LR12-019</td>
<td>MCAP</td>
<td>3:178936093</td>
<td>c.1635G&gt;T</td>
<td>p.Glu545Asp</td>
<td>3%MIPs 25%MIPs -</td>
<td>13</td>
<td>NA</td>
<td>Mosaic</td>
</tr>
</tbody>
</table>

### Helical domain (AA 517–694) *n* = 4

<table>
<thead>
<tr>
<th>Sample #</th>
<th>Diagnosis</th>
<th>Position</th>
<th>cDNA change</th>
<th>Amino acid change</th>
<th>Alternative/total alleles (% alt alleles)</th>
<th>N COSMIC</th>
<th>INH</th>
<th>Typea</th>
</tr>
</thead>
<tbody>
<tr>
<td>LR13-197</td>
<td>MEG-MDEG</td>
<td>3:178936091</td>
<td>c.1633G&gt;A</td>
<td>p.Glu545Lys</td>
<td>0%Amp 26%MIPs -</td>
<td>1371</td>
<td>De novo</td>
<td>Mosaic</td>
</tr>
<tr>
<td>LR12-019</td>
<td>MCAP</td>
<td>3:178936093</td>
<td>c.1635G&gt;T</td>
<td>p.Glu545Asp</td>
<td>3%MIPs 25%MIPs -</td>
<td>13</td>
<td>NA</td>
<td>Mosaic</td>
</tr>
</tbody>
</table>

### Linker region (AA 695–796) *n* = 10

<table>
<thead>
<tr>
<th>Sample #</th>
<th>Diagnosis</th>
<th>Position</th>
<th>cDNA change</th>
<th>Amino acid change</th>
<th>Alternative/total alleles (% alt alleles)</th>
<th>N COSMIC</th>
<th>INH</th>
<th>Typea</th>
</tr>
</thead>
<tbody>
<tr>
<td>LR09-142g</td>
<td>MCAP</td>
<td>3:178938934</td>
<td>c.2176G&gt;A</td>
<td>p.Glu726Lys</td>
<td>12%MIPs -</td>
<td>18</td>
<td>NA</td>
<td>Mosaic</td>
</tr>
<tr>
<td>LR11-072</td>
<td>MCAP</td>
<td>3:178938934</td>
<td>c.2176G&gt;A</td>
<td>p.Glu726Lys</td>
<td>4%MIPs - 3%LB, MIPs -</td>
<td>34</td>
<td>Mother negative</td>
<td>Mosaic</td>
</tr>
<tr>
<td>LR12-037</td>
<td>MCAP</td>
<td>3:178938934</td>
<td>c.2176G&gt;A</td>
<td>p.Glu726Lys</td>
<td>- 20%MIPs -</td>
<td>34</td>
<td>De novo</td>
<td>Mosaic</td>
</tr>
<tr>
<td>LR12-345</td>
<td>MCAP</td>
<td>3:178938934</td>
<td>c.2176G&gt;A</td>
<td>p.Glu726Lys</td>
<td>- 14%MIPs -</td>
<td>34</td>
<td>De novo</td>
<td>Mosaic</td>
</tr>
<tr>
<td>LR12-418</td>
<td>MCAP</td>
<td>3:178938934</td>
<td>c.2176G&gt;A</td>
<td>p.Glu726Lys</td>
<td>- 13%MIPs 4%MIPs -</td>
<td>34</td>
<td>NA</td>
<td>Mosaic</td>
</tr>
</tbody>
</table>

### Abbreviations:
- Amp, amplicon sequencing; BS, buccal swab; DMEG, dysplastic megalencephaly; FB, skin-derived fibroblasts; INH, inheritance; L, left; LB, lymphoblastoid cell line; LNSS, linear nevus sebaceous syndrome; MCAP, megalencephaly-capillary malformation syndrome; MEG, megalencephaly; MIPs, molecular inversion probes; NA, unavailable; OVG, overgrowth; R, right; Sanger, Sanger sequencing; SS, Agilent SureSelect; WES, whole-exome sequencing. Protein domains are adapted from: http://www.uniprot.org/uniprot/P42336#family_and_domains. Notes: Mutations classified as cancer hotspot mutations (most activating) are highlighted in red (13). *Novel PIK3CA mutations.* Expanded or newly published molecular results on previously published patients: LR12-001 (58) and LR09-142 (61).
- Constitutinal or apparently constitutional mutations. In patient LR12-330, blood-derived DNA was not available for analysis, therefore, the possibility of mosaicism cannot be conclusively excluded.
- Mutations were assigned conservatively constitutional when the AAP was greater than 35% in blood (lymphocyte-derived) DNA. Mutations with AAPs less than 35%, particularly in several tissue samples, were assigned as mosaic. Mutations were presumed to be constitutional when the AAP was greater than 35% in nonblood-derived DNA (particularly saliva). The numbers of COSMIC mutations are based on those present when the COSMIC database was last accessed in May 2016.
in 30 of 159 reads (19%). AAPs in the 6 patients tested by targeted NGS ranged from 1%–47%. In 2 patients (LR14-323, LR15-227), AAPs were 47% in blood-derived DNA, again suggesting constitutional mutations.

Overall, we identified 29 PIK3CA mutations in 72 individuals. We also reviewed published data on PIK3CA mutations in developmental pediatric disorders (all phenotypes except for cancer; Figure 2) (16–29). When added to our data, we find that all developmental phenotypes combined have been associated
with 41 different PIK3CA mutations involving 33 codons and include 27 recurrent mutations (i.e., seen in more than one affected individual). The three hotspot PIK3CA mutations (p.Glu542Lys, p.Glu545Lys, and p.His1047Arg) are highly recurrent (i.e., seen frequently) in pediatric disorders. We found that two “non-hotspot mutations” (p.Glu726Lys and p.Gly914Arg) are also highly recurrent in developmental phenotypes, having now been detected in more than 10 patients each.

Most of these mutations are proven or predicted to have a gain-of-function (GOF) mechanism, and oncogenic mutations at all of these amino acid sites have been seen in COSMIC (13). Published functional studies demonstrate a GOF mechanism for at least 9 of 41 mutations. For another 8 mutations, GOF has been shown for different missense mutations at the same codon (13). For example, our study identified p.Asn345Thr, p.Glu545Asp, and p.Gln546His mutations, while GOF has been shown for p.Asn345Lys, p.Glu545Lys, p.Glu545Gly, p.Gln546Lys, and p.Gln546Pro. We also detected p.Tyr1021His and p.Ala1035Thr mutations in our series, while p.Tyr1021Cys and p.Ala1035Val mutations have been found in other patients (16). Further, Glu545Ala and Glu545Lys mutations have been previously detected in individuals with undefined “Cowden-like” features as well (30).

Clinical results
The clinical features of all mutation-positive individuals are listed in Supplemental Tables 2 and 3. Clinical photographs and brain MRI images of some of these individuals are shown in Figures 3–5. Based on clinical and molecular characteristics, these patients can be segregated into three groups, as follows.

---

**Figure 2. Distribution of PIK3CA mutations by functional domain in cancer and developmental (pediatric) disorders.** Graph showing the number of published PIK3CA mutations by amino acid location in the Catalogue of Somatic Mutations in Cancer (COSMIC) database of somatic variation in cancer (shown in green; last accessed May 2016) and in children with developmental disorders comparing the megalencephaly-capillary malformation syndrome (MCAP; shown in blue) and all other developmental disorders (shown in orange). Mutations shown include those reported in this study as well as published mutations. Notes: (a) The 2-tailed P value by Fisher’s exact test was less than 0.0001, supporting that the association between MCAP and non-hotspot mutations and non-MCAP and hotspot associations is extremely statistically significant. (b) “Hotspot” mutations in this analysis are the most activating mutations in somatic tissues in cancer (p.Glu542Lys, p.Glu545Lys, p.His1047Arg) (13).
Group 1: MCAP syndrome. Most individuals with mosaic PIK3CA mutations in our cohort (n = 50) had classic features of MCAP, characterized by brain overgrowth (megalecephaly) and cutaneous vascular malformations, with variable body overgrowth, connective tissue laxity, and digital anomalies (polydactyly and syndactyly) (Figure 3) (31, 32). Several of our MCAP patients had additional noteworthy clinical findings including (a) complex structural heart defects or arrhythmias (n = 12), (b) lymphatic malformations including chylothorax and lymphedema (n = 5), (c) predisposition to thrombosis (n = 3), (d) endocrine disorders including hypothyroidism (n = 3), growth hormone deficiency (n = 2), and rhizomelic shortening of the extremities (n = 3). Six patients had episodes of hypoglycemia (Supplemental Tables 2 and 3). Interestingly, several of our MCAP patients (n = 6) had constitutional (or apparently constitutional) mutations in PIK3CA (including LR14-323, LR14-278, LR14-358, LR13-048, LR15-227, and LR15-246). Two of these patients (LR15-227 and LR15-246) had the p.Glu726Lys mutation.

Group 2: Intermediate phenotypes — capillary malformations with overgrowth and lipomatosis (CMO, CMOL). Several patients (n = 7) had novel phenotypes with features either overlapping MCAP but lacking megalencephaly or suggesting a milder variant of CLOVES. Patient LR15-238 had mild somatic hemihypertrophy and cutaneous vascular malformations (facial capillary malformation, and reticulated diffuse capillary malformations on the body), but was normocephalic. She had perinatal bradycardia, transient dilated cardiomyopathy, and a history of possible nonocclusive venous thrombosis in the internal jugular vein of unknown etiology. She had the p.Arg93Gln mutation in 19% of alleles in peripheral blood lymphocytes. Interestingly, the same mutation was identified in another patient with classic features of MCAP (LR14-323). The mutation in the latter patient appeared to be constitutional, seen in ~50% of alleles in blood- and saliva-derived DNA. Patient LR11-082 had congenital onset somatic overgrowth, diffuse congenital livedo reticularis and syndactyly, as well as severe joint laxity and multiple joint dislocations (Figure 4, A–D). She was normocephalic, and brain MRIs performed at 5 days and 6 months of age showed no cortical malformation. Interestingly, she had progressive cerebellar tonsillar ectopia and a markedly large cerebellum on her MRI at 6 months (Figure 5, A and B). This patient harbored the p.Gly106Val mutation in 34% and 40% of alleles in blood and skin, respectively. Patient LR12-131 had hemihypertrophy, capillary malformation on the philtrum and midline lower lip, a capillary malformation on the right arm, and toe syndactyly, but lacked other classic features of MCAP. He had a mosaic p.Cys378Tyr mutation in 7% of alleles in blood-derived DNA.

Patient LR12-070 had congenital hemihypertrophy, epidermal nevi, and lipomatous overgrowth of the trunk that was characteristically milder than CLOVES (Figure 4, E–G). He also had overlying capillary malformations, and macrodactyly requiring surgical excision. He harbored the p.Glu453Lys mutation in 25% of alleles in fibroblasts. This mutation was undetectable in saliva. Patient LR12-184 had diffuse asymmetric somatic overgrowth, reticulated port-wine stain, syndactyly, and connective tissue laxity. He was normocephalic, with no cortical malformations, and normal cognition. Interestingly, his brain MRI scan at age 5 months showed severe cerebellar tonsillar ectopia (Figure 5C). He subsequently developed sporadic swelling of the neck believed to be due to lymphatic fluid accumulation. He harbored the p.Glu453Lys mutation at variable levels (ranging from 5%–17% of alleles in affected skin regions, 24.5% in saliva, and 0.9% of alleles in blood-derived DNA). Patient LR12-329 had hemihypertrophy, cutaneous vascular malformations, and multiple cutaneous lipomas, but lacked classic features of MCAP or CLOVES. He had the p.Glu453Lys mutation in 6% of alleles in saliva. Patient LR13-045 also had hemihypertrophy and diffuse congenital live-
do reticularis but lacked megalencephaly. He had the p.Gly914Arg mutation in 5% and 15% of alleles in normal and affected skin-derived fibroblasts, respectively. This mutation was undetectable in saliva-derived DNA.

**Group 3: hotspot-associated phenotypes (CLOVES/DMEG).** Four patients harbored hotspot PIK3CA mutations most commonly seen in cancer. Two of these patients had a mosaic p.Glu542Lys mutation. Patient LR12-183 had a novel constellation of features including megalencephaly with diffuse cortical dysplasia, partial agenesis of the corpus callosum, dysplasia/hypoplasia of the cerebellar vermis and hemispheres, diffuse dysmyelination, abnormal high T2 signal intensities in the red nuclei and thalami bilaterally, and a very large tectum (Figure 5, D–G). He also had multiple linear sebaceous nevi on the face and trunk (Figure 4, H and I). This combination of features meets criteria for linear nevus sebaceous or Schimmelpenning syndrome (33–35). This patient’s mutation was seen in 39/139 (28%) of alleles in saliva-derived DNA. Patient LR13-197 had diffuse megalencephaly with bilateral cortical dysplasia, hippocampal dysplasia, white matter dysmyelination, and dysplastic ventricles bilaterally (Figure 5, H and I). She harbored the p.Glu542Lys mutation at 26% of alleles in saliva-derived DNA. Patients LR13-172 and LR13-265 both had very low-level mosaic p.His1047Arg mutations. Patient LR13-172 had megalencephaly with macrodactyly and harbored the mutation in 1% of alleles in blood-derived DNA (Figure 4, J and K). Patient LR13-265 had features of CLOVES syndrome with lipomatous overgrowth involving the face and trunk, prominent venous network in the abdominal wall, verrucous overgrowth of the trunk, and macrodactyly. His mutation was present in 4% of affected skin-derived DNA, and absent from several unaffected skin samples. We also obtained addi-
tional information on a previously reported patient (LR12-033) (16, 36) who demonstrates features of both dysplastic megalencephaly (effectively bilateral hemimegalencephaly) and CLOVES syndrome, making her only the second reported child with this severe combination. She had severe asymmetric overgrowth, diffuse soft tissue swelling consistent with lymphedema, asymmetric head, severe hypotonia, and possible right optic nerve hypoplasia (Supplemental Figure 5).

Constitutional PIK3CA mutations. Ten individuals in our cohort harbored constitutional (or apparently constitutional) mutations of PIK3CA that were detectable in peripheral samples (blood or saliva) at AAPs ranging from 35%–60% (LR14-323, LR01-060, LR15-076, LR13-036, LR14-278, LR14-358, LR13-048, LR15-227, LR12-330, LR15-246, and LR12-340). Interestingly, seven of these individuals had features of MCAP. However, three had diffuse brain or body overgrowth without segmental or patchy manifestations. Patient LR15-076 had diffuse megalencephaly, a small patent foramen ovale, and pectus excavatum deformity. Patient LR13-036 had a unique constellation of features including megalencephaly with rhizomelic shortening of the extremities, polydactyly, complex cardiac defects (left-sided aortic arch, ventricular septal defect, atrial dilatation), recurrent hypoglycemia, omphalocele, organomegaly, laryngeal stenosis, and recurrent infections, partially overlapping with Beckwith-Wiedemann syndrome. Patient LR12-340 had somatic asymmetry, was tall for age, with bilateral 2-3 toe syndactyly, but lacked megalencephaly and vascular anomalies. Mutations in five of these individuals were detected by clinical WES. These PIK3CA mutations were detected in blood at standard coverage (20× to 30×), suggesting constitutional mutations. Of note, blood-derived DNA was not available for analysis on four of these individuals (LR12-340, LR01-060, LR11-200, and LR12-330). Therefore, the possibility of mosaicism cannot be excluded in these patients.

Discussion
Classification of PIK3CA-associated developmental phenotypes. GOF mutations of PIK3CA have been associated with a wide spectrum of pediatric phenotypes (Figure 6). The original and best known of these is KTS, which was first defined more than a century ago based on the now classic triad of capillary malformation, venous varicosities, and limb hypertrophy (37–39). However, published photographs of KTS vary considerably in appearance. Therefore, it appears that KTS is often used as an umbrella term for any disorder with vascular malformation and overgrowth. Several recently described syndromes appear to be more specific designations, and we have predominantly used these in analyzing our phenotypic data. CLOVES is the most severe KTS-like disorder. In practice, most patients with CLOVES have all of the anomalies in the mnemonic (above), dramatic postnatal progression of the segmental overgrowth and lipomatosis, and other abnormalities (19, 40, 41). A subset of patients with most but not all of the features of CLOVES has been designated as fibroadipose overgrowth/hemihypertrophy-multiple lipomatosis, but the criteria used to
separate them are unclear (10, 21). Capillary malformation with overgrowth (CMO) is a recently defined clinical entity characterized by extensive, diffuse, reticulate capillary malformations and variable hypertrophy without major complications (42). The phenotype is much less severe than CLOVES, although some patients with apparent CMO in infancy go on to develop lipomatosis at older ages. Several other relatively mild phenotypes that overlap with CMO have also been described, including the “girth” and “RCM” types of vascular malformations with overgrowth (43). MCAP is characterized by symmetric or only mildly asymmetric brain overgrowth (megalencephaly) and cutaneous vascular malformations (predominantly capillary malformations), with variable malformations of cortical development (polymicrogyria), body overgrowth, digital anomalies (syndactyly and less often polydactyly), and connective tissue laxity (31, 32, 44, 45).

Beyond these multisystem disorders, mosaic mutations of PIK3CA have recently been identified in localized or focal forms of overgrowth with dysplasia including epidermal nevi (17), isolated macrodactyly (22), infiltrating lipomatosis (25, 46), isolated lymphatic malformations (28), and, most recently, venous malformations (29). Further, several patients with focal brain involvement, including dysplastic megalencephaly (DMEG), hemimegalencephaly (HMEG), and focal cortical dysplasia (FCD) have been identified as well (20, 27, 36). These disorders have been associated with variable but often low levels of mutations in affected tissues, and very low or undetectable levels in apparently unaffected tissues, such as peripheral blood lymphocytes. Finally, several patients have been reported with constitutional mutations of PIK3CA associated with isolated megalencephaly or unspecified “Cowden-like” features (30, 47).

Several recent reviews have presented PIK3CA-associated phenotypes as essentially a single broad spectrum (10). However, the original phenotype descriptions and the molecular data presented here argue otherwise. We view the differences between the clinical presentation as consistent and recognizable, based on our observations in the largest series of mosaic mutations of a single gene underlying developmental disorders reported to date. We propose that the PIK3CA-associated spectrum of disorders should serve as a paradigm for mosaic disorders. The clinical differences between MCAP and CLOVES (including CLOVES-DMEG) syndromes are summarized in Table 5. In MCAP, head and brain overgrowth are often the presenting signs, while body overgrowth at birth is mild (+1 to +2 SD) with stature normalizing by ~8 years (32). The cutaneous vascular malformation typically presents as congenital livedo reticularis (38). The head and body overgrowth and livedo reticularis are most often diffuse with subtle asymmetry. About 60%–70% of patients have a cortical malformation, with imaging and clinical features consistent with bilat-

---

**Figure 6. The PI3K-AKT-MTOR-related developmental brain disorders spectrum.** Diagram showing the PI3K-AKT-MTOR pathway highlighting genes associated with developmental brain disorders including PIK3CA, PIK3R2, PTEN, AKT3, MTOR, CCND2, DEPDC5, NPRL2, and NPRL3 (shown in blue), as well as TSC1 and TSC2 genes (shown in red).
eral perisylvian polymicrogyria (31, 32). Lipomas, lymphatic malformations, and macrodactyly occur, but are typically uncommon, limited in size, do not progress, and rarely require surgical intervention. Epidermal nevi are rare. In CLOVES, on the other hand, head and brain size are typically normal, although rare patients with combined CLOVES and DMEG have been observed (i.e., LR12-033). The cutaneous vascular malformations have a variable presentation from a classic port-wine stain appearance to pale pink or red capillary malformations. Body overgrowth including macrodactyly, lipomas, and lymphatic malformations are segmental and often severe with rapid postnatal progression. Epidermal nevi are common (19, 41).

Our data also show that the mutational spectrum of these disorders is different. The mutations associated with MCAP syndrome in particular are distinct from other PIK3CA-related overgrowth disorders. The mutational spectrum in MCAP is broad, with greater than 20 milder GOF mutations seen across the entire gene, whereas most mutations seen in CLOVES, KTS, DMEG, and other focal forms of overgrowth (such as isolated macrodactyly) are more often associated with the cancer hotspot mutations. Given this wide mutational spectrum, we recommend using a broad molecular strategy that tests the entire PIK3CA coding sequence, rather than targeted mutation screening of common mutations in individuals with similar phenotypes. Several individuals lacked megalencephaly, and had variable somatic overgrowth with vascular malformations, further expanding the spectrum of PIK3CA-related overgrowth. Therefore, the molecular data on PIK3CA-related overgrowth disorders to date show that they comprise a discontinuous spectrum of disorders that correlate well with the severity (strength of GOF) and distribution of the mutation, rather than a single continuous spectrum.

Diagnosis of mosaicism in developmental disorders

Genetic disorders caused by postzygotic (or mosaic) variants are increasingly recognized with the advent of deep NGS methods. The contribution of mosaic variation to developmental brain disorders has also been increasingly appreciated in various segmental disorders including HMEG, FCD, and, more recently, bilateral perisylvian polymicrogyria (16, 20, 27, 36, 48–51). However, detecting low-frequency mosaic variants in clinical practice remains challenging for several reasons. First, the availability of nonblood tissue sources (such as affected skin or brain) may be limited. Second, available diagnostic testing methods used in clinical diagnostic labs (such as Sanger sequencing or standard-depth NGS) are insufficiently sensitive to detect mosaic variation. We therefore sought to address these issues by sequencing large cohorts of samples from individuals with overgrowth in general, and MCAP in particular, using 2 independent deep-targeted NGS methods that utilize MIPs and amplicon deep sequencing. These methods have been particularly demonstrated to be effective for detecting low-frequency variants (14, 15). When combined with single-molecule tagging with multiplex targeted sequencing (smMIPs), the MIP method further increases the sensitivity for the detection of low-frequency variants, as single-molecule tagging marks sequences derived from a common progenitor molecule. Our results also challenge the current clinical practice of using blood-derived DNA for almost all clinical genetic testing. Our study suggests that a change in clinical practice is indicated for disorders suspected to be associated with somatic mosaicism, such as those involving asymmetry or other segmental or focal presentation, i.e., as a large majority of vascular malformations, overgrowth, cutaneous pigmentary abnormalities, noncancerous dysplasia (i.e., epidermal nevi).

### Table 5. Comparison of intermediate and hotspot PIK3CA gain-of-function phenotypes

<table>
<thead>
<tr>
<th>Feature/Group</th>
<th>Less-activating mutations (MCAP)</th>
<th>Hotspot mutations (CLOVES, DMEG)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brain size</td>
<td>Large</td>
<td>Normal</td>
</tr>
<tr>
<td>Malformations of cortical development</td>
<td>Normal or PMG</td>
<td>DMEG-HMEG-FCD type 2</td>
</tr>
<tr>
<td>Skin capillary malformation</td>
<td>Congenital livedo reticularis</td>
<td>Port-wine stain</td>
</tr>
<tr>
<td>Skin pigmented nevus</td>
<td>(none)</td>
<td>Epidermal nevi</td>
</tr>
<tr>
<td>Body overgrowth congenital</td>
<td>Diffuse, mild asymmetry</td>
<td>Segmental and severe</td>
</tr>
<tr>
<td>Body overgrowth postnatal</td>
<td>No</td>
<td>Severe</td>
</tr>
<tr>
<td>Lipomas-lipomatosis</td>
<td>Few and small</td>
<td>Many large and grow</td>
</tr>
<tr>
<td>Lymphatic malformation</td>
<td>Rare and small</td>
<td>Frequent and large</td>
</tr>
</tbody>
</table>

Abbreviations: CLOVES, congenital lipomatous truncal overgrowth, vascular anomalies, epidermal nevi and skeletal/spinal anomalies; DMEG, dysplastic megalencephaly; FCD, focal cortical dysplasia; HMEG, hemimegalencephaly; MCAP, megalencephaly-capillary malformation syndrome; PMG, polymicrogyria.
Summary
Our data expand on the molecular and phenotypic spectrum of PIK3CA-related developmental disorders. First, we show that the mutational spectrum in children with MCAP is broader than other PIK3CA-related overgrowth disorders. Further, we report on several atypical or expanded phenotypes, including children with body overgrowth without megalencephaly, that appear milder than CLOVES and other severe forms of somatic overgrowth. Finally, we report on constitutional mutations of PIK3CA in a subset of children, some of whom have milder features such as diffuse megalencephaly with intellectual disability, similar to the PTEN-related disorders (52). Based on our data and review of the literature, we propose a molecularly based classification of these disorders. Overall, our molecular studies in this large series show that phenotypes of mosaic disorders are impacted not only by tissue distribution and levels of mosaicism, but also by class (strength of activation) of mutation. Our series helps inform best clinical practices for diagnostic methods and tissues to sample in mosaic disorders broadly.

Methods
Sample processing. Genomic DNA was extracted from blood using the Puregene Blood Core Kit (QIAGEN), saliva using the Oragene Saliva Kit (DNA Genotek), buccal cells from Oragene Saliva Kits following centrifugation, and skin fibroblasts following a skin biopsy. Fibroblasts were grown in DMEM/F12 with glutamine, supplemented with 10% FBS and 1% PenStrep (all three from Gibco). DNA was extracted both directly and from fibroblasts after digesting with protease K in Cell Lysis Buffer (both QIAGEN) using prepIT.L2P (DNA Genotek) and ethanol precipitation. Lymphocytes were separated from whole blood using Ficoll and subsequently transformed to lymphoblastoid (LB) cell lines using Epstein-Barr virus (EBV) in RPMI 1640 media with glutamine, supplemented with 10% FBS and 1% PenStrep (all three from Gibco). Genomic DNA was then extracted using a Puregene Blood Core Kit.

Multiplex targeted sequencing using smMIPs. We designed a pool of 48 smMIP oligonucleotide probes targeting the coding sequences of PIK3CA. smMIPs tiled across a total of 3,340 bp of genomic sequence, including 100% of the 2,202 coding bp of PIK3CA. One hundred–nanogram capture reactions were performed in parallel. Massively parallel sequencing was performed with the Illumina HiSeq. Variant analysis was performed using our previously published pipeline including MIPgen and PEAR version 0.8.1 (48, 53). All missense, nonsense, and splice site variants seen at a frequency of less than 1% in public databases, in 2 or more capture events were retained for analysis. Our MIP capture method has been previously published (48).

Amplicon sequencing. Amplicon sequencing was performed using previously published methods (48). We performed locus-specific amplification of genomic DNA followed by GS Junior sequencing (Roche). We designed fusion primers containing genome-specific sequences along with distinct multiplex identifier sequences (used to differentiate samples being run together on the same plate) and sequencing adapters, to generate amplicons ranging in size from 290 to 310 bp, using Primer3Plus software. Primer sequences are provided in Supplemental Table 4. Small DNA fragments were removed with Agencourt AMPure XP (Beckman Coulter) according to the manufacturer’s protocol. All amplicons were quantified with the Quant-iT PicoGreen dsDNA reagent (Life Technologies), pooled at equimolar ratios, amplified by emulsion PCR using the GS Junior Titanium emPCR kit (Lib-A kit, Roche Applied Science), and pyrosequenced in the sense and antisense strands on a GS Junior sequencer following the manufacturers’ instructions. Data were analyzed using GS Amplicon Variant Analyzer version 3.0 software.

WES. WES produced 18.2 GB of sequencing data per proband. Mean coverage of captured regions was ~250× for proband samples, with ~99.17% covered with at least 10× coverage, an average of 89.82% of base call quality of Q30 or greater, and an overall average mean quality score of Q35. WES was performed using previously published methods (54).

Targeted NGS. This assay sequenced all exons of PIK3CA, and average coverage ranged from 320 to greater than 1,000 sequencing reads per bp. Genomic regions were captured using biotinylated RNA oligonucleotides (SureSelect), prepared in paired-end libraries with ~200-bp insert size, and sequenced on an Illumina HiSeq2500 instrument with 100-bp read lengths, in a modification of a previously reported pipeline (55). Large deletions and duplications were detected by this panel using previously published methods (56).

Statistics. We completed a pairwise comparison of mean values for alternative allele percentages in the three tissue types, where a 2-tailed t test gave P values of 0.0352, 0.0365, and 0.566 for blood versus saliva, blood versus skin, and saliva versus skin, respectively. A P value less than 0.05 was considered statistically significant.
Study approval. This study was approved by the IRB at Seattle Children's Hospital. Informed consent was obtained from subjects prior to enrollment in the study, except for subjects whose samples and clinical data were sent without identifying information. De-identified subjects were included under IRB waiver of consent, per IRB protocol at Seattle Children's Hospital. Written informed consent was provided for use of the patients' photographs in the clinical figures of this manuscript.

Author contributions
GMM designed the research study, conducted experiments, analyzed data, and wrote the manuscript. AET and AG analyzed data and assisted in writing the manuscript. VC, EAB, MK, SC, and BM conducted experiments and assisted with data analysis. CO and KP recruited subjects to this study and assisted with acquiring data. JJ, CP, and KT assisted with data acquisition and analysis. MC, IG, GKM, IKM, DC, NS, ASP, RB, ET, SB, LB, SEC, SG, MS, FS, CP, MN, SS, RS, MG, BOC, CEK, LJY, AHM, MM, MCA, LB, RC, KD, KTB, MEP, MP, LW, JR, RH, MA, KG, MC, HvE, LG, CC, HK, and JG assisted with acquiring data. JS and RG helped design the study. WBD designed the research study and wrote the manuscript.

Acknowledgments
We thank the patients, their families, and referring providers for their support of our research, and the Macrocephaly-Capillary Malformation (M-CM) Network (https://www.m-cm.net/) for their support as well. Research reported in this publication was supported by the National Institute of Neurological Disorders and Stroke (NINDS) and the National Heart, Lung and Blood Institute (NHLBI) of the National Institutes of Health under award numbers K08NS092898 (to G. Mirzaa), R01NS092772 and R01HL130996 (to W. Dobyns), by the EU Seventh Framework Programme (FP7) under the project DESIRE grant agreement N602531, E-Rare JTC 2011 (to R. Guerrini), and the Wellcome Trust under grant number 102731 (to A. Goriely). G. Mirzaa and W. Dobyns had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The funding sources have no role in the design and conduct of the study, collection, management, analysis and interpretation of the data, preparation, review or approval of the manuscript, or decision to submit the manuscript for publication.

Address correspondence to: Ghayda Mirzaa, Center for Integrative Brain Research, Seattle Children's Research Institute, 1900 9th Avenue, Mailstop C9S-10, Seattle, Washington 98101, USA. Phone: 206.884.1276; E-mail: gmirzaa@uw.edu.

386.