



Published in final edited form as:

N Engl J Med. 2014 August 21; 371(8): 733–743. doi:10.1056/NEJMoa1314432.

Somatic Mutations in Cerebral Cortical Malformations

Saumya S. Jamuar, M.R.C.P.C.H., Anh-Thu N. Lam, B.S., Martin Kircher, Ph.D., Alissa M. D'Gama, B.A., Jian Wang, Ph.D., Brenda J. Barry, M.S., Xiaochang Zhang, Ph.D., Robert Sean Hill, Ph.D., Jennifer N. Partlow, M.S., Aldo Rozzo, D.V.M., Ph.D., Sarah Servattalab, B.S., Bhaven K. Mehta, M.A., Meral Topcu, Ph.D., Dina Amrom, M.D., Eva Andermann, M.D., Ph.D., Bernard Dan, Ph.D., Elena Parrini, Ph.D., Renzo Guerrini, M.D., Ingrid E. Scheffer, M.B., B.S., Ph.D., Samuel F. Berkovic, M.D., Richard J. Leventer, M.B., B.S., Ph.D., Yiping Shen, Ph.D., Bai Lin Wu, Ph.D., A. James Barkovich, M.D., Mustafa Sahin, M.D., Ph.D., Bernard S. Chang, M.D., Michael Bamshad, M.D., Deborah A. Nickerson, Ph.D., Jay Shendure, M.D., Ph.D., Annapurna Poduri, M.D., M.P.H., Timothy W. Yu, M.D., Ph.D., and Christopher A. Walsh, M.D., Ph.D.

Division of Genetics and Genomics, Manton Center for Orphan Disease Research, Howard Hughes Medical Institute (S.S.J., A.-T.N.L., A.M.D., B.J.B., X.Z., R.S.H., J.N.P., A.R., S.S., B.K.M., T.W.Y., C.A.W.), and the Departments of Laboratory Medicine (J.W., Y.S., B.L.W.) and Neurology (M.S., A.P.), Boston Children's Hospital, the Departments of Pediatrics (S.S.J., A.-T.N.L., A.M.D., B.J.B., X.Z., R.S.H., J.N.P., A.R., S.S., B.K.M., T.W.Y., C.A.W.), Neurology (S.S.J., A.-T.N.L., A.M.D., B.J.B., X.Z., R.S.H., J.N.P., A.R., S.S., B.K.M., T.W.Y., C.A.W., M.S., A.P.), and Pathology (Y.S., B.L.W.), Harvard Medical School, the Department of Neurology, Beth Israel Deaconess Medical Center (B.S.C.), and the Department of Neurology, Massachusetts General Hospital (T.W.Y.) — all in Boston; the Department of Paediatrics, KK Women's and Children's Hospital, Singapore, Singapore (S.S.J.); the Department of Genome Sciences, University of Washington, Seattle (M.K., M.B., D.A.N., J.S.); the Department of Laboratory Medicine, Shanghai Children's Medical Center, Shanghai (J.W., Y.S.); the Division of Neurology, Department of Pediatrics, Hacettepe University School of Medicine, Sıhhiye, Ankara, Turkey (M.T.); the Neurogenetics Unit, Montreal Neurological Hospital and Institute, Department of Neurology and Neurosurgery (D.A., E.A.) and Department of Human Genetics (E.A.), McGill University, Montreal; the Department of Neurology, Hôpital Universitaire des Enfants Reine Fabiola, Université Libre de Bruxelles, Brussels (B.D.); the Pediatric Neurology Unit and Laboratories, Children's Hospital A. Meyer—University of Florence, Florence, Italy (E.P., R.G.); the Department of Medicine, University of Melbourne, Austin Health, Heidelberg (I.E.S., S.F.B.), Department of Paediatrics, Royal Children's Hospital, University of Melbourne, and the Florey Institute of Neuroscience and Mental Health, Melbourne (I.E.S.), and the Department of Neurology, Royal Children's Hospital, Murdoch Children's Research Institute, University of Melbourne, Parkville (R.J.L.) — all in Victoria, Australia; Claritas Genomics (Y.S., B.L.W.) and the Program in Medical and Population Genetics, Broad Institute of Massachusetts Institute of Technology and Harvard (C.A.W.) — both in Cambridge, MA; and the Department of Radiology

Copyright © 2014 Massachusetts Medical Society.

Address reprint requests to Dr. Walsh at the Division of Genetics and Genomics, Boston Children's Hospital, Center for Life Sciences 15062.2, 3 Black-fan Circle, Boston, MA, 02115, or at christopher.walsh@childrens.harvard.edu..

Disclosure forms provided by the authors are available with the full text of this article at NEJM.org.

and Biomedical Imaging, University of California, San Francisco (UCSF), and UCSF Benioff Children's Hospital — both in San Francisco (A.J.B.).

Abstract

BACKGROUND—Although there is increasing recognition of the role of somatic mutations in genetic disorders, the prevalence of somatic mutations in neurodevelopmental disease and the optimal techniques to detect somatic mosaicism have not been systematically evaluated.

METHODS—Using a customized panel of known and candidate genes associated with brain malformations, we applied targeted high-coverage sequencing (depth, 200×) to leukocyte-derived DNA samples from 158 persons with brain malformations, including the double-cortex syndrome (subcortical band heterotopia, 30 persons), polymicrogyria with megalencephaly (20), periventricular nodular heterotopia (61), and pachygyria (47). We validated candidate mutations with the use of Sanger sequencing and, for variants present at unequal read depths, subcloning followed by colony sequencing.

RESULTS—Validated, causal mutations were found in 27 persons (17%; range, 10 to 30% for each phenotype). Mutations were somatic in 8 of the 27 (30%), predominantly in persons with the double-cortex syndrome (in whom we found mutations in *DCX* and *LIS1*), persons with periventricular nodular heterotopia (*FLNA*), and persons with pachygyria (*TUBB2B*). Of the somatic mutations we detected, 5 (63%) were undetectable with the use of traditional Sanger sequencing but were validated through subcloning and subsequent sequencing of the subcloned DNA. We found potentially causal mutations in the candidate genes *DYNC1H1*, *KIF5C*, and other kinesin genes in persons with pachygyria.

CONCLUSIONS—Targeted sequencing was found to be useful for detecting somatic mutations in patients with brain malformations. High-coverage sequencing panels provide an important complement to whole-exome and whole-genome sequencing in the evaluation of somatic mutations in neuropsychiatric disease. (Funded by the National Institute of Neurological Disorders and Stroke and others.)

Somatic mutation, a postzygotic event, leads to two or more populations of cells with distinct genotypes in an organism, despite development from a single fertilized egg.^{1,2} Although the role of somatic mutation in cancer cells is well established,³ an analogous role for somatic mutations that occur randomly during the normal mitotic cell divisions of embryonic development — and that are therefore present in clones of cells in one or more tissues of the body — has been recognized only recently. Somatic mutations have been described in several noncancerous disorders, including the McCune–Albright syndrome,⁴ the Sturge–Weber syndrome,⁵ the Proteus syndrome,⁶ and hemimegalencephaly.^{7–9} The “obligatorily” somatic mutations that cause these diseases appear to be strongly activating mutations that are lethal when present in the germ-line. A notch lower in the hierarchy of severity are the somatic mutations found in persons with tuberous sclerosis, the double-cortex syndrome, periventricular nodular heterotopia, and bilateral megalencephalic disorders; these persons typically have milder presentations of disease than do those who carry the mutation in their germ-line.^{1,8,9} The broader relevance of somatic mutations to other neurologic conditions, such as autism and epilepsy in persons without structural brain malformations, is not known.

Recent advances in genomic technology provide unprecedented opportunities to assess germ-line mutations and have underscored the inability of traditional Sanger sequencing methods to consistently detect mosaic variants.^{6,10-14} Deep next-generation sequencing is commonly used to detect somatic mutations in tumor samples and has occasionally been used in noncancerous disorders,^{8,15} but its clinical usefulness in detecting noncancerous somatic mutations has not been established. Brain malformations are known to be associated with somatic mutations, although the prevalence of such associations is not known.¹ We tested the diagnostic sensitivity of deep next-generation sequencing for the detection of disease-associated somatic mutations present at low levels in blood samples and estimated the prevalence of mosaicism in these conditions.

Methods

Participants

The study was approved by the institutional review boards of Boston Children's Hospital and Beth Israel Deaconess Medical Center, Boston. Written informed consent was obtained from all participants or their parents or guardians. Of the 158 persons studied, 30 had the double-cortex syndrome (also called subcortical band heterotopia), 20 had polymicrogyria with megalencephaly, 61 had periventricular nodular heterotopia, and 47 had pachygyria; all conditions were diagnosed with the use of magnetic resonance imaging (MRI). Twenty persons had undergone previous genetic testing, but they had not undergone screening for mutations in all genes that have been implicated in their specific disorders.

Target Capture and Sequencing

We used the Illumina TruSeq Custom Amplicon Kit to capture all exons, intron–exon boundaries, and 10-bp flanking sequences of target genes (RefSeq database, hg19 assembly). We designed two targeted gene panels — panel 1, made up of 14 genes, and panel 2, made up of 54 genes. Gene panel 1 consisted of 10 known causative genes and 4 candidate genes and was used to test the participants with the double-cortex syndrome, polymicrogyria with megalencephaly, or periventricular nodular heterotopia; gene panel 2 included 18 known causative genes and 36 candidate genes and was used to test the participants with pachygyria (for known genes, see Table S1 in the Supplementary Appendix, available with the full text of this article at NEJM.org). To detect mosaicism of 5% or higher, libraries were prepared and sequenced in multiple batches to achieve a mean depth of at least 200× (Fig. S2 in the Supplementary Appendix). Details regarding the preparation of samples and bioinformatic methods are provided in the Supplementary Appendix.

Sanger Sequencing Validation

Sanger sequencing was used to validate all rare and protein-altering (nonsynonymous, nonsense, splice-site, frameshift, and insertion–deletion [indel]) variants in the target genes. Validated variants were tested for familial segregation when samples from parents or siblings were available. Parentage was confirmed with the use of the AmpFLSTR Identifier PCR Amplification Kit (Life Technologies).

Criteria for Pathogenicity

We defined pathogenic mutations as those that met the following criteria: they occurred in genes that, when mutated, are known to cause the disorder; they were absent from controls (i.e., samples in the Exome Sequencing Project, dbSNP, and 1000 Genomes databases); they were predicted to alter the sequence of the encoded protein (non-synonymous, nonsense, splice-site, frameshift, and indel mutations); and they were predicted with the use of in silico prediction software (SIFT¹⁶ or PolyPhen-2¹⁷) to adversely affect protein function. In addition, for germline variants, we required the variant to have arisen de novo, to segregate with the disorder, to have been inherited from a parent with somatic mosaicism, or to be characterized by a dominant or X-linked mode of inheritance, when applicable. For six probands with germline mutations, parental DNA was unavailable.

Somatic-Mutation Analysis

Because an alternate-allele read frequency of 40% or lower deviates significantly from the 50:50 ratio expected from heterozygous germline variants (Fig. S2 in the Supplementary Appendix), variants for which the alternate-allele read frequency was 40% or lower were evaluated for potential mosaicism by subcloning. Original DNA was reamplified by means of a polymerase-chain-reaction (PCR) assay, subcloned into a TA vector (TOPO TA Cloning Kit, Invitrogen), and transformed into TOP10 *Escherichia coli* cells. Multiple individual transformants were reisolated, and Sanger sequencing was performed to confirm the presence or absence of the predicted variant and to quantify the degree of mosaicism.

Results

Targeted Next-Generation Sequencing

With the use of targeted next-generation sequencing, all variants in positive control samples were successfully identified, and control somatic mutations at allele frequencies of 1% or higher were detected when the read depth was at least 1000× (Fig. S3 in the Supplementary Appendix). With a mean depth of approximately 300× (64% coverage, 200×), a mean (\pm SD) of 368 ± 153 single-nucleotide and indel variants per sample were identified. After common and synonymous variants had been filtered and excluded, 57 rare, protein-altering variants were found in 50 persons, including 35 variants with an alternate-allele read frequency higher than 40% and 22 variants with an alternate-allele read frequency of 40% or lower. All 35 variants with an alternate-allele read frequency higher than 40% were validated with the use of Sanger sequencing; 19 of these variants fulfilled the criteria for pathogenicity, and 1 was classified as a variant of uncertain significance (Tables S3 through S7 in the Supplementary Appendix). Of the 5 variants with an alternate-allele read frequency of 30 to 40%, 2 were validated with Sanger sequencing. One, in a person with periventricular nodular heterotopia (Participant PH-16001), fulfilled the criteria for pathogenicity, whereas the other, in a person with pachygyria (Participant PAC-1701), was classified as a variant of uncertain significance. Subcloning results suggested that the mutation in Participant PH-16001 was mosaic (Table S9 in the Supplementary Appendix) and the variant in Participant PAC-1701 was germline. Of 17 variants with alternate-allele read frequencies lower than 30%, 4 were validated with Sanger sequencing. Of these 4 variants, 2 were pathogenic (in Participant DC-2801, with the double-cortex syndrome, and

Participant PAC-902, with pachygyria), and 2 were inherited (in Participant BFP-801, with pachygyria, and Participant PH-23901, with periventricular nodular heterotopia); one of the inherited variants was classified as a variant of uncertain significance, and the other was classified as benign. Of the remaining 13 variants at an alternate-allele read frequency lower than 30%, 5 were validated by means of subcloning (Table 1, and Tables S3, S4, and S9 in the Supplementary Appendix). Whether validation was successful was predictable according to the depth of coverage: variants with coverage of at least 100×, especially mosaic variants, were more likely to be validated (Fig. 1A), and germline and mosaic variants with coverage of 20× or less and 60× or less, respectively, were likely to be errors in next-generation sequencing (Fig. 1A and 1B).

Subcloning of candidate mosaic mutations allowed independent confirmation and quantification of these mutations, and alternate-allele read frequencies determined by means of subcloning closely matched those defined by means of next-generation sequencing. However, Sanger sequencing misclassified mosaic mutations at both higher and lower ranges of mosaicism (Fig. 1C). Four mosaic mutations at an alternate-allele read frequency lower than 17% were completely undetectable with the use of Sanger sequencing, and a fifth (in Participant DC-2101, who had the double-cortex syndrome) showed an indistinct peak visible only on retrospective analysis; these five mutations account for 63% of all mosaic mutations identified in our study. All five of these mutations were readily validated by means of subcloning (Fig. 2A, 2B, and 2C). Although somatic variants present at high allele frequencies were readily detected with the use of Sanger sequencing, the proportion of the mutant allele was overestimated, to the extent that they were interpreted as germline mutations (Fig. 2D, and Fig. S4 in the Supplementary Appendix).

The eight mosaic mutations we identified show very strong evidence of pathogenicity (Tables S10 and S11 in the Supplementary Appendix). Three have been reported previously, one is predicted to result in a nonconservative missense substitution of an amino acid residue that is mutant in other patients, three are predicted to be null alleles (one nonsense, one splicing, and one frame-shift), and one is a missense mutation in *LIS1*; the amino acid sequence of *LIS1* is highly conserved and virtually identical between mouse and human. The predicted proportion of blood cells carrying the mutation (Table 1) was calculated by doubling the alternate-allele read frequency for heterozygous autosomal mutations in both sexes and for X-linked mutations in females. For mutations in X-linked genes in males, the predicted percentage of mutant cells was equal to the alternate-allele read frequency. Our data suggest that central nervous system phenotypes (Fig. S5 in the Supplementary Appendix) can occur in association with mutations detectable in as few as approximately 10% of peripheral-blood cells, although the proportion of mutant cells in the brain is not known.

To determine the specificity of our gene panel, we compared the number of synonymous variants (silent, noncoding, and synonymous) with the number of rare protein-altering variants in genes that were strongly associated with only one phenotype across the three phenotypes represented in gene panel 1 (Table S1 in the Supplementary Appendix). There were similar numbers of synonymous variants per person across the three phenotypes ($P = 0.82$ by Fisher's exact test) (Fig. S6 in the Supplementary Appendix), whereas rare protein-

altering mutations were strongly enriched in genes previously associated with each diagnosis ($P < 0.001$ by Fisher's exact test) (Table 1, and Fig. S6 and Table S5 in the Supplementary Appendix). For example, *DCX* mutations are known to cause the double-cortex syndrome, and in our panel, rare, protein-altering mutations in *DCX* were strongly associated with the double-cortex phenotype and not with other conditions, whereas synonymous changes in *DCX* were equally common across phenotypes. This strongly suggests that the rare protein-altering mutations are pathogenic.

Using gene panel 1, we detected mutations in 30% of persons with the double-cortex syndrome; 20% of the mutations were mosaic (Table 2, and Table S11 in the Supplementary Appendix). A known recurrent *DCX* mutation (R186C) was present in the germline in Participant DC-601 and was somatic in Participant DC-4601. As expected, the germline mutation was associated with a thick band of heterotopia both anteriorly and posteriorly, whereas the mosaic mutation, with an alternate-allele read frequency of approximately 5% (approximately 10% mutant cells), was associated with a milder phenotype, with a much thinner band of heterotopia predominantly limited to anterior regions (Fig. S5 and Tables S3 and S4 in the Supplementary Appendix). Among persons with polymicrogyria with megalencephaly, we detected no mosaic variants, but we did detect known pathogenic germline variants in two (10%) of them: *AKT3* (R465W) and *PIK3CA* (A1035V). We detected pathogenic mutations in *FLNA* in eight persons with periventricular nodular heterotopia (13%): seven of the mutations were germline, and one was mosaic (S1449fs in Participant PH-16001; alternate-allele read frequency, 35%).

Using gene panel 2, we detected pathogenic variants in eight persons with pachygyria (17%), one of whom (Participant PAC-902) had a pathogenic allele in *TUBB2B* in the mosaic state (R380P; alternate-allele read frequency, 23%) (Table 1 and Fig. 2D, and Table S5 in the Supplementary Appendix).¹⁸ MRI showed frontal pachygyria; parietal, occipital, and temporal polymicrogyria; a small dysplastic cerebellum; a hypoplastic pons; and hypoplastic optic nerves (Fig. 3A); these features were consistent with a *TUBB2B* phenotype.¹⁹ This participant had previously undergone an extensive evaluation, including whole-exome sequencing, that did not uncover any potential variants. Review of the data from whole-exome sequencing showed the mosaic variant, but it had been filtered out as a result of low coverage (4×).

Mutations in *DYNC1H1* and New Candidate Genes

Two participants (BFP-601 and PMG-17401, both of whom had pachygyria) had de novo mutations in *DYNC1H1* (E561G and R3344Q) (Table S5 in the Supplementary Appendix). In a parallel study, we detected de novo mutations in *DYNC1H1* in two other participants (PAC-1601 and LIS-8201, both of whom had pachygyria) through whole-exome sequencing, which allowed a further delineation of the radiographic phenotype of this disorder (Fig. 3B and 3C). The four persons had strikingly similar MRI findings, with posterior-predominant pachygyria and corpus callosum abnormalities (Fig. 3B, and Table S12 in the Supplementary Appendix). Mutations in *DYNC1H1* have been associated with peripheral neuropathy²⁰ and more recently with brain malformations, predominantly diffuse polymicrogyria with a thickened corpus callosum.²¹ However, a person with the same

mutation as Participant PMG-17401 (R3344Q) also had posterior-predominant pachygyria.²¹ These four new cases further support the role of *DYNC1H1* mutations in brain malformations and suggest that posterior-predominant pachygyria is an imaging hallmark of conditions associated with mutations in this gene. We also found a previously unreported variant in *KIF5C* (A268S) and variants in three candidate genes that bear further study — *KIF7* (G94D), *KIF1A* (R18W), and *KIF26A* (Q455R) (Table S5 in the Supplementary Appendix). All variants were predicted by PolyPhen-2 to be “probably damaging” and were absent from the control population, but identification of additional alleles in future studies will be required before the implicated variants can be considered to be causal (see the Supplementary Appendix).

Discussion

Using a targeted approach that involved deep, high-coverage, next-generation sequencing, we identified likely causative mutations in 17% of study participants and found that a substantial fraction of these mutations (30%) were mosaic, with mutation present in some but not all peripheral-blood cells and presumably similarly mosaic in brain tissue. Our data suggest that a subset of persons with somatic mutations affecting the brain have the mutation at detectable levels in blood, in as few as 5% of DNA reads, corresponding to a heterozygous mutation in 10% of cells. However, 63% of mosaic mutations detected by means of next-generation sequencing were undetectable with direct Sanger sequencing, which suggests that detection requires the high sequence coverage afforded by next-generation-sequencing panels (ideally 500×, producing alternate-allele coverage of >20×, assuming the allele is present in at least 10% of cells), since clinical sequencing or conventional whole-exome sequencing typically involves coverage of 40 to 80×. Sanger sequencing has previously been noted to have a threshold of detection of approximately 15 to 20%.¹⁴ We observed that mosaic mutations present at higher levels (60 to 70% of cells) tended to look like germline variants when Sanger sequencing was used. Therefore, Sanger chemistry appears to have limitations on both sides of the spectrum of detection: it is insensitive to low levels of mosaicism, and it miscalls high levels of mosaicism as being constitutive.

We did not functionally validate all germline and somatic mutations, but our criteria for pathogenicity were conservative. Of the 27 mutations detected, 12 were previously reported to be pathogenic (Table S11 in the Supplementary Appendix), and another 10 were functionally null mutations (frameshift, essential splice-site, or nonsense) in genes previously determined to cause brain malformation when mutated. For the remaining 5 mutations, several lines of evidence strongly support pathogenicity. First, they occurred in a previously implicated gene. Second, they were very rare or absent from control databases. Third, they were predicted by in silico prediction software to be pathogenic. Fourth, all mutations affected highly conserved amino acid residues in the respective proteins. Fifth, there was high coverage across these mutations, and they were validated with the use of either bulk Sanger sequencing or subcloning followed by colony sequencing. Last, deleterious mutations were strongly correlated with phenotype, such that, for example, deleterious mutations in *DCX* were common among patients with diseases caused by *DCX* but absent in other patients. Similarly, deleterious mutations in *LIS1*, *TUBA1A*, *FLNA*,

AKT3, and *PIK3CA* were seen only in persons with the gene-specific phenotype, which supported causal roles for these mutations.

We identified novel germline mutations in *DYNC1H1* and *KIF5C*, as well as in *KIF7*, *KIF1A*, and *KIF26A*. *DYNC1H1* has been implicated in polymicrogyria,²¹ although our four participants with *DYNC1H1* mutations had posterior-predominant pachygyria.²¹ Polymicrogyria is characterized by multiple small gyri,²² whereas pachygyria (“few gyri”) and lissencephaly (“smooth brain”) are characterized by fewer gyri. *DYNC1H1* forms a complex with the *LIS1* protein, and the *DYNC1H1* mutant phenotype is remarkably similar to the *LIS1* mutant phenotype in its posterior predominance.

Although we focused on identifying somatic mutations that affect the brain but are detectable at low levels in blood cells, some important somatic mutations affecting the brain appear to be completely undetectable in blood. Mosaic tetrasomy 12p in the Pallister–Killian syndrome is typically undetectable in blood but is readily detectable in skin fibroblasts and buccal swabs.²³ We recently reported an *AKT3* point mutation and a chromosome 1q copy-number gain associated with hemimegalencephaly, but we could not detect these mutations in blood.⁷ Using our panel, we detected the *AKT3* mutation in DNA from brain but not blood, despite more than 3000× coverage of the *AKT3* gene from blood DNA (data not shown). Identification of certain highly toxic mutations may ultimately be possible only in affected tissue. In a recent study, somatic mosaicism was identified in 23% of patients with the Cornelia de Lange syndrome by analyzing DNA from buccal swabs.²⁴ The presence of a mosaic mutation was confirmed with DNA from saliva and buccal swabs from one participant in our study (Participant DC-401, who had the double-cortex syndrome), and a higher proportion of the mutant allele was present in the buccal swab than in blood (Table S10 in the Supplementary Appendix); however, further study is required to determine the distribution of somatic variants among blood, skin, saliva, and brain.

Whole-exome sequencing has an important place in the evaluation of neurodevelopmental abnormalities,^{25,26} but it can miss somatic variants if read depth or the alternate-allele read frequency is low, as evidenced by the fact that in our study, the *TUBB2B* mutation was missed with the use of whole-exome sequencing in one participant with pachygyria. Until whole-exome sequencing with very high coverage is routinely available, gene panels permit a higher depth of coverage and cost-efficient detection of somatic variants.²⁷ Future applications of sequencing panels may allow the successful interrogation of other disorders with known mosaicism and provide an understanding of other disorders with high rates of de novo mutations, such as autism spectrum disorders and other neuropsychiatric disorders.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Supported by grants from the National Institute of Neurological Disorders and Stroke (R01NS079277 and R01NS035129, to Dr. Walsh; and K23NS069784, to Dr. Poduri), the National Institute of Mental Health (1RC2MH089952, to Dr. Walsh), the University of Washington Center for Mendelian Genomics (HG006493), the Manton Center for Orphan Disease Research, the National Institute of General Medical Sciences (T32GM07753, to

Ms. D'Gama), and the Allen Foundation (to Dr. Walsh). Dr. Walsh is an Investigator of the Howard Hughes Medical Institute.

We thank the participants and their families, as well as our collaborators from the Brain Malformation Study Group for the referrals (a complete list of investigators in the Brain Malformation Study Group is provided in the Supplementary Appendix).

References

1. Poduri A, Evrony GD, Cai X, Walsh CA. Somatic mutation, genomic variation, and neurological disease. *Science*. 2013; 341:1237758. [PubMed: 23828942]
2. Biesecker LG, Spinner NB. A genomic view of mosaicism and human disease. *Nat Rev Genet*. 2013; 14:307–20. [PubMed: 23594909]
3. Watson IR, Takahashi K, Futreal PA, Chin L. Emerging patterns of somatic mutations in cancer. *Nat Rev Genet*. 2013; 14:703–18. [PubMed: 24022702]
4. Weinstein LS, Shenker A, Gejman PV, Merino MJ, Friedman E, Spiegel AM. Activating mutations of the stimulatory G protein in the McCune–Albright syndrome. *N Engl J Med*. 1991; 325:1688–95. [PubMed: 1944469]
5. Shirley MD, Tang H, Gallione CJ, et al. Sturge–Weber syndrome and port-wine stains caused by somatic mutation in GNAQ. *N Engl J Med*. 2013; 368:1971–9. [PubMed: 23656586]
6. Lindhurst MJ, Sapp JC, Teer JK, et al. A mosaic activating mutation in AKT1 associated with the Proteus syndrome. *N Engl J Med*. 2011; 365:611–9. [PubMed: 21793738]
7. Poduri A, Evrony GD, Cai X, et al. Somatic activation of AKT3 causes hemispheric developmental brain malformations. *Neuron*. 2012; 74:41–8. [PubMed: 22500628]
8. Rivière JB, Mirzaa GM, O'Roak BJ, et al. De novo germline and postzygotic mutations in AKT3, PIK3R2 and PIK3CA cause a spectrum of related megalencephaly syndromes. *Nat Genet*. 2012; 44:934–40. [PubMed: 22729224]
9. Lee JH, Huynh M, Silhavy JL, et al. De novo somatic mutations in components of the PI3K-AKT3-mTOR pathway cause hemimegalencephaly. *Nat Genet*. 2012; 44:941–5. [PubMed: 22729223]
10. Yang Y, Muzny DM, Reid JG, et al. Clinical whole-exome sequencing for the diagnosis of mendelian disorders. *N Engl J Med*. 2013; 369:1502–11. [PubMed: 24088041]
11. de Ligt J, Willemsen MH, van Bon BW, et al. Diagnostic exome sequencing in persons with severe intellectual disability. *N Engl J Med*. 2012; 367:1921–9. [PubMed: 23033978]
12. Opitz JM, Jorde LB. Hamartoma syndromes, exome sequencing, and a protean puzzle. *N Engl J Med*. 2011; 365:661–3. [PubMed: 21793737]
13. Musunuru K, Pirruccello JP, Do R, et al. Exome sequencing, ANGPTL3 mutations, and familial combined hypolipidemia. *N Engl J Med*. 2010; 363:2220–7. [PubMed: 20942659]
14. Rohlin A, Wernersson J, Engwall Y, Wiklund L, Björk J, Nordling M. Parallel sequencing used in detection of mosaic mutations: comparison with four diagnostic DNA screening techniques. *Hum Mutat*. 2009; 30:1012–20. [PubMed: 19347965]
15. Pagnamenta AT, Lise S, Harrison V, et al. Exome sequencing can detect pathogenic mosaic mutations present at low allele frequencies. *J Hum Genet*. 2012; 57:70–2. [PubMed: 22129557]
16. Ng PC, Henikoff S. Predicting deleterious amino acid substitutions. *Genome Res*. 2001; 11:863–74. [PubMed: 11337480]
17. Adzhubei IA, Schmidt S, Peshkin L, et al. A method and server for predicting damaging missense mutations. *Nat Methods*. 2010; 7:248–9. [PubMed: 20354512]
18. Cushion TD, Dobyns WB, Mullins JG, et al. Overlapping cortical malformations and mutations in TUBB2B and TUBA1A. *Brain*. 2013; 136:536–48. [PubMed: 23361065]
19. Jaglin XH, Poirier K, Saillour Y, et al. Mutations in the beta-tubulin gene TUBB2B result in asymmetrical polymicrogyria. *Nat Genet*. 2009; 41:746–52. [PubMed: 19465910]
20. Willemsen MH, Vissers LE, Willem-sen MA, et al. Mutations in DYNC1H1 cause severe intellectual disability with neuronal migration defects. *J Med Genet*. 2012; 49:179–83. [PubMed: 22368300]

21. Poirier K, Lebrun N, Broix L, et al. Mutations in TUBG1, DYNC1H1, KIF5C and KIF2A cause malformations of cortical development and microcephaly. *Nat Genet.* 2013; 45:639–47. [PubMed: 23603762]
22. Liu JS. Molecular genetics of neuronal migration disorders. *Curr Neurol Neurosci Rep.* 2011; 11:171–8. [PubMed: 21222180]
23. Jamuar S, Lai A, Unger S, Nishimura G. Clinical and radiological findings in Pallister-Killian syndrome. *Eur J Med Genet.* 2012; 55:167–72. [PubMed: 22387057]
24. Huisman SA, Redeker EJ, Maas SM, Mannens MM, Hennekam RC. High rate of mosaicism in individuals with Cornelia de Lange syndrome. *J Med Genet.* 2013; 50:339–44. [PubMed: 23505322]
25. Yu TW, Chahrour MH, Coulter ME, et al. Using whole-exome sequencing to identify inherited causes of autism. *Neuron.* 2013; 77:259–73. [PubMed: 23352163]
26. Allen AS, Berkovic SF, Cossette P, et al. De novo mutations in epileptic encephalopathies. *Nature.* 2013; 501:217–21. [PubMed: 23934111]
27. Rehm HL. Disease-targeted sequencing: a cornerstone in the clinic. *Nat Rev Genet.* 2013; 14:295–300. [PubMed: 23478348]

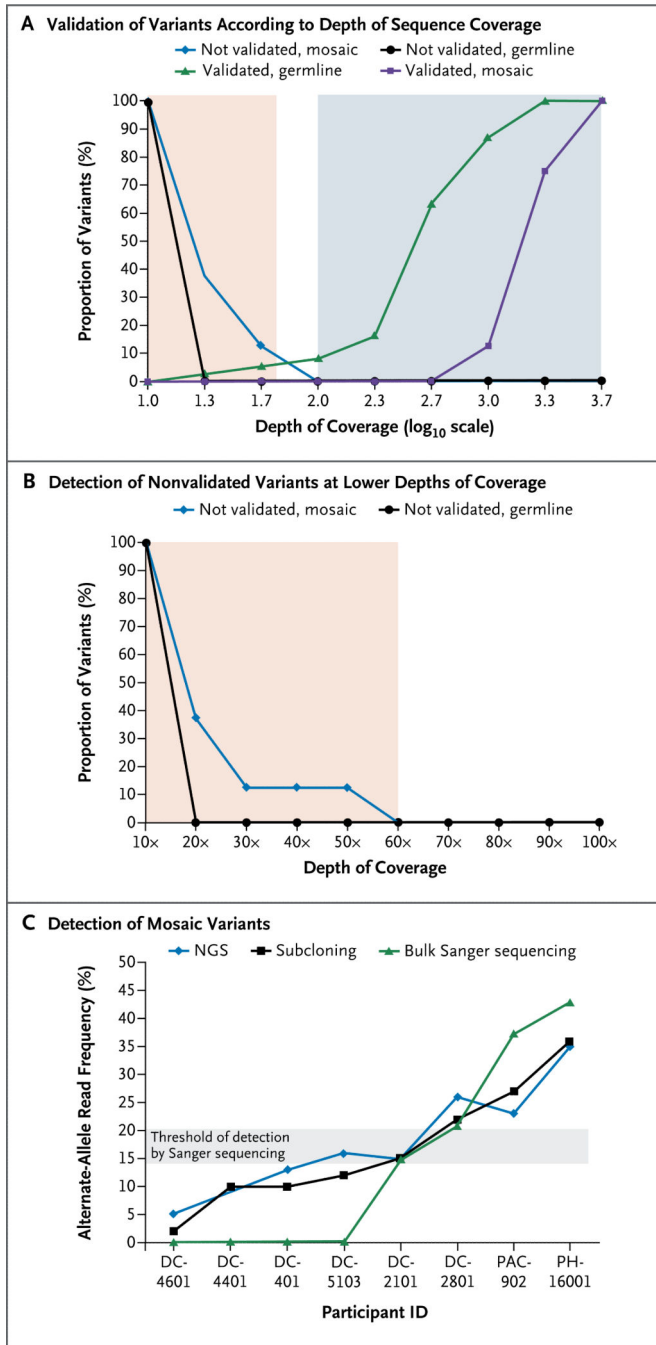


Figure 1. Specificity Analysis of the Deep Targeted-Sequencing Technique

Panel A shows the validation rate versus the depth of sequence coverage for germline and mosaic variants that were and were not validated successfully. Variants detected at a depth of more than 100x (blue shaded area) were more likely to be validated. Panel B is an enlarged view of the pink shaded and unshaded areas in Panel A. Variants detected at a depth of less than 60x (pink shaded area) were less likely to be validated than those detected at a depth of greater than 60x. This was especially true of mosaic variants. Panel C shows the proportion of mosaic variants detected by next-generation sequencing (NGS), which was

similar to that detected by subcloning. When Sanger sequencing was used, the mosaic variants at allele frequencies below 17% (in Participants DC-4601, DC-4401, DC-401, and DC-5103, all of whom had the double-cortex syndrome) were not detected, and mutations present at high allele frequencies (at approximately 35% in Participant PH-16001, who had periventricular nodular heterotopia, and approximately 23% in Participant PAC-902, who had pachygyria) resembled germline mutations. The mosaic variant in Participant DC-2101 was missed in the initial analysis and was detected as a small peak only on directed examination of the Sanger chromatogram.

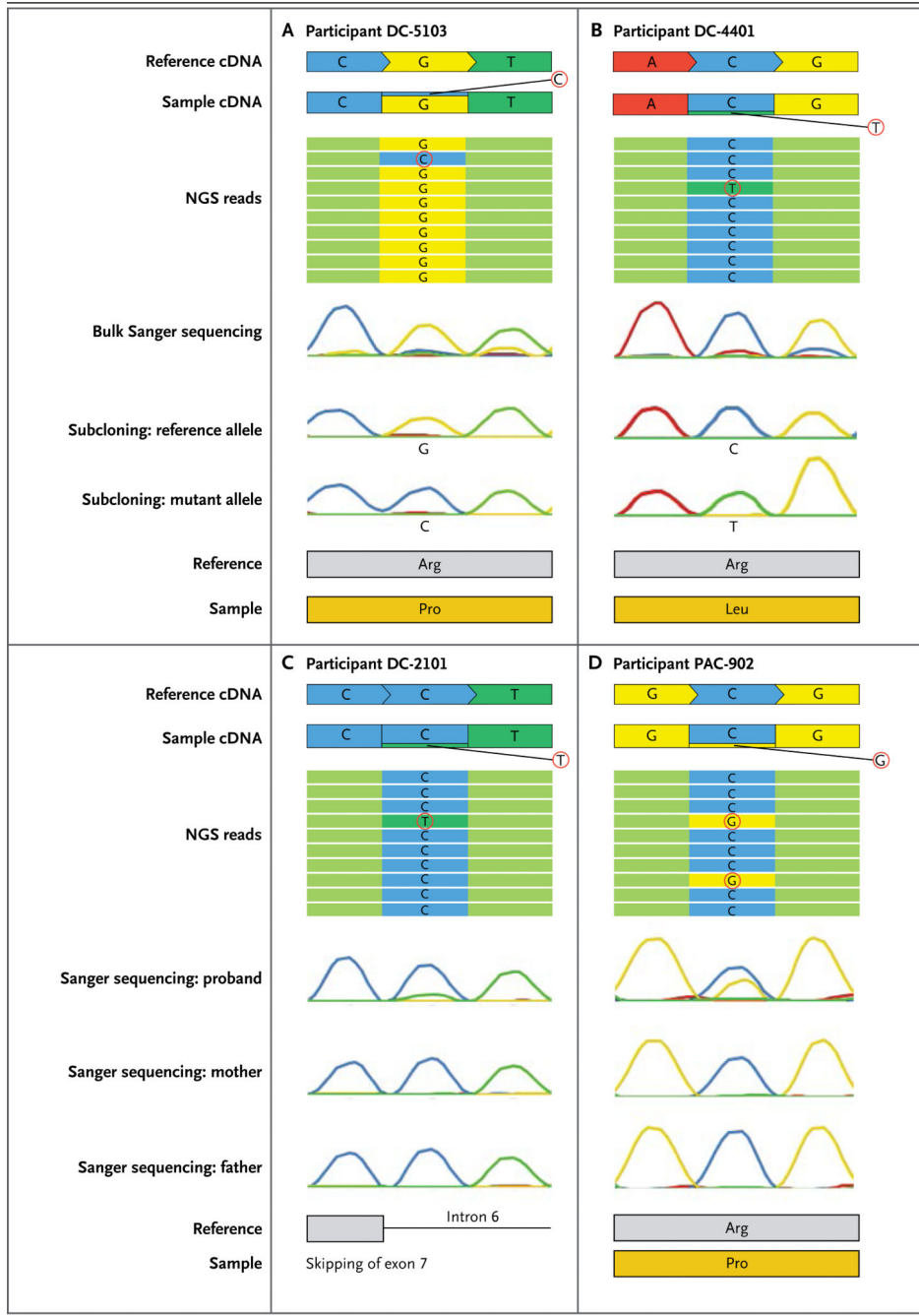


Figure 2. Detection of Variants by Means of NGS and Sanger Sequencing

The figure shows NGS reads aligned with Sanger sequencing chromatograms from bulk DNA amplification and after subcloning for Participant DC-5103 (Panel A) and Participant DC-4401 (Panel B); the variants (circled nucleotides) were detected in a fraction of the reads but were missed when targeted Sanger sequencing was used. Subcloning detected the variants in a proportion of clones, and the proportion of NGS reads was strongly correlated with the proportion of clones containing a mosaic read. NGS reads aligned with Sanger sequencing chromatograms from bulk DNA amplification of samples from Participant

DC-2101 (Panel C) and Participant PAC-902 (Panel D) and their respective parents show that the mosaic variant (circled nucleotide) was detected as a smaller peak, as compared with the reference allele, on Sanger sequencing and arose de novo (i.e., it was absent from the parents). For the splicing mutation in Participant DC-2101, the prediction software did not enable a choice from among the potential splice-acceptor sites; therefore, no amino acid is depicted.

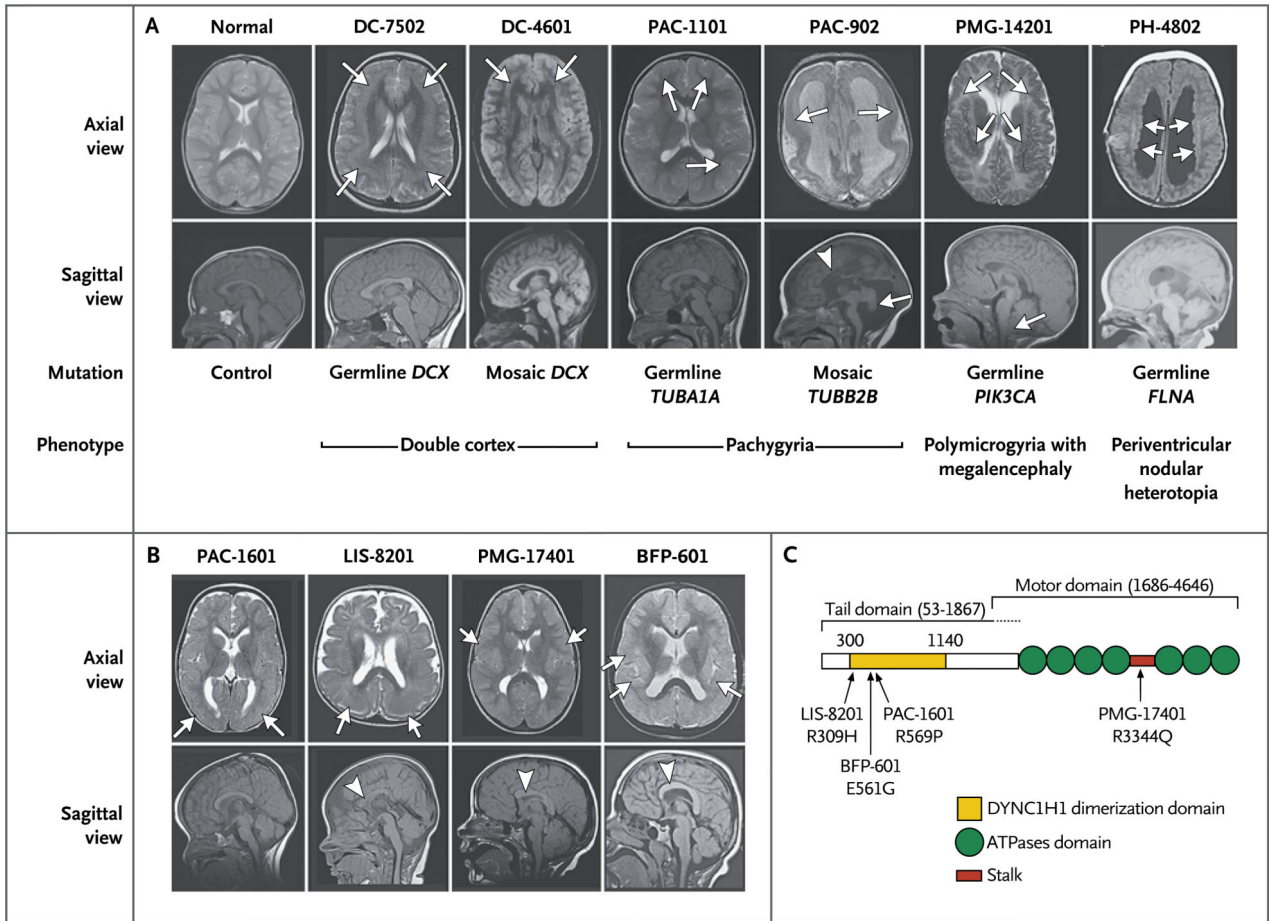


Figure 3. MRI Features of Brain Malformations Associated with Germline and Mosaic Variants, and Positions of Variants within the *DYNC1H1* Protein

Panel A shows axial and midline sagittal MRI scans of the head of a person without neurologic abnormalities; two persons with the double-cortex syndrome (white arrows), associated with germline variants in one (Participant DC-7502) and a mosaic variant in the other (Participant DC-4601); two participants with pachygyria (white arrows), associated with germline variants in one (Participant PAC-1101) and a mosaic variant in the other (Participant PAC-902); a person with polymicrogyria (white arrows) and megalencephaly associated with a germline variant (Participant PMG-14201); and a person with periventricular nodular heterotopia (white arrows) associated with germline variants (Participant PH-4802). In addition, the sagittal image in Participant PAC-902 shows dysgenesis of the corpus callosum (arrowhead) and hypoplastic cerebellar vermis (white arrow), and the sagittal image of Participant 14201 shows cerebellar tonsillar herniation (white arrow) due to mass effect from the enlarged cerebral hemispheres. In Panel B, axial and midline sagittal MRI scans in the four persons with de novo mutations in *DYNC1H1* show posterior-predominant pachygyria (white arrows in axial views in Participants PAC-1601 and LIS-8201), thick, irregular cortex in the perisylvian region (white arrows in scans from Participants PMG-17401 and BFP-601), and a mildly dysmorphic corpus callosum (arrowheads in sagittal views in Participants LIS-8201, PMG-1704, and BFP-601).

Panel C is a schematic representation of the DYNC1H1 protein, showing the positions of the alterations (black arrows) detected in our study.

Table 1

Details of Mosaic Mutations.*

Participant Number	Phenotype	Sex	Gene	Variant	Alternate-Allele Read Frequency %	Mutant Cell Frequency % [†]	Protein Change	Effect Predicted by SIFT or PolyPhen-2 [‡]	Previously Reported
DC-4601 [§]	Double-cortex syndrome	F	<i>DCX</i>	ChrX:110644367:C→T	5	10	R186C	Damaging	Yes
DC-4401 [§]	Double-cortex syndrome	M	<i>DCX</i>	ChrX:110653322:G→A	9	9	R78L	Damaging	Yes
DC-401 [§]	Double-cortex syndrome	F	<i>LISI</i>	Chr7:2569382:A→T	13	26	K64X	NA	No
DC-2101	Double-cortex syndrome	F	<i>DCX</i>	ChrX:110544972:G→A	15	30	Splicing	NA	No
DC-5103	Double-cortex syndrome	F	<i>LISI</i>	Chr7:2583480:G→C	16	32	R342P	Damaging	No
PAC-902	Pachygyria	M	<i>TUBB2B</i>	Chr6:3225184:G→C	23	46	R380P	Damaging	Yes [¶]
DC-2801	Double-cortex syndrome	F	<i>LISI</i>	Chr7:2579901:G→A	26	52	Splicing	NA	Yes
PH-16001	Periventricular nodular heterotopia	F	<i>FLNA</i>	ChrX:153587482:delG	35	70	S1449Pfs	NA	No

* Participants are listed in order of alternate-allele read frequency.

[†] Data are the proportion of blood cells predicted to carry the mutations, calculated as double the alternate-allele read frequency for heterozygous autosomal mutations in either sex or for X-linked mutations in females and as equal to the alternate-allele read frequency for genes that are X-linked in males.[‡] The software does not give predictions for nonsense, frameshift, or splicing mutations. NA denotes not applicable.[§] Parental DNA was not available.[¶] A different amino acid alteration at this position has been reported previously.

Table 2

Phenotypes and Distributions of Types of Pathogenic Mutations Detected.

Phenotype	Participants	All	Mutations	Germline	Mosaic	Mosaic Mutations Missed by Sanger Sequencing
	no.					no. of participants (%)
Double-cortex syndrome	30	9 (30)	3 (10)	6 (20)	5/6 (83)	
Polymicrogyria with megalencephaly	20	2 (10)	2 (10)	0	—	
Periventricular nodular heterotopia	61	8 (13)	7 (11)	1 (2)	0/1	
Pachygyria	47	8 (17)	7 (15)	1 (2)	0/1	
Total	158	27 (17)	19 (12)	8 (5)*	5/8 (63)	

* Mosaic mutations account for 30% of the 27 mutations detected in the study participants.