REVIEW ARTICLE
https://doi.org/10.1038/s41593-018-0257-3

nature neuroscience

Somatic mosaicism and neurodevelopmental disease

Alissa M. D'Gama[1,2,3] and Christopher A. Walsh[1,2,3]*

Traditionally, we have considered genetic mutations that cause neurodevelopmental diseases to be inherited or de novo germline mutations. Recently, we have come to appreciate the importance of de novo somatic mutations, which occur postzygotically and are thus present in only a subset of the cells of an affected individual. The advent of next-generation sequencing and single-cell sequencing technologies has shown that somatic mutations contribute to normal and abnormal human brain development. Somatic mutations are one important cause of neuronal migration and brain overgrowth disorders, as suggested by visible focal lesions. In addition, somatic mutations contribute to neurodevelopmental diseases without visible lesions, including epileptic encephalopathies, intellectual disability, and autism spectrum disorder, and may contribute to a broad range of neuropsychiatric diseases. Studying somatic mutations provides insight into the mechanisms underlying human brain development and neurodevelopmental diseases and has important implications for diagnosis and treatment.

From germline to somatic mutations

Historically, studies of inherited mutations associated with human disease used linkage analysis and candidate gene approaches to identify disease-causing mutations in families with monogenic disease and genome-wide association studies to identify common variants associated with complex disease. An inherited germline mutation is present in one or both parents and in all of the cells of an affected individual. Thus, the mutation can be identified by testing DNA from any tissue of the parent(s) carrying the mutation or the affected individual (Fig. 1a).

The last 15 to 20 years have increasingly highlighted the importance of de novo germline mutations in human disease. Although cytogenetics was traditionally used to identify microscopically visible de novo chromosomal abnormalities, the advent of high-resolution microarrays and next-generation sequencing (NGS) technologies allowed a clearer understanding of the full range of de novo mutations. De novo mutations are undetectable in either parent of an affected individual. Usually, a de novo germline mutation arises during gametogenesis in one of the parents and is present in all cells of the affected individual. Hence, the mutation can be identified from any tissue of the affected individual (Fig. 1b). The number of de novo germ cell mutations increases with age of mothers (~0.37 per year) and fathers (~1.51 per year)[1], and a father's age at conception is associated with risk of autism spectrum disorder (ASD) and schizophrenia (SCZ)[2]. Every individual's genome carries approximately one de novo germline mutation in the exome, the protein-coding region of the genome[3].

Given that the contribution of de novo mutations to a disease will be increased if there is a large mutational target (that is, dominant mutations in many genes are associated with the disease) and if the mutations have a large negative effect on survival and reproduction[4], it is not surprising that de novo mutations are associated with many neurodevelopmental diseases. There are many genes critical for brain development and neuronal function, and individuals with neurodevelopmental diseases often do not reproduce[5,6], placing the mutations under strong negative selection and necessitating continual de novo appearance of mutations for the diseases to remain present in the human population.

Recently, there has been growing evidence for de novo somatic mutations, which have been traditionally studied in cancer[7], in neurodevelopmental diseases (Table 1). A somatic mutation arises postzygotically during embryonic development or postnatal life and is present only in daughter cells of the originally mutated cell. In principle, 'somatic mosaicism' refers to mutations that arise in somatic cells and are not present in germ cells, 'gonosomal mosaicism' refers to mutations present in a subset of somatic and germ cells (Fig. 1c), and 'germline mosaicism' refers to mutations that arise in germ cells and are not present in somatic cells (Fig. 1d)[8]. In practice, it is difficult or impossible to tease apart the differences. In this review, we use 'germline mutation' to refer to a mutation that is present in all of the cells of a monozygotic individual and 'somatic mutation' to refer to a mutation that is present in only a subset of the cells of a monozygotic individual.

While germline mutations can be identified from any tissue of an affected individual, somatic mutations are more difficult to detect. At one extreme, if a somatic mutation occurs in the zygote or during the first mitotic divisions, such that all of the cells that contribute to the embryo carry the mutation, it can be identified from any tissue (Fig. 1e). If a somatic mutation occurs early in development, it may be present in a high percentage of cells across multiple tissues and identified relatively easily (Fig. 1f). If a somatic mutation occurs late in development, it may be present in a percentage of cells in only one tissue and identified only if that tissue is tested, which is generally not feasible for brain tissue (Fig. 1g). At the other extreme, if a somatic mutation occurs in a postmitotic cell like a neuron, it can be identified only if that cell is tested (Fig. 1h). Box 1 discusses current methods to detect and validate somatic mutations in the human brain.

How do somatic mutations lead to neurodevelopmental disease?

The human brain is vulnerable to somatic mutation, especially during neurogenesis when approximately 10^5 neurons per min are generated

[1]Division of Genetics and Genomics, Manton Center for Orphan Disease, and Howard Hughes Medical Institute, Boston Children's Hospital, Boston, MA, USA. [2]Departments of Neurology and Pediatrics, Harvard Medical School, Boston, MA, USA. [3]Broad Institute of MIT and Harvard, Cambridge, MA, USA. *e-mail: christopher.walsh@childrens.harvard.edu

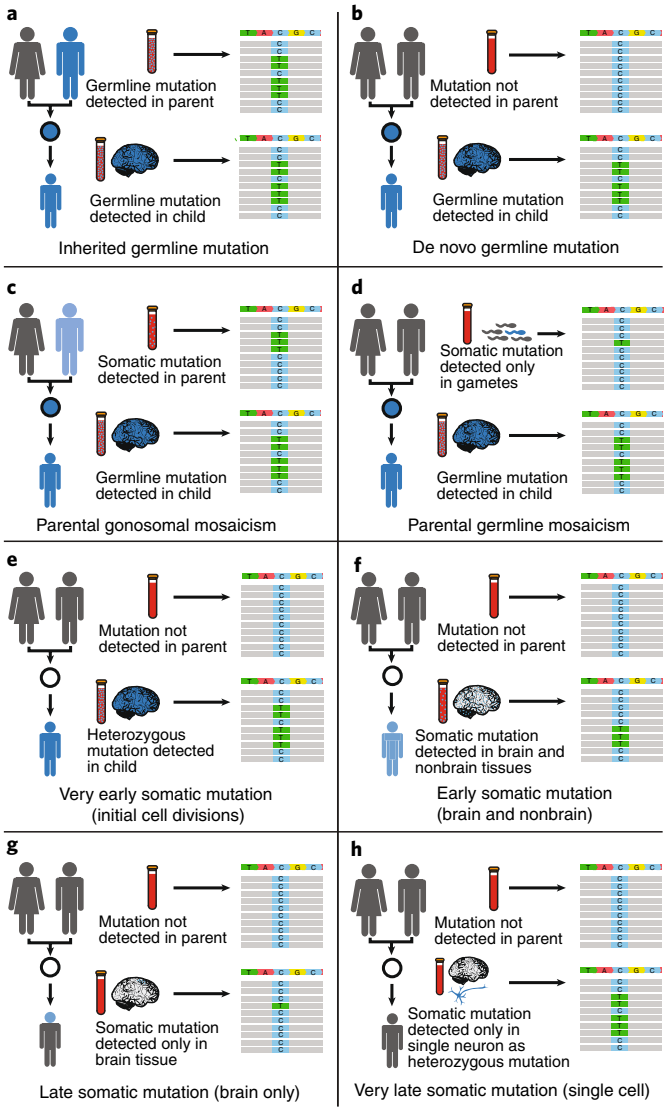

Fig. 1 | Types of mutations and their detection using NGS. a, An inherited germline mutation is detectable in all tissues of the parent and child at AAF = 50%. **b**, A de novo germline mutation is not detectable in the parent and is detectable in all tissues of the child at AAF = 50%. **c**, Parental gonosomal mosaicism is detectable in some tissues of the parent at AAF <50%, and in all tissues of the child at AAF = 50%. **d**, Parental germline mosaicism is detectable in gametes of the parent (in this case, in sperm at AAF < 50%), and in all tissues of the child at AAF = 50%. **e**, A very early somatic mutation, which occurs in the first cell divisions such that all cells that contribute to the embryo carry the mutation, is not detectable in the parent and is detectable in all tissues of the child at AAF = 50%. **f**, An early somatic mutation, which is present in brain and nonbrain tissues, is not detectable in the parent and is detectable in some tissues of the child at AAF <50%. **g**, A late somatic mutation, which is present only in the brain, is not detectable in the parent and is detectable only in brain tissue of the child at AAF < 50%. **h**, A very late somatic mutation, which occurs in a postmitotic neuron, is not detectable in the parent and is detectable only in that neuron of the child, which requires single-cell sequencing. In all panels, blue represents the mutation and darker shades indicate increasing degree of mosaicism.

from an initial population of progenitors[9,10]. The somatic mutation rate is high during neurogenesis (~5.1 single-nucleotide variants (SNVs) per day per progenitor)[11], and somatic mutations that occur during this period may lead to neurodevelopmental diseases. Ultimately, the human brain contains about 86 billion neurons[9], and each neuron continues to accumulate somatic mutations approximately linearly with age (~23 SNVs per year per neuron)[12], which may contribute to neurodegenerative diseases. In addition to causing neurological disease, it has been hypothesized that somatic mutations in the human brain contribute to generating neuronal diversity[13]. Box 2 discusses somatic mosaicism in the normal human brain.

To understand how somatic mutations lead to neurodevelopmental disease, we must consider the developmental time period and progenitor cell in which the somatic mutation occurs, as well as the effect of the mutation on the original cell: if it is neutral or so damaging that the cell is selected against, the mutation will not cause disease. Thus, we assume a somewhat damaging mutation that allows the cell to survive. If the mutation occurs in an irrelevant lineage or very late in a relevant lineage, it is unlikely to cause disease; for example, a somatic mutation in a neuronal gene that occurs in a hematopoietic stem cell or in a postmitotic neuron. If the mutation occurs in a relevant lineage and nonmalignantly causes proliferation (or malignantly transforms the cell), it can cause disease; as discussed below, somatic mutations activating a growth pathway in neural progenitor cells lead to abnormal brain overgrowth[14]. If the mutation does not have a proliferative effect, it can still cause disease if it occurs early in a relevant lineage and disrupts a gene critical for that tissue's development and/or function; as discussed below, somatic mutations in many genes important for brain development and neuronal function have recently been associated with ASD[15–18].

We must also consider the different classes of somatic mutation that occur in the human brain, which range from SNVs, insertion–deletions, and microsatellite instabilities to copy-number variants (CNVs), large structural variants, chromosomal aneuploidies, and mobile element insertions. Somatic SNVs and insertion–deletions are relatively frequent (~1,500 somatic SNVs per neuron[19]) and have been identified in many neurodevelopmental diseases; as discussed below, somatic SNVs that disrupt the function of ion channels lead to epileptic encephalopathies[20–24]. Somatic expansion of trinucleotide repeats is surprisingly common in certain neurological diseases; for example, somatic instability of the Fragile X trinucleotide repeat has been demonstrated in human brain[25]. Although somatic structural variants are rarer (<0.5 somatic CNV per neuron[26,27]), somatic chromosomal aneuploidies were some of the first somatic variants associated with human disease, as they were microscopically visible. Trisomies of chromosomes 13, 18, and 21 are the only autosomal trisomies compatible with life in the germline state; however, trisomies of many chromosomes have been identified in the somatic state[28–34]. Recently, somatic mobile element insertions, specifically of long interspersed nuclear element-1 (L1), have been associated with neurodevelopmental diseases. Rett syndrome, which is caused by mutations in *MECP2*, was the first disease implicated. MECP2 acts as a transcriptional repressor of the LINE1 promoter in neural progenitors, and somatic L1 insertions are increased in the brains of Rett syndrome patients carrying *MECP2* mutations[35].

Finally, we must consider the different genetic mechanisms by which a somatic mutation can lead to a neurodevelopmental disease phenotype. A first mechanism involves obligatory somatic mutation diseases, which are caused by mutations seen only in the somatic state, as the germline state is incompatible with life. For example, Proteus syndrome, characterized by multiple tissue overgrowths, is caused by an activating somatic point mutation in *AKT1*[36]; congenital lipomatous overgrowth with vascular, epidermal, and skeletal anomalies syndrome is caused by activating somatic mutations in the same pathway, specifically point mutations in *PIK3CA*[37]; and Sturge–Weber syndrome, characterized by port-wine stain, leptomeningeal angiomas, and seizures, is caused by an activating somatic point mutation in *GNAQ*[38]. As discussed below, hemimegalencephaly (HME), in which a cerebral hemisphere is abnormally enlarged, has recently been discovered to be an obligatory somatic mutation disease;

Table 1 | Neurodevelopmental diseases associated with somatic mutations

Disease	Gene(s)	Mechanism	References
Autism spectrum disorder	Many genes	Not resolved, likely germline and milder somatic mutations, potential two-hit model	15–18,117
Double-cortex syndrome	*DCX*, *LIS1*	Germline and milder somatic mutations	49–51
Epileptic encephalopathy (including Dravet and Ohtahara syndromes)	*CDKL5*, *GABRA1*, *GABRG2*, *GRIN2B*, *KCNQ2*, *PCHD19*, *SCN1A*, *SCN2A*, *SCN8A*	Germline and milder somatic mutations	20–22,24,73–77
Focal cortical dysplasia	*MTOR*, *PIK3CA*;	Obligatory somatic mutations	58,60
	DEPDC5, *NPRL2*, *NPRL3*, *TSC1*, *TSC2*	Two-hit model: inherited germline and demonstrated or inferred second-hit somatic mutations	63–67
Fragile X syndrome	*FMR1*	Germline and milder somatic mutations	102,140
Hemimegalencephaly	*AKT3*, *MTOR*, *PIK3CA*;	Obligatory somatic mutations	54,59
	DEPDC5, *PTEN*, *TSC2*	Two-hit model: inherited germline and demonstrated or inferred second hit somatic mutations	14,58,62
Intellectual disability	Many genes	Germline and milder somatic mutations	122,123
Neurofibromatosis type 1	*NF1*	Two-hit model: inherited germline and second hit somatic mutations	41,42,140
Neurofibromatosis type 2	*NF2*	Germline and milder somatic mutations	141,142
Periventricular nodular heterotopia	*FLN1*	Germline and milder somatic mutations	47,48
Rett syndrome	*MECP2*	Germline and milder somatic mutations in females, obligatory somatic mutations in males, association with increased L1 retrotransposition	35,100,101,140
Schizophrenia	Many genes	Not resolved, associated with increased L1 retrotransposition	130,131
Tuberous sclerosis	*TSC1*, *TSC2*	Two-hit model: inherited germline and second hit somatic mutations	43–45

in fact, de novo germline mutations in some of the same alleles that cause HME in the somatic state have been rarely reported and lead to death in the first years of life[14,39].

In contrast to the 'single-hit' mechanism causing obligatory somatic mutation diseases, a second mechanism involves the 'two-hit' model of Knudson, in which an individual inherits a mutation in one allele of a recessive oncogene, and a somatic mutation of the second allele leads to disease manifestations, such as tissue overgrowth or cancer[40]. In the neurocutaneous disease neurofibromatosis type 1 (NF1), which is characterized by café-au-lait spots, optic gliomas, and peripheral nervous system tumors (neurofibromas), affected individuals carry a germline mutation in one allele of *NF1*, and somatic mutations of the second allele lead to neurofibromas[41,42]. Similarly, in tuberous sclerosis complex (TSC), which is characterized by facial and skin lesions, seizures, intellectual disability, renal and cardiac tumors, and cortical tubers, affected individuals carry a germline mutation in *TSC1* or *TSC2*, and somatic mutations of the second allele have been identified in non-nervous system tumors and cortical tubers[43–45].

A third mechanism involves diseases that can be caused by both germline and somatic mutations, with the somatic mutations often manifesting as milder phenotypes than the germline mutations. In periventricular nodular heterotopia (PVNH), newborn neurons do not migrate to their proper positions in the cerebral cortex but instead accumulate next to the ventricles. Germline mutations in X-linked *FLNA* cause PVNH in females and are usually lethal in males[46,47]. However, somatic mutations in *FLNA* cause milder disease in both females and males[47,48].

Neuronal migration disorders

Somatic mutations have been known to cause neuronal migration disorders like PVNH for almost two decades, because magnetic resonance imaging (MRI) of these disorders demonstrated incomplete disruption of brain development, suggesting a causative mechanism that would lead to focal effects. Lissencephaly, which is characterized by a smooth brain on MRI, and subcortical band heterotopia (also known as double-cortex syndrome), which is characterized by a 'double cortex' on MRI, highlight the relationship between alternate allele frequency (AAF) and phenotype severity (Fig. 2). Lissencephaly is caused by germline mutations in *LIS1* in males and females and by germline mutations in X-linked *DCX* in males[49,50]. Double-cortex syndrome is caused by germline mutations in *DCX* in females (due to random X-chromosome inactivation) and by somatic mutations in *DCX* and *LIS1* in males and females[49,50]. As the AAF of the causative mutation increases, from as low as 5% in a female with a somatic *DCX* mutation to 50% in females with germline *DCX* mutations and up to 100% in males with germline *DCX* mutations, there is a corresponding increase in the severity seen on MRI, from mild to severe double-cortex syndrome to complete lissencephaly[51].

Somatic mutations were initially suspected to be associated with neuronal migration disorders in 5–10% of patients and recently in up to 30% of patients using NGS. Targeted deep sequencing (≥200×) on leukocyte-derived DNA from 158 patients with brain malformations identified germline mutations in 19 patients and somatic mutations in 8 patients[51]. Most of the somatic mutations were undetectable by Sanger sequencing, which has a minimum detection threshold of ~15% AAF (below this, somatic mutations are difficult to distinguish from background) and an upper threshold of ~30–35% AAF (above this, somatic mutations are difficult to distinguish from germline mutations)[51,52]. Of note, these studies used DNA extracted from clinically accessible tissues. Thus, for some neurodevelopmental diseases associated with somatic mutations, the mutations occur relatively early in embryonic development, before gastrulation, and are present in brain and nonbrain tissues.

Box 1 | Methods to detect and validate somatic mutations in human brain

Given that somatic mutations are present in only a subset of cells, their identification requires sequencing of DNA extracted from bulk tissue or from sorted pooled cells at high depth of coverage[14,51,60] or sequencing of DNA extracted from single cells at relatively shallow coverage[12,19]. For bulk human brain tissue and sorted pooled cells (for example, NeuN$^+$ cells), our ability to detect a somatic mutation depends on the percentage of cells carrying the mutation and the depth of coverage. For example, mutations in the mTOR pathway causing FCD that are present in 2% of cells are difficult to detect at standard depths of coverage (30–100×), but possible to detect at high depths of coverage (300–1,000×). Currently, coverage depth needs to be balanced with the size of the target region (i.e., targeted sequencing, WES, or WGS) due to the cost of sequencing. Deep sequencing of bulk human brain tissue has been successfully used to identify many somatic mutations causing neurodevelopmental diseases, which is not surprising if disease-causing somatic mutations need to be present in a threshold percentage of cells to disrupt neuronal development and function[14,51,60]. Deep sequencing of sorted pooled cells will likely identify additional disease-causing somatic mutations limited to specific cell lineages. For single neurons and non-neuronal cells, whole genome amplification or clonal expansion is necessary to generate sufficient DNA for sequencing[144,145]. These approaches have been successfully used to analyze private and clonal somatic mutations in the developing and aging human brain[11,12,19,26,27,68]. Finally, biological validation from unamplified source material and/or technical validation from amplified source material (for single cells) are critical to discriminate true somatic mutations from false positives[132]. Validation initially involved relatively low-throughput methods, including subcloning followed by Sanger colony sequencing and mass spectrometry[51,59], and now involves high-throughput methods like digital droplet PCR and high-coverage resequencing[143,146]. Table 3 provides further details on the advantages and challenges of current methods.

Box 2 | Somatic mosaicism in the normal human brain

Recent studies have demonstrated that somatic mutations occur in the normal human brain from development through aging, and somatic mutations are hypothesized to contribute to the generation of neuronal diversity. One of the most exciting findings has been the demonstration of somatic L1 insertions in human brain, as L1 repeats comprise ~25% of our genome[35] and were initially considered 'junk DNA'. Several studies using single-cell approaches have demonstrated that L1 retrotransposition occurs in human neurons at ~≤1 somatic insertion per neuron, although the rate remains controversial[68,147–149]. The impact of L1 insertions on normal and abnormal human brain development (as discussed in the main text, L1 insertions have already been associated with Rett syndrome and SCZ) is an important area for future studies. In addition, single-cell WGS studies have shown that somatic aneuploidy is rare in human neurons, while somatic CNVs are relatively more common and include both private and clonal CNVs, with one study reporting that 13–41% of human cortical neurons carry a large CNV[26,27,150]. In terms of somatic SNVs, our initial study using single-cell sorting, multiple displacement amplification, and WGS estimated ~800–2,000 somatic SNVs per neuron in the adult human brain[19], and we recently showed that each neuron accumulates somatic mutations approximately linearly with age[12]. Another recent study using WGS of clonal-cell populations to analyze single-cell genomes from fetal human brain demonstrated higher rates of somatic SNVs during early embryogenesis and especially neurogenesis (~5.1 SNVs per day per progenitor) compared to adulthood, estimating 200–400 somatic SNVs per progenitor cell at 15–21 weeks gestation[11]. Overall, a neuron has ~300–900 somatic SNVs at birth, and then continues to accumulate somatic SNVs at a rate of ~23 SNVs per year per neuron in the prefrontal cortex and at ~40 SNVs per year per neuron in the dentate gyrus of the hippocampus[11,12]. The number of somatic SNVs in aged human brain is thus much higher than that in infant human brain, and the number of somatic SNVs in neurodegenerative disease brain is higher than that in normal human brain[12]. Future studies need to further investigate the association between somatic mutation, aging, and neurodegenerative disease.

Focal cortical dysplasia and HME

Focal cortical dysplasia (FCD), which is characterized by a small region of abnormal cortex[53], and HME, which is characterized by abnormal enlargement of a cerebral hemisphere[54], have long been hypothesized to be due to somatic mutations[55] given the focal, unilateral lesions seen on MRI. FCD and HME are important causes of intractable pediatric epilepsy, and patients often require surgical resection of abnormal brain tissue for attempted seizure control, providing access to brain tissue for research[56,57]. Direct study of resected abnormal brain tissue has revealed somatic activating mutations in *MTOR* and in genes encoding positive regulators of the mTOR pathway (*AKT3* and *PIK3CA*) in FCD, HME, and related megalencephaly syndromes[54,58–61], as well as loss-of-function (LOF) mutations in multiple genes encoding negative regulators of the mTOR pathway (*DEPDC5*, *NPRL2*, *NPRL3*, *PTEN*, *TSC1*, and *TSC2*)[14,62–67]. Knudson's two-hit model is hypothesized to apply to the latter category, and several cases with both germline and somatic mutations in negative regulators have been described, including *TSC2* mutations in isolated HME[14,64].

Overall, the AAF for somatic mutations associated with FCD ranges from ~1% to 10% within the lesion (with the mutation generally absent outside the lesion) and from ~3% to 30% in HME[14]. Interestingly, the somatic mutations causing these diseases appear to lead to a proliferative advantage, as studies of functionally silent somatic mutations in normal human brain and nonproliferative disease-causing somatic mutations in abnormal human brain have demonstrated that mutations at ≥5% AAF are generally detectable in brain and nonbrain tissues[19,51]. However, in essentially all FCD and HME cases in which brain and nonbrain tissue samples are available, the disease-causing somatic mutations are detected only in brain tissue[14]. Thus, for some neurodevelopmental diseases associated with somatic mutations, the mutations occur relatively late in embryonic development and are detectable only if brain tissue is tested, which has implications for clinical diagnosis in cases where surgical resection is not required.

Using NGS, single-cell sequencing, and mouse models, recent studies have demonstrated that FCD and HME lie on a disease continuum, with the resulting phenotype dependent on the developmental time and progenitor cell in which the disease-causing somatic mutation occurs (Fig. 3). NGS studies have shown that the average AAF associated with FCD is substantially lower than that associated with HME, but there is overlap[14]. Single-cell sequencing of neuronal and non-neuronal cells isolated from abnormal brain tissue of FCD and HME patients with disease-causing somatic mutations demonstrated that the mutations are always present in neurons but variably present in glia, suggesting that abnormal activation of the mTOR pathway in the neuronal lineage is necessary for both FCD and HME pathogenesis[14,26,68]. Moreover, in the two FCD patients with the lowest AAFs in bulk sequencing, somatic

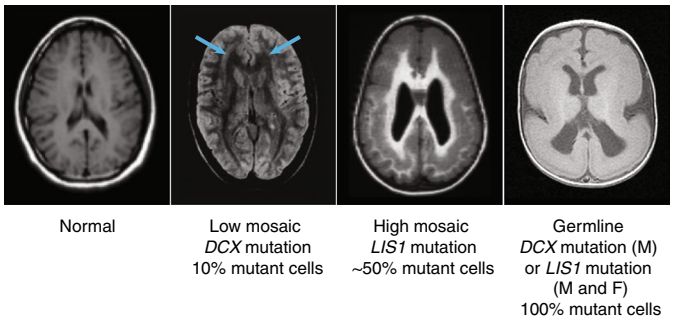

Fig. 2 | Somatic mutations in neuronal migration disorders. MRIs of (from left to right) a control individual with a normal brain, an affected individual with mild double-cortex syndrome and a somatic mutation in *DCX* present in 10% of cells, an affected individual with severe double-cortex syndrome and a somatic mutation in *LIS1* present in ~50% of cells, and an affected individual with lissencephaly, which results from a germline mutation in *DCX* in males or in *LIS1* in males and females. Blue arrows point to the small anterior subcortical band heterotopia. Note that the severity of the phenotype increases as the percentage of cells carrying the mutation increases[51]. M, male; F, female.

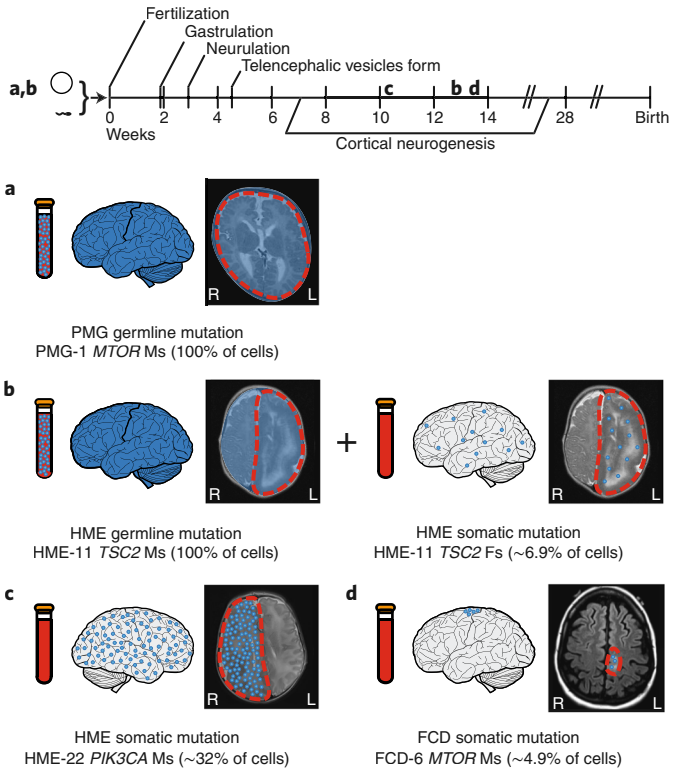

Fig. 3 | FCD and HME represent a continuum, with lesion differences reflecting the time and place of origin of the mutation. a, Germline mutations occur before fertilization and are detectable in the brain and a clinically accessible blood sample. Germline activating mutations in the mTOR pathway can lead to megalencephaly, as seen in case PMG-1 with a de novo germline *MTOR* mutation. **b**, Two-hit germline and somatic LOF mutations in negative regulators of the mTOR pathway can lead to FCD and HME. In some cases, such as HME-11 with two *TSC2* mutations, both a germline and a somatic mutation leading to HME have been identified. The germline mutation was detectable in brain and blood, whereas the somatic mutation occurred later during embryonic development and was detectable only in brain. **c**, Somatic activating mutations in positive regulators of the mTOR pathway can also lead to these diseases. Mutations present at a higher AAF, suggesting that they arose earlier during cortical neurogenesis, appear more likely to lead to HME; for example, a somatic activating point mutation in *PIK3CA* was identified in ~32% of cells in the abnormal hemisphere of case HME-22. **d**, Mutations present at a lower AAF, suggesting that they arose later during cortical neurogenesis, appear more likely to lead to FCD; for example, a somatic activating point mutation in *MTOR* was identified in ~4.9% of the cells in the abnormal cortical tissue of case FCD-6. PMG, polymicrogyria; Ms, missense mutation; Fs, frameshift mutation. Figure and legend adapted with permission from ref. [14], Cell Press.

mutations were limited to neurons, suggesting that in some cases the mutations occur after segregation of neuronal and glial lineages. Mouse models of FCD and HME further suggested that somatic mutations activating the mTOR pathway in dorsal cerebral cortical progenitors cause both FCD and HME and that activation in the excitatory neuron lineage is necessary and in some cases sufficient for disease pathogenesis. In utero electroporation of mutant *MTOR* or *AKT3* leads to cytomegalic neurons, abnormal neuronal migration, and seizures[60,69]. mTOR inhibition with rapamycin suppressed seizures in the *MTOR* model[60] and rescued migration when administered prenatally (but not postnatally) in the *AKT3* model[69]. Expression of mutant *PIK3CA* in different cell lineages using Cre-recombination leads to varied phenotypes. Prenatal expression in neural progenitors, including expression limited to the dorsal telencephalic lineage, leads to megalencephaly, cytomegalic neurons, abnormal cortical lamination, and seizures (not all aspects were analyzed in each model)[14,70], and PI3K-signaling inhibition suppressed seizures in one model[70]. Prenatal expression limited to the interneuron lineage leads only to a subtle decrease in cortical interneuron number[14]. Most recently, neuron-specific *DEPDC5* knockout using Cre-recombination was shown to lead to megalencephaly, cytomegalic neurons, and seizures[71].

Nonlesional epileptic encephalopathies

In addition to causing neurodevelopmental diseases with visible lesions, somatic mutations also contribute to diseases without visible lesions, but less is known about these diseases because such mutations are not independently manifest. Epileptic encephalopathies are associated with inherited and de novo germline mutations, and they have also been associated with somatic mosaicism in multiple genes (*CDKL5*, *GABRA1*, *GABRG2*, *GRIN2B*, *KCNQ2*, *PCHD19*, *SCN1A* (in discordant monozygotic twins), and *SCN2A*) in affected children and with germline or gonosomal mosaicism in multiple genes (*KCNQ2*, *SCN1A*, *SCN2A*, *SCN8A*) in asymptomatic or less-severely affected parents of affected children[20–24,72–77]. Epilepsy associated with mutations in X-linked *PCHD19* is an interesting example, because only females with germline mutations and males with somatic mutations are affected, while hemizygous males are generally unaffected[78]. Thus, the abnormal phenotype requires that the gene be disabled in some, but not all, neurons, recently suggested as being due to abnormal cell sorting when normal and abnormal neurons are present[79].

SCN1A mutations, which are associated with Dravet syndrome, and *SCN2A* mutations, which are associated with Ohtahara syndrome, have been identified in the mosaic state in both parents and affected children[20–24]. In one study, 8.6% of presumed de novo germline *SCN1A* mutations in 174 children with Dravet syndrome were discovered to be mosaic in a parent, and mosaic parents without epileptic symptoms had significantly lower AAFs than those with symptoms[23]. In another study, monozygotic twins discordant for Dravet syndrome were found to be discordant for an *SCN1A* mutation, suggesting that the mutation occurred very early, at the premorula stage[22].

Although the somatic mutations in the cases described above occurred relatively early and were detected in clinically accessible

Table 2 | Summary of four ASD studies analyzing WES data for mosaic mutations

Study	Number of families	Number of mosaic mutations	Validation method	Validation rate (on subset)	Burden in cases vs. controls	Percent of de novo mutations	Percent contribution to simplex ASD risk
Freed et al.[16]	2,388	221	SSV, phasing, pyroseq	87% (phasing), 82% (pyroseq)	All classes of mosaic mutations	5.4	5.1
Dou et al.[15]	2,361	1,248	ultra-deep amplicon reseq, phasing and WGS	53% (reseq)	Missense and LOF mosaic mutations with AAF> 20%	-	3.4
Lim et al.[18]	5,947	468	subcloning and SCS, pyroseq, targeted PCR and MiSeq reseq	84.8–97%	damaging non-synonymous mosaic mutations in critical exons of prenatally expressed genes	7.5	-
Krupp et al.[17]	2,264	470	smMIPs	54% (at least 2 callers)	synonymous mosaic mutations	22	3–4

Pyroseq, pyrosequencing; reseq, resequencing; SCS, Sanger colony sequencing; smMIPs, single-molecule molecular-inversion probes, MiSeq, Illumina MiSeq sequencing instrument; WGS, whole-genome sequencing.

tissues, the hypothesis that brain-only somatic mutations lead to some cases of nonlesional epilepsy has not been adequately tested. Even with access to brain tissue, the nonlesional nature of many epilepsy mutations makes them hard to detect, since one does not know which brain region to sample. If such somatic mutations occurred late enough to be restricted to specific cell lineages, isolating those lineages might be required to detect the mutations. For example, mTOR pathway activation in the interneuron lineage does not lead to a visibly abnormal brain, but appears to disrupt interneuron migration to the cortex and reduce the seizure threshold[14,80]. Thus, a somatic mutation in a progenitor cell of the ganglionic eminences could disrupt neuronal circuitry and contribute to epileptogenesis, but might be difficult to detect even in bulk brain DNA.

Autism spectrum disorder

ASD is a common neurodevelopmental disorder characterized by deficits in social interaction and communication, as well as restrictive and repetitive behaviors[81–83]. ASD shows high heritability (70–90%[84,85]), but is highly genetically heterogeneous. Germline de novo CNVs and SNVs seem to be a significant cause of risk, while autosomal recessive, X-linked, and common variants all contribute to ASD risk as well[3,86–99].

Although several syndromic diseases that are sometimes associated with autistic symptoms—including TSC, NF1, Fragile X, and Rett syndrome[100–105]—can be caused by somatic mutations, until recently, there were only case reports of somatic mutations in idiopathic ASD and one larger study suggesting an association between low-level mosaic aneuploidy and ASD using fluorescence in-situ hybridization[106–113]. Several initial studies using whole-exome sequencing (WES) to identify de novo mutations in large ASD cohorts noted variants consistent with somatic mosaicism in an affected child or gonosomal mosaicism in a parent[88,92,94]. In one study of 209 simplex families, 3.5% of reported de novo mutations appeared somatic[94]. However, the initial WES calling algorithms were not designed to detect somatic mutations, and the relatively low coverage of the exome presents challenges to detect them.

In the past year, several groups have reanalyzed WES data from simplex ASD families, mainly from the Simons Simplex Collection (SSC), to detect somatic mutations, and they demonstrated that somatic mutations in affected children contribute to ASD risk in up to 3–5% of all simplex cases (Table 2). Freed and Pevsner analyzed WES data from 2,388 SSC families, finding that mosaic mutations account for 5.4% of de novo mutations and estimating that 33% of mosaic mutations contribute to 5.1% of simplex ASD diagnoses[16]. In their callset, all classes of mosaic mutations were substantially enriched in children with ASD compared to unaffected siblings. Dou et al. analyzed WES data from 2,361 SSC families for mosaicism in both children with ASD and parents, estimating that 65.8% of missense or LOF postzygotic mutations with high AAFs (>20%) in children with ASD and 53.4% of parental missense or LOF postzygotic mutations with low AAFs transmitted to affected children contribute to 3.4% and 2.6% of simplex ASD diagnoses, respectively[15]. These mutations tend to be enriched in LOF-constrained exons (i.e., exons in which there are fewer LOF mutations in public databases like ExAC (Exome Aggregation Consortium, http://exac.broadinstitute.org) than expected based on mutation rates), and the new ASD risk genes identified tend to have higher expression in the cerebellum. Lim et al. analyzed 5,947 families from the SSC and the Autism Sequencing Consortium, finding that postzygotic mutations account for 7.5% of de novo mutations[18]. Damaging nonsynonymous postzygotic mutations in critical exons of prenatally expressed genes were enriched in people with ASD compared to controls, and genes with these mutations were enriched for expression in the amygdala. Intriguingly, people with ASD with mosaic mutations tended to have higher IQs than those with germline mutations. Finally, Krupp et al. analyzed 2,264 SSC families for mosaicism in both children with ASD and their parents, finding that mosaic mutations account for 22% of de novo SNVs and that transmitted parental mosaic mutations account for 6.8% of presumed de novo mutations in children[17]. In their callset, synonymous mosaic mutations were substantially enriched in ASD cases compared to controls, and enriched for impacts on splicing. They estimate that mosaic mutations contribute to 3–4% of ASD risk in simplex families, with 2% due to synonymous mosaic mutations. Although these studies used overlapping WES datasets, their final callsets for somatic mutations were quite different, suggesting that further analysis—and ideally a consensus on calling, filtering, and validation pipelines for somatic mutations—is needed to resolve the differences. The studies also used relatively low-coverage (60×) WES data, which is extremely insensitive for the detection of somatic mutations; higher-coverage data would likely reveal a larger number of somatic mutations with lower AAFs. Taken together, the studies demonstrate that somatic mutations clearly contribute to simplex ASD risk and occur in both previously implicated ASD risk genes, such as *SCN2A*[114], as well as in new candidate genes. In the future, a model in which different combinations of germline and somatic mutations contribute to ASD risk may help explain the incomplete penetrance and variable expressivity observed in ASD and other neuropsychiatric disorders.

Table 3 | Comparison of methods to study somatic mutations in human brain

	Advantages	Challenges
DNA source		
Single cells	Detection: highest power (able to detect somatic mutations at single-cell resolution and clonal mutations at AAF > 0.1%[19]).Validation: technical validation is relatively simple.Application: somatic mutations can be used as lineage markers to study normal and abnormal brain development.	Detection: requires isolation of single cells, which is relatively complex, and (i) whole-genome amplification or (ii) clonal expansion via somatic cell nuclear transfer (relatively inefficient) to generate sufficient DNA for sequencing. Sequencing errors arising during these processes (false positives) must be discriminated from somatic mutations (true positives).Validation: private somatic mutations (only detected in one cell) can only be technically validated. (If clonal, the somatic mutation may be biologically validated in unamplified bulk tissue or pooled cells.)Application: private somatic mutations are unlikely to be pathogenic.
Pooled cells	Detection: medium power (able to detect somatic mutations restricted to specific cell lineages that would be present at relatively low AAFs, and thus more difficult to detect, in bulk tissue). Does not necessarily require WGA.Validation: biological validation is possible.Application: clonal somatic mutations are more likely to be pathogenic. Distribution provides information on critical cell types for disease pathogenesis.	Detection: requires isolation of pooled cell populations, which is relatively complex. Fluorescence-activated nuclear sorting requires extensive antibody optimization to ensure purity of pooled cell populations (currently limited by number of validated antibodies).Validation: biological validation is relatively complex.
Bulk tissue	Detection: DNA extraction from bulk tissue is relatively simple. Does not require WGA.Validation: biological validation is possible.Application: clonal somatic mutations are more likely to be pathogenic. Pathogenic somatic mutations may be confirmed clinically (usually using targeted deep sequencing in a CLIA lab).	Detection: lowest power (dependent on depth of sequencing; becomes cost-prohibitive to accurately detect low AAFs <1%[14]).Validation: biological validation is relatively complex.Application: no information on distribution of somatic mutations.
Detection methods		
Sanger sequencing	Relatively simple to perform and analyze.	Lower detection threshold of ~15% AAF (somatic mutations become hard to discriminate from background) and upper detection threshold of ~30–35% AAF (somatic mutations become hard to discriminate from germline mutations)[51]. Does not provide quantitative AAF.
Targeted NGS (including WES)	Higher depth of coverage (and thus more power to detect AAF > 1%[14]) compared to WGS for given cost. Provides quantitative AAF.	Relatively complex to perform and analyze, especially for unmatched samples. No information on genomic loci not in target region. Unable to accurately detect structural variants.
WGS	Able to accurately detect range of somatic variation (from SNVs to structural variants) across all genomic loci. Provides quantitative AAF.	Relatively complex to perform and analyze, especially for unmatched samples. Lower depth of coverage (and thus less power to detect low AAFs) compared to targeted NGS for given cost.
Validation methods		
Targeted NGS	Independent biological validation. Sensitivity depends on depth of resequencing, >0.1%[19]. Provides quantitative AAF.	Relatively complex to perform.
Digital droplet PCR	Independent biological validation. Sensitivity to AAF ≥ 0.001%[143]. Provides quantitative AAF.	Relatively complex to perform.
Sanger sequencing	Relatively simple to perform.	Private somatic mutations only able to be validated in DNA from WGA or clonal expansion (technical validation).

In terms of DNA source: if the source is a single cell, DNA is extracted from a single cell sorted from a small piece of brain tissue; if the source is pooled cells, DNA is extracted from a specific cell population (for example, NeuN$^+$ cells) sorted from a small piece of brain tissue and pooled before extraction; and if the source is bulk tissue, DNA is extracted directly from a small piece of brain tissue, which thus contains DNA from all of the cell types within that piece of brain tissue. A private somatic mutation is a mutation that is uniquely identified in a single cell. In terms of complexity, 'simple to perform' refers to a method that requires minimal time (<1 week), expertise, and/or cost (<$100), such as a Sanger sequencing validation experiment, while 'complex to perform' refers to a method that require significant time (>1 week), expertise, and/or cost (>$100), such as an NGS sequencing experiment.

Recent WES studies have used DNA extracted mainly from blood to identify somatic mutations that occurred relatively early in development, but there is also interest in testing DNA extracted from postmortem brain tissue to identify somatic mutations that occurred later[115]. An initial study used targeted deep sequencing of DNA extracted from 55 postmortem ASD brains to detect SNVs in 78 known ASD candidate genes, identifying deleterious somatic mutations in two people with ASD and one with Fragile X premutation[25]. Histopathological studies of postmortem ASD brains have also demonstrated 'patches' of disorganization in the prefrontal and temporal cortex of some people with ASD[116,117], which have been proposed to represent visible lesions due to somatic mutations. Although further whole-genome sequencing studies are needed to quantify the rates and characteristics of somatic mutations in ASD brain, such mutations have the potential to provide insight into critical cell types or brain regions for ASD pathogenesis[118].

Intellectual disability

Syndromic intellectual disability (ID), such as Cornelia de Lange syndrome, which is characterized by ID, growth retardation, facial dysmorphism, and limb reduction defects, has been associated with somatic mutations in case reports[119,120]. Similarly to

ASD, severe ID, which occurs in 0.5% of newborns, is genetically heterogeneous, including contributions from de novo CNVs and SNVs, autosomal recessive variants, and X-linked variants[121]. In a study using whole-genome sequencing of 50 trios to identify de novo mutations in severe ID, three de novo SNVs in candidate ID genes were validated as mosaic[122]. Recently, additional analysis of this dataset revealed that 6.5% of a subset of presumed de novo germline mutations actually validated as mosaic mutations[123]. An additional four presumed de novo germline mutations were inherited from a mosaic parent. More broadly, mosaic structural abnormalities were recently detected in ~1% of 1,303 children with developmental disorders[124]. Thus, as with ASD, careful analysis of NGS data is revealing that somatic mutations represent an under-recognized source of de novo variation in ID, but future NGS studies with higher coverage are needed to fully explore somatic mutations in ID.

Schizophrenia

As we uncover a role for somatic mutations in ASD and ID, it is of interest to investigate whether such mutations play a role in other neuropsychiatric diseases like SCZ, which also has high heritability (~80%[125]). Early studies using fluorescence in-situ hybridization suggested an increase in somatic aneuploidy of chromosomes 1, 18, and X in SCZ compared to control brains[126,127], and a recent study using comparative genomic hybridization arrays revealed somatic CNVs across multiple chromosomes—including regions containing SCZ candidate genes—in SCZ compared to control brains[128]. In contrast, a larger study analyzing DNA extracted from blood using SNP arrays noted no difference in mosaic CNVs in SCZ cases compared to controls[129]. Recently, two studies implicated somatic L1 retrotransposition in SCZ. The first study used quantitative real-time PCR to show substantially increased L1 content in bulk prefrontal cortex and neuronal cells of SCZ cases compared to controls[130]. The second, larger study used L1 amplification and sequencing to similarly note a substantial increase in intragenic novel L1 insertions in the dorsolateral prefrontal cortex of SCZ cases compared to controls[131]. Additional studies are needed to investigate somatic mutations in SCZ and other neuropsychiatric diseases, which is a major focus of the Brain Somatic Mosaicism Network recently formed by the NIMH[132].

Clinical implications

Somatic mutations are not routinely tested for clinically and are technically difficult to detect. Currently, standard WES is becoming clinically integrated and successfully diagnosing Mendelian disorders[133]. Deep NGS (at least 200×, ideally 500×) in an affected child and both parents may more optimally detect somatic mutations[51], limited by the tissue(s) available[14]. Identifying a disease-causing somatic mutation provides a diagnostic explanation and has implications for family planning. Recent studies have demonstrated that up to 5–10% of seemingly de novo germline mutations are in fact somatic mutations, which are associated with a low recurrence risk similar to the population risk, and that a similar percentage are due to parental germline or gonosomal mosaic mutations, which accumulate with age and may be associated with a higher recurrence risk[123,134].

Studying somatic mutations is also important for developing mechanism-based treatments for neurodevelopmental diseases. As with cancer[7], somatic mutations can point to critical signaling pathways to target. FCD and HME are largely caused by somatic mutations that abnormally activate the mTOR pathway. Thus, mTOR pathway inhibitors such as everolimus are attractive candidates for these diseases, and a clinical trial evaluating everolimus for FCD and TSC patients is ongoing (NCT02451696).

Conclusions and future directions

We have recently come to appreciate the importance of somatic mutation in neurodevelopmental disease using NGS and single-cell sequencing methods. Future studies need to optimize methods for detecting and validating somatic mutations, currently associated with several challenges (Table 3), to fully explore the rates and characteristics of somatic mutations across the genome in normal and diseased human brain. Hopefully, such studies will identify the threshold for disease (since we know somatic mutations as low as 1% AAF can cause disease[14]) and uncover genetic etiologies for currently unsolved cases, which may include mutations in introns or noncoding regions, mutations in exons below current detection thresholds, or epigenetic mechanisms. In addition, future studies need to leverage identified somatic mutations to further our understanding of normal and abnormal brain development. Integrated analysis of the genome and transcriptome of single cells and in situ analysis of somatic mutations may illuminate critical cell types and brain regions for pathogenesis[135,136]. Functional studies of somatic mutations using genome editing in animal models and human induced pluripotent stem cells (hiPSCs) may elucidate underlying mechanisms. hiPSCs are an especially attractive model: in contrast to static human brain samples, hiPSCs allow dynamic investigation of the molecular and cellular processes that lead to abnormal brain development, including in three dimensions using cerebral organoids[137], and they retain the genetic signature, including somatic mutations, of the somatic cell they were derived from[138]. hiPSC models have already provided insight into monogenic neurodevelopmental disorders, particularly Rett syndrome, as well as complex disorders, including ASD, associated with somatic mutations[139]. In the coming years, we believe that somatic mutations will further emerge as a powerful tool for understanding human brain development and disease, improving our ability to diagnose and treat patients with a range of neurological diseases.

Received: 16 January 2018; Accepted: 21 September 2018;
Published online: 22 October 2018

References

1. Jónsson, H. et al. Parental influence on human germline de novo mutations in 1,548 trios from Iceland. *Nature* **549**, 519–522 (2017).
2. Kong, A. et al. Rate of de novo mutations and the importance of father's age to disease risk. *Nature* **488**, 471–475 (2012).
3. O'Roak, B. J. et al. Exome sequencing in sporadic autism spectrum disorders identifies severe de novo mutations. *Nat. Genet.* **43**, 585–589 (2011).
4. Veltman, J. A. & Brunner, H. G. De novo mutations in human genetic disease. *Nat. Rev. Genet.* **13**, 565–575 (2012).
5. Larsen, F. W. & Mouridsen, S. E. The outcome in children with childhood autism and Asperger syndrome originally diagnosed as psychotic. A 30-year follow-up study of subjects hospitalized as children. *Eur. Child Adolesc. Psychiatry* **6**, 181–190 (1997).
6. Power, R. A. et al. Fecundity of patients with schizophrenia, autism, bipolar disorder, depression, anorexia nervosa, or substance abuse vs their unaffected siblings. *JAMA Psychiatry* **70**, 22–30 (2013).
7. Philp, A. J. et al. The phosphatidylinositol 3′-kinase p85alpha gene is an oncogene in human ovarian and colon tumors. *Cancer Res.* **61**, 7426–7429 (2001).
8. Biesecker, L. G. & Spinner, N. B. A genomic view of mosaicism and human disease. *Nat. Rev. Genet.* **14**, 307–320 (2013).
9. Azevedo, F. A. et al. Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain. *J. Comp. Neurol.* **513**, 532–541 (2009).
10. Workman, A. D., Charvet, C. J., Clancy, B., Darlington, R. B. & Finlay, B. L. Modeling transformations of neurodevelopmental sequences across mammalian species. *J. Neurosci.* **33**, 7368–7383 (2013).
11. Bae, T. et al. Different mutational rates and mechanisms in human cells at pregastrulation and neurogenesis. *Science* **359**, 550–555 (2018).
This study used clonal cell expansion of fetal human neuronal progenitor cells to estimate the somatic mutation rate during neurogenesis.
12. Lodato, M. A. et al. Aging and neurodegeneration are associated with increased mutations in single human neurons. *Science* **359**, 555–559 (2018).
This study used whole-genome sequencing of single human neurons to demonstrate that somatic SNVs in human neurons accumulate linearly with age.

13. Poduri, A., Evrony, G. D., Cai, X. & Walsh, C. A. Somatic mutation, genomic variation, and neurological disease. *Science* **341**, 1237758 (2013).
14. D'Gama, A. M. et al. Somatic mutations activating the mTOR pathway in dorsal telencephalic progenitors cause a continuum of cortical dysplasias. *Cell Rep.* **21**, 3754–3766 (2017).
15. Dou, Y. et al. Postzygotic single-nucleotide mosaicisms contribute to the etiology of autism spectrum disorder and autistic traits and the origin of mutations. *Hum. Mutat.* **38**, 1002–1013 (2017).
This study, along with refs 16–18, analyzed WES of people with ASD to estimate the contribution of somatic mutationsto ASD risk.
16. Freed, D. & Pevsner, J. The contribution of mosaic variants to autism spectrum disorder. *PLoS Genet.* **12**, e1006245 (2016).
This study, along with refs. 15,17,18, analyzed WES of people with ASD to estimate the contribution of somatic mutations to ASD risk.
17. Krupp, D. R. et al. Exonic mosaic mutations contribute risk for autism spectrum disorder. *Am. J. Hum.Genet.* **101**, 369–390 (2017).
This study, along with refs. 15,16,18, analyzed WES of people with ASD to estimate the contribution of somatic mutations to ASD risk.
18. Lim, E. T. et al. Rates, distribution and implications of postzygotic mosaic mutations in autism spectrum disorder. *Nat. Neurosci.* **20**, 1217–1224 (2017).
This study, along with refs 15–17, analyzed WES of people with ASD to estimate the contribution of somatic mutations to ASD risk.
19. Lodato, M. A. et al. Somatic mutation in single human neurons tracks developmental and transcriptional history. *Science* **350**, 94–98 (2015).
This study used whole-genome sequencing of single human neurons to estimate the number of somatic mutations in each postmitotic neuron.
20. Gennaro, E. et al. Somatic and germline mosaicisms in severe myoclonic epilepsy of infancy. *Biochem. Biophys. Res. Commun.* **341**, 489–493 (2006).
21. Nakamura, K. et al. Clinical spectrum of *SCN2A* mutations expanding to Ohtahara syndrome. *Neurology* **81**, 992–998 (2013).
22. Vadlamudi, L. et al. Timing of de novo mutagenesis–a twin study of sodium-channel mutations. *N. Engl. J. Med.* **363**, 1335–1340 (2010).
23. Xu, X. et al. Amplicon resequencing identified parental mosaicism for approximately 10% of "de novo" *SCN1A* mutations in children with Dravet syndrome. *Hum. Mutat.* **36**, 861–872 (2015).
24. Zerem, A. et al. Paternal germline mosaicism of a *SCN2A* mutation results in Ohtahara syndrome in half siblings. *Eur. J. Paediatr. Neurol.* **18**, 567–571 (2014).
25. D'Gama, A. M. et al. Targeted DNA sequencing from autism spectrum disorder brains implicates multiple genetic mechanisms. *Neuron* **88**, 910–917 (2015).
26. Cai, X. et al. Single-cell, genome-wide sequencing identifies clonal somatic copy-number variation in the human brain. *Cell Rep.* **8**, 1280–1289 (2014).
27. McConnell, M. J. et al. Mosaic copy number variation in human neurons. *Science* **342**, 632–637 (2013).
This study demonstrated that mosaic CNVs are common in human neurons.
28. Crowe, C. A., Schwartz, S., Black, C. J. & Jaswaney, V. Mosaic trisomy 22: a case presentation and literature review of trisomy 22 phenotypes. *Am. J. Med. Genet.* **71**, 406–413 (1997).
29. Daber, R. et al. Mosaic trisomy 17: variable clinical and cytogenetic presentation. *Am. J. Med. Genet. A.* **155A**, 2489–2495 (2011).
30. Gérard-Blanluet, M. et al. Mosaic trisomy 9 and lobar holoprosencephaly. *Am. J. Med. Genet.* **111**, 295–300 (2002).
31. Laus, A. C. et al. Karyotype/phenotype correlation in partial trisomies of the long arm of chromosome 16: case report and review of literature. *Am. J. Med. Genet. A.* **158A**, 821–827 (2012).
32. Pangalos, C. et al. Understanding the mechanism(s) of mosaic trisomy 21 by using DNA polymorphism analysis. *Am. J. Hum. Genet.* **54**, 473–481 (1994).
33. Patel, C. et al. Mosaic trisomy 1q: The longest surviving case. *Am. J. Med. Genet. A.* **149A**, 1795–1800 (2009).
34. Schrander-Stumpel, C. T. et al. Mosaic tetrasomy 8p in two patients: clinical data and review of the literature. *Am. J. Med. Genet.* **50**, 377–380 (1994).
35. Muotri, A. R. et al. L1 retrotransposition in neurons is modulated by MeCP2. *Nature* **468**, 443–446 (2010).
This study showed the first association between a neurodevelopmental disease, Rett syndrome, and increased somatic L1 insertions.
36. Lindhurst, M. J. et al. A mosaic activating mutation in *AKT1* associated with the Proteus syndrome. *N. Engl. J. Med.* **365**, 611–619 (2011).
37. Kurek, K. C. et al. Somatic mosaic activating mutations in *PIK3CA* cause CLOVES syndrome. *Am. J. Hum. Genet.* **90**, 1108–1115 (2012).
38. Shirley, M. D. et al. Sturge-Weber syndrome and port-wine stains caused by somatic mutation in *GNAQ*. *N. Engl. J. Med.* **368**, 1971–1979 (2013).
39. Kingsmore, S. et al. Exome sequencing reveals de novo germline mutation of the mammalian target of rapamycin (*MTOR*) in a patient with megalencephaly and intractable seizures. *Journal of Genomes and Exomes* **2**, 63–72 (2013).
40. Knudson, A. G. Jr. Mutation and cancer: statistical study of retinoblastoma. *Proc. Natl. Acad. Sci. USA* **68**, 820–823 (1971).
41. Maertens, O. et al. Comprehensive *NF1* screening on cultured Schwann cells from neurofibromas. *Hum. Mutat.* **27**, 1030–1040 (2006).
42. Garcia-Linares, C. et al. Dissecting loss of heterozygosity (LOH) in neurofibromatosis type 1-associated neurofibromas: Importance of copy neutral LOH. *Hum. Mutat.* **32**, 78–90 (2011).
43. Crino, P. B., Aronica, E., Baltuch, G. & Nathanson, K. L. Biallelic TSC gene inactivation in tuberous sclerosis complex. *Neurology* **74**, 1716–1723 (2010).
44. Qin, W. et al. Analysis of TSC cortical tubers by deep sequencing of *TSC1*, *TSC2* and *KRAS* demonstrates that small second-hit mutations in these genes are rare events. *Brain Pathol.* **20**, 1096–1105 (2010).
45. Tyburczy, M. E. et al. A shower of second hit events as the cause of multifocal renal cell carcinoma in tuberous sclerosis complex. *Hum. Mol. Genet.* **24**, 1836–1842 (2015).
46. Sheen, V. L. et al. Mutations in the X-linked filamin 1 gene cause periventricular nodular heterotopia in males as well as in females. *Hum. Mol. Genet.* **10**, 1775–1783 (2001).
47. Guerrini, R. et al. Germline and mosaic mutations of *FLN1* in men with periventricular heterotopia. *Neurology* **63**, 51–56 (2004).
48. Parrini, E., Mei, D., Wright, M., Dorn, T. & Guerrini, R. Mosaic mutations of the *FLN1* gene cause a mild phenotype in patients with periventricular heterotopia. *Neurogenetics* **5**, 191–196 (2004).
49. Sicca, F. et al. Mosaic mutations of the *LIS1* gene cause subcortical band heterotopia. *Neurology* **61**, 1042–1046 (2003).
50. Gleeson, J. G. et al. Somatic and germline mosaic mutations in the doublecortin gene are associated with variable phenotypes. *Am. J. Hum. Genet.* **67**, 574–581 (2000).
51. Jamuar, S. S. et al. Somatic mutations in cerebral cortical malformations. *N. Engl. J. Med.* **371**, 733–743 (2014).
52. Rohlin, A. et al. Parallel sequencing used in detection of mosaic mutations: comparison with four diagnostic DNA screening techniques. *Hum. Mutat.* **30**, 1012–1020 (2009).
53. Blümcke, I. et al. The clinicopathologic spectrum of focal cortical dysplasias: a consensus classification proposed by an ad hoc Task Force of the ILAE Diagnostic Methods Commission. *Epilepsia* **52**, 158–174 (2011).
54. Poduri, A. et al. Somatic activation of *AKT3* causes hemispheric developmental brain malformations. *Neuron* **74**, 41–48 (2012)**.**
This study, along with refs 59,61, used direct study of human brain tissue to identify a genetic cause for HME.
55. Hua, Y. & Crino, P. B. Single cell lineage analysis in human focal cortical dysplasia. *Cereb. Cortex* **13**, 693–699 (2003).
56. Blumcke, I. et al. Histopathological findings in brain tissue obtained during epilepsy surgery. *N. Engl. J. Med.* **377**, 1648–1656 (2017).
57. Harvey, A. S., Cross, J. H., Shinnar, S., Mathern, G. W. & Taskforce, I. P. E. S. S. Defining the spectrum of international practice in pediatric epilepsy surgery patients. *Epilepsia* **49**, 146–155 (2008).
58. Jansen, L. A. et al. PI3K/AKT pathway mutations cause a spectrum of brain malformations from megalencephaly to focal cortical dysplasia. *Brain* **138**, 1613–1628 (2015).
59. Lee, J. H. et al. De novo somatic mutations in components of the PI3K-AKT3-mTOR pathway cause hemimegalencephaly.*Nat. Genet.* **44**, 941–945 (2012).
This study, along with refs 54,61, used direct study of human brain tissue to identify a genetic cause for HME.
60. Lim, J. S. et al. Brain somatic mutations in *MTOR* cause focal cortical dysplasia type II leading to intractable epilepsy. *Nat. Med.* **21**, 395–400 (2015).
61. Rivière, J. B. et al. De novo germline and postzygotic mutations in *AKT3, PIK3R2* and *PIK3CA* cause a spectrum of related megalencephaly syndromes. *Nat. Genet.* **44**, 934–940 (2012).
This study, along with refs 54,59, used direct study of human brain tissue to identify a genetic cause for HME.
62. D'Gama, A. M. et al. mTOR pathway mutations cause hemimegalencephaly and focal cortical dysplasia. *Ann. Neurol.* **77**, 720–725 (2015).
63. Lim, J. S. et al. Somatic mutations in *TSC1* and *TSC2* cause focal cortical dysplasia. *Am. J. Hum. Genet.* **100**, 454–472 (2017).
64. Baulac, S. et al. Familial focal epilepsy with focal cortical dysplasia due to *DEPDC5* mutations. *Ann. Neurol.* **77**, 675–683 (2015).
65. Sim, J. C. et al. Familial cortical dysplasia caused by mutation in the mammalian target of rapamycin regulator *NPRL3*. *Ann. Neurol.* **79**, 132–137 (2016).
66. Scheffer, I. E. et al. Mutations in mammalian target of rapamycin regulator *DEPDC5* cause focal epilepsy with brain malformations. *Ann. Neurol.* **75**, 782–787 (2014).
67. Weckhuysen, S. et al. Involvement of GATOR complex genes in familial focal epilepsies and focal cortical dysplasia. *Epilepsia* **57**, 994–1003 (2016).

68. Evrony, G. D. et al. Single-neuron sequencing analysis of L1 retrotransposition and somatic mutation in the human brain. *Cell* **151**, 483–496 (2012).
69. Baek, S. T. et al. An AKT3-FOXG1-reelin network underlies defective migration in human focal malformations of cortical development. *Nat. Med.* **21**, 1445–1454 (2015).
70. Roy, A. et al. Mouse models of human PIK3CA-related brain overgrowth have acutely treatable epilepsy. *eLife* **4**, e12703 (2015).
71. Yuskaitis, C. J. et al. A mouse model of DEPDC5-related epilepsy: neuronal loss of Depdc5 causes dysplastic and ectopic neurons, increased mTOR signaling, and seizure susceptibility. *Neurobiol. Dis.* **111**, 91–101 (2018).
72. Allen, A. S. et al. De novo mutations in epileptic encephalopathies. *Nature* **501**, 217–221 (2013).
73. Carvill, G. L. et al. Targeted resequencing in epileptic encephalopathies identifies de novo mutations in *CHD2* and *SYNGAP1*. *Nat. Genet.* **45**, 825–830 (2013).
74. Depienne, C. et al. Sporadic infantile epileptic encephalopathy caused by mutations in *PCDH19* resembles Dravet syndrome but mainly affects females. *PLoS Genet.* **5**, e1000381 (2009).
75. Veeramah, K. R. et al. Exome sequencing reveals new causal mutations in children with epileptic encephalopathies. *Epilepsia* **54**, 1270–1281 (2013).
76. Weckhuysen, S. et al. *KCNQ2* encephalopathy: emerging phenotype of a neonatal epileptic encephalopathy. *Ann. Neurol.* **71**, 15–25 (2012).
77. Stosser, M. B. et al. High frequency of mosaic pathogenic variants in genes causing epilepsy-related neurodevelopmental disorders. *Genet. Med.* **20**, 403–410 (2018).
78. Perez, D., Hsieh, D. T. & Rohena, L. Somatic mosaicism of *PCDH19* in a male with early infantile epileptic encephalopathy and review of the literature. *Am. J. Med. Genet. A.* **173**, 1625–1630 (2017).
79. Pederick, D. T. et al. Abnormal cell sorting underlies the unique X–linked inheritance of *PCDH19* epilepsy. *Neuron* **97**, 59–66 (2018).
80. Fu, C. et al. GABAergic interneuron development and function is modulated by the *Tsc1* gene. *Cereb. Cortex* **22**, 2111–2119 (2012).
81. American Psychiatric Association & DSM-5 Task Force. *Diagnostic and Statistical Manual of Mental Disorders: DSM-5*. (American Psychiatric Association, Washington, D.C., 2013).
82. Autism and Developmental Disabilities Monitoring Network Surveillance Year 2010 Principal Investigators. Prevalence of autism spectrum disorder among children aged 8 years - autism and developmental disabilities monitoring network, 11 sites, United States, 2010. *MMWR Surveill. Summ.* **63**, 1–21 (2014).
83. Kanner, L. Autistic disturbances of affective contact. *Acta Paedopsychiatr.* **35**, 100–136 (1968).
84. Bailey, A. et al. Autism as a strongly genetic disorder: evidence from a British twin study. *Psychol. Med.* **25**, 63–77 (1995).
85. Steffenburg, S. et al. A twin study of autism in Denmark, Finland, Iceland, Norway and Sweden. *J. Child Psychol. Psychiatry* **30**, 405–416 (1989).
86. De Rubeis, S. et al. Synaptic, transcriptional and chromatin genes disrupted in autism. *Nature* **515**, 209–215 (2014).
87. Iossifov, I. et al. The contribution of de novo coding mutations to autism spectrum disorder. *Nature* **515**, 216–221 (2014).
88. Iossifov, I. et al. De novo gene disruptions in children on the autistic spectrum. *Neuron* **74**, 285–299 (2012).
89. Levy, D. et al. Rare de novo and transmitted copy-number variation in autistic spectrum disorders. *Neuron* **70**, 886–897 (2011).
90. Lim, E. T. et al. Rare complete knockouts in humans: population distribution and significant role in autism spectrum disorders. *Neuron* **77**, 235–242 (2013).
91. Neale, B. M. et al. Patterns and rates of exonic de novo mutations in autism spectrum disorders. *Nature* **485**, 242–245 (2012).
92. O'Roak, B. J. et al. Recurrent de novo mutations implicate novel genes underlying simplex autism risk. *Nat. Commun.* **5**, 5595 (2014).
93. O'Roak, B. J. et al. Multiplex targeted sequencing identifies recurrently mutated genes in autism spectrum disorders. *Science* **338**, 1619–1622 (2012).
94. O'Roak, B. J. et al. Sporadic autism exomes reveal a highly interconnected protein network of de novo mutations. *Nature* **485**, 246–250 (2012).
95. Pinto, D. et al. Functional impact of global rare copy number variation in autism spectrum disorders. *Nature* **466**, 368–372 (2010).
96. Sanders, S. J. et al. Multiple recurrent de novo CNVs, including duplications of the 7q11.23 Williams syndrome region, are strongly associated with autism. *Neuron* **70**, 863–885 (2011).
97. Sanders, S. J. et al. De novo mutations revealed by whole-exome sequencing are strongly associated with autism. *Nature* **485**, 237–241 (2012).
98. Sebat, J. et al. Strong association of de novo copy number mutations with autism. *Science* **316**, 445–449 (2007).
99. Yu, T. W. et al. Using whole-exome sequencing to identify inherited causes of autism. *Neuron* **77**, 259–273 (2013).
100. Bourdon, V. et al. Evidence of somatic mosaicism for a *MECP2* mutation in females with Rett syndrome: diagnostic implications. *J. Med. Genet.* **38**, 867–871 (2001).
101. Clayton-Smith, J., Watson, P., Ramsden, S. & Black, G. C. Somatic mutation in *MECP2* as a non-fatal neurodevelopmental disorder in males. *Lancet* **356**, 830–832 (2000).
102. Helderman-van den Enden, A. T. et al. Monozygotic twin brothers with the fragile X syndrome: different CGG repeats and different mental capacities. *J. Med. Genet.* **36**, 253–257 (1999).
103. Tinschert, S. et al. Segmental neurofibromatosis is caused by somatic mutation of the neurofibromatosis type 1 (*NF1*) gene. *Eur. J. Hum. Genet.* **8**, 455–459 (2000).
104. Verhoef, S. et al. High rate of mosaicism in tuberous sclerosis complex. *Am. J. Hum. Genet.* **64**, 1632–1637 (1999).
105. Vogt, J. et al. Monozygotic twins discordant for neurofibromatosis type 1 due to a postzygotic *NF1* gene mutation. *Hum. Mutat.* **32**, E2134–E2147 (2011).
106. Castermans, D. et al. Position effect leading to haploinsufficiency in a mosaic ring chromosome 14 in a boy with autism. *Eur. J. Hum. Genet.* **16**, 1187–1192 (2008).
107. Havlovicova, M. et al. A girl with neurofibromatosis type 1, atypical autism and mosaic ring chromosome 17. *Am. J. Med. Genet. A.* **143A**, 76–81 (2007).
108. Kakinuma, H., Ozaki, M., Sato, H. & Takahashi, H. Variation in GABA-A subunit gene copy number in an autistic patient with mosaic 4 p duplication (p12p16). *Am. J. Med. Genet. B. Neuropsychiatr. Genet.* **147B**, 973–975 (2008).
109. Meyer, K. J., Axelsen, M. S., Sheffield, V. C., Patil, S. R. & Wassink, T. H. Germline mosaic transmission of a novel duplication of *PXDN* and *MYT1L* to two male half-siblings with autism. *Psychiatr. Genet.* **22**, 137–140 (2012).
110. Oliveira, G. et al. Partial tetrasomy of chromosome 3q and mosaicism in a child with autism. *J. Autism Dev. Disord.* **33**, 177–185 (2003).
111. Papanikolaou, K. et al. A case of partial trisomy of chromosome 8p associated with autism. *J. Autism Dev. Disord.* **36**, 705–709 (2006).
112. Sauter, S. et al. Autistic disorder and chromosomal mosaicism 46,XY[123]/46,XY,del(20)(pter –> p12.2)[10]. *Am. J. Med. Genet. A.* **120A**, 533–536 (2003).
113. Yurov, Y. B. et al. Unexplained autism is frequently associated with low-level mosaic aneuploidy. *J. Med. Genet.* **44**, 521–525 (2007).
114. Hoischen, A., Krumm, N. & Eichler, E. E. Prioritization of neurodevelopmental disease genes by discovery of new mutations. *Nat. Neurosci.* **17**, 764–772 (2014).
115. Wintle, R. F. et al. A genotype resource for postmortem brain samples from the Autism Tissue Program. *Autism Res.* **4**, 89–97 (2011).
116. Casanova, M. F. et al. Focal cortical dysplasias in autism spectrum disorders. *Acta Neuropathol. Commun.* **1**, 67 (2013).
117. Stoner, R. et al. Patches of disorganization in the neocortex of children with autism. *N. Engl. J. Med.* **370**, 1209–1219 (2014).
118. Willsey, A. J. et al. Coexpression networks implicate human midfetal deep cortical projection neurons in the pathogenesis of autism. *Cell* **155**, 997–1007 (2013).
119. Castronovo, P. et al. Somatic mosaicism in Cornelia de Lange syndrome: a further contributor to the wide clinical expressivity? *Clin. Genet.* **78**, 560–564 (2010).
120. Gervasini, C. et al. Molecular characterization of a mosaic *NIPBL* deletion in a Cornelia de Lange patient with severe phenotype. *Eur. J. Med. Genet.* **56**, 138–143 (2013).
121. Vissers, L. E., Gilissen, C. & Veltman, J. A. Genetic studies in intellectual disability and related disorders. *Nat. Rev. Genet.* **17**, 9–18 (2016).
122. Gilissen, C. et al. Genome sequencing identifies major causes of severe intellectual disability. *Nature* **511**, 344–347 (2014).
123. Acuna-Hidalgo, R. et al. Post-zygotic point mutations are an underrecognized source of de novo genomic variation. *Am. J. Hum. Genet.* **97**, 67–74 (2015).
This study reanalyzed whole-genome sequencing from an ID cohort to demonstrate that somatic mutations are under-recognized and miscategorized as germline de novo mutations.
124. King, D. A. et al. Mosaic structural variation in children with developmental disorders. *Hum. Mol. Genet.* **24**, 2733–2745 (2015).
125. Cannon, T. D., Kaprio, J., Lönnqvist, J., Huttunen, M. & Koskenvuo, M. The genetic epidemiology of schizophrenia in a Finnish twin cohort. A population-based modeling study. *Arch. Gen. Psychiatry* **55**, 67–74 (1998).
126. Yurov, Y. B. et al. The schizophrenia brain exhibits low-level aneuploidy involving chromosome 1. *Schizophr. Res.* **98**, 139–147 (2008).
127. Yurov, Y. B., Vostrikov, V. M., Vorsanova, S. G., Monakhov, V. V. & Iourov, I. Y. Multicolor fluorescent in situ hybridization on post-mortem brain in schizophrenia as an approach for identification of low-level chromosomal aneuploidy in neuropsychiatric diseases. *Brain Dev.* **23**(Suppl 1), S186–S190 (2001).

128. Sakai, M. et al. Assessment of copy number variations in the brain genome of schizophrenia patients. *Mol. Cytogenet.* **8**, 46 (2015).
129. Ruderfer, D. M. et al. Mosaic copy number variation in schizophrenia. *Eur. J. Hum. Genet.* **21**, 1007–1011 (2013).
130. Bundo, M. et al. Increased L1 retrotransposition in the neuronal genome in schizophrenia. *Neuron* **81**, 306–313 (2014).
This study used L1 copy number quantification to demonstrate increased L1 insertions in neurons from schizophrenia patients.
131. Doyle, G. A. et al. Analysis of LINE-1 elements in DNA from postmortem brains of individuals with schizophrenia. *Neuropsychopharmacology* **42**, 2602–2611 (2017).
132. McConnell, M. J. et al. Intersection of diverse neuronal genomes and neuropsychiatric disease: the Brain Somatic Mosaicism Network. *Science* **356**, eaal1641 (2017).
133. Yang, Y. et al. Clinical whole-exome sequencing for the diagnosis of mendelian disorders. *N. Engl. J. Med.* **369**, 1502–1511 (2013).
134. Campbell, I. M. et al. Parental somatic mosaicism is underrecognized and influences recurrence risk of genomic disorders. *Am. J. Hum. Genet.* **95**, 173–182 (2014).
This study screened children with genomic disorders and their parents to demonstrate that parental somatic mosaicism is under-recognized and that mutations initially categorized as germline de novo mutations are sometimes detectable in DNA from parental blood.
135. Dey, S. S., Kester, L., Spanjaard, B., Bienko, M. & van Oudenaarden, A. Integrated genome and transcriptome sequencing of the same cell. *Nat. Biotechnol.* **33**, 285–289 (2015).
136. Janiszewska, M. et al. In situ single-cell analysis identifies heterogeneity for *PIK3CA* mutation and *HER2* amplification in *HER2*-positive breast cancer. *Nat. Genet.* **47**, 1212–1219 (2015).
137. Mariani, J. et al. *FOXG1*-dependent dysregulation of GABA/glutamate neuron differentiation in autism spectrum disorders. *Cell* **162**, 375–390 (2015).
138. Abyzov, A. et al. Somatic copy number mosaicism in human skin revealed by induced pluripotent stem cells. *Nature* **492**, 438–442 (2012).
139. Hollingsworth, E. W. et al. iPhemap: an atlas of phenotype to genotype relationships of human iPSC models of neurological diseases. *EMBO Mol. Med.* **9**, 1742–1762 (2017).
140. Lehner, T., Miller, B. L. & State, M. W. *Genomics, Circuits, and Pathways in Clinical Neuropsychiatry*. (Academic Press, Boston, MA, USA, 2016).
141. Bijlsma, E. K., Wallace, A. J. & Evans, D. G. Misleading linkage results in an NF2 presymptomatic test owing to mosaicism. *J. Med. Genet.* **34**, 934–936 (1997).
142. Halliday, D. et al. Genetic severity score predicts clinical phenotype in NF2. *J. Med. Genet.* **54**, 657–664 (2017).
143. Heyries, K. A. et al. Megapixel digital PCR. *Nat. Methods* **8**, 649–651 (2011).
144. Dean, F. B., Nelson, J. R., Giesler, T. L. & Lasken, R. S. Rapid amplification of plasmid and phage DNA using Phi 29 DNA polymerase and multiply-primed rolling circle amplification. *Genome Res.* **11**, 1095–1099 (2001).
145. Zong, C., Lu, S., Chapman, A. R. & Xie, X. S. Genome-wide detection of single-nucleotide and copy-number variations of a single human cell. *Science* **338**, 1622–1626 (2012).
146. Hiatt, J. B., Pritchard, C. C., Salipante, S. J., O'Roak, B. J. & Shendure, J. Single molecule molecular inversion probes for targeted, high-accuracy detection of low-frequency variation. *Genome Res.* **23**, 843–854 (2013).
147. Erwin, J. A. et al. L1-associated genomic regions are deleted in somatic cells of the healthy human brain. *Nat. Neurosci.* **19**, 1583–1591 (2016).
148. Evrony, G. D., Lee, E., Park, P. J. & Walsh, C. A. Resolving rates of mutation in the brain using single-neuron genomics. *eLife* **5**, e12966 (2016).
149. Upton, K. R. et al. Ubiquitous L1 mosaicism in hippocampal neurons. *Cell* **161**, 228–239 (2015).
150. Gole, J. et al. Massively parallel polymerase cloning and genome sequencing of single cells using nanoliter microwells. *Nat. Biotechnol.* **31**, 1126–1132 (2013).

Acknowledgements

The authors thank members of the Walsh laboratory for helpful discussions. A.M.D. was supported by the NIGMS (T32GM007753). C.A.W. was supported by the NINDS (R01NS079277), the NIMH (U01MH106883) through the Brain Somatic Mosaicism Network, the Allen Discovery Center program through The Paul G. Allen Frontiers Group, and the Manton Center for Orphan Disease Research. C.A.W. is an Investigator of the Howard Hughes Medical Institute.

Competing interests

The authors declare no competing interests.

Additional information

Reprints and permissions information is available at www.nature.com/reprints.

Correspondence should be addressed to C.A.W.