UC San Diego
UC San Diego Previously Published Works

Title

Genomic Mosaicism of the Brain: Origin, Impact, and Utility.

Permalink

Journal

Authors

Graham, Jared

Breuss, Martin

Schlachetzki, Johannes

et al.

Publication Date

2024-06-01

DOI

www.neurosci.cn
www.springer.com/12264

REVIEW

Genomic Mosaicism of the Brain: Origin, Impact, and Utility

Jared H. Graham[1] · **Johannes C. M. Schlachetzki**[2] ·
Xiaoxu Yang[3,4] · **Martin W. Breuss**[1]

Abstract Genomic mosaicism describes the phenomenon where some but not all cells within a tissue harbor unique genetic mutations. Traditionally, research focused on the impact of genomic mosaicism on clinical phenotype—motivated by its involvement in cancers and overgrowth syndromes. More recently, we increasingly shifted towards the plethora of neutral mosaic variants that can act as recorders of cellular lineage and environmental exposures. Here, we summarize the current state of the field of genomic mosaicism research with a special emphasis on our current understanding of this phenomenon in brain development and homeostasis. Although the field of genomic mosaicism has a rich history, technological advances in the last decade have changed our approaches and greatly improved our knowledge. We will provide current definitions and an overview of contemporary detection approaches for genomic mosaicism. Finally, we will discuss the impact and utility of genomic mosaicism.

Keywords Genomics · Brain development · Brain homeostasis · Genomic mosaicism

✉ Martin W. Breuss
martin.breuss@cuanschutz.edu

1 Department of Pediatrics, Section of Clinical Genetics and Metabolism, University of Colorado School of Medicine, Aurora 80045-2581, CO, USA

2 Department of Cellular and Molecular Medicine, University of California San Diego, La Jolla 92093-0021, San Diego, CA, USA

3 Department of Neurosciences, University of California San Diego, La Jolla 92093-0021, San Diego, CA, USA

4 Rady Children's Institute for Genomic Medicine, San Diego 92123, CA, USA

Introduction

In the last ten to fifteen years, research efforts increasingly focused on genomic mosaicism—the phenomenon where individual or entire lineages of cells within a tissue harbor genetic mutations that are not present in every cell from one individual [1]. These genetic mutations are generally a consequence of non-fidelity during DNA replication or repair and can be caused by intrinsic or extrinsic mutational mechanisms [2–4]. If such an event occurs early in embryogenesis, the resulting variant will be present in multiple organ and tissue types throughout the body; however, if the variant occurs later after cell fate has already been determined, it will only be detected in that specific lineage [5, 6]. In the case of the nervous system, a mutation that arises from neural stem cells lining the neural tube early in development [7–9] would be present throughout all or many cells within the brain [5]. Alternatively, if a mutation occurs in a terminally differentiated neuron, that mutation will only be present in this one cell and persist for the remaining lifetime of the neuron [10].

Historically, studies devoted to mosaicism predominantly focused on its impact on human disease [1, 11]. This view is motivated by overgrowth syndromes and cancers that harbor mosaic 'driver' mutations that result in increased proliferation and are, consequently, positively selected. In some cases, they simply would be incompatible with life if present in every cell of our body: examples are the overgrowth disorder Proteus syndrome or focal brain malformations [12–17]. On the flip side, disorders where classical *de novo* mutations are typically causative—such as autism spectrum disorders (ASD) or constitutive malformations of cortical development (MCD)—can also be caused by mosaic mutations [18, 19]. While reported examples are often related to systemic and quite abundant mosaicism, the potential role of

⚫ Springer

760

Neurosci. Bull. June, 2024, 40(6):759–776

this phenomenon if 'hidden' within a tissue at lower levels is intriguing; this has been discussed and explored specifically for neurological and neuropsychiatric disorders by the NIH-sponsored Brain Somatic Mosaicism Network (BSMN) [20].

More recently, however, we increasingly appreciate mosaicism as a natural phenomenon that occurs throughout development and in every aging cell [21, 22]. This view highlights the utility of genomic mosaicism in understanding lineage development and aging homeostasis of cells—through lineage tracing and mutational signature analysis [23–26]. Although the vast majority of mosaic variants may not be drivers of disease, they can nevertheless help us to understand pathologies as readouts of mutational exposures or clonal distributions.

Here, we will provide an overview of these topics, definitions that we currently employ to categorize and understand mosaicism, and a summary of the most commonly used approaches to detect genomic mosaicism. Finally, we will also contrast the use of 'natural' with 'engineered' mosaicism and highlight their advantages and disadvantages. While we will provide a discussion of more general aspects of genomic mosaicism, most examples will focus on the mammalian brain. For a further discussion of this phenomenon in other tissue contexts or from a more technical perspective, please refer to reviews authored by us and others [27–33].

Also of note, our categorizations focus entirely on the nuclear genome and ignore genetic variation that is present in the mitochondria [34]. Because of the many copies of the mitochondrial genome in every cell and the phenomenon of heteroplasmy, each cell is mosaic for mutations in this organelle. While this is an important field of study, for this review, we will largely ignore this phenomenon. We point interested readers to excellent reviews and manuscripts of interest [34–38].

Categorization of Mosaic Variants

While genomic mosaicism itself is defined straightforwardly by the presence of a genetic variant in some but not all cells within a collection, it can be further categorized in different ways. For instance, in clinical genetics, the patterns and impact of mosaic mutations are important [39], whereas in reproductive genetics the risk of transmission to the next generation is central [27, 40]. In the latter context, a mosaic variant can be inherited by the next generation if it arose before the specification of primordial germ cells—the progenitors of all germ cells—or within their lineage; recurrence risk within a family is consequently a result of the level of mosaicism that is present in the parental gonads [6, 27, 41].

For developmental analyses, the classification of genomic mosaicism can derive from the timing of mutations within a lineage. For instance, if a somatic mutation arises in a proliferating cell, all daughters of this cell will inherit the variant [42]. This is often referred to as 'clonal' mosaicism, and it is contrasted with 'private' mosaicism which is only present within one cell and is often a consequence of cellular aging [43]. An alternative but related classification distinguishes between 'developmental' and 'aging' mutations (Fig. 1). The former are obligatory clonal mosaic variants, whereas the latter are private if they occur in a terminally differentiated, postmitotic cell, such as a neuron. However, an aging mutation in a proliferating cell would fulfill the definition of being clonal. While the theoretical distinction of these classes is relatively straightforward, experimentally, it is limited by the detection sensitivity of the employed analytical method. Therefore, the concept of 'detectable' clonal mosaicism is sometimes employed for the analysis of entire or larger subsets of tissues, typically synonymous with early developmental or extensively selected mosaic variants [42, 44].

An important feature of mosaic variants when analyzed in the context of tissues or a collection of cells is the Allelic Fraction (AF; also used with a preceding variant, VAF, alternate, AAF, or minor, MAF). This metric describes the fraction of mutant alleles relative to all detected alleles. For instance, a germline heterozygous mutation is detected at an AF of 0.5 in every tissue. In contrast, mosaic variants are detected at lower AFs that reflect their abundance: a mutation that occurs in one of two diploid cells would be present at an AF of 0.25, whereas one that occurs in one of 50 diploid cells at an AF of 0.01. To obtain the fraction of cells harboring a mosaic mutation, for diploid cells, one has to multiply the AF by two; for haploid cells, the AF is equal to the cellular fraction.

Types of Mosaic Variants

Independent of their timing, mosaic variants are also distinguished based on the type of mutations (Fig. 2); and these are largely similar to variants that are encountered as germline mutations [45–49]. The most commonly detected and conceptually simplest type of mosaic variants is mosaic single nucleotide variants (mSNVs). They encompass any variant that exchanges one base for another and are typically a result of DNA damage that is not or inadequately repaired [22, 50]. While all possible transitions and transversions can be observed in the genome, intrinsic and extrinsic mutational mechanisms result in significant biases in their rates [51]. For instance, the transition from a cytosine to a thymidine (denoted as C>T) is often encountered at high frequencies, as it results from the deamination of a methylated cytosine [52]; this, in turn, is read as uracil during replication, fixing this change in the genome. As cytosine methylation and the subsequent deamination are a relatively common occurrence
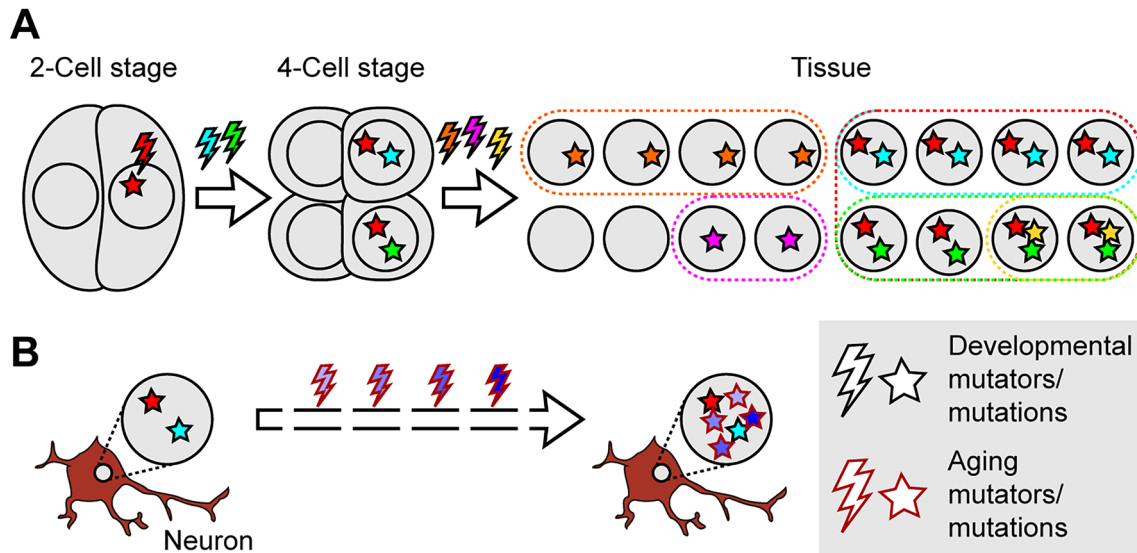
Fig. 1 Developmental and aging mosaicism. **A** In development, mutations that occur at very early stages are transmitted to daughter lineages. Subsequent mutations further distinguish distinct lineages or sub-lineages. Developmental mutations are obligatory clonal mosaic variants. **B** Differentiated cells (here exemplified by a postmitotic neuron) already carry developmental mutations and accumulate additional aging mutations. In postmitotic cells, these aging mutations are obligatory private mosaic variants. Note that developmental and aging mutators are primarily distinguished by their timing, but may share intrinsic or extrinsic mutagenic stressors.

in the human genome, this is an often encountered mosaic mutation type. While these concepts have been described in much more detail for various cancers [51, 53–55], they are now an integral part of any mosaicism research.

The burden estimates of mSNVs vary depending on the experimental design and employed method. A study by Bae and colleagues proposed that neuronal progenitors accumulate 5.1 mSNVs per neuronal progenitor per day which culminated in a newborn neuron harboring 300-900 mSNVs within one year of birth [21]. This was in agreement with studies by Lodato and colleagues who also noted that postmitotic neurons accumulated dozens of more mutations per year with some variability depending on the brain region [26]. While these estimates are subject to potential technical artifacts that may increase false-positives and suffer from false-negative mSNVs, they are comparable to mosaic mutation accumulation in clones isolated from somatic proliferating tissues [56]. An interesting comparison benchmark for these rates is the human per-generation mutation rate, which mostly reflects mutations accumulated in the lineage of the egg and sperm [27, 57, 58]. Here, the male contribution increases by approximately 1.5 variants per year, which is comparable to the rates found for clones within the seminiferous tubules [56, 59].

The functional consequence of mSNVs is dependent on their location, and their interpretation can be challenging if they do not directly impact the function, folding, or expression of a protein product [60, 61]. In addition to the ambiguity of the mSNV's effect, it also has to be put into the context

of its abundance and presence in various tissues, which can further modify the functional impact (a consideration that also applies to other types of mosaic mutations). While the most abundant type genome-wide, most mSNVs will have no or limited impact on phenotypes.

A variant type that is conceptually closely related to mSNVs is small mosaic insertions or deletions (mInDels). They range by definition from 1 to 50 base pairs in size and are often the result of polymerase slippage or imperfect repair [62]. In normal development, they likely occur an order of magnitude less frequently than mSNVs based on population frequency and mosaicism assessments by us and others [23, 44, 59, 63]; of note, these analyses may underestimate their frequency due to increased technical challenges in detecting mInDels compared to mSNVs. Their functional impact is higher than for mSNVs when found in coding sequences, as they often result in a frameshift mutation that results in the premature termination of a protein.

From a mosaicism perspective, short tandem repeat expansions and contractions (mSTRΔs) represent a largely unexplored territory [46]. They are generally a result of polymerase slippage in mitosis, although unequal crossing over in meiosis also plays a role in their origin; yet, the latter is irrelevant for somatic mSTRΔs, as they occur in non-meiotic tissues, such as the brain. While we previously reported genome-wide mSTRΔs in the sperm of fathers that were identified leveraging variation data from offspring [6], a more comprehensive analysis of these mosaic mutations across tissues has not been performed—mainly

762

Neurosci. Bull. June, 2024, 40(6):759–776

Name	Abbreviation	Size	
Mosaic Single Nucleotide Variant	mSNV	1	
Mosaic Insertion or Deletion	mInDel	1-50	
Mosaic Short Tandem Repeat Variation	mSTRΔ	Variable	
Mosaic Structural Variant	mSV	>50	
- Copy Number Variant	mCNV	>50	
- Aneuploidies	-	Chromosomes	
- Copy Neutral Loss of Heterozygosity	mCN-LOH	>50 - chromosomes	
- Retrotransposition	-	Typically 1,000s	

Fig. 2 Types and Scale of Mosaic Variants. Small mosaic variant types like mSNVs, mIndels, or mSTRΔs are the most common types of mosaic genetic variation. However, larger mosaic variants can be grouped as mSVs. Each type of observed mosaic variant is illustrated in this figure with an example. mSNV: a T to G base pair substitution; mInDel: a one-base pair deletion; mSTRΔ: a one unit CAG expansion; mCNV: two examples for a genomic tandem duplication and a deletion; aneuploidies: duplication on one chromosome; mCN-LOH: duplication of a part of the green haplotype while partially losing the blue haplotype; retrotransposition: insertion of the red retroviral mRNA sequence into the locus.

due to technical challenges when using typical sequencing techniques [64]. Thus, their frequency is largely unknown, despite their potential health impact being well-understood for a range of neurological disorders [65, 66]. For instance, the classical repeat disorder Huntington's disease has been reported by Telenius and colleagues to exhibit tissue-specific instability—or mosaicism—in the brain decades ago [67]. Based on their work, a mosaic increase in expansions may worsen local tissue phenotypes.

A less frequent form of mosaic variants is mosaic structural variants (mSVs). They can be copy-neutral or result in copy number-variation (mCNV); mCNVs, like germline CNVs, are defined as genomic intervals with deletions or duplications—as small as a few hundred base pairs or as large as entire chromosomes, which are referred to as aneuploidies [68]. Depending on the type of mSV and the employed detection method they can be challenging to detect; yet, they were successfully identified in individual cells or the context of disease if abundantly present [69–72].

Furthermore, a recent study suggests that as many as 10% of neurons may carry mCNVs and in some cases complex karyotypes [73]. While they are significantly rarer events than mSNVs or mInDels, they may have a much larger potential impact on function, largely depending on their size, affected genomic regions, and whether they result in a change of copy number.

One interesting sub-class of mSVs that has received increased recent scrutiny is mosaic copy number neutral losses of heterozygosity (mCN-LOHs) [74–76]. Loss of heterozygosity refers to a phenomenon where, instead of carrying paternal and maternal genomic material, two copies of one or the other are present. These are referred to as 'copy number neutral', as there are still two alleles, and the region is still considered diploid, despite receiving two copies from one parent [49]. These mSVs can range in size from very small genomic intervals to entire chromosomes. mCN-LOH variants have been studied in the field of cancer genetics and the context of clonal selection—as a LOH event could alter a

heterozygous cancer-driving variant to a homozygous state [49, 77, 78]. However, in cases of severe mutations, mCN-LOHs may also be selected to express milder phenotypes, as demonstrated by an intriguing study by Lee and colleagues [79].

There is one additional variant type—technically a subclass of mSVs—that leads to genomic mosaicism and is significantly enriched in brain tissue: retrotransposition of mobile elements. The most commonly studied type is LINE-1 or L1: L1-related sequences comprise 17% of the human genome and the L1 family contains the most active transposable element in the human genome [80, 81]. These retrotransposons can create mosaic populations by inserting DNA sequences in random locations in the genomes of different populations of cells. While largely quiescent in most somatic cells, it has been shown that LINE-1 elements are highly active in developing neuronal progenitors [82, 83].

There has been disagreement regarding the frequency of these events, and current estimates range from <0.6 to 13.7 insertions per neuron [84, 85]. As a potentially relatively large insertion (thousands of base pairs), these events may impact gene expression significantly; this is compounded by their potential to interfere with splicing if integrated into an intronic region [85]. In addition, they are also interesting functionally from an evolutionary viewpoint as they quickly enable sequences to be transcribed and expressed [86]; however, whether an analogous mechanism is important in the context of brain mosaicism remains currently unexplored.

Detection of Mosaic Variants

For this review, we distinguish three distinct types of mosaicism detection (Fig. 3A, Table 1): (1) visualization of mutations in tissues or on the level of individual cells;

Fig. 3 Types and scale of mosaicism detection approaches. **A** Mosaic mutations in a subpopulation of cells may be detected by three theoretical approaches: (1) through direct visualization of mutations employing FISH or chromosome spreads, (2) through bulk analysis of genomic material, or (3) through assessment of genomic material at the level of single cells. **B** Biological insights obtained from mosaicism analysis are heavily dependent on the scale of sampling. For instance, mosaicism may be detected from an entire tissue like the neocortex, microdissection, or microdissections, all of which provide distinct information due to their drastically different scale.
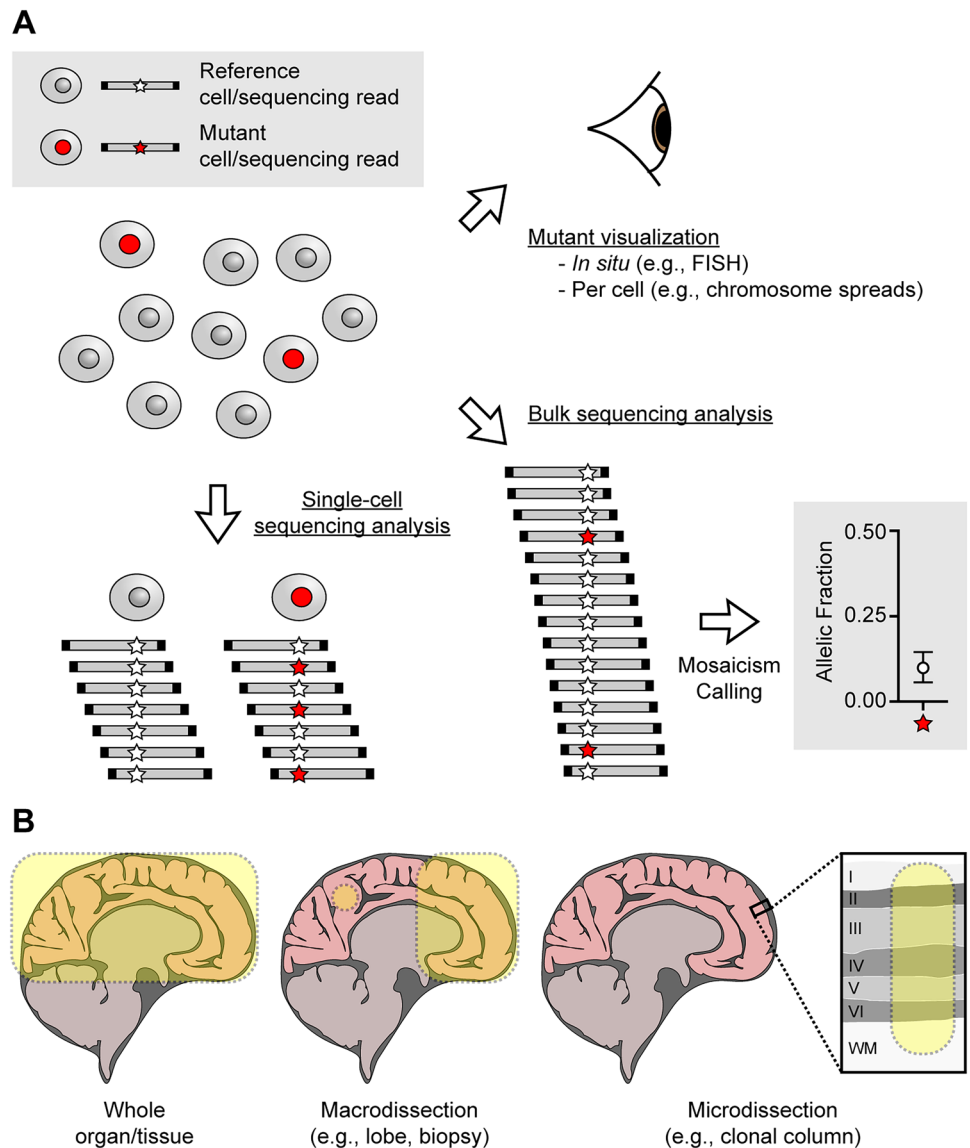
764

Neurosci. Bull. June, 2024, 40(6):759–776

Table 1 Approaches to discover novel mosaic variants

Approach	Resolution	Advantages	Limitations
Top-down – bulk tissue analysis	–	Comprehensive for abundant variants	Sensitivity is limited by read-depth; and highly dependent on the sampling strategy
Methods (Detection)			
DNA microarrays	mSV	Low cost; high depth and sensitivity	Limited resolution based on array density
Whole-exome or panel sequencing	mSNV	Lower cost than genome sequencing; higher read-depth and sensitivity in practice	No analysis outside the exome (e.g., introns) or panel; capture/enrichment can cause artifacts
Whole-genome sequencing	mSNV-mSV	Comprehensive variant detection	Highest cost results in lower read-depth and sensitivity in practice
RNA sequencing	mSNV	Lower cost than genomes; captures functional information	No analysis outside of expressed genes; RNA editing/processing may introduce artifacts or false positives
Top down – clonal biopsy analysis –		Highly sensitive for clonal variants; the sweet spot for many applications	Less comprehensive than bulk tissue analysis and less sensitive than single-cell analysis; relies on a clonal architecture of tissues for maximum efficiency
Bottom Up – Single-Cell Analysis –		Sensitivity down to a single-cell level	Lower sampling due to cost (one analysis per cell) can prevent comprehensive sampling; amplification artifacts
Methods (Amplification)			
Enzymatic whole-genome amplification	–	Possible on extracted cells or nuclei from primary tissue that was frozen	Introduces amplification artifacts and consequently false positives
Clonal expansion	–	Relatively cheaper; higher fidelity to the endogenous replication machinery	Cell culture artifacts; require living cells that can clonally expand

(2) bulk genomic DNA analysis where mutant alleles are detected using specialized mosaicism detection tools; and (3) single-cell genomic DNA analysis. For the latter two approaches, the most common choice of technology is direct sequencing analysis. Mosaicism analysis can be performed for the detection of unknown or already known variants; this is true across the three mentioned types.

In the realm of visualization of genomic mutations, chromosomal karyotyping—in use since the 1950s—is a way of imaging entire chromosomes following an arrest in metaphase [87]. With the addition of techniques such as G-banding, it is possible to identify mutations at the partial-arm or whole chromosome level [88]. While often used in clinical genetics to infer karyotypes or large-scale structural variants for a patient, these methods inherently work on a single-cell level. Thus, this renders them excellent tools for understanding genomic mosaicism if a sufficient number of cells are assayed. Indeed, this approach allows for the identification of constitutive as well as mosaic deletions, duplications, or translocations of sufficient size [88, 89].

Employing RNA or DNA fluorescent probes, Fluorescent *in Situ* Hybridization (FISH) allows the sequence-specific detection of chromosomal locations in interphase and metaphase [90–92]. Thus, using FISH, researchers and clinicians can assay specific intervals on chromosomes for aneuploidies [93, 94]. FISH can also be used in the context of a tissue [95]; combining single-molecule FISH with allele-specific probes theoretically also allows the assessment of mSNVs if they are located in expressed genes [96, 97]. While not currently employed, further technological advances and the combination with super-resolution microscopy may eventually enable the direct detection of genomic mosaic variants.

While the previous methods are useful for interrogating mSVs visually at low throughput, they are generally not very scalable. In the 2000s and 2010s, DNA microarrays were popularized as a higher-resolution method to interrogate DNA copy numbers [98, 99]. Array comparative genome hybridization (aCGH) panels were designed so that an experimental sample would be compared against a diploid reference sample, and these arrays were further improved by the addition of single nucleotide polymorphism (SNP) genotyping panels that could also identify LOH. This approach vastly improved the ability to detect mCNVs and mCN-LOH down to a resolution of approximately 50 kb at AFs as low as 0.01 from bulk samples [100–102]. Such SNP genotyping data was leveraged effectively to detect these types of

mosaic variants in blood, which allowed the detection of clonal hematopoiesis and its impact on neurological disorders [71, 78, 102, 103]. While this approach still has limited resolution as compared to next-generation sequencing technologies, DNA microarrays remain relevant due to the still competitive cost for large-scale genomic analyses.

Detection of Genomic Mosaicism from Tissues

The advent of next-generation sequencing (NGS) fundamentally changed our approach to detecting mosaicism [104]. As NGS inherently is a method that assesses the sequence of individual DNA molecules, it is uniquely suited for mosaicism research—and it allows the detection of mSNVs. The theoretical limit of sensitivity for mosaic variant detection is set by the depth of sequencing and statistical considerations of random sampling. For instance, employing simple binomial calculations, a variant present at 0.1 AF in a non-limiting DNA sample will be picked up in at least one read 65.13% of the time when sequencing at 10×, or 99.99% at 100×. Similarly, a variant present at 0.01 AF will be picked up only 9.56% of the time at 10×, or 63.40% at 100×.

As the employed read-depth is typically cost-limited by the sequenced genomic space, whole-genome sequencing (WGS) is generally performed at lower depths (~30–60×) than whole-exome sequencing (WES; ~50–100×) or targeted panels (often >1,000×) [105, 106]. Due to decreasing sequencing costs, these numbers are very much moving targets, and we and others have employed deeper sequencing to understand mosaicism at higher sensitivities [23, 107–110]. This approach is especially powerful when assessing unbiased bulk mosaicism on a tissue level or within a microdissection (Fig. 3B). When dealing with microdissections, instead, it is common to perform regular or even shallower sequencing on a larger number of samples [24, 56, 111]. An alternative approach is mosaicism detection from RNA sequencing data [112]. Conceptually similar to detecting mosaicism from exome sequencing, there are biological and technical complications that need to be considered carefully, such as the level of transcription from areas containing mosaic variants, potential splice variants that may lead to uncalled mosaicism, or RNA editing that may result in false-positives. However, RNA sequencing analysis may reveal additional functional details of the impact of mosaic variants as DNA is transcribed to mRNA or even downstream (Table 1).

While there is a strong correlation between read depth and sensitivity, the above-provided sensitivity calculations assumed that one mutant read is sufficient to detect mosaicism. In practice, NGS and its computational processing have inherent error rates that impair our ability to detect mosaic variants [29]. Thus, in addition to sequencing at sufficient depth, mosaicism analysis also requires specialized analytical pipelines [113]. As unbiased mosaicism analysis from bulk sequencing samples was pioneered by the cancer research field, variants were typically expected to be positively selected, consequently at relatively high AFs, and only present in the tumor but not 'normal' control tissue. Thus, algorithms were initially designed to detect mosaicism for this purpose specifically [17].

More recently, the focus of mosaicism research has shifted to include lower abundance mosaic variants (<0.05 AF) and those that are shared among tissues. In many cases, these modern pipelines still include classical tools, such as Mutect2 [114, 115], but also employ additional classifiers that provide a secondary level of evaluation and increase specificity (i.e., reduce the number of false-positive mosaic variants) [116–118]. Alternatively, some programs provide both variant detection and classification, such as our previously developed tool MosaicHunter [119]. All of these pipelines typically have areas of strengths and weaknesses, and they might require different experimental designs (e.g., a tumor-normal comparison). Therefore, many analytical pipelines employ a combination of methods to improve sensitivity, specificity, or both. If a variant is known (or a group of variants is routinely seen), these approaches are often replaced by more specialized pipelines (e.g., for drivers of clonal hematopoiesis) [120]. It is important to note that these analytical approaches are rapidly evolving in parallel with sequencing technologies and computational innovations. For instance, duplex sequencing can significantly improve the specificity of any detected variants but comes at the cost of increased sequencing depth requirements [121, 122].

An important part of many mosaicism detection experiments from bulk samples is the subsequent validation of candidate variants through orthogonal approaches. We want to focus on two of the most common here: targeted amplicon sequencing and droplet-digital polymerase chain reaction (ddPCR); we do, however, acknowledge that there are many others, such as subcloning of amplified products, Multiplex-Ligation Probe Amplification (MLPA), denaturing high-performance liquid chromatography (DHPLC), and so forth [19, 123, 124]. The use of targeted amplicons that are subsequently sequenced at high depth is a commonly used approach for validation and quantification by us and others [6, 125, 126]. ddPCR enables the genotyping of single molecules through lipid droplet partition [127]. While the latter has inherent advantages, such as being independent of NGS approaches and highly sensitive, it is also relatively expensive when not used to test the same variants repeatedly and less scalable. Similar to the unbiased detection methods, the use of validation approaches is dependent on the specific question, the number of variants tested, and similar considerations.

766

Neurosci. Bull. June, 2024, 40(6):759–776

Detection of Genomic Mosaicism from Individual Cells

Finally, the detection of mosaicism can also be performed on a single-cell or single-nucleus level (Table 1). Technological advances now allow the high-fidelity amplification of a single genome and the detection of mSNVs [128, 129]. This has been used to great effect for neurons to understand development, aging, and disease or in cardiomyocytes [25, 26, 130]. Similarly, employing whole-genome amplification, mSVs can be detected from a single cell; this has been extensively studied in sperm but also in neurons [73, 131–134]. A variation of this approach is the analysis of a clone (e.g., crypts in the gut) *in situ*, where it is possible to isolate tissues that are mono- or low-level polyclonal. This allows an understanding of mosaicism across phylogenies but also within human tissues to understand mutation rates in such clones [24, 56, 135]. A different and interesting addition to single-cell or single-nucleus technologies is the combination with functional readouts, such as single-cell RNA sequencing [136].

If a cell type is capable of clonal expansion (e.g., skin cells, neural progenitors, cancer cells), it is possible to expand individual cells and sequence the resulting population as a representation of the genomic mosaicism present in the founder cell [21, 137, 138]. This takes advantage of the superior amplification of genomic DNA by the cellular machinery. For cell types that are not inherently available for clonal expansion, it is possible to perform nuclear transfer into proliferation-competent donors; this has been demonstrated for postmitotic neurons in mice [139]. However, this is a complicated process that has not yet been successfully applied to human neurons.

While it is inherently attractive to assess mosaicism on the level of a single cell, there are some technical limitations to employing this approach. First, the amplification of genomic material from a single cell or nucleus is error-prone and may result in a larger number of false-positive mosaic variants [140]. There are, however, strategies to remedy this, such as the genotyping of additional material for confirmation of clonal mosaicism or the restriction to 'phased' haplotypes [110, 141]. Here, the assumption is that mosaic variants should be restricted to the one parental haplotype where it originally arose. If a mosaic mutation is instead found across the two, it should be considered an artifact, as it is exceedingly unlikely for the same mosaic mutation to occur twice within a sample or cell [29, 116]. Similarly, clonal expansion—if possible—may suffer from cell culture artifacts, such as mutations acquired after isolation from primary tissues or selection of certain genotypes. Second, independent of possible errors, there is an additional conceptual limitation for employing single-cell analysis. We often refer to this approach as 'bottom-up'—in contrast to the bulk-based 'top-down'. While single cells offer the highest sensitivity of mosaic variant detection, they also provide a less comprehensive picture of clonal mosaicism depending on the sampling strategy and the number of assessed cells.

Impact and Utility of Genomic Mosaicism

As discussed in the preceding sections, various types of mosaic variation can be detected through different technical and conceptual approaches. However, why do we want to detect mosaicism in the first place? We propose that mosaicism detection can serve three fundamental purposes. (1) Genomic mosaicism may have a direct impact on observed phenotypes (Fig. 4A); this can be either due to positive selection of the mutation and a change in clonality or due to a dominant phenotype. Thus, the detection of these mutations may aid our understanding of disease pathology or enable treatment in the future. In addition, somatic mosaic variants can also be used for their utility: (2) as clonal lineage marks to understand normal development or mutational rates (Fig. 4B); (3) as a readout of environmental processes that induce certain types or patterns of mutations at developmental or past cellular time points (Fig. 4C). We will discuss each of these three with a focus on the brain subsequently; however, we will not provide a comprehensive discussion of each point for brain mosaicism. Thus, we want to draw attention to other excellent reviews that highlight these concepts in the brain, especially in the context of diseases [10, 18, 72, 142–146].

Mosaic variants within the nervous system have been identified as drivers of neurological disease in several instances. For instance, focal cortical dysplasia (FCD) and hemimegalencephaly (HME) are classical mosaicism-driven neurological diseases; they are characterized by dysmorphism and hyperexcitability of a small region of the cortex or an entire cerebral hemisphere, respectively [147]. Analysis of brain tissues from FCD and HME patients has revealed activating somatic mutations in the *mTOR-AKT3-PI3KCA* pathway as well as loss-of-function mutations in genes that are negative regulators [13, 16, 17, 148–151]. Importantly, these somatic mutations were mostly found exclusively in brain tissue, suggesting that they were acquired later in neurodevelopment [152]. Conceptually, these disorders are part of a spectrum where the exact phenotypic presentation depends on the timing of the driver mutation and its abundance within the tissue. Mechanistically, they likely represent a combination of positively selected and dominant mechanisms, as there is evidence of associated overgrowth syndromes and a patch of hyperexcitable dysplastic cells may induce drug-resistant epilepsy within a network of neurons.

In other neurological disorders, such as Alzheimer's disease, schizophrenia, or ASD, it is speculated that mosaic variants may also contribute to or exacerbate the overall phenotype [71, 107, 108, 153–157]. While the contribution of

Fig. 4 Impact and utility of natural mosaicism. **A** Mosaic mutations may act as a driver of disease. Clones harboring mosaic mutations can be positively selected for continued expansion and proliferation which may directly result in disease. Alternatively, mosaic mutations may exhibit a dominant phenotype. Note that these two scenarios are not mutually exclusive. **B** Natural mosaicism marks cellular lineages and can be used for lineage reconstruction or clonal analysis. For instance, in this example, distinct clones are marked by Neutral Mutation (NM) 1 and NM 3, whereas NM 3 marks a sub-clonal lineage in combination with NM 1. **C** Natural mosaicism can be used as a molecular readout of the microenvironment that cells are encountering. Exposure to different environmental mutagens such as reactive oxygen species or toxins can lead to very specific 'mutational signatures'.
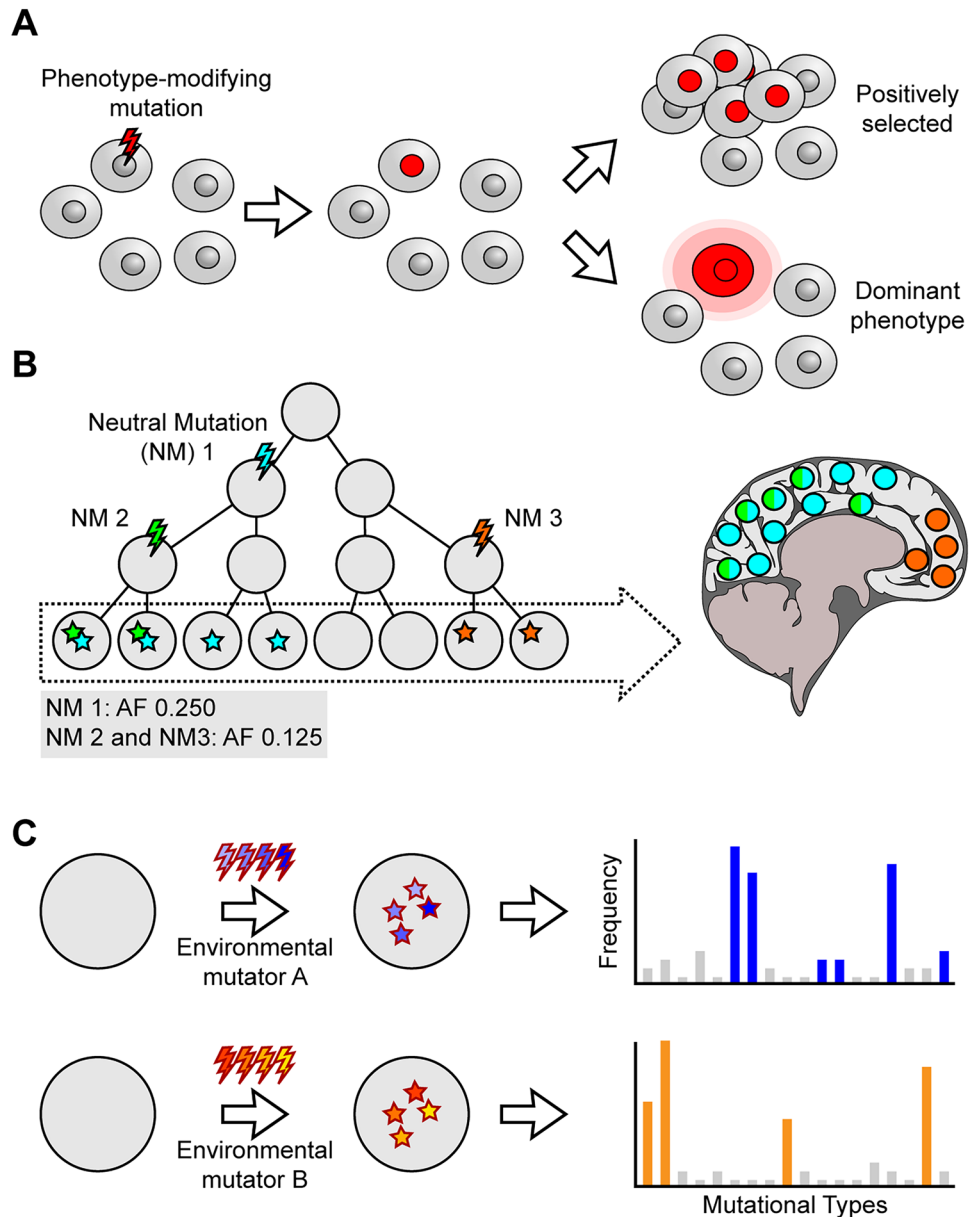
mosaicism to such disorders is clear in cases where it is also detectable from blood (especially if a known disease mutation), the direct contribution of 'cryptic' (i.e., brain-specific) mosaicism has been more elusive. Unlike the clear disease mutations found at lower abundance (often below 0.10 or even 0.05 AF) in FCD [151], the minimum abundance of causative mutations in other disorders remains unclear. Thus, this is still an exciting ongoing field of research.

Utility of Natural Genomic Mosaicism for Lineage Tracing

Just as the timing of mosaic mutations in development is important to understand disease pathogenesis, researchers can also utilize neutral somatic mutations to study developmental processes and clonal lineages directly (Fig. 4B). Indeed, analysis of naturally occurring somatic mutations as lineage 'barcodes' has been utilized to study embryonic development and more generally cellular lineages of humans [24, 158]. We and others have similarly employed this framework to specifically understand these in the brain [23, 25, 110, 138]. An important addition to such studies—independent of the use of bulk or single-cell analyses—is the restriction of analysis to specific cell types (often neurons). This is typically achieved through fluorescence-activated nuclear sorting, as most human brain samples are frozen immediately after collection, which significantly complicates cellular sorting [159].

What are some of the insights that have been derived from employing mosaicism for lineage analysis? For instance,

work by us and others has focused on the spread of clones that can be distinguished by mosaic variants [23, 110]. In 2015, the Walsh Lab sequenced 36 single-cell neurons to identify thousands of somatic mutations [25]. From these, they were able to reconstruct a lineage tree and identify points of divergence. As an extension, Bizzotto and colleagues performed high-depth sequencing on multiple human tissues to identify mSNVs [110]. They concluded that at the onset of gastrulation, there exists a pool of approximately 170 cells with 50 to 100 founders committed to the forebrain. These analyses also revealed that the spread of clones across the cortical surface is largely inverse-correlated with the observed abundance, although there are exceptions to this. A recent study by us further revealed that within the neocortex—in contrast to the overall patterning of the neural tube—clones are first separated along the left-right axis before anterior-posterior [23]; however, the same left-right separation did not extend to the hindbrain. Focusing on the neocortical hemispheres, we further proposed a neural progenitor founder pool of approximately 90 to 200 cells at the time of left-right separation.

Clonal analysis that focused on lineages of defined cell types further confirmed previous findings from rodent models for the first time in humans [23, 110, 136, 160–163]. These studies also suggested the existence of developmental bottlenecks or restrictions that can modify the contributions of early lineages within similar tissues in the absence of selection. Together, while these studies described some aspects of neurodevelopment comprehensively, there are still many remaining questions that will require careful experimental (or sampling) design and analyses; these include the migratory and developmental patterns of interneurons, the lineages and clonality of microglia, and the developmental trajectories of non-neocortical brain regions.

Utility of Natural Genomic Mosaicism to Map Mutational Histories

Finally, genomic mosaicism can also be a useful biomarker of cellular environment or stressors (Fig. 4C). The rate or frequency of mosaic mutations may reflect endogenous or exogenous mutators. For instance, the somatic mutation rate is significantly increased in neurons of individuals with mutations in the DNA repair machinery during aging [26]. The mutation burden can also be increased in seemingly healthy individuals in both germ cells and neurons due to alterations in the same or similar pathways [108, 164]. This approach even enables a distinction of repair fidelity across development: for instance, the earliest cell divisions in an embryo appear to show increased mutation rates compared to the latter, possibly due to the inheritance of the repair machinery through the egg cell [6, 44, 165].

Beyond the frequency of mutations, the observed types vary based on the mutator as well. Here, mSNVs are mainly analyzed and categorized from the perspective of the pyrimidine base (i.e., cytosine or thymidine) and the newly acquired mutations: thus, there are six possible substitutions, three for each. We mentioned above C>T substitutions and that they mainly derive from the deamination of methylated cytosine [52]; similarly, other cellular processes may drive different types of mutations. These patterns mainly derive from cancer studies but have been widely applied across genomic mosaicism research [166].

In addition to the described six categories of mutational types, more recently, the context of a mutation—the neighboring bases—has also been considered. This represented a significant innovation, as the two bases immediately adjacent to the mutated base allow for a finer dissection of molecular mutation mechanisms [167–170]. This is achieved through the statistical isolation of 'mutational signatures' which represent distinct potential drivers of mutations. While some of these turned out to be artifactual, others have been directly connected to internal and environmental mutagens, such as DNA replication, ultraviolet A light radiation, tobacco smoke, or certain chemotherapeutics [171–173]. An important limitation to assessing mutational signatures is the requirement for a sufficient number of observed mutations. Thus, this approach is mainly applicable to large collections of cancer genome data, or studies that leverage the individual genomes of cells like those centered on neurons [51]. In certain situations, lower numbers may be sufficient if driven by specific mutations, as demonstrated by a study focused on transgenerational mutation rates [164]. If this method can be implemented, it opens a window into the experienced environment of cells during development or in the context of disease.

For instance, neuroinflammation is a common symptom of many neurodegenerative disorders and is associated with a dysregulation of redox balance in the brain [174, 175]. A higher level of reactive oxygen or nitrogen species can result in elevated rates of somatic mutations in individual cells or their lineages present in the brain [176–178]. Analyzing the rates and types of somatic mutations in neurotypical versus diseased individuals provides information on the disease environment and may even identify contributing factors. In a recent example, this approach was applied to Alzheimer's disease, confirming an inflammatory environment that causes oxidative DNA damage in neuronal nuclei [179]. The human genome effectively acts as a tape recorder of its environment; when combined with developmental lineage analysis it is possible to also resolve the embryonic environment. However, an efficient implementation requires the generation of large data sets that allow for the stratification of distinct developmental stages at higher resolution. The National Institutes of Health have recognized this limitation

and have started the Somatic Mosaicism across Human Tissues (SMaHT) network which was initiated in 2022 and plans to provide a database of human genomic mosaicism and related technological toolboxes.

Utility of Engineered Genomic Mosaicism

Conceptually, both lineage analysis and mutational signatures represent a fascinating conundrum. For both, adult tissues—often from deceased individuals—are employed to understand embryonic processes that often occurred decades ago. While this allows insights into early development in the context of an organism—humans—that is otherwise intractable for such studies, it comes with some caveats. First, lineage reconstructions require many assumptions, including that mutations and lineages are neutrally selected and do not disappear; while this is mostly an appropriate approximation, it has been demonstrated to be inaccurate in some cases, especially in the context of the first cell divisions [110, 138, 180]. Second, due to the naturally determined rate of mutation and technical issues, lineage trees often remain incomplete. Thus, while these approaches allowed interesting and fundamental insights into neurodevelopment despite these limitations, there are clear advantages when mosaicism can be engineered rather than passively detected.

Lineage tracing through engineered mosaicism has a long history in neuroscience research. Conceptually, the goal is to mark the genome of a subpopulation with a permanent change that can be detected at a later time point [181, 182]. For instance, classical lineage tracing experiments in the brain by Walsh and Cepko employed retroviral vectors where a subset of cells was labeled based on their location at the time of injection [183–185]. Depending on the vector, the readout of the mosaic change was based on a visual phenotype or a direct readout of a known genetic sequence. The advent of advanced mouse genetics enabled the now classical lineage tracing with a fluorophore or otherwise active reporter protein upon irreversible activation by a Cre recombinase (Fig. 5A) [181, 182]. More complex reporters include the use of combinatorial multi-fluorophore systems to differentiate individual lineages in parallel or Mosaic Analysis with Double Markers that sparsely label individual daughters of a single cell [186, 187]. The most significant drawback of this system for lineage tracing is the limited resolution employing fluorophores.
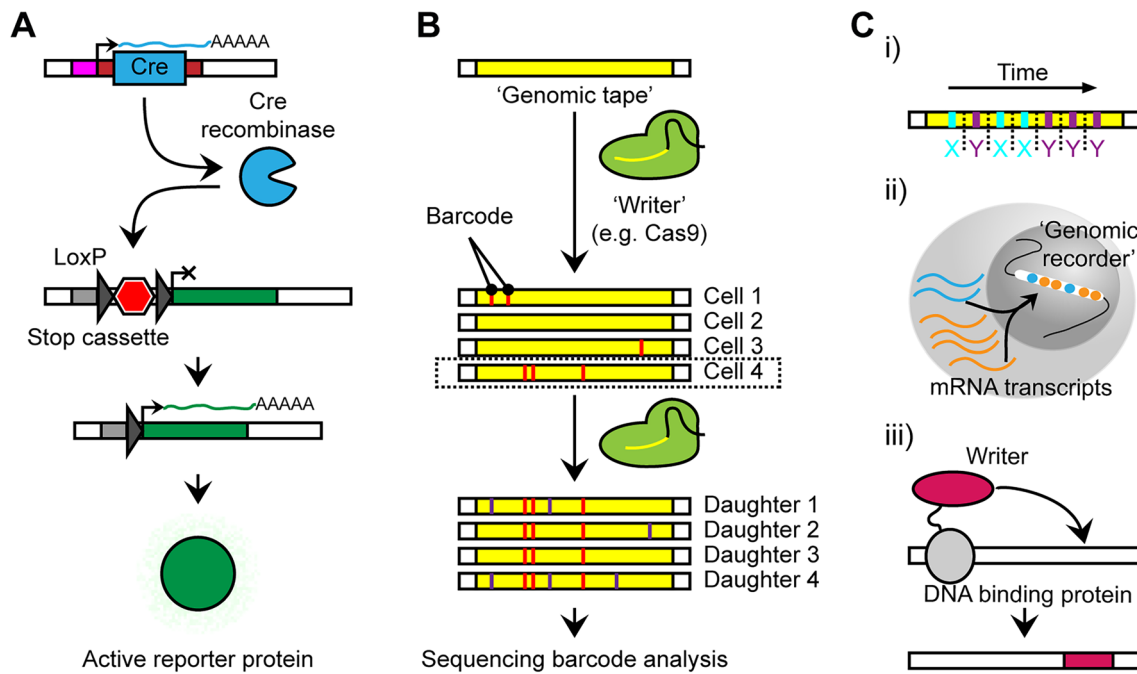

Fig. 5 Types of engineered mosaicism. **A** One of the most utilized methods to track lineages involves the use of fluorescent markers, such as GFP. One possible configuration employs a stop cassette which is flanked by Loxp sites and prevents the transcription of GFP. When Cre recombinase is expressed from a lineage-defined locus, the stop cassette is removed and the GFP is expressed in this cell and its daughters. **B** A more recent innovation used to track lineages employs 'genomic writers' (e.g., Cas9) that are targeted to a defined locus, often denoted as 'genomic tape'. Here these writers can introduce either random or defined mutations that act as genomic barcodes to distinguish cells and their lineages. These barcodes are subsequently read through targeted sequencing. **C** Using writers in combination with other systems, it is possible to further encode the temporal resolution of defined signals (i), the expression status of a cell (ii), or protein binding to genomic regions (iii).

A potential solution is the direct readout of genomic alterations that can distinguish many lineages in parallel based on combinatorial sequence variation. The fundamental idea of such systems is the use of a genomic 'writer' with a controlled or known expression that can introduce genomic changes in a locus that contains 'genomic tape' (Fig. 5B). In 2016, a novel method named GESTALT (genome editing of synthetic target arrays for lineage tracing) utilized CRISPR/Cas9 to barcode cells in this way throughout early development [188]. This technique allows for the sequencing at a later developmental point and the reconstruction of a lineage tree by analyzing the barcoded sequence. A fundamental limitation of GESTALT is the introduction of random insertions and deletions that may interfere with each other and complicate lineage reconstruction. This was addressed by work from the same laboratory through the use of sequential genome editing in a proof-of-concept in cell lines [189]. A similar approach has also been employed in mice [190].

Employing an orthogonal but related concept, Kalhor and colleagues devised a homing CRISPR method in which homing guide RNAs were designed to act on their binding region [191]. The Cas9 enzyme can then introduce a variety of different types of mutations to the binding site of the homing guide RNA to act as cellular barcodes. They applied this method to the developing murine brain, assessing barcodes from the left and right sides of the forebrain, midbrain, and hindbrain. Their results found that commitment to the anterior-posterior axis occurs before the lateral axis; however, they did not resolve this pattern in the neocortex itself. The homing CRISPR technique shows the feasibility of utilizing engineered mosaicism to identify and track lineage, both spatially and temporally. While technically different, the use of Cre recombinase on the complex engineered 'Polylox' locus follows a similar logic [192]: here, combinatorial recombination distinguishes distinct lineages, as was demonstrated in the hematopoietic system. Common to all these techniques is the ability to retrieve the combinatorial and lineage-defining loci through direct sequencing as the genomic tape is known *a priori*. Following the retrieval of the engineered mosaic marks, specialized algorithms allow the reconstruction of lineages similar to genome-wide natural mosaicism.

The two biggest advantages of using such methods are the tunable resolution of lineages based on the activity of the writer and the reduced sequencing cost due to the known mutated genomic tape. However, these described methods are inherently unable to reflect any 'cellular states' other than their initiation condition (e.g., expression of a Cre recombinase). In response to this limitation, several alternative approaches have been developed that allow the recording of such states (Fig. 5C). Fundamentally, these methods aim to reflect features such as gene expression or protein binding to chromatin as a permanent record in the genome.

For instance, a study by Chen and colleagues proposes the driving of multiple writers with distinct signatures by distinct enhancers [193]; this would enable to recording of a temporally resolved sequence of a predetermined number of input signals and use them for lineage tracing. Going one step further, bacterial systems can store copies of expressed RNA in their genome [194–196]. While currently only employed to record highly expressed RNAs, this has the potential to also mark lineages by variable integration of these transcripts. However, these systems have not yet been translated into eukaryotes or employed to distinguish clones of cells. Thus, a genome-wide recorder of transcriptional activity is currently unavailable in mammals. Finally, a method named 'Calling Cards', developed by the Mitra laboratory provides a distinct recording of cellular state [197, 198]. This method marks binding sites of transcription factors through the integration of a permanent transposon into the genome. Therefore, the genome-wide binding of a protein of interest can be assessed based on a customizable genomic scar.

While engineered mosaicism has advantages over natural mosaicism in terms of lineage tracing, there are important applications of the latter in model organisms. For instance, a recent study by Uchimura and colleagues tracked natural mosaicism in somatic and germ cell lineages for lineage tracing [199]. Importantly, their approach allowed them to retrieve features of mutation rates and signatures, which allows the inferral of the mutational environment during development. Moreover, the use of natural mosaicism allows the avoidance of potentially complex breeding strategies to introduce the necessary tunable genomic writers and the genomic tape. As sequencing costs continue to decrease, the most important advantage of engineered mosaicism is its superior resolution. However, it is conceivable to modulate mutation rates through the use of chemical mutagens or genetic backgrounds with impaired genome repair mechanisms. While currently not optimized, the bioinformatic analysis could be adapted for model organisms as needed.

Conclusions

Driven by technological advances, genomic mosaicism research in the brain has progressed significantly in the last decade. This review provides a bird's eye view of current trends in this field, including conceptual definitions, current methodological approaches, and an overview of the impact and utility of genomic mosaicism. As sequencing is one of the main drivers of current discovery, this field benefits from the rapidly decreasing costs of next-generation and—prospectively—third-generation sequencing. This allows researchers interested in lineage tracing questions in the

brain to more easily apply genomic mosaicism approaches. We hope this review will act as a primer for interested parties and allow wider adoption of the here-described concepts.

References

1. Biesecker LG, Spinner NB. A genomic view of mosaicism and human disease. Nat Rev Genet 2013, 14: 307–320.
2. Seplyarskiy VB, Sunyaev S. The origin of human mutation in light of genomic data. Nat Rev Genet 2021, 22: 672–686.
3. Petljak M, Alexandrov LB. Understanding mutagenesis through delineation of mutational signatures in human cancer. Carcinogenesis 2016, 37: 531–540.
4. Chatterjee N, Walker GC. Mechanisms of DNA damage, repair, and mutagenesis. Environ Mol Mutagen 2017, 58: 235–263.
5. Bizzotto S, Walsh CA. Genetic mosaicism in the human brain: From lineage tracing to neuropsychiatric disorders. Nat Rev Neurosci 2022, 23: 275–286.
6. Breuss MW, Antaki D, George RD, Kleiber M, James KN, Ball LL. Autism risk in offspring can be assessed through quantification of male sperm mosaicism. Nat Med 2020, 26: 143–150.
7. Kriegstein A, Alvarez-Buylla A. The glial nature of embryonic and adult neural stem cells. Annu Rev Neurosci 2009, 32: 149–184.
8. Götz M, Huttner WB. The cell biology of neurogenesis. Nat Rev Mol Cell Biol 2005, 6: 777–788.
9. Stiles J, Jernigan TL. The basics of brain development. Neuropsychol Rev 2010, 20: 327–348.
10. Miller MB, Reed HC, Walsh CA. Brain somatic mutation in aging and Alzheimer's disease. Annu Rev Genomics Hum Genet 2021, 22: 239–256.
11. Yang X, Yang C, Zheng X, Xiong L, Tao Y, Wang M, *et al*. MosaicBase: A knowledgebase of postzygotic mosaic variants in noncancer disease-related and healthy human individuals. Genomics Proteomics Bioinform 2020, 18: 140–149.
12. Lindhurst MJ, Sapp JC, Teer JK, Johnston JJ, Finn EM, Peters K, *et al*. A mosaic activating mutation in *AKT1* Associated with the *Proteus* syndrome. N Engl J Med 2011, 365: 611–619.
13. Lim JS, Kim WI, Kang HC, Kim SH, Park AH, Park EK, *et al*. Brain somatic mutations in MTOR cause focal cortical dysplasia type II leading to intractable epilepsy. Nat Med 2015, 21: 395–400.
14. D'Gama AM, Geng Y, Couto JA, Martin B, Boyle EA, LaCoursiere CM, *et al*. Mammalian target of rapamycin pathway mutations cause hemimegalencephaly and focal cortical dysplasia. Ann Neurol 2015, 77: 720–725.
15. Poduri A, Evrony GD, Cai X, Elhosary PC, Beroukhim R, Lehtinen MK, *et al*. Somatic activation of AKT3 causes hemispheric developmental brain malformations. Neuron 2012, 74: 41–48.
16. Lee JH, Huynh M, Silhavy JL, Kim S, Dixon-Salazar T, Heiberg A, *et al*. De novo somatic mutations in components of the PI3K-AKT3-mTOR pathway cause hemimegalencephaly. Nat Genet 2012, 44: 941–945.
17. Chung C, Yang X, Bae T, Vong KI, Mittal S, Donkels C, *et al*. Comprehensive multi-omic profiling of somatic mutations in malformations of cortical development. Nat Genet 2023, 55: 209–220.
18. D'Gama AM. Somatic mosaicism and autism spectrum disorder. Genes 2021, 12: 1699.
19. Jamuar SS, Lam ATN, Kircher M, D'Gama AM, Wang J, Barry BJ, *et al*. Somatic mutations in cerebral cortical malformations. N Engl J Med 2014, 371: 733–743.
20. McConnell MJ, Moran JV, Abyzov A, Akbarian S, Bae T, Cortes-Ciriano I, *et al*. Intersection of diverse neuronal genomes and neuropsychiatric disease: The brain somatic mosaicism network. Science 2017, 356: eaal1641.
21. Bae T, Tomasini L, Mariani J, Zhou B, Roychowdhury T, Franjic D, *et al*. Different mutational rates and mechanisms in human cells at pregastrulation and neurogenesis. Science 2018, 359: 550–555.
22. Brazhnik K, Sun S, Alani O, Kinkhabwala M, Wolkoff AW, Maslov AY, *et al*. Single-cell analysis reveals different age-related somatic mutation profiles between stem and differentiated cells in human liver. Sci Adv 2020, 6: eaax2659.
23. Breuss MW, Yang X, Schlachetzki JCM, Antaki D, Lana AJ, Xu X, *et al*. Somatic mosaicism reveals clonal distributions of neocortical development. Nature 2022, 604: 689–696.
24. Coorens THH, Moore L, Robinson PS, Sanghvi R, Christopher J, Hewinson J, *et al*. Extensive phylogenies of human development inferred from somatic mutations. Nature 2021, 597: 387–392.
25. Lodato MA, Woodworth MB, Lee S, Evrony GD, Mehta BK, Karger A, *et al*. Somatic mutation in single human neurons tracks developmental and transcriptional history. Science 2015, 350: 94–98.
26. Lodato MA, Rodin RE, Bohrson CL, Coulter ME, Barton AR, Kwon M, *et al*. Aging and neurodegeneration are associated with increased mutations in single human neurons. Science 2018, 359: 555–559.
27. Breuss MW, Yang X, Gleeson JG. Sperm mosaicism: Implications for genomic diversity and disease. Trends Genet 2021, 37: 890–902.
28. Jaiswal S. Clonal hematopoiesis and nonhematologic disorders. Blood 2020, 136: 1606–1614.
29. Dou Y, Gold HD, Luquette LJ, Park PJ. Detecting somatic mutations in normal cells. Trends Genet 2018, 34: 545–557.
30. Happle R. Mosaicism in human skin. Arch Dermatol 1993, 129: 1460.
31. Gajecka M. Unrevealed mosaicism in the next-generation sequencing era. Mol Genet Genomics 2016, 291: 513–530.
32. Martincorena I. Somatic mutation and clonal expansions in human tissues. Genome Med 2019, 11: 35.
33. Olafsson S, Anderson CA. Somatic mutations provide important and unique insights into the biology of complex diseases. Trends Genet 2021, 37: 872–881.
34. Anne Hahn. The cellular mitochondrial genome landscape in disease. Trends Cell Biol 2019, 29: 227–240.

35. Park CB, Larsson NG. Mitochondrial DNA mutations in disease and aging. J Cell Biol 2011, 193: 809–818.

36. Cote-L'Heureux A, Maithania YNK, Franco M, Khrapko K. Are some mutations more equal than others? eLife 2023, 12: e87194.

37. Sanchez-Contreras M, Sweetwyne MT, Tsantilas KA, Whitson JA, Campbell MD, Kohrn BF, et al. The multi-tissue landscape of somatic mtDNA mutations indicates tissue-specific accumulation and removal in aging. eLife 2023, 12: e83395.

38. Soong NW, Hinton DR, Cortopassi G, Arnheim N. Mosaicism for a specific somatic mitochondrial DNA mutation in adult human brain. Nat Genet 1992, 2: 318–323.

39. Martínez-Glez V, Tenorio J, Nevado J, Gordo G, Rodríguez-Laguna L, Feito M, et al. A six-attribute classification of genetic mosaicism. Genet Med 2020, 22: 1743–1757.

40. Bernkopf M, Abdullah UB, Bush SJ, Wood KA, Ghaffari S, Giannoulatou E, et al. Personalized recurrence risk assessment following the birth of a child with a pathogenic de novo mutation. Nat Commun 2023, 14: 853.

41. Jónsson H, Sulem P, Arnadottir GA, Pálsson G, Eggertsson HP, Kristmundsdottir S, et al. Multiple transmissions of de novo mutations in families. Nat Genet 2018, 50: 1674–1680.

42. Machiela MJ, Chanock SJ. The ageing genome, clonal mosaicism and chronic disease. Curr Opin Genet Dev 2017, 42: 8–13.

43. Lodato MA, Walsh CA. Genome aging: Somatic mutation in the brain links age-related decline with disease and nominates pathogenic mechanisms. Hum Mol Genet 2019, 28: R197–R206.

44. Yang X, Breuss MW, Xu X, Antaki D, James KN, Stanley V, et al. Developmental and temporal characteristics of clonal sperm mosaicism. Cell 2021, 184: 4772-4783.e15.

45. Ku CS, Loy EY, Salim A, Pawitan Y, Chia KS. The discovery of human genetic variations and their use as disease markers: Past, present and future. J Hum Genet 2010, 55: 403–415.

46. Fan H, Chu JY. A brief review of short tandem repeat mutation. Genomics Proteomics Bioinform 2007, 5: 7–14.

47. Weischenfeldt J, Symmons O, Spitz F, Korbel JO. Phenotypic impact of genomic structural variation: Insights from and for human disease. Nat Rev Genet 2013, 14: 125–138.

48. Nesta AV, Tafur D, Beck CR. Hotspots of human mutation. Trends Genet 2021, 37: 717–729.

49. O'Keefe C, McDevitt MA, Maciejewski JP. Copy neutral loss of heterozygosity: A novel chromosomal lesion in myeloid malignancies. Blood 2010, 115: 2731–2739.

50. Wang D, Kreutzer DA, Essigmann JM. Mutagenicity and repair of oxidative DNA damage: Insights from studies using defined lesions. Mutat Res 1998, 400: 99–115.

51. Alexandrov LB, Stratton MR. Mutational signatures: The patterns of somatic mutations hidden in cancer genomes. Curr Opin Genet Dev 2014, 24: 52–60.

52. Ehrlich M, Norris KF, Wang RY, Kuo KC, Gehrke CW. DNA cytosine methylation and heat-induced deamination. Biosci Rep 1986, 6: 387–393.

53. Wang Z, Zhang T, Wu W, Wu L, Li J, Huang B, et al. Detection and localization of solid tumors utilizing the cancer-type-specific mutational signatures. Front Bioeng Biotechnol 2022, 10: 883791.

54. Koutros S, Rao N, Moore LE, Nickerson ML, Lee D, Zhu B, et al. Targeted deep sequencing of bladder tumors reveals novel associations between cancer gene mutations and mutational signatures with major risk factors. Clin Cancer Res 2021, 27: 3725–3733.

55. Petljak M, Alexandrov LB, Brammeld JS, Price S, Wedge DC, Grossmann S, et al. Characterizing mutational signatures in human cancer cell lines reveals episodic APOBEC mutagenesis. Cell 2019, 176: 1282-1294.e20.

56. Moore L, Cagan A, Coorens THH, Neville MDC, Sanghvi R, Sanders MA, et al. The mutational landscape of human somatic and germline cells. Nature 2021, 597: 381–386.

57. Nachman MW, Crowell SL. Estimate of the mutation rate per nucleotide in humans. Genetics 2000, 156: 297–304.

58. Abecasis GR, Altshuler D, Auton A, Brooks LD, Durbin RM, et al. A map of human genome variation from population-scale sequencing. Nature 2010, 467: 1061–1073.

59. Jónsson H, Sulem P, Kehr B, Kristmundsdottir S, Zink F, Hjartarson E, et al. Parental influence on human germline de novo mutations in 1548 trios from Iceland. Nature 2017, 549: 519–522.

60. Katsonis P, Koire A, Wilson SJ, Hsu TK, Lua RC, Wilkins AD, et al. Single nucleotide variations: Biological impact and theoretical interpretation. Protein Sci 2014, 23: 1650–1666.

61. Lord J, Baralle D. Splicing in the diagnosis of rare disease: Advances and challenges. Front Genet 2021, 12: 689892.

62. Montgomery SB, Goode DL, Kvikstad E, Albers CA, Zhang ZD, Mu XJ, et al. The origin, evolution, and functional impact of short insertion-deletion variants identified in 179 human genomes. Genome Res 2013, 23: 749–761.

63. Domogala DD, Gambin T, Zemet R, Wu CW, Schulze KV, Yang Y, et al. Detection of low-level parental somatic mosaicism for clinically relevant SNVs and indels identified in a large exome sequencing dataset. Hum Genomics 2021, 15: 72.

64. Mousavi N, Shleizer-Burko S, Yanicky R, Gymrek M. Profiling the genome-wide landscape of tandem repeat expansions. Nucleic Acids Res 2019, 47: e90.

65. Chintalaphani SR, Pineda SS, Deveson IW, Kumar KR. An update on the neurological short tandem repeat expansion disorders and the emergence of long-read sequencing diagnostics. Acta Neuropathol Commun 2021, 9: 98.

66. Depienne C, Mandel JL. 30 years of repeat expansion disorders: What have we learned and what are the remaining challenges? Am J Hum Genet 2021, 108: 764–785.

67. Telenius H, Kremer B, Goldberg YP, Theilmann J, Andrew SE, Zeisler J, et al. Somatic and gonadal mosaicism of the Huntington disease gene CAG repeat in brain and sperm. Nat Genet 1994, 6: 409–414.

68. Riggs ER, Andersen EF, Cherry AM, Kantarci S, Kearney H, Patel A, et al. Correction: Technical standards for the interpretation and reporting of constitutional copy-number variants: A joint consensus recommendation of the American college of medical genetics and genomics (ACMG) and the clinical genome resource (ClinGen). Genet Med 2021, 23: 2230.

69. Abyzov A, Mariani J, Palejev D, Zhang Y, Haney MS, Tomasini L, et al. Somatic copy number mosaicism in human skin revealed by induced pluripotent stem cells. Nature 2012, 492: 438–442.

70. McConnell MJ, Lindberg MR, Brennand KJ, Piper JC, Voet T, Cowing-Zitron C, et al. Mosaic copy number variation in human neurons. Science 2013, 342: 632–637.

71. Sherman MA, Rodin RE, Genovese G, Dias C, Barton AR, Mukamel RE, et al. Large mosaic copy number variations confer autism risk. Nat Neurosci 2021, 24: 197–203.

72. Maury EA, Walsh CA. Somatic copy number variants in neuropsychiatric disorders. Curr Opin Genet Dev 2021, 68: 9–17.

73. Sun C, Kathuria K, Emery SB, Kim B, Burbulis IE, Shin JH, et al. Mapping the Complex Genetic Landscape of Human Neurons. bioRxiv 2023: https://doi.org/10.1101/2023.1103.1107.531594.

74. Zilina O, Koltsina M, Raid R, Kurg A, Tonisson N, Salumets A. Somatic mosaicism for copy-neutral loss of heterozygosity and DNA copy number variations in the human genome. BMC Genomics 2015, 16: 703.

75. Kim H, Suyama M. Genome-wide identification of copy neutral loss of heterozygosity reveals its possible association with

spatial positioning of chromosomes. Hum Mol Genet 2023, 32: 1175–1183.

76. Garcia-Linares C, Fernández-Rodríguez J, Terribas E, Mercadé J, Pros E, Benito L, *et al*. Dissecting loss of heterozygosity (LOH) in neurofibromatosis type 1-associated neurofibromas: Importance of copy neutral LOH. Hum Mutat 2011, 32: 78–90.

77. Jozwiak J, Jozwiak S, Wlodarski P. Possible mechanisms of disease development in tuberous sclerosis. Lancet Oncol 2008, 9: 73–79.

78. Loh PR, Genovese G, McCarroll SA. Monogenic and polygenic inheritance become instruments for clonal selection. Nature 2020, 584: 136–141.

79. Lee S, Shin CH, Lee J, Jeong SD, Hong CR, Kim JD, *et al*. Somatic uniparental disomy mitigates the most damaging EFL1 allele combination in Shwachman-Diamond syndrome. Blood 2021, 138: 2117–2128.

80. Faulkner GJ, Garcia-Perez JL. L1 mosaicism in mammals: Extent, effects, and evolution. Trends Genet 2017, 33: 802–816.

81. Suarez NA, Macia A, Muotri AR. LINE-1 retrotransposons in healthy and diseased human brain. Dev Neurobiol 2018, 78: 434–455.

82. Coufal NG, Garcia-Perez JL, Peng GE, Yeo GW, Mu Y, Lovci MT, *et al*. L1 retrotransposition in human neural progenitor cells. Nature 2009, 460: 1127–1131.

83. Zhu X, Zhou B, Pattni R, Gleason K, Tan C, Kalinowski A, *et al*. Machine learning reveals bilateral distribution of somatic L1 insertions in human neurons and glia. Nat Neurosci 2021, 24: 186–196.

84. Upton KR, Gerhardt DJ, Jesuadian JS, Richardson SR, Sánchez-Luque FJ, Bodea GO, *et al*. Ubiquitous L1 mosaicism in hippocampal neurons. Cell 2015, 161: 228–239.

85. Evrony GD, Cai X, Lee E, Cai X, Elhosary PC, Lehmann HS, *et al*. Single-neuron sequencing analysis of L1 retrotransposition and somatic mutation in the human brain. Cell 2012, 151: 483–496.

86. Jordan IK. Evolutionary tinkering with transposable elements. Proc Natl Acad Sci U S A 2006, 103: 7941–7942.

87. Hirschhorn K. Chromosome identification. Annu Rev Med 1973, 24: 67–74.

88. Dolan M. The role of the Giemsa stain in cytogenetics. Biotech Histochem 2011, 86: 94–97.

89. Zhang Y, Zhong M, Zheng D. Chromosomal mosaicism detected by karyotyping and chromosomal microarray analysis in prenatal diagnosis. J Cell Mol Med 2021, 25: 358–366.

90. Bayani J, Squire JA. Fluorescence *in situ* hybridization (FISH). Curr Protocols Cell Biol 2004, 23(1): 22–24.

91. Knoll JH, Lichter P. *In situ* hybridization to metaphase chromosomes and interphase nuclei. Curr Protocol Human Genet 2005, 45(1): 4–3.

92. Hu L, Ru K, Zhang L, Huang Y, Zhu X, Liu H, *et al*. Fluorescence *in situ* hybridization (FISH): An increasingly demanded tool for biomarker research and personalized medicine. Biomark Res 2014, 2: 3.

93. Cantú ES, Eicher DJ, Pai GS, Donahue CJ, Harley RA. Mosaic vs. nonmosaic trisomy 9: Report of a liveborn infant evaluated by fluorescence in situ hybridization and review of the literature. Am J Med Genet 1996, 62: 330–335.

94. Weise A, Liehr T. Fluorescence *in situ* hybridization for prenatal screening of chromosomal aneuploidies. Expert Rev Mol Diagn 2008, 8: 355–357.

95. Reichard KK, Hall BK, Corn A, Foucar MK, Hozier J. Automated analysis of fluorescence *in situ* hybridization on fixed, paraffin-embedded whole tissue sections in B-cell lymphoma. Mod Pathol 2006, 19: 1027–1033.

96. Raj A, van den Bogaard P, Rifkin SA, van Oudenaarden A, Tyagi S. Imaging individual mRNA molecules using multiple singly labeled probes. Nat Methods 2008, 5: 877–879.

97. Urbanek MO, Krzyzosiak WJ. Discriminating RNA variants with single-molecule allele-specific FISH. Mutat Res Rev Mutat Res 2017, 773: 230–241.

98. Pawel Stankiewicz. Use of array CGH in the evaluation of dysmorphology, malformations, developmental delay, and idiopathic mental retardation. Curr Opin Genet Dev 2007, 17: 182–192.

99. Coughlin CR 2nd, Scharer GH, Shaikh TH. Clinical impact of copy number variation analysis using high-resolution microarray technologies: Advantages, limitations and concerns. Genome Med 2012, 4: 80.

100. Scott SA, Cohen N, Brandt T, Toruner G, Desnick RJ, Edelmann L. Detection of low-level mosaicism and placental mosaicism by oligonucleotide array comparative genomic hybridization. Genet Med 2010, 12: 85–92.

101. Vogel I, Vestergaard EM, Lildballe DL, Christensen R, Hoseth GE, Petersen AC, *et al*. Placental mosaicism in the era of chromosomal microarrays. Eur J Med Genet 2020, 63: 103778.

102. Loh PR, Genovese G, Handsaker RE, Finucane HK, Reshef YA, Palamara PF, *et al*. Insights into clonal haematopoiesis from 8, 342 mosaic chromosomal alterations. Nature 2018, 559: 350–355.

103. Maury EA, Sherman MA, Genovese G, Gilgenast TG, Rajarajan P, Flaherty E*, et al.* Schizophrenia-associated somatic copy number variants from 12,834 cases reveal contribution to risk and recurrent, isoform-specific NRXN1 disruptions. medRxiv 2022: https://doi.org/10.1101/2021.1112.1124.21268385.

104. Quail MA, Smith M, Coupland P, Otto TD, Harris SR, Connor TR, *et al*. A tale of three next generation sequencing platforms: Comparison of Ion Torrent, Pacific biosciences and Illumina MiSeq sequencers. BMC Genomics 2012, 13: 341.

105. Bewicke-Copley F, Arjun Kumar E, Palladino G, Korfi K, Wang J. Applications and analysis of targeted genomic sequencing in cancer studies. Comput Struct Biotechnol J 2019, 17: 1348–1359.

106. Griffith M, Miller CA, Griffith OL, Krysiak K, Skidmore ZL, Ramu A, *et al*. Optimizing cancer genome sequencing and analysis. Cell Syst 2015, 1: 210–223.

107. Rodin RE, Dou Y, Kwon M, Sherman MA, D'Gama AM, Doan RN, *et al*. The landscape of somatic mutation in cerebral cortex of autistic and neurotypical individuals revealed by ultra-deep whole-genome sequencing. Nat Neurosci 2021, 24: 611.

108. Bae T, Fasching L, Wang Y, Shin JH, Suvakov M, Jang Y, *et al*. Analysis of somatic mutations in 131 human brains reveals aging-associated hypermutability. Science 2022, 377: 511–517.

109. Wang Y, Bae T, Thorpe J, Sherman MA, Jones AG, Cho S, *et al*. Comprehensive identification of somatic nucleotide variants in human brain tissue. Genome Biol 2021, 22: 92.

110. Bizzotto S, Dou Y, Ganz J, Doan RN, Kwon M, Bohrson CL, *et al*. Landmarks of human embryonic development inscribed in somatic mutations. Science 2021, 371: 1249–1253.

111. Ellis P, Moore L, Sanders MA, Butler TM, Brunner SF, Lee-Six H, *et al*. Reliable detection of somatic mutations in solid tissues by laser-capture microdissection and low-input DNA sequencing. Nat Protoc 2021, 16: 841–871.

112. Rockweiler NB, Ramu A, Nagirnaja L, Wong WH, Noordam MJ, Drubin CW, *et al*. The origins and functional effects of postzygotic mutations throughout the human life span. Science 2023, 380: eabn7113.

113. Koboldt DC. Best practices for variant calling in clinical sequencing. Genome Med 2020, 12: 91.

114. Cibulskis K, Lawrence MS, Carter SL, Sivachenko A, Jaffe D, Sougnez C, *et al*. Sensitive detection of somatic point mutations

in impure and heterogeneous cancer samples. Nat Biotechnol 2013, 31: 213–219.

115. Benjamin D, Sato T, Cibulskis K, Getz G, Stewart C, Lichtenstein L. Calling somatic SNVs and indels with Mutect2. bioRxiv 2019, https://doi.org/10.1101/861054.

116. Dou Y, Kwon M, Rodin RE, Cortés-Ciriano I, Doan R, Luquette LJ, *et al*. Accurate detection of mosaic variants in sequencing data without matched controls. Nat Biotechnol 2020, 38: 314–319.

117. Yang X, Xu X, Breuss MW, Antaki D, Ball LL, Chung C, *et al*. Control-independent mosaic single nucleotide variant detection with DeepMosaic. Nat Biotechnol 2023, 41: 870–877.

118. Sahraeian SME, Liu R, Lau B, Podesta K, Mohiyuddin M, Lam HYK. Deep convolutional neural networks for accurate somatic mutation detection. Nat Commun 2019, 10: 1041.

119. Huang AY, Zhang Z, Ye AY, Dou Y, Yan L, Yang X, *et al*. MosaicHunter: Accurate detection of postzygotic single-nucleotide mosaicism through next-generation sequencing of unpaired, trio, and paired samples. Nucleic Acids Res 2017, 45: e76.

120. Kessler MD, Damask A, O'Keeffe S, Banerjee N, Li D, Watanabe K, *et al*. Common and rare variant associations with clonal haematopoiesis phenotypes. Nature 2022, 612: 301–309.

121. Schmitt MW, Kennedy SR, Salk JJ, Fox EJ, Hiatt JB, Loeb LA. Detection of ultra-rare mutations by next-generation sequencing. Proc Natl Acad Sci USA 2012, 109: 14508–14513.

122. Abascal F, Harvey LMR, Mitchell E, Lawson ARJ, Lensing SV, Ellis P, *et al*. Somatic mutation landscapes at single-molecule resolution. Nature 2021, 593: 405–410.

123. van Veghel-Plandsoen MM, Wouters CH, Kromosoeto JNR, den Ridder-Klünnen MC, Halley DJJ, van den Ouweland AMW. Multiplex ligation-depending probe amplification is not suitable for detection of low-grade mosaicism. Eur J Hum Genet 2011, 19: 1009–1012.

124. Liu W, Smith DI, Rechtzigel KJ, Thibodeau SN, James CD. Denaturing high performance liquid chromatography (DHPLC) used in the detection of germline and somatic mutations. Nucleic Acids Res 1998, 26: 1396–1400.

125. Doan RN, Miller MB, Kim SN, Rodin RE, Ganz J, Bizzotto S, *et al*. MIPP-Seq: Ultra-sensitive rapid detection and validation of low-frequency mosaic mutations. BMC Med Genomics 2021, 14: 47.

126. Xu X, Yang X, Wu Q, Liu A, Yang X, Ye AY, *et al*. Amplicon resequencing identified parental mosaicism for approximately 10% of *de novo* SCN1A mutations in children with dravet syndrome. Hum Mutat 2015, 36: 861–872.

127. Zhou B, Haney MS, Zhu X, Pattni R, Abyzov A, Urban AE. Detection and quantification of mosaic genomic DNA variation in primary somatic tissues using ddPCR: Analysis of mosaic transposable-element insertions, copy-number variants, and single-nucleotide variants. Methods Mol Biol 2018, 1768: 173–190.

128. Gonzalez-Pena V, Natarajan S, Xia Y, Klein D, Carter R, Pang Y, *et al*. Accurate genomic variant detection in single cells with primary template-directed amplification. Proc Natl Acad Sci USA 2021, 118: e2024176118.

129. Chen M, Song P, Zou D, Hu X, Zhao S, Gao S, *et al*. Comparison of multiple displacement amplification (MDA) and multiple annealing and looping-based amplification cycles (MALBAC) in single-cell sequencing. PLoS One 2014, 9: e114520.

130. Choudhury S, Huang AY, Kim J, Zhou Z, Morillo K, Maury EA, *et al*. Somatic mutations in single human cardiomyocytes reveal age-associated DNA damage and widespread oxidative genotoxicity. Nat Aging 2022, 2: 714–725.

131. Cai X, Evrony GD, Lehmann HS, Elhosary PC, Mehta BK, Poduri A, *et al*. Single-cell, genome-wide sequencing identifies clonal somatic copy-number variation in the human brain. Cell Rep 2014, 8: 1280–1289.

132. Davis Bell A, Mello CJ, Nemesh J, Brumbaugh SA, Wysoker A, McCarroll SA. Insights into variation in meiosis from 31, 228 human sperm genomes. Nature 2020, 583: 259–264.

133. Chronister WD, Burbulis IE, Wierman MB, Wolpert MJ, Haakenson MF, Smith ACB, *et al*. Neurons with complex karyotypes are rare in aged human neocortex. Cell Rep 2019, 26: 825-835.e7.

134. Rohrback S, April C, Kaper F, Rivera RR, Liu CS, Siddoway B, *et al*. Submegabase copy number variations arise during cerebral cortical neurogenesis as revealed by single-cell whole-genome sequencing. Proc Natl Acad Sci U S A 2018, 115: 10804–10809.

135. Cagan A, Baez-Ortega A, Brzozowska N, Abascal F, Coorens THH, Sanders MA, *et al*. Somatic mutation rates scale with lifespan across mammals. Nature 2022, 604: 517–524.

136. Huang AY, Li P, Rodin RE, Kim SN, Dou Y, Kenny CJ, *et al*. Parallel RNA and DNA analysis after deep sequencing (PRDD-seq) reveals cell type-specific lineage patterns in human brain. Proc Natl Acad Sci USA 2020, 117: 13886–13895.

137. Jang Y, Fasching L, Bae T, Tomasini L, Schreiner J, Szekely A, *et al*. Efficient reconstruction of cell lineage trees for cell ancestry and cancer. Nucleic Acids Res 2023, 51: e57.

138. Fasching L, Jang Y, Tomasi S, Schreiner J, Tomasini L, Brady MV, *et al*. Early developmental asymmetries in cell lineage trees in living individuals. Science 2021, 371: 1245–1248.

139. Hazen JL, Faust GG, Rodriguez AR, Ferguson WC, Shumilina S, Clark RA, *et al*. The complete genome sequences, unique mutational spectra, and developmental potency of adult neurons revealed by cloning. Neuron 2016, 89: 1223–1236.

140. Abyzov A, Tomasini L, Zhou B, Vasmatzis N, Coppola G, Amenduni M, *et al*. One thousand somatic SNVs per skin fibroblast cell set baseline of mosaic mutational load with patterns that suggest proliferative origin. Genome Res 2017, 27: 512–523.

141. Bohrson CL, Barton AR, Lodato MA, Rodin RE, Luquette LJ, Viswanadham VV, *et al*. Linked-read analysis identifies mutations in single-cell DNA-sequencing data. Nat Genet 2019, 51: 749–754.

142. D'Gama AM, Walsh CA. Somatic mosaicism and neurodevelopmental disease. Nat Neurosci 2018, 21: 1504–1514.

143. Jourdon A, Fasching L, Scuderi S, Abyzov A, Vaccarino FM. The role of somatic mosaicism in brain disease. Curr Opin Genet Dev 2020, 65: 84–90.

144. Poduri A, Evrony GD, Cai X, Walsh CA. Somatic mutation, genomic variation, and neurological disease. Science 2013, 341: 1237758.

145. Iourov IY, Vorsanova SG, Kurinnaia OS, Kutsev SI, Yurov YB. Somatic mosaicism in the diseased brain. Mol Cytogenet 2022, 15: 45.

146. Thorpe J, Osei-Owusu IA, Avigdor BE, Tupler R, Pevsner J. Mosaicism in human health and disease. Annu Rev Genet 2020, 54: 487–510.

147. Jansen LA, Mirzaa GM, Ishak GE, O'Roak BJ, Hiatt JB, Roden WH, *et al*. PI3K/AKT pathway mutations cause a spectrum of brain malformations from megalencephaly to focal cortical dysplasia. Brain 2015, 138: 1613–1628.

148. Rivière JB, Mirzaa GM, O'Roak BJ, Beddaoui M, Alcantara D, Conway RL, *et al*. De novo germline and postzygotic mutations in AKT3, PIK3R2 and PIK3CA cause a spectrum of related megalencephaly syndromes. Nat Genet 2012, 44: 934–940.

149. Sim JC, Scerri T, Fanjul-Fernández M, Riseley JR, Gillies G, Pope K, *et al*. Familial cortical dysplasia caused by mutation in the mammalian target of rapamycin regulator NPRL3. Ann Neurol 2016, 79: 132–137.

150. Scheffer IE, Heron SE, Regan BM, Mandelstam S, Crompton DE, Hodgson BL, *et al*. Mutations in mammalian target of rapamycin

regulator*DEPDC5*cause focal epilepsy with brain malformations. Ann Neurol 2014, 75: 782–787.

151. Baldassari S, Ribierre T, Marsan E, Adle-Biassette H, Ferrand-Sorbets S, Bulteau C, *et al*. Dissecting the genetic basis of focal cortical dysplasia: A large cohort study. Acta Neuropathol 2019, 138: 885–900.

152. D'Gama AM, Woodworth MB, Hossain AA, Bizzotto S, Hatem NE, LaCoursiere CM, *et al*. Somatic mutations activating the mTOR pathway in dorsal telencephalic progenitors cause a continuum of cortical dysplasias. Cell Rep 2017, 21: 3754–3766.

153. Beck JA, Poulter M, Campbell TA, Uphill JB, Adamson G, Geddes JF, *et al*. Somatic and germline mosaicism in sporadic early-onset Alzheimer's disease. Hum Mol Genet 2004, 13: 1219–1224.

154. Krupp DR, Barnard RA, Duffourd Y, Evans SA, Mulqueen RM, Bernier R, *et al*. Exonic mosaic mutations contribute risk for autism spectrum disorder. Am J Hum Genet 2017, 101: 369–390.

155. Lim ET, Uddin M, De Rubeis S, Chan Y, Kamumbu AS, Zhang X, *et al*. Rates, distribution and implications of postzygotic mosaic mutations in autism spectrum disorder. Nat Neurosci 2017, 20: 1217–1224.

156. Freed D, Pevsner J. The contribution of mosaic variants to autism spectrum disorder. PLoS Genet 2016, 12: e1006245.

157. Dou Y, Yang X, Li Z, Wang S, Zhang Z, Ye AY, *et al*. Postzygotic single-nucleotide mosaicisms contribute to the etiology of autism spectrum disorder and autistic traits and the origin of mutations. Hum Mutat 2017, 38: 1002–1013.

158. Spencer Chapman M, Ranzoni AM, Myers B, Williams N, Coorens THH, Mitchell E, *et al*. Lineage tracing of human development through somatic mutations. Nature 2021, 595: 85–90.

159. Nott A, Schlachetzki JCM, Fixsen BR, Glass CK. Nuclei isolation of multiple brain cell types for omics interrogation. Nat Protoc 2021, 16: 1629–1646.

160. Bergles DE, Richardson WD. Oligodendrocyte development and plasticity. Cold Spring Harb Perspect Biol 2015, 8: a020453.

161. Bayraktar OA, Fuentealba LC, Alvarez-Buylla A, Rowitch DH. Astrocyte development and heterogeneity. Cold Spring Harb Perspect Biol 2014, 7: a020362.

162. Yuan B, Wang M, Wu X, Cheng P, Zhang R, Zhang R, *et al*. Identification of *de novo* mutations in the Chinese autism spectrum disorder cohort *via* whole-exome sequencing unveils brain regions implicated in autism. Neurosci Bull 2023, 7: 1–12.

163. Li X, Liu G, Yang L, Li Z, Zhang Z, Xu Z, *et al*. Decoding cortical glial cell development. Neurosci Bull 2021, 37: 440–460.

164. Kaplanis J, Ide B, Sanghvi R, Neville M, Danecek P, Coorens T, *et al*. Genetic and chemotherapeutic influences on germline hypermutation. Nature 2022, 605: 503–508.

165. Gao Z, Moorjani P, Sasani TA, Pedersen BS, Quinlan AR, Jorde LB, *et al*. Overlooked roles of DNA damage and maternal age in generating human germline mutations. Proc Natl Acad Sci USA 2019, 116: 9491–9500.

166. Helleday T, Eshtad S, Nik-Zainal S. Mechanisms underlying mutational signatures in human cancers. Nat Rev Genet 2014, 15: 585–598.

167. Alexandrov LB, Kim J, Haradhvala NJ, Huang MN, Ng AWT, Wu Y, *et al*. The repertoire of mutational signatures in human cancer. Nature 2020, 578: 94–101.

168. Alexandrov LB, Nik-Zainal S, Wedge DC, Campbell PJ, Stratton MR. Deciphering signatures of mutational processes operative in human cancer. Cell Rep 2013, 3: 246–259.

169. Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio SAJR, Behjati S, Biankin AV, *et al*. Signatures of mutational processes in human cancer. Nature 2013, 500: 415–421.

170. Koh G, Degasperi A, Zou X, Momen S, Nik-Zainal S. Mutational signatures: Emerging concepts, caveats and clinical applications. Nat Rev Cancer 2021, 21: 619–637.

171. Zhivagui M, Hoda A, Valenzuela N, Yeh YY, Dai J, He Y, *et al*. DNA damage and somatic mutations in mammalian cells after irradiation with a nail polish dryer. Nat Commun 2023, 14: 276.

172. Alexandrov LB, Ju YS, Haase K, Van Loo P, Martincorena I, Nik-Zainal S, *et al*. Mutational signatures associated with tobacco smoking in human cancer. Science 2016, 354: 618–622.

173. Pich O, Muiños F, Lolkema MP, Steeghs N, Gonzalez-Perez A, Lopez-Bigas N. The mutational footprints of cancer therapies. Nat Genet 2019, 51: 1732–1740.

174. Aguilera G, Colín-González AL, Rangel-López E, Chavarría A, Santamaría A. Redox signaling, neuroinflammation, and neurodegeneration. Antioxid Redox Signal 2018, 28: 1626–1651.

175. Tovar-y-Romo LB, Penagos-Puig A, Ramírez-Jarquín JO. Endogenous recovery after brain damage: Molecular mechanisms that balance neuronal life/death fate. J Neurochem 2016, 136: 13–27.

176. McKinnon PJ. Maintaining genome stability in the nervous system. Nat Neurosci 2013, 16: 1523–1529.

177. Verheijen BM, Vermulst M, van Leeuwen FW. Somatic mutations in neurons during aging and neurodegeneration. Acta Neuropathol 2018, 135: 811–826.

178. Lodato MA, Ziegenfuss JS. The two faces of DNA oxidation in genomic and functional mosaicism during aging in human neurons. Front Aging 2022, 3: 991460.

179. Miller MB, Huang AY, Kim J, Zhou Z, Kirkham SL, Maury EA, *et al*. Somatic genomic changes in single Alzheimer's disease neurons. Nature 2022, 604: 714–722.

180. Coorens THH, Oliver TRW, Sanghvi R, Sovio U, Cook E, Vento-Tormo R, *et al*. Inherent mosaicism and extensive mutation of human placentas. Nature 2021, 592: 80–85.

181. Ma J, Shen Z, Yu YC, Shi SH. Neural lineage tracing in the mammalian brain. Curr Opin Neurobiol 2018, 50: 7–16.

182. Hsu YC. Theory and practice of lineage tracing. Stem Cells 2015, 33: 3197–3204.

183. Walsh C, Cepko CL. Clonal dispersion in proliferative layers of developing cerebral cortex. Nature 1993, 362: 632–635.

184. Walsh C, Cepko CL. Clonally related cortical cells show several migration patterns. Science 1988, 241: 1342–1345.

185. Walsh C, Cepko CL. Widespread dispersion of neuronal clones across functional regions of the cerebral cortex. Science 1992, 255: 434–440.

186. Livet J, Weissman TA, Kang H, Draft RW, Lu J, Bennis RA, *et al*. Transgenic strategies for combinatorial expression of fluorescent proteins in the nervous system. Nature 2007, 450: 56–62.

187. Zong H, Espinosa JS, Su HH, Muzumdar MD, Luo L. Mosaic analysis with double markers in mice. Cell 2005, 121: 479–492.

188. McKenna A, Findlay GM, Gagnon JA, Horwitz MS, Schier AF, Shendure J. Whole-organism lineage tracing by combinatorial and cumulative genome editing. Science 2016, 353: aaf7907.

189. Choi J, Chen W, Minkina A, Chardon FM, Suiter CC, Regalado SG, *et al*. A time-resolved, multi-symbol molecular recorder via sequential genome editing. Nature 2022, 608: 98–107.

190. Bowling S, Sritharan D, Osorio FG, Nguyen M, Cheung P, Rodriguez-Fraticelli A, *et al*. An engineered CRISPR-Cas9 mouse line for simultaneous readout of lineage histories and gene expression profiles in single cells. Cell 2020, 181: 1410-1422.e27.

191. Kalhor R, Kalhor K, Mejia L, Leeper K, Graveline A, Mali P, *et al*. Developmental barcoding of whole mouse via homing CRISPR. Science 2018, 361: eaat9804.

192. Pei W, Feyerabend TB, Rössler J, Wang X, Postrach D, Busch K, *et al*. Polylox barcoding reveals haematopoietic stem cell fates realized *in vivo*. Nature 2017, 548: 456–460.

193. Chen W, Choi J, Nathans JF, Agarwal V, Martin B, Nichols E, *et al*. Multiplex genomic recording of enhancer and signal transduction activity in mammalian cells. bioRxiv 2021: https://doi.org/10.1101/2021.1111.1105.467434.

194. Lear SK, Shipman SL. Molecular recording: transcriptional data collection into the genome. Curr Opin Biotechnol 2023, 79: 102855.

195. Schmidt F, Cherepkova MY, Platt RJ. Transcriptional recording by CRISPR spacer acquisition from RNA. Nature 2018, 562: 380–385.

196. Bhattarai-Kline S, Lear SK, Fishman CB, Lopez SC, Lockshin ER, Schubert MG, *et al*. Recording gene expression order in DNA by CRISPR addition of retron barcodes. Nature 2022, 608: 217–225.

197. Moudgil A, Wilkinson MN, Chen X, He J, Cammack AJ, Vasek MJ, *et al*. Self-reporting transposons enable simultaneous read-out of gene expression and transcription factor binding in single cells. Cell 2020, 182: 992-1008.e21.

198. Wang H, Johnston M, Mitra RD. Calling cards for DNA-binding proteins. Genome Res 2007, 17: 1202–1209.

199. Uchimura A, Matsumoto H, Satoh Y, Minakuchi Y, Wakayama S, Wakayama T, *et al*. Early embryonic mutations reveal dynamics of somatic and germ cell lineages in mice. Genome Res 2022, 32: 945–955.