

Review

Genomic Mosaicism Formed by Somatic Variation in the Aging and Diseased Brain

Isabel Costantino ^{1,2,†}, Juliet Nicodemus ^{1,2,†} and Jerold Chun ^{1,*}

¹ Translational Neuroscience Initiative, Sanford Burnham Prebys Medical Discovery Institute, La Jolla, CA 92037, USA; icostantino@sbpdiscovery.org (I.C.); jnicodemus@sbpdiscovery.org (J.N.)

² Neurosciences Graduate Program, School of Medicine, University of California San Diego, La Jolla, CA 92093, USA

* Correspondence: jchun@sbpdiscovery.org; Tel.: +1-(858)-646-3100

† These authors contributed equally to this work.

Abstract: Over the past 20 years, analyses of single brain cell genomes have revealed that the brain is composed of cells with myriad distinct genomes: the brain is a genomic mosaic, generated by a host of DNA sequence-altering processes that occur somatically and do not affect the germline. As such, these sequence changes are not heritable. Some processes appear to occur during neurogenesis, when cells are mitotic, whereas others may also function in post-mitotic cells. Here, we review multiple forms of DNA sequence alterations that have now been documented: aneuploidies and aneusomies, smaller copy number variations (CNVs), somatic repeat expansions, retrotransposons, genomic cDNAs (gencDNAs) associated with somatic gene recombination (SGR), and single nucleotide variations (SNVs). A catch-all term of DNA content variation (DCV) has also been used to describe the overall phenomenon, which can include multiple forms within a single cell's genome. A requisite step in the analyses of genomic mosaicism is ongoing technology development, which is also discussed. Genomic mosaicism alters one of the most stable biological molecules, DNA, which may have many repercussions, ranging from normal functions including effects of aging, to creating dysfunction that occurs in neurodegenerative and other brain diseases, most of which show sporadic presentation, unlinked to causal, heritable genes.

Keywords: Alzheimer's disease; amyotrophic lateral sclerosis; aneuploidy; copy number variation; Parkinson's disease; repeat expansion; retrotransposons; single-nucleotide variation; somatic gene recombination; somatic variation



Citation: Costantino, I.; Nicodemus, J.; Chun, J. Genomic Mosaicism Formed by Somatic Variation in the Aging and Diseased Brain. *Genes* **2021**, *12*, 1071. <https://doi.org/10.3390/genes12071071>

Academic Editors: Concetta Crisafulli and Marco Calabrò

Received: 2 June 2021

Accepted: 12 July 2021

Published: 14 July 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Beginning with the writings of Ramón y Cajal around the start of the 20th century [1,2], a striking characteristic of the vertebrate nervous system was identified through documentation of the enormous diversity and complexity of its constituent cells, which increases with age [3,4], wherein no two cells are truly identical. The molecular mechanisms underlying this diversity remain only partially known; however, they must extend beyond morphology to physiological and functional diversity, as epitomized by the complex visual system [5–8] and single-cell physiologies [9]. Elements of cellular and functional diversity likely underlie most, if not all, of the brain's activities including those with temporal stability, such as long-term memory [10–12]. These properties of the brain have implicated possible diverse changes to molecules involved with stable, biological information storage, particularly affecting DNA.

The notion that genomic DNA sequence changes—distinct from epigenetic changes that do not alter sequence [13]—might underlie the normal complexity of the nervous system emerged in the 1960s during theoretical explanations of generating antibody diversity [14], which initiated discussions and the search for evidence that DNA sequences could vary amongst cells from the same brain. This possibility has now been born-out in many studies that demonstrated genomically mosaic brains are composed of cells with

distinct DNA sequences. Critically, these DNA sequence changes occur somatically rather than affecting the germline and are not heritable; thus, the term “genomic” rather than “genetic” [15] mosaicism is used to underscore somatic alterations lacking heritability.

It is notable that analyses of single-cell genomes have required hand-in-hand technology development, which in part explains the prolonged gap between initial hypothesis generation in the 1960s and actual identification of genomic mosaicism. Twenty years ago, the first evidence for pervasive genomic changes in the brain was reported through chromosomal aneuploidies and aneusomies [16]. In the ensuing years, a range of DNA sequence forms, altered within single cells, has also been reported, from the largest form as aneuploidies and aneusomies, to smaller copy number variants (CNVs), somatic repeat expansion, retrotransposons, genomic cDNAs (gencDNAs), and single nucleotide variations (SNVs) (Figure 1). These elements are reviewed here as they relate to the normal, aging, and neurodegenerative diseased brain. The critical role for technology that is continuously improving is also discussed, along with limitations. Genomic mosaicism appears to change with aging and disease, which has both mechanistic and therapeutic implications.

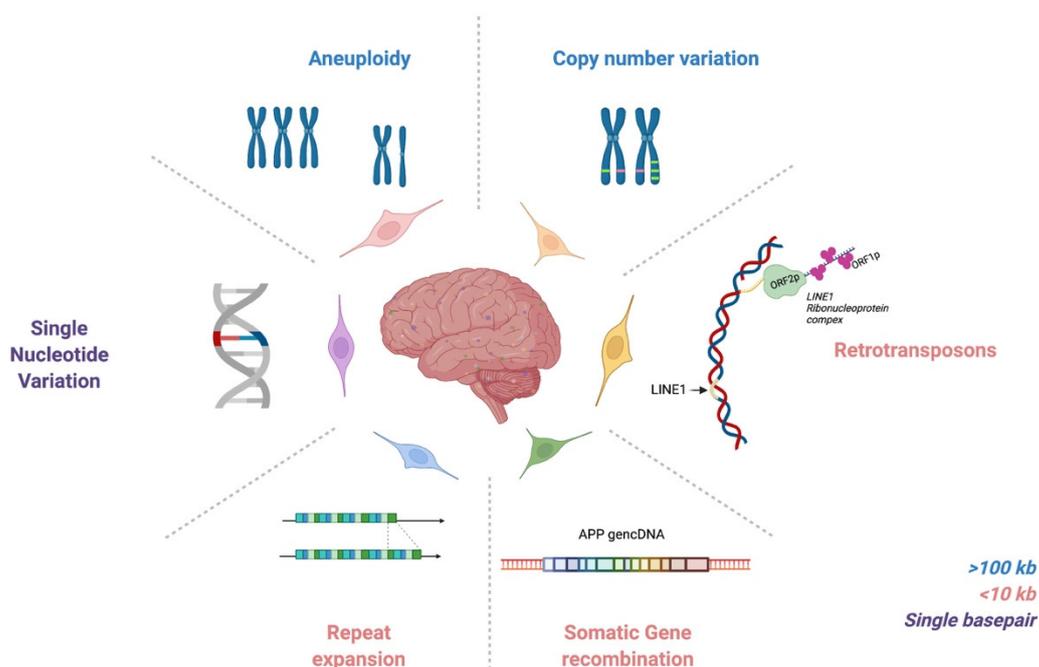


Figure 1. Neural somatic DNA content variation (DCV) encompasses many forms and sizes of somatic changes. Aneuploidies/aneusomies and copy number variations (CNVs) typically reflect large genomic gains/losses that may affect many hundreds of genes. Retrotransposons and somatic gene recombination are reverse transcriptase-mediated insertions of nucleic acids into the genome. Repeat expansions reflect increases in length of unstable, small repeating sequences. Single nucleotide variations (SNVs) are the smallest, but most frequent, DNA changes. They encompass changes from one base to another or insertions/deletions of a single base. Together, these somatic changes create genomic mosaicism throughout the human brain.

2. DNA Content Variation & Aneuploidy

2.1. DNA Content Variation & Aneuploidy in Health and Aging

The first form of neural somatic variation identified was mosaic aneuploidy among mitotic neural progenitor cells [16]. Aneuploidy is defined as the gain and/or loss of chromosomes from the euploid complement. Aneusomy—where a partial chromosomal assessment is made without knowledge of all other chromosomes—was identified in postmitotic neurons of mouse and human using spectral karyotyping and fluorescence in situ hybridization (FISH) [17–21]. The literature tends to be imprecise about the use of the terms “aneuploidy” vs. “aneusomy,” and for the remainder of this review, they will be used synonymously, with the proviso that they in fact have distinct meanings. In the

adult brain, FISH studies estimated that ~10% of cells are aneuploid [21–26]. Iourov et al. utilized FISH and interphase chromosome-specific multicolor banding to establish a rate of aneuploidy in 12 non-diseased brains. While different chromosomes had comparative levels of aneusomy (0.4–1.2%), mean levels varied widely amongst chromosomes and individuals, with chromosome loss more common than gain [21]. However, ~10% may be an overestimate of aneuploidy. Knouse et al. utilized single cell whole genome sequencing (scWGS) of 89 cells from the frontal lobe of four individuals (48–70 years). Of those, only two cells were aneuploid (Chr22 monosomy and Chr18 trisomy). Therefore, they concluded that the aneuploidy rate in the adult human brain is 2.2% [27]. This rate of aneuploidy has since been corroborated in other scWGS studies of non-diseased brains [28], with the caveat of very limited sample sizes relative to the 170 billion cells of the human brain.

Aneuploidy, along with other substantial somatic variants, such as retrotransposons and copy number variants, can be captured by DNA content variation (DCV) [25,29]. DCV appears widespread within the non-diseased brain. A greater DNA content in cortical neurons as compared to cortical glia, cerebellar cells, and lymphocytes, has been demonstrated using propidium iodide (PI) flow cytometry [29]. Slide based cytometry of multiple cortical regions from healthy individuals revealed that 11.5% of neurons show increased DNA content above diploid level, with no major regional differences within the cortex (non-cortical areas such as the cerebellum were not examined) [30]. The frequency of neurons with high DCV declines with age, indicating that increased DCV might compromise neuronal viability in the aging brain or that high DCV may predispose individuals to aging-related neurodegenerative diseases [30], a possibility supported by increased DCV observed in sporadic Alzheimer's disease (SAD) [31].

There are multiple mechanisms through which somatic DCV can occur in the brain. DCV and aneuploidy can have a developmental origin, resulting from mitotic failures in development. Alternatively, postmitotic neurons in the adult brain can synthesize excess DNA de novo through dysfunctional neuronal DNA repair [32–34]. Many questions about the prevalence and mechanism of formation of DCV and aneuploidy in the normal, aging, and diseased brain remain, requiring further examination through multimodal experimental approaches. A compilation of found somatic variants by disease type and experimental methodology is available in Appendix A.

2.2. DNA Content Variation & Aneuploidy in Neurodegenerative Disease

Changes in DCV and aneuploidy have been reported in multiple neurodegenerative diseases, including ataxia-telangiectasia (A-T) and Alzheimer's disease (AD) [21,25,35,36]. A-T is an autosomal recessive primary immunodeficiency disease caused by mutations in ataxia-telangiectasia mutated kinase (*ATM*), a gene critical in preventing deleterious non-homologous end joining [37]. Patients present in childhood with multiple nervous and immune system disorders, including progressive cerebellar ataxia, oculocutaneous telangiectasia, immunodeficiencies, metabolic diseases, and increased susceptibility to malignancies [38]. Iourov et al. observed a 2–3-fold increase in stochastic aneuploidy across different chromosomes in neurons of the cerebellum and cortex of the A-T brain using FISH [24]. Furthermore, the cerebellum in A-T demonstrated a 5–20-fold increase in DNA double-strand breaks and aneuploidy affecting chromosomes 14, 7, and X. These data, combined with previous research in *ATM* knockout mice demonstrating aneuploidy in 40% of the mouse brain, support the hypothesis that *ATM* deficiency may lead to reduced developmental clearance of aneuploid neural cells [39]. Iourov et al. thus proposed that A-T is a disease of selective mosaic genomic instability, with neuronal aneuploidization in the A-T brain as a secondary genetic mechanism contributing to A-T brain pathology [21,24,40].

Iourov et al. also examined the rates of aneuploidy in AD [21]. AD is the most common cause of dementia. It is characterized by progressive cognition and memory deficits, with classic histopathological hallmarks of cortical atrophy, neuronal cell death, and stereotyped amyloid beta plaque and neurofibrillary tangle accumulation [41–43]. While general aneuploidy in the AD brain was not found to be significantly increased compared to

controls, chromosome 21 aneuploidy (both monosomy and trisomy) was reported to be dramatically increased (10-fold), which lends support for the hypothesis that sporadic AD may be associated with trisomy 21 [21,35]. However, these data contrast with both FISH and scWGS studies, in which there was no observed selective gain of chromosome 21 in neurons of AD patients compared to controls [28,31,44]; technical differences including limited sampling could account for these differences.

Arendt et al., examined cortical neurons in AD patients and controls via slide-based cytometry and found DNA content was increased 2–3-fold in AD patients compared to controls. Furthermore, DCV was highest in regions of selective vulnerability, such as the entorhinal cortex, and corresponded with earlier age of death in AD patients (60–70 years old versus 80–90 years old) [45]. This phenomenon is also reported in controls, in which DCV decreases with age [30]. One hypothesis is that increased DCV contributes to higher disease burden, therefore resulting in an earlier age of death in AD patients. Increases in DCV in AD cortices as compared to cerebellum and controls have also been seen in PI flow cytometry and scWGS [31].

3. Copy Number Variation

3.1. Somatic Copy Number Variation in Health & Aging

Somatic CNV occurs when the genome experiences gains and/or losses that change the number of copies of a particular gene amongst cells within the same individual. The presence of unique or shared CNVs across cells or tissues indicates that they likely arose within different stages of development (Figure 2). Similar to aneuploidies, the methodology used for detection can greatly alter the conclusions drawn about brain CNVs. Detected CNV size minimum and acceptable rates of variability in read depth may contribute to false calls and differences in reported CNV frequency [46,47]. With strict parameters, somatic CNVs > 1 Mb are estimated in ~9% of brain cells [46]. Stringent reanalysis of multiple single-cell datasets estimates that ~7% of neural genomes contain somatic CNVs ranging in size from 2.9 to 159.1 Mb with significant interindividual variability (0–24% of genomes containing somatic CNVs) [48]. Compared to non-neurons, neurons have more and larger CNVs that cover more territory within the genome [48,49].

Brain and regional CNVs and CNV rates can be identified through comparative analyses of brain regions and peripheral tissues. One method is array comparative genomic hybridization (CGH), where genomic DNA of experimental samples and references are labeled with different fluorochromes and hybridized onto an array of immobilized DNA targets. Differences in fluorescence intensities from a reference are used to identify genomic gains and/or losses [50,51]. An array-CGH study of CNVs > 50 kB from healthy subjects (including brain cortex, pons, and cerebellum) showed a single CNV in pons, but not the cerebellum or cortex within an individual [52]. Array-CGH analysis also revealed high heterogeneity of CNVs within tissues of a single individual. Of the 75% of CNVs that were CNS-specific (versus lymphocytes), 43% were unique to the frontal lobe and 32% to the cerebellum [53]. Across brain regions, deletions were more common than duplications in both young and aged adults [17,46,54]. However, the cerebellum was shown to have significantly larger deletions than the frontal lobe [53]. This is in agreement with the finding of increased DCV in the frontal lobe compared to the cerebellum [29], but the detected CNV gains [17,46,48,54,55] do not directly explain the estimated gain of ~250 Mb in the DNA content shown by Westra et al. [29]. As with aneuploidies, aged individuals have fewer neurons with CNVs, indicating that the presence of CNVs may affect the fitness of cells [48,56]. However, significantly more CNVs have been found in the aged brain compared to blood within the same individual [53], indicating that selective pressure against genomic mutations may still be less in the brain than in other tissues.

The genomic location of somatic CNVs provide clues about the mechanism of generation. CNVs and their boundaries are enriched in repetitive sequences (SINES/LINES, repeat elements, and noncoding RNAs) and telomeres [46]. Enrichment hotspots have been

described in multiple studies, including potential enrichment within long genes such as *GPC6*, *NRNX3*, and *RBFOX1* [27,48].

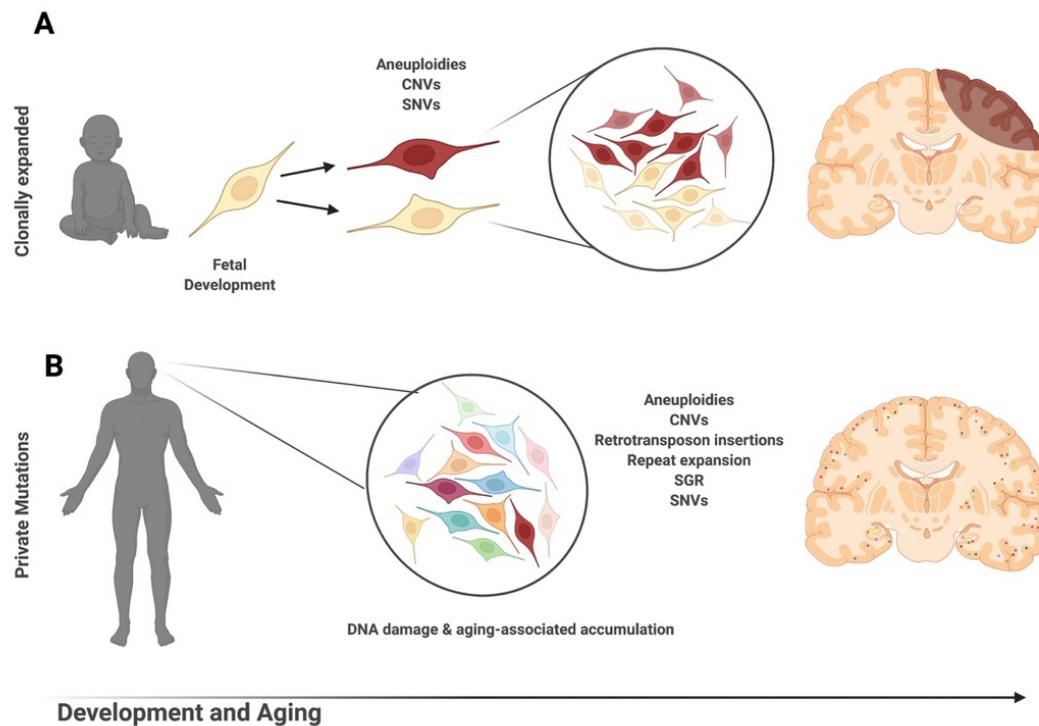


Figure 2. Clonal versus private somatic mutations in the human brain. (A) Somatic mutations that occur at different timepoints within embryogenesis or neurogenesis will result in clonal mutations shared by the brain and other body tissues or between brain cells of a single lineage. Clonal somatic aneuploidies, copy number variations (CNVs), and single-nucleotide variations (SNVs) have been described between the brain and peripheral lymphocytes (occurring early in embryogenesis), between the brain and other ectodermal-derived tissues (occurring after trilaminar disc formation), and between separate brain regions (occurring during neurogenesis). (B) Private or unique somatic mutations are unique to a single neural cell and accumulate throughout one’s lifespan. These somatic events may be independent of replication, as evidenced by their presence in post-mitotic cells. Aneuploidies, CNVs, retrotransposon insertions, repeat expansions, somatic gene recombination (SGR), and SNVs have all been described as occurring in post-mitotic neurons.

3.2. Somatic Copy Number Variation in Neurodegenerative Disease

Increased copy number of neurodegenerative disease-related genes, even sparsely present within a brain, may have a significant impact on pathogenesis or heterogenous presentation of neurodegenerative diseases.

Parkinson’s disease (PD) and multiple system atrophy (MSA) are α -synucleinopathies [57]. PD is characterized by resting tremor, rigidity, bradykinesia, and other signs and symptoms associated with loss of dopaminergic neurons in the substantia nigra (SN) and other regions [58]. MSA is a rare and fatal disease characterized by combinations of parkinsonism, autonomic failure, ataxia, and pyramidal tract dysfunction associated with glial inclusions and cell loss [59]. In both familial and sporadic PD, copy number changes in *SNCA* have been described [60]. *SNCA* is located within a fragile genomic region that may be especially vulnerable to breaks during DNA replication [61]. This presents the intriguing question of whether somatic copy number changes may underlie sporadic disease that would be undetectable by standard lymphocyte-based genomic testing. One study of two individuals with early-onset PD using paired buccal and leukocyte samples demonstrated no *SNCA* gains in leukocytes, but a duplicate or triplicate copy number was found in 82% and 43% of oral mucosal cells [62]. The brain was not assessed in this study. Oral mucosa and brain tissue are of ectodermal origin as opposed to lymphocytes, which are of mesodermal origin, indicating that clonal, disease-relevant genomic changes

in early development may be obscured when using blood as the predominant sample for genotyping.

Within dopaminergic neurons of the SN, there is evidence of somatic *SNCA* gains via FISH in PD patients [63]. While *SNCA* copy number increases were only observed in a small population of dopaminergic neurons (<1%), this could still be highly significant. SN dopaminergic neurons are significantly depleted in end-stage disease. Therefore, the ~1% of affected dopaminergic neurons may reflect only the surviving neurons within the region. The number of *SNCA* gains was negatively associated with onset age. Within this same study, the control case with the highest gains (1.87% of dopaminergic neurons) demonstrated incidental α -synuclein pathology with rare Lewy neurites. In a separate study, FISH analysis of cingulate cortex neurons demonstrated increased *SNCA* copy gains in MSA (2.8%) and PD (2.15%) compared to controls (1.12%) [64]. Furthermore, synuclein inclusions were more common in cells with CNVs than without in PD (22.1% vs. 5.7% neuromelanin positive) and in MSA (31.2% vs. 15.9% in olig2 positive). Whole-genome amplification and sequencing of two MSA cases in the same study found only one pontine neuron with a gain including the *SNCA* region [64]. Interestingly, one study in diseased brainstem and putamen (MSA) found non-neurons to have nearly all gains (>95%) compared to ~45% in neurons [64].

Demonstrating the very different results that may occur across techniques, array-CGH used to detect high level copy number changes of PD-relevant genes (*SNCA* included) detected no evidence of copy number changes [63]. A genome array study of frontal lobe in PD patients also demonstrated no specific gains or losses in *SNCA* but did detect CNVs unique to PD brains (versus controls) in other PD candidate genes (*BCL2*, *NRSN1*, and *RYR2*) [54].

In AD, there is also interest in the copy number of disease-related genes, especially *APP* (amyloid precursor protein), since increased copies have been linked to Down syndrome-associated dementia [65]. In a study of 1511 whole brain exomes, a single AD patient was found with a triplication of the *APP* locus [66]. Increases in the *APP* copy number in AD cortical nuclei versus control using qPCR and PNA-FISH with *APP* probes have been reported, with no associated increase in chromosome 21 [31]. Therefore, it appears that a non-aneuploidy mechanism is driving *APP* CNVs in these neurons. However, targeted enrichment of AD-related genes in AD entorhinal cortex, with a sensitivity threshold of 10% of cells, found no CNVs of *APP*, *PSEN1*, *PSEN2*, or *MAPT* [67]. If increased copies are below 10% or involve partial *APP* sequences (<1 Mb), it may be undetectable using array or sequencing methods without ultra-high depth.

CNV changes have been further investigated in amyotrophic lateral sclerosis (ALS). ALS is a fatal neurodegenerative disease characterized by progressive muscle weakness because of the loss of spinal and cortical motor neurons [68]. Pamphlett et al. examined copy number differences between paired brain and blood samples and discovered 410 brain-specific CNVs (deletions > amplifications), with 121 found exclusively in ALS. Of these 121 ALS-specific CNVs, 24 were rare and overlapped in genic or promoter regions. Although no CNVs were identified in genes known to cause familial or sporadic ALS, CNVs were found in genes correlated with ALS pathogenesis, including *ATG7*, *GRIK1*, *GRIK2*, *FOXO3*, and *GGTLC2* [69]. In addition, three of these genes were previously noted in germline ALS copy number studies: *CSMD1*, *CNTN4*, and *GGTLC2* [70–72].

4. Retrotransposons

Retrotransposons compose ~40% of the human genome [73]. Part of a broader group of transposable elements, retrotransposons utilize a “copy and paste” mechanism to integrate into the genome via an RNA intermediate. This mechanism has allowed active retrotransposons to increase in copy number across the genome during evolution [74]. Retrotransposons are further categorized as either autonomous or non-autonomous retrotransposons. Autonomous retrotransposons are composed of long terminal repeat (LTR) retrotransposons and non-LTR-retrotransposons, so named because of the presence or

absence of flanking LTRs. LTR-retrotransposons, such as human endogenous retroviruses (HERVs), comprise about 8% of the human genome [73,75]. Of the 31 HERV subfamilies, HERVk (HML2) is capable of unfixed endogenous retrovirus insertions [76]. Non-LTR retrotransposons, also known as Long Interspersed Nuclear Elements (LINEs), comprise about 17% of the human genome. LINE1 encoded proteins also mediate mobilization of Short Interspersed Nuclear Elements (SINEs), including Alu and SVA elements, by recognizing, binding, and integrating non-autonomous SINEs into the genome [77]. SINEs comprise 13% of the human genome [73]. Somatic reinsertion of retrotransposons can mediate functional changes through many mechanisms, including insertional mutagenesis, loss of function mutations, premature transcript termination, alternative splicing, chromatin alterations, gene silencing, promoter effects, changes in mRNA localization, processed pseudogene formation, and DNA damage [78–85].

Multiple endogenous defense mechanisms counteract the expression, function, and reinsertion of retrotransposons in the mammalian genome. Indeed, a minority of retrotransposons are expressed, and, of those, an even smaller minority appear to be functional because of mutations that eliminate open reading frames (ORFs) or produce inactivated translated proteins [86]. These defense mechanisms occur at both the transcriptional and post-transcriptional level: chromatin modification, DNA methylation, PIWI-interacting RNAs, RNA interference mechanisms, post-transcriptional degradation of the target TE transcript via siRNAs, and autophagy [87,88]. Of note, autophagy failure is implicated as the underlying disease pathology in multiple neurodegenerative diseases including ALS, PD, and AD (reviewed in: [89]).

4.1. Retrotransposons in the Normal Brain

Retrotransposons may play a role in normal brain physiology and development. However, more studies are required to verify functionality, especially in the human brain. LINE1 reactivation accompanies dopaminergic neuron maturation in post-mitotic somatic cell trans-differentiation models [90]. LINE1 inhibition also has been reported to impair this trans-differentiation potential. LINE1 reactivation and reinsertion may therefore create lineage-specific genomic mosaicism crucial to cell identity specification, providing support for the hypothesis that shared patterns of somatic mutations shape cellular identity [14]. LINE1 retroinsertion has also been implicated in early life and memory formation. Increased LINE1 copy number has been correlated with induced early-life stress [91]. Inhibition of LINE1 retrotransposition in the adult hippocampus impairs long-term memory formation in mouse models [92]. Furthermore, LINE1 copy numbers are increased in healthy human hippocampal neurons, although the exact prevalence is debated [93–96]. It has been speculated that retrotransposition may represent a form of neuronal plasticity, in which permanent genomic changes occur in response to experiences.

HERV expression is tightly regulated during development, and increased transcription of HERVk is detrimental for the development and function of cortical neurons in human-pluripotent-stem-cell-based systems [97]. LINE1 is also implicated as a heritable genetic contributor to somatic mosaicism via CNV formation. Rearrangements within inherited LINE1s can result in the deletion of proximal genomic regions, suggesting that LINE1-associated genomic regions are hotspots for somatic copy number variants in the brain [98]. LINE1 retrotransposition is normally tightly regulated through chromatin accessibility and transcription factors. Cell division promotes LINE1 retrotransposition because of the breakdown of the nuclear envelope, although it is not required [99,100]. The 5'UTR of LINE1 contains transcription factor binding sites for YY1, RUNX, and SRY (e.g., SOX2) families that mediate LINE1 expression [101–103]. SOX2 downregulation induces LINE1 expression [104]. Interestingly, SOX2 expression is decreased in developing, aging, and AD brains [105,106]. While LINE1 retrotransposition can occur in post-mitotic neurons and LINE1 expression is known to increase with age, the prevalence of LINE1 retrotransposition in the aged brain is unknown [107].

4.2. Somatic Retrotransposition in Neurodegenerative Disease

Somatic LINE1 retrotransposition has been indicated in several pathogenic processes, including neurodegenerative disorders, autoimmune disorders, and cancer [93,108–111]. In particular, LINE1 has been implicated in A-T, Rett syndrome (RTT), frontotemporal lobar degeneration (FTLD), ALS, and AD. These diseases can be divided into two categories: (1) diseases in which LINE1 expression/retrotransposition is increased because of mutations in genes that regulate retrotransposons and (2) diseases in which there is aging associated neurological degeneration and retrotransposon copy number/expression.

The first category includes A-T and RTT. LINE1 copy number is increased in A-T brains compared with healthy controls [112,113]. This increase is directly due to ATM deficiency, with *ATM* knock-out mice and *ATM*-deficient neural progenitor cells demonstrating significant increases in LINE1 retrotransposition [112]. Similar increases in LINE1 copy number have been demonstrated in RTT. RTT is a progressive neurodevelopmental disorder in which patients between 6–18 months of age develop severe impairments, including loss of speech and purposeful hand use, microcephaly, seizures, and autistic-like behaviors [114]. RTT is caused by mutation of the X-linked gene methyl CpG binding protein 2 (*MECP2*), leading to abnormal epigenetic regulation and LINE1 retrotransposition [109]. Somatic LINE1 insertions are significantly increased in both cortical neurons and non-brain tissues of RTT patients as compared to non-disease controls, with higher numbers of retrotranspositions in the brain compared to other tissues [113,115].

4.3. Aging-Associated Neurological Degeneration & Retrotransposon Copy Number/Expression

Aging-associated neurological degeneration and increased retrotransposon copy number/expression occur in FTLD, ALS, and AD. FTLD is the second most common cause of early onset (≤ 65 years) dementia and is marked by degeneration of the frontal and temporal lobes [116,117]. Forty percent of patients with FTLD have cytoplasmic inclusions of TDP-43, which normally functions in the repair of double strand breaks by non-homologous end joining [118]. TDP-43 cytoplasmic inclusions are also found in patients with ALS, AD, and other neurodegenerative disorders [119]. Transposable element transcripts are extensively bound by TDP-43, an association which is reduced in FTLD patients [120].

LINE1 has also been implicated in ALS. The first indication of retrotransposon activity in ALS patients came from the detection of increased levels of reverse transcriptase in non-brain compartments of ALS patients, at levels comparable to those found in human immunodeficiency virus (HIV)-infected patients [121,122]. The source and significance of this increased reverse transcriptase are still unknown. Tam et al. analyzed the transcriptomes of 148 ALS cortical tissue samples and identified three ALS transcriptomic clusters marked by (1) oxidative and proteotoxic stress (61%), (2) glial activation (19%), and (3) high levels of retrotransposon expression and signatures of *TARDBP*/TDP-43 dysfunction (20%) [123]. Loss of nuclear TDP-43 in neurons of ALS patients is also associated with decondensation of LINE retrotransposons and increased LINE1 DNA content [124]. HERVk has also been implicated in the pathophysiology of ALS, with HERVk *pol* transcripts increased in patients with ALS [125]. Sequencing of these ALS HERVk *pol* transcripts revealed several actively transcribed loci in the HML-2 and HML-3 subfamilies. In addition, HERVk reverse transcriptase (RT) protein was selectively expressed in neurons and was significantly increased in ALS patients compared to controls [125]. In order to examine whether HERVk expression could be causative for ALS, rather than correlative, Li et al. generated a transgenic mouse model overexpressing HERVk *env* in neurons. This mouse model developed progressive and specific loss of upper/lower motor neurons and showed double stranded DNA damage and nucleolar dysfunction, leading to a 50% mortality rate by 10 months [126]. Expression of HERVk in human neurons in vitro also caused toxicity, with decreases in cell number and neurite retraction [126]. However, further studies have presented conflicting data about whether HERVk transcripts are significantly elevated in ALS as compared to controls [126,127].

AD tau pathology is associated with increased retrotransposon expression, including HERVs and LINE1, in the human brain, as well as chromatin relaxation at selected HERV and LINE1 loci [128]. Mutant Tau *Drosophila melanogaster* also had significantly increased retrotransposon expression, indicating a causative effect of tau pathology on retrotransposon expression [128]. However, quantitative PCR of LINE1 revealed no difference in the LINE1 copy number in the brain and blood between AD patients and the aged-matched control group without dementia [129]. It remains unclear if the increased expression or reduced sequestration of retrotransposons in FTLN, ALS, or AD results in increased somatic insertion, genomic instability, or inflammation, thus contributing to disease progression. Therefore, further research on this topic is needed.

5. Somatic Repeat Expansion

Somatic repeat expansions are accumulations of short tandem repeat sequences (1–12 base pairs). These sequences can expand or contract over time through DNA polymerase slippage, non-allelic homologous recombination, or errors in DNA repair [130]. Several familial neurodegenerative diseases are caused by repeat expansions, including Huntington's disease (HD), ALS, and spinocerebellar ataxia type 1 (SCA1); however, the presence of different repeat numbers across cells of a given individual may contribute to differential disease pathogenesis and phenotype.

HD is caused by an autosomal dominant CAG expansion in the Huntingtin gene (*HTT*). Between 50–70% of disease onset and severity is determined by the number of CAG repeats in *HTT* [131]. The CAG repeat is prone to expansion between successive generations leading to decreased age of onset and increased severity of disease, also known as genetic anticipation [132]. Although repeat expansion during early development may play a critical role in age of onset, instability during adulthood is key to disease progression. In HD, somatic expansion occurs most prominently in the brain compared to other body tissues, occurring preferentially in brain regions that are hardest hit by degeneration [133–135]. Sequencing of repeat length in grey matter indicates that neurons bear the brunt of repeat expansion [136,137]. Because neurons are post-mitotic, somatic repeat expansions are hypothesized to occur as a result of DNA damage rather than errors during DNA replication. The high energetic demands of neurons result in the production of high levels of reactive oxygen species, which endanger DNA integrity by chemically modifying bases and creating single and double stranded breaks [138]. Repeat sequences of DNA are prone to mutation because of the formation of secondary structures when the strands are separated either during replication, transcription, or DNA damage repair [139,140]. Genome-wide association studies (GWAS) identified DNA mismatch repair genes as significant modifiers of HD age of onset, in agreement with earlier work in mouse models [141–144]. Mechanisms of age-dependent CAG expansion have been demonstrated in neurons after terminal differentiation, demonstrating that somatic expansion does not require active cell division [145].

SCA1 is an autosomal dominant disorder caused by a CAG repeat in the *ATXN1* gene. SCA1 typically presents in the third to fourth decade of life and is characterized by progressive gait impairment, difficulties in speech and swallowing, and cerebellar atrophy [146]. As with HD, there is anticipation between generations, and expansion is thought to occur because of failure in DNA repair mechanisms [147]. A comprehensive quantitative analysis of CAG repeats across a range of CNS regions and peripheral tissues demonstrated similarly high levels of instability in cortex and neostriatum as those seen in HD. As opposed to HD, where these areas are correlated with disease severity, this pattern does not hold in SCA1 [148].

GGGGCC repeat expansions within the *C9ORF72* gene are present in ~35% of familial and ~6% of sporadic ALS and ~25% of familial and ~5% of sporadic FTLN, making it the most common known genetic cause of ALS and FTLN [149]. The pathogenicity of this hexanucleotide repeat has been attributed to both loss and gain-of-function mechanisms [150,151]. The presence of a large expansion may alter *C9ORF72* transcript expression

because of mRNA quadruplex structures or epigenetic changes in accessibility produced by methylation. This can result in reduced C9orf72 protein, which plays important homeostatic roles in nucleocytoplasmic shuttling, endosomal transport, autophagy, and stress granules. Toxic gain-of-function mechanisms are implicated in both the bidirectionally transcribed mRNA and the translated protein products of C9ORF72. Repeat-containing sense or antisense RNA may cause cellular stress by sequestering essential RNA-binding proteins, therefore impeding RNA processing. C9orf72 dipeptide repeat proteins are synthesized from the sense and antisense transcripts through repeat associated non-ATG translation. Inclusions positive for dipeptide repeat proteins are found in CNS tissue from germline expansion-positive ALS and FTLN patients and are thought to disrupt nucleocytoplasmic transport and RNA processing [150,151].

Somatic expansion of the C9ORF72 GGGGCC repeat has been found in different tissues of individuals with a germline repeat expansion in both ALS and FTLN [152–154]. One hypothesis for the development of phenotypically heterozygous ALS/FTLN is selective vulnerability of cortical or spinal cells to somatic expansion of the hexanucleotide repeat, even within individuals not harboring a germline expansion. In a study of ALS spinal cord in patients without an expansion found in blood, this hypothesis was not supported as no somatic expansion of the C9ORF72 hexanucleotide was found [155]. Further studies are needed to support or disprove this hypothesis.

6. Somatic Gene Recombination

A novel form of somatic DNA changes in the brain, somatic gene recombination (SGR), was recently reported [156]. SGR was first described outside the CNS as somatic recombination of immunoglobulin genes, termed V(D)J recombination, in proliferating lymphocytes [157]. Bushman et al. reported increased amyloid precursor protein (APP) gene copy number in single neurons from SAD brains [31], leading to further examination of somatic changes affecting APP in AD. The presence of somatic, mosaic neuronal recombination and reinsertion of APP in the brain [156,158] could account for the copy number increases previously observed [31]. These recombined somatic copies were dubbed “gencDNAs,” so-called because the genomic intron-less sequences resembled complementary DNAs (cDNAs), thus the term genomic cDNAs (gencDNAs). APP gencDNAs were identified in frontal lobe neurons from both SAD and age-matched controls, including brain-specific splice variants (APP-751 and APP-695). In SAD, these APP gencDNAs were more numerous (~3–5-fold higher based on DNA in situ hybridization). Furthermore, long-read sequencing of APP gencDNA amplicons revealed greater sequence diversity in SAD brains, including frequent single base changes, insertions, and deletions. This included identification of 11 SNVs known to cause autosomal dominant familial forms of AD, but that were present in SAD neurons. The gencDNAs identified were often truncated APP species containing intra-exonic junctions (IEJs) where non-sequential exons are joined at non-canonical splice sites. IEJs have since been described in both DNA and RNA-based datasets, indicating potential transcription of these novel-spliced sequences [158,159]. Supporting the genomic presence of APP gencDNAs, novel insertion sites of the APP 5' or 3' untranslated regions into chromosomes 1, 2, 9, 10, and 12 were identified [158]. In addition, reads spanning APP exon-exon junctions with mate reads on chromosomes 1, 3, 5, 6, and 13 were also reported. An independent report documented APP gencDNAs in human AD hippocampus [160].

These findings are not without some controversy. Through a report using non-overlapping techniques, scWGS to 45× depth was unsuccessful in identifying gencDNAs in non-diseased and SAD neurons [161]. This study was limited to <10 neurons per brain and inconsistently identified germline pseudogenes within individuals, highlighting technical challenges that may impact gencDNA identification. The presence of a contaminated pull-down dataset was identified, highlighting the difficulty of library preparation and sequencing analysis without contamination [161]; however, repetition through 10 independent experiments that produced datasets devoid of contamination by

two groups confirmed the original findings [158]. In addition, plasmid contamination could not explain the 10 genomic integration sites of *APP* sequences, nor the presence of 11 SNV familial mutations only in AD brain samples.

The creation of gencDNAs via SGR is hypothesized to require three main elements: gene transcription, reverse transcription of expressed RNA transcripts by an endogenous RT, and DNA strand breaks to allow for “retro-insertion” [156]. This proposed mechanism is supported by cell culture models of *APP* gencDNA formation in Chinese hamster ovary cells, a cell line with endogenous RT activity. Formation of gencDNAs required induction of DNA damage and was prevented with application of RT inhibitors [156]. The exact identity of endogenous RTs involved in *APP* SGR is unknown.

gencDNAs are reminiscent of germline processed pseudogenes that are classified as inactive evolutionary relics: they lack introns, show reduced sequence homology to the parent genes, are found in a non-wildtype gene location, and require similar components for reinsertion (gene transcription, RNA intermediates, RT activity, and DNA strand breaks). However, there are distinct and important differences. Processed pseudogenes are found at precise locations within the germline, which are stable during development and cell proliferation. In contrast, gencDNAs are somatic (not germline) and believed to occur in postmitotic neurons leading to diversity in location, number, and form amongst cells of the brain. “Somatic processed pseudogenes” have been described in proliferating cancer cells, contrasting with gencDNAs that occur in postmitotic cells. Additional similarities and differences between gencDNAs and processed pseudogenes have been reviewed [162].

Copy number changes of *APP* in the form of reinserted copies could be especially disease-relevant in AD. *APP* locus duplication causes autosomal dominant-early onset AD [163], and Down syndrome-associated triplication of the *APP* locus results in early onset dementia [65]. Additional copy number changes, in the form of gencDNAs, could result in SAD following a similar *APP* overexpressing disease process. It has been proposed that the protein products of variant *APP* sequences could result in more heterogenous biochemical processes than are utilized in classic A β peptide generation by γ - and β -secretases [164]. These varied products may escape therapeutic targeting by monoclonal antibodies raised against stereotyped A β peptide sequences, while secretase inhibitors might not be required for small products encoded by *APP* gencDNAs. Validation and clarification of SGR remain for future studies.

7. Somatic Single Nucleotide Variations

7.1. Somatic Single Nucleotide Variations in Development & Aging

The smallest and most common form of somatic genome changes affects single base pairs, SNVs. SNVs include transitions (purine to purine, pyrimidine to pyrimidine), transversions (purine \leftrightarrow pyrimidine), insertions, and deletions (the latter two are known as indels). SNVs in neurons accumulate because of endogenous and exogenous sources. Errors in DNA replication during development lead to shared SNVs in a cell’s downstream lineage. SNVs unique to single postmitotic cells can arise spontaneously secondary to DNA damage, such as through endogenous oxidative damage or exogenous factors such as radiation and smoking, or might arise through other processes such as SGR [158,165,166].

By birth, our brain cells already have accumulated several hundred clonal SNVs arising from proliferating progenitor cells [166]. During neurogenesis, \sim 5.1 SNVs accumulate per day per progenitor; this means that within a given cell division, \sim 8.6 SNVs occur per progenitor [167]. Three percent of SNVs found during development affect protein-coding sequences or gene regulation, and thus are likely to have functional consequences [167]. However, this same study found that SNVs are depleted in areas of open chromatin indicating efficient DNA repair in areas of active transcription. These clonal mutations are retained and can be detected into adulthood [168] without noticeable de-enrichment [169]. In adulthood, both clonal and private SNVs can be detected throughout the brain. In the adult prefrontal cortex, each neuron contains \sim 1500 somatic SNVs, with 60% of genomes containing at least one clonal SNV [168]. Compared to other tissues, the brain has a

lower SNV rate than tissues with continual turn over or higher exposure to environmental mutagens (skin, lung, liver, small intestine, etc.) [170].

Age positively correlates with brain SNV load. When comparing average SNV prevalence per base pair in frontal lobe from those <10 years versus those >40 years, there was a 5.7-fold increase in SNVs [171]. In single neurons within the prefrontal cortex, an estimated 23 SNVs accumulate each year. By age 80, individual genomes of these neurons would contain approximately 2500 SNVs each [172]. The SNV rate in the brain does not appear to be uniform. Dentate gyrus neurons accumulated nearly two-fold higher (~40) SNVs per year [172]. Age was more strongly correlated with SNV load in basal ganglia, nucleus accumbens, hypothalamus, and hippocampus than the frontal lobe [170]. Together, this suggests that different regions of the brain are more vulnerable to SNV accumulation, which could stem from a number of factors (differential oxidative damage, DNA repair, exposure to environmental mutagens, or SGR processes). As with clonal SNVs detected during neurogenesis, there was no correlation between actively transcribed regions (by expression level or heterochromatin markers) and expected numbers of SNVs [170,173]. Open chromatin likely has enhanced DNA repair compared to more closed regions. It is notable that these estimates are based upon small sample sizes [168,172].

7.2. Somatic Single Nucleotide Variation in Neurodegenerative Disease

Many neurodegenerative disorders, including PD and AD, have a small minority of patients for which a single mutated allele causes autosomal dominant disease. The possibility that sporadic versions of these diseases, in which no germline mutant can be detected, could be caused by somatic mutations to risk genes is intriguing. Depending on when these mutations arise, they may escape detection in lymphocyte-based genotyping. One family has been described for a mosaic mutation causing inherited AD [174]. A *PSEN1* mutation (P436Q) was found in 8% of peripheral lymphocytes and 14% of cells in the cortex of an individual diagnosed with early-onset progressive parkinsonian syndrome with dementia. One child carried this mutation within the germline and had an even earlier presentation of progressive cerebellar syndrome with dementia, an indication that the parent mutation occurred before gastrulation [174]. A similar mosaic mutation resulting in inherited disease has been reported in ALS [175]. The index case demonstrated a *FUS* mutation at a low level of mosaicism difficult to detect in blood and saliva via whole exome sequencing, with higher allele fractions in hair. CNS tissue was unavailable for sequencing. The son of the index case had juvenile onset, and sequencing of blood demonstrated a heterozygous *FUS* mutation [175].

Brain whole exome sequencing is a common method used to interrogate somatic SNVs with a high likelihood of functional impact without a bias to particular genes. One such study sequenced exomes of matched hippocampus and blood samples from patients with SAD, vascular dementia, or no neurodegenerative disease [176]. While the vast majority of SNVs in SAD were detected in both the brain and blood (~97%), an average of 575 mutations were found in the brain only. No known pathogenic mutations were found in familial genes; however, non-pathogenic mutations were seen in the familial gene *PSEN1*, genes identified via GWAS (*BIN1*, *ABCA7*, and *PICALM*), as well as in genes related to SAD or A β processing [176]. A second study pairing brain and blood used laser capture to enrich for neurons within the hippocampal formation [160]. In SAD, the authors estimated that somatic SNVs within neurons increase at a rate of 0.53 per exome per year. When extrapolated to the genome, this would mean 22 SNVs per year within a whole genome [160]. This estimate is roughly half the ~40 SNVs per year predicted in isolated dentate gyrus [172], which may reflect differences in methodology. Finally, this study found one SAD patient with a pathogenic mutation in *PIN1* and 14 out of 52 SAD patients with at least one putatively pathogenic brain somatic mutation in pathways associated with tau phosphorylation [160]. Such low-level SNVs within the hippocampal formation may contribute to tau aggregation and propagation in SAD. When comparing SNV rates between disease and control groups, no difference was found in multiple studies [160,169,176].

One method to detect SNVs at low allele frequencies is to increase the sequencing depth via targeted enrichment of genes of interest. Genes must be selected a priori and are typically chosen for their direct connection to autosomal dominant disease or demonstration of increased risk in sporadic disease. One study examining *APP*, *PSEN1*, *PSEN2*, and *MAPT* in entorhinal cortex from 72 SAD and 58 non-AD controls found and validated three possibly damaging low-frequency SNVs: *MAPT* Q124K at a frequency of 1.1%, *PSEN2* S130L at a frequency of 1.6%, and *MAPT* S735A at a frequency of 0.7% [67]. A recent study targeting PD-associated genes in SN, frontal lobe, cerebellum, and blood from synucleinopathies and controls did not detect somatic SNVs in PD-associated genes [177]. Two studies using high-resolution melting curve analysis of *SNCA* amplicons demonstrated no evidence for low-level mosaicism in the coding regions of *SNCA* across multiple brain regions in PD, MSA, and Lewy body dementia (LBD), another α -synucleinopathy. This technique is limited to low-frequency SNVs and could miss high frequency SNVs that would occur early in embryogenesis [178,179]. A larger panel of 56 genes associated with neurodegenerative disorders (including AD and LBD) and 46 genes associated with cancer demonstrated no difference in SNV rates between brain regions or between gene panels in a cohort of SAD, LBD, and control patients. Seven SNVs were found in neurodegenerative disease genes (mean allele frequency 0.82%). Within a single individual, an SNV in neurodegenerative disease-associated gene *TAF15* was found across multiple brain regions at different frequencies (4.37–9.77%). The authors used a computational model of brain development to estimate that each individual has 10^5 – 10^6 pathologically mutated cells, and up to ~10% of all humans will have SNVs within neurodegenerative genes present diffusely across the brain [180].

Autosomal dominant AD is typically characterized by early disease onset (≤ 65 years) [181]. Thus, early-onset AD patients who are negative for germline SNVs are good candidates to screen for potential disease-causing somatic SNVs. Ultra-deep sequencing of eleven genes in brain from 445 early-onset SAD patients identified nine candidate SNVs at a frequency of between 0.22% to 10.8% [182]. No pathogenic mutations were found in familial genes, but mutations in other AD-related genes may have contributed to disease in these patients. One brain-specific mutation in *CD55* was later validated in a late-onset SAD patient in a study of AD-related genes in early- and late-onset SAD patients [183]. This mutation was present in the temporal cortex at an allele frequency of 0.4% and may contribute to AD pathogenesis via its role as a neuroprotective complement regulator [184].

RNA-seq has also been used to identify disease-causing mutations in AD brain. In a study examining multiple brain regions, 104 genes were found to have disease-causing SNVs in AD, converging on genes associated with the cytoskeleton, autism, and intellectual disability. SNVs in the autism-associated gene *ADNP* (activity-dependent neuroprotective protein) occurred more frequently in AD than in controls, with increased SNV frequency associated with increased tau burden [185]. The causal roles for identified SNVs in AD remain to be demonstrated.

8. Future Technologies, Research, & Therapeutics

Multiple technological challenges persist in detecting somatic variants, especially in post-mitotic populations such as neurons. Somatic variants, especially those that occur during aging, may occur at low frequencies, making their detection within whole genomes technologically and analytically difficult (reviewed in: [186,187]). While next generation sequencing (NGS) technologies have revolutionized genomics research, the error rates of NGS approaches (0.1–1%) remain a concerning problem [188]. Furthermore, DNA contamination, PCR induced error, misclassification, and DNA damage can lead to false positive reporting of mosaicism [189]. New sequencing and bioinformatics tools are actively being developed to help address these limitations and improve detection of low frequency events.

O2n-seq utilizes single-strand DNA circularization to create two different copies of one original molecule in a pair of paired-end reads [190]. If a variant is supported by only one DNA copy, an error must have occurred at the site and thus this sequence is discarded during subsequent data analysis. Only variants supported by both DNA copies are treated as true variants. This strategy may improve detection of de novo, low-frequency mutations on NGS platforms by eliminating sequencing errors, improving data efficiency, and reducing library bias seen with other methods. Other sequencing methods, such as duplex sequencing, Multiple Independent Primer PCR Sequencing, RePlo, and REBELseq, also promise to increase the detection of low-frequency events and reduce amplification bias/PCR-induced sequencing artifacts [191–195].

Detection of somatic variants is often accomplished through scWGS. Genome sequencing on a single-cell level requires massive amplification of genomes combined with a high sequencing depth. Often analysis is limited to small numbers of cells (<20 cells per sample) [168,172]. Techniques such as multiple displacement amplification (MDA) can provide sufficient amplification with a high-fidelity polymerase [47]. Sequencing typically requires at least 30X sequencing coverage. Bottleneck sequencing (BotSeqS) is an alternative to single-cell sequencing. Fragmented genomes are labeled with sequencing adapters and then diluted before PCR amplification to create a bottleneck allowing efficient random sampling of the genomic templates [171]. This method requires that the SNV is present on both the positive and negative DNA strands. Another methodology for expanding the genome to sufficient levels for sequencing is clonal expansion, where proliferating cells can be isolated and grown in culture [167]. This technique requires proliferating cells and thus cannot be applied to terminally differentiated, postmitotic neurons. However, it benefits from the use of cellular replication machinery, which has higher fidelity than ex vivo methods.

Analysis of larger sample numbers has been made possible through the use of RNA-sequencing and/or whole exome-sequencing datasets [169,170,173]. The Genome-Tissue Expression project has provided an excellent resource of matched RNA-seq and whole exome data from individuals. Successfully and accurately calling SNVs from RNA-seq data requires overcoming a high false discovery rate (representing mutations seen in RNA but not confirmed in matched whole-exome data) and a high false negative rate (representing mutations seen in whole-exome data not confirmed in RNA). A high false-positive rate can obscure true SNVs present at low allele-frequencies. It can be overcome by only considering genomic regions where both alleles are present in RNA-seq, using robust pipelines to remove artifacts and validating SNVs using whole-exome blood samples [170]. These analyses also limit detection to genes with sufficient expression but could provide a helpful guide to SNVs that may produce a functional impact within a cell.

Identifying novel insertion sites within human brain cells is an important component of validating somatic insertions into the genome. PCR methods, such as FLEA-PCR (flanking sequence exponential anchored-PCR) and pulldown enrichments, have provided insights into somatically acquired insertion sites in cancers [196,197]. However, as with other PCR-based methods, they are vulnerable to amplification and ligation artifacts. Furthermore, they are benefited by the clonal expansion characteristic of cancerous malignancy. A recent study leveraged long-read sequencing to identify somatic transposable element insertions in the *Drosophila melanogaster* head and midgut [198]. They applied Oxford Nanopore Technologies long-read sequencing to bulk genomic DNA from pooled midguts, which undergo clonal expansion, and heads from 60 individuals and sequenced to 85X coverage. Long-read sequencing allowed the identification of putative somatic integrations, as a single continuous sequencing read could fully contain the transposable element, insertion site, and validated target site duplication. This method of non-amplification-based long read sequencing analyzed only insertions supported by a single read (“singletons”) as potentially somatic. The frequency of transposition between gut and head samples could not be directly compared since somatic transposition in only a few cells of the head was below the level of detection, as compared to the clonally expanded midgut. The *Drosophila*

head contains ~100,000 cells [199] and the haploid genome is ~180 Mbps [200]. By contrast, a single human brain has ~170 billion cells [201,202], with ~3100 Mbps average genome per cell (that may be significantly larger from cell-to-cell) [25,29,31].

These technologies further highlight the importance of crosstalk between fields within biomedical research. Many of these sequencing technologies were developed for the detection of novel, low-frequency somatic variants in cancer biology. Adapting these and other cancer technologies could provide increasingly fruitful results towards the understanding and treatment of sporadic neurodegenerative diseases. Cancer biology may also provide examples of genomic mosaicism understudied in the context of brain aging and neurodegeneration. Extrachromosomal circular DNA (ecDNA) has been characterized in tumor tissues [203]. These DNA species are excised from the linear genome, although they can potentially arise from any part of the genome as a consequence of DNA damage and repair. It is hypothesized that their origin is non-random and might arise from specific genomic hotspots [204]. Of interest, ecDNAs modulate gene copy numbers and transcription rates by different molecular means [205,206]. ecDNAs have also been found in somatically mosaic patterns in normal tissues, such as differentiated muscle and brain [204,207]. There is scant literature on human brain ecDNAs; however, several recent reviews have postulated that they may have a role in brain aging and neurodegeneration [208,209]. Technical and bioinformatics tools have been developed for genomic in-depth characterizations of ecDNAs [210]. These tools, combined with optimized wet lab approaches, may help to facilitate future investigations of ecDNAs within young, aged, and diseased brain.

Potential therapeutics for neurodegenerative diseases produced by somatic variants are already being investigated. Recent retrotransposon research in ALS and AD has led to multiple clinical trials on the safety and efficacy of HIV antiretroviral regimens for reducing clinical symptoms (NCT02868580, NCT02437110, NCT04500847, NCT04552795, NCT03706885). In addition, antisense oligonucleotide therapies could prove promising for modulating neurodegeneration impacted by somatic copy number variants or repeat expansion [211,212].

Identification of neural somatic variants, and therefore their targeted treatment, is hampered by our current genotyping strategies. Clinical identification of genetic variants relies on the identification of germline mutations in peripheral lymphocytes thus missing any brain-specific somatic variants. There is a critical need to identify alternative methods for identifying neural somatic mutations in patients. As previous studies have highlighted, because of the common embryonic origin of buccal and neural tissue, buccal swabs can provide insight into clonal, somatic variants that occur during embryogenesis. CNS-derived cerebrospinal fluid and blood exosomes are an additional peripheral source of neurodegenerative disease biomarkers and provide potential surrogate markers for neural somatic mutations [213]. Future imaging technologies may allow in vivo assessment of DCV towards classifying and possibly diagnosing disease states and progression. Collaborative multi-omic efforts, such as those orchestrated by the Allen Brain Institute and the Brain Initiative Cell Census Network, are promising avenues through which neural somatic variants can be identified. While tremendous advances have been made over the last two decades, further technological advances and collaborative efforts are important next steps towards understanding neural somatic variants and their role in multi-factorial neurodegenerative diseases.

Author Contributions: Conceptualization, I.C. and J.N.; writing—original draft preparation, I.C., J.N. and J.C.; writing—review and editing, I.C., J.N. and J.C.; visualization, I.C. and J.N.; supervision, J.C.; funding acquisition, I.C. and J.C. All authors have read and agreed to the published version of the manuscript.

Funding: The research was funded by the National Institute on Aging and the National Institute of Mental Health of the National Institutes of Health under award numbers R01AG071465, R01AG065541, U01MH114828-03 (Chun); T32AG066596 (Costantino). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Acknowledgments: We thank Linnea Ransom for scientific input and conceptual design. We thank Danielle Jones for administrative assistance.

Conflicts of Interest: J.C. is a co-founder of Mosaic Pharmaceuticals.

Abbreviations

AD—Alzheimer’s disease; ALS—amyotrophic lateral sclerosis; A-T—Ataxia-telangiectasia; CGH—Comparative genomic hybridization; CNV—Copy number variation; DCV—DNA content variation; ecDNA—Extrachromosomal circular DNA; FTLN—Frontotemporal lobar degeneration; FISH—Fluorescence in situ hybridization; GWAS—Genome-wide association studies; HD—Huntington’s disease; HIV—Human immunodeficiency virus; IEJ—Intra-exonic junction; LBD—Lewy body dementia; LINE—Long interspersed nuclear element; LTR—long terminal repeat; MSA—Multiple system atrophy; PD—Parkinson’s disease; PI—propidium iodide; RT—Reverse transcriptase; RTT—Rett syndrome; SAD—Sporadic Alzheimer’s disease; scWGS—Single-cell whole genome sequencing; SGR—Somatic gene recombination; SINE—Short interspersed nuclear element; SNV—Single nucleotide variant.

Appendix A

A comprehensive list of studies examining somatic variation in neurodegenerative diseased tissue can be found in [Table A1](#).

Table A1. Summary of studies examining somatic mutations in neurodegenerative disease.

Disease	Somatic Variation	DNA Source	Technique	Somatic Finding	Reference	
AD	DCV	16 SAD (7M:9F) & 16 ND (7M:9F); mean age 75 yo. 5 postmortem cortical regions	Slide-based cytometry	↑ DC in all AD cortical regions	Arendt et al., 2015 [45]	
	CNV	32 SAD (14M:18F, 62–101 yo.) & 40 ND (15M:25F, 17–103 yo.) postmortem FR, CBLM	qPCR; PNA-FISH; flow cytometry	↑ DC & ↑ <i>APP</i> copy number in SAD	Bushman et al., 2015 [31]	
		20 SAD, 20 PD/LBD, & 14 ND; mean age 80–84 yo. 5 postmortem brain regions	Whole exome sequencing	1 <i>APP</i> gain in AD case	Keogh et al., 2018 [66]	
	RTsp	72 SAD (61–105 yo.) & 58 non-AD (18–97 yo.) postmortem ERC	Gene enrichment & amplicon sequencing	No CNVs in <i>APP</i> , <i>PSEN1</i> , <i>PSEN2</i> , <i>MAPT</i>	Sala Frigerio et al., 2015 [67]	
		422 SAD & 201 ND postmortem FR	RNA-seq	↑ LINE1 & HERVk expression	Guo et al., 2018 [128]	
	SGR	7 SAD (1M:6F, 72–88 yo.) & 6 ND (3M:3F, 80–94 yo.) postmortem FR, CBLM	Amplicon sequencing; <i>APP</i> exonic pulldown; DISH	↑ <i>APP</i> gencDNAs in AD NeuN+ FR	Lee et al., 2018 [156]	
		52 SAD (16M:36F, 70–96 yo.) & 11 ND (7M:4F, 57–89 yo.) paired blood & postmortem laser-captured HIF	Whole exome sequencing	<i>APP</i> IEJs in AD HPC	Park et al., 2019 [160]	
	SNVs	2 related individuals (58 yo. F, 39 yo. F) blood & postmortem CTX	Allele-specific oligonucleotide hybridization	Mosaic <i>PSEN1</i> SNV in mother; germline heterozygous <i>PSEN1</i> SNV in child	Beck et al., 2004 [174]	
		17 SAD (5M:12F), 2 VD (1M:1F), & 2 ND (2M) 46–94 yo. paired blood & postmortem HPC, CBLM	Whole exome sequencing	AD brain-specific SNVs in AD assoc. genes	Parcerisas et al., 2014 [176]	
		72 SAD (61–105 yo.) & 58 non-AD (18–97 yo.) postmortem ERC	Gene enrichment & amplicon sequencing	2 <i>MAPT</i> & 1 <i>PSEN2</i> SNVs in AD & non-AD	Sala Frigerio et al., 2015 [67]	
372 EOAD (>66 years old at diagnosis), 73 LOAD, 1 FAD, & 52 ND blood & postmortem brain		smMIP assay & amplicon sequencing	9 candidate SNVs of benign/unknown sig.	Nicolas et al., 2018 [182]		
ALS	CNV	32 SALS (22M:10F; 47–84 yo.) & ND (18M:6F) blood & postmortem brain	Microarray	24 CNVs in genic/promoter regions	Pamphlett et al., 2011 [69]	
		25 SALS, 3 FALS (mean 63 yo.) & 12 ND (mean 60 yo.) 4 CTX regions	RTqPCR	↑ HERVk pol expression in ALS	Douville et al., 2010 [125]	
	RE	148 SALS, 11 other neurologic disease, & 17 ND postmortem CTX	RNA-seq	↑ LINE1 expression in ALS	Tam et al., 2019 [123]	
		19 SALS (9M:7F:3unknown; 50–79 yo.), <i>C9ORF72</i> expansion negative; postmortem spinal cord ALS with or without <i>C9ORF72</i> repeat expansion blood, CNS, non-neural tissues	RepeatPrimer PCR & amplicon size genotyping	No somatic expansion of <i>C9ORF72</i> repeat	Ross et al., 2019 [155]	
	SNV	2 related individuals (33 yo. M, 50 yo. F, living) blood & saliva	Southern blot	Intra-individual variation of <i>C9ORF72</i> repeat	Buchman et al., 2013 [152], Dols-Icardo et al., 2014 [153], Nordin et al., 2015 [154]	
			Whole exome sequencing	<i>FUS</i> mosaic SNV in mother; germline heterozygous <i>FUS</i> SNV in child	Hisahara et al., 2021 [175]	
	A-T	RTsp	7 A-T & 7 ND, 8–28 yo. laser capture of postmortem HPC	Taqman-based qPCR for ORF2 sequence	↑ LINE1 copy number in A-T	Coufal et al., 2011 [112]
			4 A-T, 2 RT, 72 other, & 20 ND postmortem neural & non-neural tissue	Whole-genome sequencing	↑ LINE1 copy number in A-T cortex	Jacob-Hirsch et al., 2018 [113]

Table A1. Cont.

Disease	Somatic Variation	DNA Source	Technique	Somatic Finding	Reference
FTLD	RTsp	FTLD & ND brain	Crosslinking-immunoprecipitation sequencing	↓ binding of RTsp by TDP-43 in FTLD brains	Li et al., 2012 [120]
	RE	FTLD with or without <i>C9ORF72</i> repeat expansion blood, CNS, non-neural tissues	Southern blot	Intra-individual variation of <i>C9ORF72</i> repeat	Buchman et al., 2013 [152], Dols-Icardo et al., 2014 [153], Nordin et al., 2015 [154]
HD	RE	3 HD (27–40 yo.) postmortem striatum	PCR amplification; small-pool PCR	↑ CAG repeats in striatal cells	Kennedy et al., 2003 [133]
		5 HD (3M:2F, 40–64 yo.) postmortem striatum & TC	PCR amplification & Southern blot	↑ CAG repeats in neurons vs. glia	Shelbourne et al., 2007 [136]
		24 HD young onset (20–41yo.) & 24 old onset (40–81 yo.) postmortem CTX & CBLM	Small-pool PCR	↑ repeat size assoc. with young onset	Swami et al., 2009 [134]
		7 adult-onset HD (2M:5F, 39–66 yo.) & 1 juvenile-onset HD (1M, 6 yo.), postmortem CNS & PNS tissue	Repeat length genotyping	↑ <i>ATXN1</i> CAG repeats in brain tissues	Mouro Pinto et al., 2020 [148]
MSA	CNV	5 MSA (55–76 yo.) & 30 ND (59–94 yo.) postmortem SN dopaminergic neurons	FISH	↑ <i>SNCA</i> copy number in MSA	Mokretar et al., 2018 [63]
		18 MSA (5M:13F, 52–82 yo.) & 17 ND (10M:7F, 59–92 yo.), postmortem cingulate CTX, CBLM	FISH; whole-genome sequencing	↑ <i>SNCA</i> copy number in MSA	Perez-Rodriguez et al., 2019 [64]
PD	CNV	8 PD (5M:3F, 63–81yo.) & 26 ND (18M:8F, 44–85yo.) postmortem FR	Microarray	CNVs detected in PD candidate genes (not <i>SNCA</i>)	Pamphlett et al., 2012 [54]
		2 PD living donors (40 yo. M, 23yo. M): mucosal cells, blood	FISH	↑ in 4p22.1 locus of <i>SNCA</i> in PD mucosa	Perandones et al., 2014 [62]
		41 PD (56–83 yo.) & 30 ND (59–92 yo.) postmortem SN dopaminergic neurons	FISH	↑ <i>SNCA</i> copy number in PD	Mokretar et al., 2018 [63]
		26 PD (20M:6F, 60–83 yo.) & 18 ND (10M:7F, 59–92 yo.) cingulate CTX & CBLM	FISH; whole-genome sequencing	↑ <i>SNCA</i> copy number in PD	Perez-Rodriguez et al., 2019 [64]
	SNV	28 PD (17M:11F, 62–90 yo.) postmortem SN & CBLM	HRM analysis of amplicons	No SNVs within <i>SNCA</i>	Proukakis et al., 2013 [178]
		511 idiopathic PD (age of onset 61 yo.) postmortem CBLM, SN, FR	HRM analysis of amplicons	No SNVs within <i>SNCA</i>	Proukakis et al., 2014 [179]
RTT	RTsp	20 PD/LBD (67–91 yo.) & 15 ND (64–97) blood & 4 postmortem brain regions	Gene enrichment panels	Brain-specific SNVs in neurodegenerative genes	Keogh et al., 2018 [180]
		25 sporadic PD (21M:4F, 55–88 yo.), 1 familial PD, & 12 ND (4M:8F, 69–104 yo.); postmortem SN	Gene enrichment panel; ddPCR	No disease-relevant SNVs detected	Leija-Salazar et al., 2020 [177]
RTT	RTsp	4 A-T, 2 RT, 72 other, & 20 ND postmortem neural & non-neural tissue	Whole-genome sequencing	↑ LINE1 copy number in RTT	Jacob-Hirsch et al., 2018 [113]
		5 RTT (5F, 17–21 yo.) & 5 ND (5F, 16–25 yo.) postmortem CNS & peripheral tissue	PCR-based targeted bulk sequencing	↑ LINE1 insertions in CTX neurons	Zhao et al., 2019 [115]
SCA1	RE	1 SCA1 postmortem CNS & peripheral tissue	Repeat length genotyping	↑ <i>ATXN1</i> CAG repeats in brain tissues	Mouro Pinto et al., 2020 [148]

AD = Alzheimer's disease; ALS = amyotrophic lateral sclerosis; A-T = ataxia telangiectasia; CBLM = cerebellum; CNV = copy number variation; CNS = central nervous system; CTX = cortex; DCV = DNA content variation; ddPCR = digital lobe PCR; DISH = DNA in situ hybridization; EOAD = early-onset Alzheimer's disease; ERC = entorhinal cortex; FALS = familial amyotrophic lateral sclerosis; FISH = fluorescent in situ hybridization; FR = frontal cortex; FTLD = frontotemporal lobar degeneration; HD = Huntington's disease; HIF = hippocampal formation; HPC = hippocampus; HRM = high-resolution melting curve; IEJs = intra-exonic junctions; LBD = Lewy Body Dementia; LOAD = late-onset Alzheimer's disease; MSA = multiple system atrophy; ND = non-diseased; OB = olfactory bulb; PD = Parkinson's disease; PNA-FISH = peptide nucleic acid fluorescence in situ hybridization; PNS = peripheral nervous system; RE = repeat expansion; RTqPCR = quantitative reverse transcription PCR; RTsp = retrotransposons; RTT = Rett's syndrome; SALS = sporadic amyotrophic lateral sclerosis; SCA1 = spinocerebellar ataxia type 1; SGR = somatic gene recombination; smMIP = single-molecule molecular inversion probes; SNV = single-nucleotide variation; TC = temporal cortex; VD = vascular dementia; yo = age at time of death (years old).

References

1. Cajal, S.R.y. *Studies on the Cerebral Cortex*; University of Madrid: Madrid, Spain, 1901; Volume 1.
2. Cajal, S.R.y. *Histologie du Système Nerveux de L'homme et des Vertébrés*; Maloine: Paris, French, 1909; ISBN 9782011774125.
3. de Castro, F.; López-Mascaraque, L.; De Carlos, J.A. Cajal: Lessons on brain development. *Brain Res. Rev.* **2007**, *55*, 481–489. [[CrossRef](#)]
4. Cajal, S.R.y. *Studies on Vertebrate Neurogenesis*; Thomas, C.C., Ed.; Springfield: Virginia, VA, USA, 1960.
5. Hubel, D.; Wiesel, T.N. Ferrier lecture—Functional architecture of macaque monkey visual cortex. *Proc. R. Soc. London Ser. B Biol. Sci.* **1977**, *198*, 1–59. [[CrossRef](#)]
6. Hubel, D.H.; Wiesel, T.N. Receptive fields of single neurones in the cat's striate cortex. *J. Physiol.* **1959**, *148*, 574–591. [[CrossRef](#)]
7. Hubel, D.; Wiesel, T. David Hubel and Torsten Wiesel. *Neuron* **2012**, *75*, 182–184. [[CrossRef](#)] [[PubMed](#)]
8. Hubel, D.H.; Wiesel, T.N. Brain Mechanisms of Vision. *Sci. Am.* **1979**, *241*, 150–162. [[CrossRef](#)]
9. Sakmann, B.; Neher, E. Patch Clamp Techniques for Studying Ionic Channels in Excitable Membranes. *Annu. Rev. Physiol.* **1984**, *46*, 455–472. [[CrossRef](#)]
10. Squire, L.R. Memory for relations in the short term and the long term after medial temporal lobe damage. *Hippocampus* **2017**, *27*, 608–612. [[CrossRef](#)]
11. Squire, L.R. On the course of forgetting in very long-term memory. *J. Exp. Psychol. Learn. Mem. Cogn.* **1989**, *15*, 241–245. [[CrossRef](#)] [[PubMed](#)]
12. Squire, L.; Slater, P.; Chace, P. Retrograde amnesia: Temporal gradient in very long term memory following electroconvulsive therapy. *Science* **1975**, *187*, 77–79. [[CrossRef](#)] [[PubMed](#)]
13. Felsenfeld, G. A Brief History of Epigenetics. *Cold Spring Harb. Perspect. Biol.* **2014**, *6*, a018200. [[CrossRef](#)] [[PubMed](#)]
14. Dreyer, W.J.; Gray, W.R.; Hood, L. The Genetic, Molecular, and Cellular Basis of Antibody Formation: Some Facts and a Unifying Hypothesis. *Cold Spring Harb. Symp. Quant. Biol.* **1967**, *32*. [[CrossRef](#)]
15. Gayon, J. From Mendel to epigenetics: History of genetics. *Comptes R. Biol.* **2016**, *339*, 225–230. [[CrossRef](#)]
16. Rehen, S.K.; McConnell, M.J.; Kaushal, D.; Kingsbury, M.A.; Yang, A.H.; Chun, J. Chromosomal variation in neurons of the developing and adult mammalian nervous system. *Proc. Natl. Acad. Sci. USA* **2001**, *98*. [[CrossRef](#)]
17. McConnell, M.J.; Lindberg, M.R.; Brennand, K.J.; Piper, J.C.; Voet, T.; Cowing-Zitron, C.; Shumilina, S.; Lasken, R.S.; Vermeesch, J.R.; Hall, I.M.; et al. Mosaic Copy Number Variation in Human Neurons. *Science* **2013**, *342*, 632–637. [[CrossRef](#)]
18. Kingsbury, M.A.; Friedman, B.; McConnell, M.J.; Rehen, S.K.; Yang, A.H.; Kaushal, D.; Chun, J. Aneuploid neurons are functionally active and integrated into brain circuitry. *Proc. Natl. Acad. Sci. USA* **2005**, *102*. [[CrossRef](#)] [[PubMed](#)]
19. Rehen, S.K. Constitutional Aneuploidy in the Normal Human Brain. *J. Neurosci.* **2005**, *25*. [[CrossRef](#)]
20. Yurov, Y.B.; Iourov, I.Y.; Vorsanova, S.G.; Liehr, T.; Kolotii, A.D.; Kutsev, S.I.; Pellestor, F.; Beresheva, A.K.; Demidova, I.A.; Kravets, V.S.; et al. Aneuploidy and Confined Chromosomal Mosaicism in the Developing Human Brain. *PLoS ONE* **2007**, *2*. [[CrossRef](#)]
21. Iourov, I.Y.; Vorsanova, S.G.; Liehr, T.; Yurov, Y.B. Aneuploidy in the normal, Alzheimer's disease and ataxia-telangiectasia brain: Differential expression and pathological meaning. *Neurobiol. Dis.* **2009**, *34*, 212–220. [[CrossRef](#)]
22. Yurov, Y.B.; Iourov, I.Y.; Monakhov, V.V.; Soloviev, I.V.; Vostrikov, V.M.; Vorsanova, S.G. The Variation of Aneuploidy Frequency in the Developing and Adult Human Brain Revealed by an Interphase FISH Study. *J. Histochem. Cytochem.* **2005**, *53*. [[CrossRef](#)]
23. Iourov, I.Y.; Liehr, T.; Vorsanova, S.G.; Kolotii, A.D.; Yurov, Y.B. Visualization of interphase chromosomes in postmitotic cells of the human brain by multicolour banding (MCB). *Chromosom. Res.* **2006**, *14*. [[CrossRef](#)]
24. Iourov, I.Y.; Vorsanova, S.G.; Liehr, T.; Kolotii, A.D.; Yurov, Y.B. Increased chromosome instability dramatically disrupts neural genome integrity and mediates cerebellar degeneration in the ataxia-telangiectasia brain. *Hum. Mol. Genet.* **2009**, *18*. [[CrossRef](#)] [[PubMed](#)]
25. Bushman, D.M.; Chun, J. The genomically mosaic brain: Aneuploidy and more in neural diversity and disease. *Semin. Cell Dev. Biol.* **2013**, *24*, 357–369. [[CrossRef](#)]
26. Yurov, Y.B.; Vorsanova, S.G.; Liehr, T.; Kolotii, A.D.; Iourov, I.Y. X chromosome aneuploidy in the Alzheimer's disease brain. *Mol. Cytogenet.* **2014**, *7*. [[CrossRef](#)]
27. Knouse, K.A.; Wu, J.; Whittaker, C.A.; Amon, A. Single cell sequencing reveals low levels of aneuploidy across mammalian tissues. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 13409–13414. [[CrossRef](#)] [[PubMed](#)]
28. van den Bos, H.; Spierings, D.C.J.; Taudt, A.S.; Bakker, B.; Porubský, D.; Falconer, E.; Novoa, C.; Halsema, N.; Kazemier, H.G.; Hoekstra-Wakker, K.; et al. Single-cell whole genome sequencing reveals no evidence for common aneuploidy in normal and Alzheimer's disease neurons. *Genome Biol.* **2016**, *17*. [[CrossRef](#)]
29. Westra, J.W.; Rivera, R.R.; Bushman, D.M.; Yung, Y.C.; Peterson, S.E.; Barral, S.; Chun, J. Neuronal DNA content variation (DCV) with regional and individual differences in the human brain. *J. Comp. Neurol.* **2010**, *518*, 3981–4000. [[CrossRef](#)]
30. Fischer, H.G.; Morawski, M.; Brückner, M.K.; Mittag, A.; Tarnok, A.; Arendt, T. Changes in neuronal DNA content variation in the human brain during aging. *Aging Cell* **2012**, *11*, 628–633. [[CrossRef](#)]
31. Bushman, D.M.; Kaeser, G.E.; Siddoway, B.; Westra, J.W.; Rivera, R.R.; Rehen, S.K.; Yung, Y.C.; Chun, J. Genomic mosaicism with increased amyloid precursor protein (APP) gene copy number in single neurons from sporadic Alzheimer's disease brains. *eLife* **2015**, *4*. [[CrossRef](#)]

32. Copani, A.; Hoozemans, J.J.M.; Caraci, F.; Calafiore, M.; Van Haastert, E.S.; Veerhuis, R.; Rozemuller, A.J.M.; Aronica, E.; Sortino, M.A.; Nicoletti, F. DNA Polymerase-beta Is Expressed Early in Neurons of Alzheimer's Disease Brain and Is Loaded into DNA Replication Forks in Neurons Challenged with beta-Amyloid. *J. Neurosci.* **2006**, *26*. [[CrossRef](#)] [[PubMed](#)]
33. Andriani, G.A.; Vijg, J.; Montagna, C. Mechanisms and consequences of aneuploidy and chromosome instability in the aging brain. *Mech. Ageing Dev.* **2017**, *161*. [[CrossRef](#)]
34. Chow, H.; Herrup, K. Genomic integrity and the ageing brain. *Nat. Rev. Neurosci.* **2015**, *16*, 672–684. [[CrossRef](#)] [[PubMed](#)]
35. Geller, L.N.; Potter, H. Chromosome missegregation and trisomy 21 mosaicism in Alzheimer's disease. *Neurobiol. Dis.* **1999**, *6*, 167–179. [[CrossRef](#)]
36. Kingsbury, M.A.; Yung, Y.C.; Peterson, S.E.; Westra, J.W.; Chun, J. Aneuploidy in the normal and diseased brain. *Cell. Mol. Life Sci.* **2006**, *63*, 2626–2641. [[CrossRef](#)]
37. Balmus, G.; Pilger, D.; Coates, J.; Demir, M.; Sczaniecka-Clift, M.; Barros, A.C.; Woods, M.; Fu, B.; Yang, F.; Chen, E.; et al. ATM orchestrates the DNA-damage response to counter toxic non-homologous end-joining at broken replication forks. *Nat. Commun.* **2019**, *10*, 87. [[CrossRef](#)] [[PubMed](#)]
38. Amirifar, P.; Ranjouri, M.R.; Yazdani, R.; Abolhassani, H.; Aghamohammadi, A. Ataxia-telangiectasia: A review of clinical features and molecular pathology. *Pediatr. Allergy Immunol.* **2019**, *30*. [[CrossRef](#)] [[PubMed](#)]
39. McConnell, M.J. Failed Clearance of Aneuploid Embryonic Neural Progenitor Cells Leads to Excess Aneuploidy in the Atm-Deficient But Not the Trp53-Deficient Adult Cerebral Cortex. *J. Neurosci.* **2004**, *24*. [[CrossRef](#)]
40. McKinnon, P.J. ATM and the Molecular Pathogenesis of Ataxia Telangiectasia. *Annu. Rev. Pathol. Mech. Dis.* **2012**, *7*, 303–321. [[CrossRef](#)]
41. Braak, H.; Braak, E. Neuropathological staging of Alzheimer-related changes. *Acta Neuropathol.* **1991**, *82*, 239–259. [[CrossRef](#)]
42. Selkoe, D.J.; Hardy, J. The amyloid hypothesis of Alzheimer's disease at 25 years. *EMBO Mol. Med.* **2016**, *8*, 595–608. [[CrossRef](#)]
43. Jack, C.R.; Bennett, D.A.; Blennow, K.; Carrillo, M.C.; Dunn, B.; Haeberlein, S.B.; Holtzman, D.M.; Jagust, W.; Jessen, F.; Karlawish, J.; et al. NIA-AA Research Framework: Toward a biological definition of Alzheimer's disease. *Alzheimer's Dement.* **2018**, *14*, 535–562. [[CrossRef](#)]
44. Westra, J.W.; Barral, S.; Chun, J. A Reevaluation of Tetraploidy in the Alzheimer's Disease Brain. *Neurodegener. Dis.* **2009**, *6*. [[CrossRef](#)]
45. Arendt, T.; Brückner, M.K.; Lösche, A. Regional mosaic genomic heterogeneity in the elderly and in Alzheimer's disease as a correlate of neuronal vulnerability. *Acta Neuropathol.* **2015**, *130*, 501–510. [[CrossRef](#)]
46. Knouse, K.A.; Wu, J.; Amon, A. Assessment of megabase-scale somatic copy number variation using single-cell sequencing. *Genome Res.* **2016**, *26*, 376–384. [[CrossRef](#)] [[PubMed](#)]
47. Gole, J.; Gore, A.; Richards, A.; Chiu, Y.J.; Fung, H.L.; Bushman, D.; Chiang, H.I.; Chun, J.; Lo, Y.H.; Zhang, K. Massively parallel polymerase cloning and genome sequencing of single cells using nanoliter microwells. *Nat. Biotechnol.* **2013**, *31*, 1126–1132. [[CrossRef](#)] [[PubMed](#)]
48. Chronister, W.D.; Burbulis, I.E.; Wierman, M.B.; Wolpert, M.J.; Haakenson, M.F.; Smith, A.C.B.; Kleinman, J.E.; Hyde, T.M.; Weinberger, D.R.; Bekiranov, S.; et al. Neurons with Complex Karyotypes Are Rare in Aged Human Neocortex. *Cell Rep.* **2019**, *26*, 825–835. [[CrossRef](#)] [[PubMed](#)]
49. Rohrback, S.; April, C.; Kaper, F.; Rivera, R.R.; Liu, C.S.; Siddoway, B.; Chun, J. Submegabase copy number variations arise during cerebral cortical neurogenesis as revealed by single-cell whole-genome sequencing. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 10804–10809. [[CrossRef](#)]
50. Lockwood, W.W.; Chari, R.; Chi, B.; Lam, W.L. Recent advances in array comparative genomic hybridization technologies and their applications in human genetics. *Eur. J. Hum. Genet.* **2006**, *14*, 139–148. [[CrossRef](#)]
51. Cheung, S.W.; Bi, W. Novel applications of array comparative genomic hybridization in molecular diagnostics. *Expert Rev. Mol. Diagn.* **2018**, *18*, 531–542. [[CrossRef](#)]
52. Piotrowski, A.; Bruder, C.E.G.; Andersson, R.; De Ståhl, T.D.; Menzel, U.; Sandgren, J.; Poplawski, A.; Von Tell, D.; Crasto, C.; Bogdan, A.; et al. Somatic mosaicism for copy number variation in differentiated human tissues. *Hum. Mutat.* **2008**, *29*, 1118–1124. [[CrossRef](#)]
53. Villela, D.; Suemoto, C.K.; Leite, R.; Pasqualucci, C.A.; Grinberg, L.T.; Pearson, P.; Rosenberg, C. Increased DNA Copy Number Variation Mosaicism in Elderly Human Brain. *Neural Plast.* **2018**, *2018*. [[CrossRef](#)]
54. Pamphlett, R.; Morahan, J.M.; Luquin, N.; Yu, B. An approach to finding brain-situated mutations in sporadic Parkinson's disease. *Park. Relat. Disord.* **2012**, *18*, 82–85. [[CrossRef](#)]
55. Cai, X.; Evrony, G.D.; Lehmann, H.S.; Elhosary, P.C.; Mehta, B.K.; Poduri, A.; Walsh, C.A. Single-Cell, Genome-wide Sequencing Identifies Clonal Somatic Copy-Number Variation in the Human Brain. *Cell Rep.* **2014**, *8*, 1280–1289. [[CrossRef](#)]
56. Peterson, S.E.; Yang, A.H.; Bushman, D.M.; Westra, J.W.; Yung, Y.C.; Barral, S.; Mutoh, T.; Rehen, S.K.; Chun, J. Aneuploid Cells Are Differentially Susceptible to Caspase-Mediated Death during Embryonic Cerebral Cortical Development. *J. Neurosci.* **2012**, *32*, 16213–16222. [[CrossRef](#)] [[PubMed](#)]
57. Alegre-Abarrategui, J.; Brimblecombe, K.R.; Roberts, R.F.; Velentza-Almpani, E.; Tilley, B.S.; Bengoa-Vergniory, N.; Proukakis, C. Selective vulnerability in α -synucleinopathies. *Acta Neuropathol.* **2019**, *138*, 681–704. [[CrossRef](#)] [[PubMed](#)]
58. Jankovic, J.; Tan, E.K. Parkinson's disease: Etiopathogenesis and treatment. *J. Neurol. Neurosurg. Psychiatry* **2020**, *91*, 795–808. [[CrossRef](#)] [[PubMed](#)]

59. Wenning, G.K.; Colosimo, C.; Geser, F.; Poewe, W. Multiple system atrophy. *Lancet Neurol.* **2004**, *3*, 93–103. [[CrossRef](#)]
60. Ahn, T.B.; Kim, S.Y.; Kim, J.Y.; Park, S.S.; Lee, D.S.; Min, H.J.; Kim, Y.K.; Kim, S.E.; Kim, J.M.; Kim, H.J.; et al. α -Synuclein gene duplication is present in sporadic Parkinson disease. *Neurology* **2008**, *70*, 43–49. [[CrossRef](#)]
61. Functammasan, A.; Walsh, E.; Chiaromonte, F.; Eckert, K.A.; Makova, K.D. A genome-wide analysis of common fragile sites: What features determine chromosomal instability in the human genome? *Genome Res.* **2012**, *22*, 993–1005. [[CrossRef](#)]
62. Perandones, C.; Giugni, J.C.; Calvo, D.S.; Raina, G.B.; De Jorge Lopez, L.; Volpini, V.; Zabetian, C.P.; Mata, I.F.; Caputo, M.; Corach, D.; et al. Mosaicism of Alpha-synuclein gene rearrangements: Report of two unrelated cases of early-onset parkinsonism. *Park. Relat. Disord.* **2014**, *20*, 558–561. [[CrossRef](#)]
63. Mokretar, K.; Pease, D.; Taanman, J.-W.; Soenmez, A.; Ejaz, A.; Lashley, T.; Ling, H.; Gentleman, S.; Houlden, H.; Holton, J.L.; et al. Somatic copy number gains of α -synuclein (SNCA) in Parkinson's disease and multiple system atrophy brains. *Brain* **2018**, *141*. [[CrossRef](#)]
64. Perez-Rodriguez, D.; Kalyva, M.; Leija-Salazar, M.; Lashley, T.; Tarabichi, M.; Chelban, V.; Gentleman, S.; Schottlaender, L.; Franklin, H.; Vasmatzis, G.; et al. Investigation of somatic CNVs in brains of synucleinopathy cases using targeted SNCA analysis and single cell sequencing. *Acta Neuropathol. Commun.* **2019**, *7*. [[CrossRef](#)]
65. Wiseman, F.K.; Al-Janabi, T.; Hardy, J.; Karmiloff-Smith, A.; Nizetic, D.; Tybulewicz, V.L.J.; Fisher, E.M.C.; Strydom, A. A genetic cause of Alzheimer disease: Mechanistic insights from Down syndrome. *Nat. Rev. Neurosci.* **2015**, *16*, 564–574. [[CrossRef](#)]
66. Keogh, M.J.; Wei, W.; Wilson, I.; Coxhead, J.; Ryan, S.; Rollinson, S.; Griffin, H.; Kurzawa-Akanbi, M.; Santibanez-Koref, M.; Talbot, K.; et al. Genetic compendium of 1511 human brains available through the UK Medical Research Council Brain Banks Network Resource. *Genome Res.* **2017**, *27*, 165–173. [[CrossRef](#)]
67. Sala Frigerio, C.; Lau, P.; Troakes, C.; Deramecourt, V.; Gele, P.; Van Loo, P.; Voet, T.; De Strooper, B. On the identification of low allele frequency mosaic mutations in the brains of Alzheimer's disease patients. *Alzheimer's Dement.* **2015**, *11*, 1265–1276. [[CrossRef](#)] [[PubMed](#)]
68. van Es, M.A.; Hardiman, O.; Chio, A.; Al-Chalabi, A.; Pasterkamp, R.J.; Veldink, J.H.; van den Berg, L.H. Amyotrophic lateral sclerosis. *Lancet* **2017**, *390*, 2084–2098. [[CrossRef](#)]
69. Pamphlett, R.; Morahan, J.M.; Luquin, N.; Yu, B. Looking for differences in copy number between blood and brain in sporadic amyotrophic lateral sclerosis. *Muscle Nerve* **2011**, *44*, 492–498. [[CrossRef](#)] [[PubMed](#)]
70. Cronin, S.; Blauw, H.M.; Veldink, J.H.; van Es, M.A.; Ophoff, R.A.; Bradley, D.G.; van den Berg, L.H.; Hardiman, O. Analysis of genome-wide copy number variation in Irish and Dutch ALS populations. *Hum. Mol. Genet.* **2008**, *17*, 3392–3398. [[CrossRef](#)] [[PubMed](#)]
71. Shoichet, S.A.; Waibel, S.; Shoichet, S.A.; Waibel, S.; Endruhn, S.; Sperfeld, A.D.; Vorwerk, B.; Müller, I.; Erdogan, F.; Ludolph, A.C.; et al. Identification of candidate genes for sporadic amyotrophic lateral sclerosis by array comparative genomic hybridization. *Amyotroph. Lateral Scler.* **2009**, *10*, 162–167. [[CrossRef](#)]
72. Wain, L.V.; Pedroso, I.; Landers, J.E.; Breen, G.; Shaw, C.E.; Leigh, P.N.; Brown, R.H.; Tobin, M.D.; Al-Chalabi, A. The Role of Copy Number Variation in Susceptibility to Amyotrophic Lateral Sclerosis: Genome-Wide Association Study and Comparison with Published Loci. *PLoS ONE* **2009**, *4*, e8175. [[CrossRef](#)]
73. Lander, E.; Linton, L.; Birren, B.; Nusbaum, C.; Zody, M.; Baldwin, J. Initial sequencing and analysis of the human genome. *Nature* **2001**, *409*. [[CrossRef](#)]
74. Goodier, J.L. Restricting retrotransposons: A review. *Mob. DNA* **2016**, *7*. [[CrossRef](#)]
75. Weiss, R.A. Human endogenous retroviruses: Friend or foe? *APMIS* **2016**, *124*, 4–10. [[CrossRef](#)]
76. Wildschutte, J.H.; Williams, Z.H.; Montesion, M.; Subramanian, R.P.; Kidd, J.M.; Coffin, J.M. Discovery of unfixed endogenous retrovirus insertions in diverse human populations. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, E2326–E2334. [[CrossRef](#)] [[PubMed](#)]
77. Dewannieux, M.; Esnault, C.; Heidmann, T. LINE-mediated retrotransposition of marked Alu sequences. *Nat. Genet.* **2003**, *35*. [[CrossRef](#)] [[PubMed](#)]
78. Cordaux, R.; Batzer, M.A. The impact of retrotransposons on human genome evolution. *Nat. Rev. Genet.* **2009**, *10*. [[CrossRef](#)] [[PubMed](#)]
79. Baillie, J.K.; Barnett, M.W.; Upton, K.R.; Gerhardt, D.J.; Richmond, T.A.; De Sapio, F.; Brennan, P.M.; Rizzu, P.; Smith, S.; Fell, M.; et al. Somatic retrotransposition alters the genetic landscape of the human brain. *Nature* **2011**, *479*. [[CrossRef](#)]
80. Faulkner, G.J.; Billon, V. L1 retrotransposition in the soma: A field jumping ahead. *Mob. DNA* **2018**, *9*. [[CrossRef](#)]
81. Kines, K.J.; Sokolowski, M.; deHaro, D.L.; Christian, C.M.; Belancio, V.P. Potential for genomic instability associated with retrotranspositionally-incompetent L1 loci. *Nucleic Acids Res.* **2014**, *42*. [[CrossRef](#)]
82. Saleh, A.; Macia, A.; Muotri, A.R. Transposable Elements, Inflammation, and Neurological Disease. *Front. Neurol.* **2019**, *10*. [[CrossRef](#)]
83. Estácio, M.R.H.; Gallegos, J.; Dekmezian, M.; Lu, Y.; Liang, S.; Issa, J.-P.J. SINE Retrotransposons Cause Epigenetic Reprogramming of Adjacent Gene Promoters. *Mol. Cancer Res.* **2012**, *10*, 1332–1342. [[CrossRef](#)] [[PubMed](#)]
84. Sorek, R. Alu-Containing Exons are Alternatively Spliced. *Genome Res.* **2002**, *12*, 1060–1067. [[CrossRef](#)]
85. Elbarbary, R.A.; Lucas, B.A.; Maquat, L.E. Retrotransposons as regulators of gene expression. *Science* **2016**, *351*, aac7247. [[CrossRef](#)]
86. Mills, R.E.; Bennett, E.A.; Iskow, R.C.; Devine, S.E. Which transposable elements are active in the human genome? *Trends Genet.* **2007**, *23*, 183–191. [[CrossRef](#)] [[PubMed](#)]

87. Guo, H.; Chitiprolu, M.; Gagnon, D.; Meng, L.; Perez-Iratxeta, C.; Lagace, D.; Gibbins, D. Autophagy supports genomic stability by degrading retrotransposon RNA. *Nat. Commun.* **2014**, *5*, 5276. [[CrossRef](#)] [[PubMed](#)]
88. Klein, S.J.; O'Neill, R.J. Transposable elements: Genome innovation, chromosome diversity, and centromere conflict. *Chromosom. Res.* **2018**, *26*, 5–23. [[CrossRef](#)] [[PubMed](#)]
89. Fujikake, N.; Shin, M.; Shimizu, S. Association Between Autophagy and Neurodegenerative Diseases. *Front. Neurosci.* **2018**, *12*. [[CrossRef](#)]
90. Della Valle, F.; Thimma, M.P.; Caiazzo, M.; Pulcrano, S.; Celii, M.; Adroub, S.A.; Liu, P.; Alanis-Lobato, G.; Broccoli, V.; Orlando, V. Transdifferentiation of Mouse Embryonic Fibroblasts into Dopaminergic Neurons Reactivates LINE-1 Repetitive Elements. *Stem Cell Rep.* **2020**, *14*. [[CrossRef](#)]
91. Bedrosian, T.A.; Quayle, C.; Novaresi, N.; Gage, F.H. Early life experience drives structural variation of neural genomes in mice. *Science* **2018**, *359*. [[CrossRef](#)]
92. Bachiller, S.; del-Pozo-Martín, Y.; Carrión, Á.M. L1 retrotransposition alters the hippocampal genomic landscape enabling memory formation. *Brain Behav. Immun.* **2017**, *64*. [[CrossRef](#)]
93. Coufal, N.G.; Garcia-Perez, J.L.; Peng, G.E.; Yeo, G.W.; Mu, Y.; Lovci, M.T.; Morell, M.; O'Shea, K.S.; Moran, J.V.; Gage, F.H. L1 retrotransposition in human neural progenitor cells. *Nature* **2009**, *460*. [[CrossRef](#)]
94. Evrony, G.D.; Cai, X.; Lee, E.; Hills, L.B.; Elhosary, P.C.; Lehmann, H.S.; Parker, J.J.; Atabay, K.D.; Gilmore, E.C.; Poduri, A.; et al. Single-Neuron Sequencing Analysis of L1 Retrotransposition and Somatic Mutation in the Human Brain. *Cell* **2012**, *151*. [[CrossRef](#)]
95. Upton, K.R.; Gerhardt, D.J.; Jesuadian, J.S.; Richardson, S.R.; Sánchez-Luque, F.J.; Bodea, G.O.; Ewing, A.D.; Salvador-Palomeque, C.; van der Knaap, M.S.; Brennan, P.M.; et al. Ubiquitous L1 Mosaicism in Hippocampal Neurons. *Cell* **2015**, *161*. [[CrossRef](#)]
96. Evrony, G.D.; Lee, E.; Park, P.J.; Walsh, C.A. Resolving rates of mutation in the brain using single-neuron genomics. *eLife* **2016**, *5*. [[CrossRef](#)]
97. Padmanabhan Nair, V.; Liu, H.; Ciceri, G.; Jungverdorben, J.; Frishman, G.; Tchieu, J.; Cederquist, G.Y.; Rothenaigier, I.; Schorpp, K.; Klepper, L.; et al. Activation of HERV-K(HML-2) disrupts cortical patterning and neuronal differentiation by increasing NTRK3. *Cell Stem Cell* **2021**. [[CrossRef](#)] [[PubMed](#)]
98. Erwin, J.A.; Paquola, A.C.M.; Singer, T.; Gallina, I.; Novotny, M.; Quayle, C.; Bedrosian, T.A.; Alves, F.I.A.; Butcher, C.R.; Herdy, J.R.; et al. L1-associated genomic regions are deleted in somatic cells of the healthy human brain. *Nat. Neurosci.* **2016**, *19*. [[CrossRef](#)] [[PubMed](#)]
99. Kubo, S.; Seleme, M.d.C.; Soifer, H.S.; Perez, J.L.G.; Moran, J.V.; Kazazian, H.H.; Kasahara, N. L1 retrotransposition in nondividing and primary human somatic cells. *Proc. Natl. Acad. Sci. USA* **2006**, *103*. [[CrossRef](#)] [[PubMed](#)]
100. Xie, Y.; Mates, L.; Ivics, Z.; Izsvák, Z.; Martin, S.L.; An, W. Cell division promotes efficient retrotransposition in a stable L1 reporter cell line. *Mob. DNA* **2013**, *4*. [[CrossRef](#)]
101. Kuwabara, T.; Hsieh, J.; Muotri, A.; Yeo, G.; Warashina, M.; Lie, D.C.; Moore, L.; Nakashima, K.; Asashima, M.; Gage, F.H. Wnt-mediated activation of NeuroD1 and retro-elements during adult neurogenesis. *Nat. Neurosci.* **2009**, *12*. [[CrossRef](#)] [[PubMed](#)]
102. Athanikar, J.N. A YY1-binding site is required for accurate human LINE-1 transcription initiation. *Nucleic Acids Res.* **2004**, *32*. [[CrossRef](#)]
103. Yang, N. An important role for RUNX3 in human L1 transcription and retrotransposition. *Nucleic Acids Res.* **2003**, *31*. [[CrossRef](#)] [[PubMed](#)]
104. Muotri, A.R.; Chu, V.T.; Marchetto, M.C.N.; Deng, W.; Moran, J.V.; Gage, F.H. Somatic mosaicism in neuronal precursor cells mediated by L1 retrotransposition. *Nature* **2005**, *435*. [[CrossRef](#)] [[PubMed](#)]
105. Carrasco-Garcia, E.; Moreno-Cugnon, L.; Garcia, I.; Borrás, C.; Revuelta, M.; Izeta, A.; Lopez-Lluch, G.; de Pancorbo, M.M.; Vergara, I.; Vina, J.; et al. SOX2 expression diminishes with ageing in several tissues in mice and humans. *Mech. Ageing Dev.* **2019**, *177*, 30–36. [[CrossRef](#)]
106. Sarlak, G.; Vincent, B. The Roles of the Stem Cell-Controlling Sox2 Transcription Factor: From Neuroectoderm Development to Alzheimer's Disease? *Mol. Neurobiol.* **2016**, *53*, 1679–1698. [[CrossRef](#)]
107. Macia, A.; Widmann, T.J.; Heras, S.R.; Ayllon, V.; Sanchez, L.; Benkaddour-Boumzaouad, M.; Muñoz-Lopez, M.; Rubio, A.; Amador-Cubero, S.; Blanco-Jimenez, E.; et al. Engineered LINE-1 retrotransposition in nondividing human neurons. *Genome Res.* **2017**, *27*, 335–348. [[CrossRef](#)]
108. Bundo, M.; Kato, T.; Iwamoto, K. Estimation of LINE-1 Copy Number in the Brain Tissue and Isolated Neuronal Nuclei. In *Genomic Mosaicism in Neurons and Other Cell Types*; Humana Press: New York, NY, USA, 2017.
109. Muotri, A.R.; Marchetto, M.C.N.; Coufal, N.G.; Oefner, R.; Yeo, G.; Nakashima, K.; Gage, F.H. L1 retrotransposition in neurons is modulated by MeCP2. *Nature* **2010**, *468*. [[CrossRef](#)] [[PubMed](#)]
110. Reilly, M.T.; Faulkner, G.J.; Dubnau, J.; Ponomarev, I.; Gage, F.H. The Role of Transposable Elements in Health and Diseases of the Central Nervous System. *J. Neurosci.* **2013**, *33*, 17577–17586. [[CrossRef](#)] [[PubMed](#)]
111. Thomas, C.A.; Tejwani, L.; Trujillo, C.A.; Negraes, P.D.; Herai, R.H.; Mesci, P.; Macia, A.; Crow, Y.J.; Muotri, A.R. Modeling of TREX1-Dependent Autoimmune Disease using Human Stem Cells Highlights L1 Accumulation as a Source of Neuroinflammation. *Cell Stem Cell* **2017**, *21*. [[CrossRef](#)]

112. Coufal, N.G.; Garcia-Perez, J.L.; Peng, G.E.; Marchetto, M.C.N.; Muotri, A.R.; Mu, Y.; Carson, C.T.; Macia, A.; Moran, J.V.; Gage, F.H. Ataxia telangiectasia mutated (ATM) modulates long interspersed element-1 (L1) retrotransposition in human neural stem cells. *Proc. Natl. Acad. Sci. USA* **2011**, *108*. [[CrossRef](#)] [[PubMed](#)]
113. Jacob-Hirsch, J.; Eyal, E.; Knisbacher, B.A.; Roth, J.; Cesarkas, K.; Dor, C.; Farage-Barhom, S.; Kunik, V.; Simon, A.J.; Gal, M.; et al. Whole-genome sequencing reveals principles of brain retrotransposition in neurodevelopmental disorders. *Cell Res.* **2018**, *28*. [[CrossRef](#)]
114. Amir, R.E.; Van den Veyver, I.B.; Wan, M.; Tran, C.Q.; Francke, U.; Zoghbi, H.Y. Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nat. Genet.* **1999**, *23*. [[CrossRef](#)]
115. Zhao, B.; Wu, Q.; Ye, A.Y.; Guo, J.; Zheng, X.; Yang, X.; Yan, L.; Liu, Q.-R.; Hyde, T.M.; Wei, L.; et al. Somatic LINE-1 retrotransposition in cortical neurons and non-brain tissues of Rett patients and healthy individuals. *PLoS Genet.* **2019**, *15*. [[CrossRef](#)] [[PubMed](#)]
116. Chen, H.; Kankel, M.W.; Su, S.C.; Han, S.W.S.; Ofengeim, D. Exploring the genetics and non-cell autonomous mechanisms underlying ALS/FTLD. *Cell Death Differ.* **2018**, *25*. [[CrossRef](#)]
117. Young, J.J.; Lavakumar, M.; Tampi, D.; Balachandran, S.; Tampi, R.R. Frontotemporal dementia: Latest evidence and clinical implications. *Ther. Adv. Psychopharmacol.* **2018**, *8*. [[CrossRef](#)] [[PubMed](#)]
118. Mitra, J.; Guerrero, E.N.; Hegde, P.M.; Liachko, N.F.; Wang, H.; Vasquez, V.; Gao, J.; Pandey, A.; Taylor, J.P.; Kraemer, B.C.; et al. Motor neuron disease-associated loss of nuclear TDP-43 is linked to DNA double-strand break repair defects. *Proc. Natl. Acad. Sci. USA* **2019**, *116*. [[CrossRef](#)]
119. Cohen, T.J.; Lee, V.M.Y.; Trojanowski, J.Q. TDP-43 functions and pathogenic mechanisms implicated in TDP-43 proteinopathies. *Trends Mol. Med.* **2011**, *17*. [[CrossRef](#)] [[PubMed](#)]
120. Li, W.; Jin, Y.; Prazak, L.; Hammell, M.; Dubnau, J. Transposable Elements in TDP-43-Mediated Neurodegenerative Disorders. *PLoS ONE* **2012**, *7*. [[CrossRef](#)]
121. Steele, A.J.; Al-Chalabi, A.; Ferrante, K.; Cudkowicz, M.E.; Brown, R.H.; Garson, J.A. Detection of serum reverse transcriptase activity in patients with ALS and unaffected blood relatives. *Neurology* **2005**, *64*. [[CrossRef](#)]
122. McCormick, A.L.; Brown, R.H.; Cudkowicz, M.E.; Al-Chalabi, A.; Garson, J.A. Quantification of reverse transcriptase in ALS and elimination of a novel retroviral candidate. *Neurology* **2008**, *70*. [[CrossRef](#)] [[PubMed](#)]
123. Tam, O.H.; Rozhkov, N.V.; Shaw, R.; Kim, D.; Hubbard, I.; Fennessey, S.; Propp, N.; Fagegaltier, D.; Harris, B.T.; Ostrow, L.W.; et al. Postmortem Cortex Samples Identify Distinct Molecular Subtypes of ALS: Retrotransposon Activation, Oxidative Stress, and Activated Glia. *Cell Rep.* **2019**, *29*. [[CrossRef](#)]
124. Liu, E.Y.; Russ, J.; Cali, C.P.; Phan, J.M.; Amlie-Wolf, A.; Lee, E.B. Loss of Nuclear TDP-43 Is Associated with Decondensation of LINE Retrotransposons. *Cell Rep.* **2019**, *27*. [[CrossRef](#)]
125. Douville, R.; Liu, J.; Rothstein, J.; Nath, A. Identification of active loci of a human endogenous retrovirus in neurons of patients with amyotrophic lateral sclerosis. *Ann. Neurol.* **2011**, *69*. [[CrossRef](#)]
126. Li, W.; Lee, M.-H.; Henderson, L.; Tyagi, R.; Bachani, M.; Steiner, J.; Campanac, E.; Hoffman, D.A.; von Geldern, G.; Johnson, K.; et al. Human endogenous retrovirus-K contributes to motor neuron disease. *Sci. Transl. Med.* **2015**, *7*, 307ra153. [[CrossRef](#)] [[PubMed](#)]
127. Mayer, J.; Harz, C.; Sanchez, L.; Pereira, G.C.; Maldener, E.; Heras, S.R.; Ostrow, L.W.; Ravits, J.; Batra, R.; Meese, E.; et al. Transcriptional profiling of HERV-K(HML-2) in amyotrophic lateral sclerosis and potential implications for expression of HML-2 proteins. *Mol. Neurodegener.* **2018**, *13*. [[CrossRef](#)] [[PubMed](#)]
128. Guo, C.; Jeong, H.-H.; Hsieh, Y.-C.; Klein, H.-U.; Bennett, D.A.; De Jager, P.L.; Liu, Z.; Shulman, J.M. Tau Activates Transposable Elements in Alzheimer's Disease. *Cell Rep.* **2018**, *23*. [[CrossRef](#)]
129. Protasova, M.S.; Gusev, F.E.; Grigorenko, A.P.; Kuznetsova, I.L.; Rogae, E.I.; Andreeva, T.V. Quantitative Analysis of L1 Retrotransposons in Alzheimer's Disease and Aging. *Biochem. Mosc.* **2017**, *82*, 962–971. [[CrossRef](#)]
130. Khristich, A.N.; Mirkin, S.M. On the wrong DNA track: Molecular mechanisms of repeat-mediated genome instability. *J. Biol. Chem.* **2020**, *295*, 4134–4170. [[CrossRef](#)]
131. Andrew, S.E.; Goldberg, Y.P.; Kremer, B.; Telenius, H.; Theilmann, J.; Adam, S.; Starr, E.; Squitieri, F.; Lin, B.; Kalchman, M.A.; et al. The relationship between trinucleotide (CAG) repeat length and clinical features of Huntington's disease. *Nat. Genet.* **1993**, *4*, 398–403. [[CrossRef](#)]
132. Kovtun, I.V.; McMurray, C.T. Trinucleotide expansion in haploid germ cells by gap repair. *Nat. Genet.* **2001**, *27*, 407–411. [[CrossRef](#)]
133. Kennedy, L.; Evans, E.; Chen, C.M.; Craven, L.; Detloff, P.J.; Ennis, M.; Shelbourne, P.F. Dramatic tissue-specific mutation length increases are an early molecular event in Huntington disease pathogenesis. *Hum. Mol. Genet.* **2003**, *12*, 3359–3367. [[CrossRef](#)] [[PubMed](#)]
134. Swami, M.; Hendricks, A.E.; Gillis, T.; Massood, T.; Mysore, J.; Myers, R.H.; Wheeler, V.C. Somatic expansion of the Huntington's disease CAG repeat in the brain is associated with an earlier age of disease onset. *Hum. Mol. Genet.* **2009**, *18*. [[CrossRef](#)] [[PubMed](#)]
135. Telenius, H.; Kremer, B.; Goldberg, Y.P.; Theilmann, J.; Andrew, S.E.; Zeisler, J.; Adam, S.; Greenberg, C.; Ives, E.J.; Clarke, L.A.; et al. Somatic and gonadal mosaicism of the Huntington disease gene CAG repeat in brain and sperm. *Nat. Genet.* **1994**, *6*, 409–414. [[CrossRef](#)] [[PubMed](#)]

136. Shelbourne, P.F.; Keller-McGandy, C.; Bi, W.L.; Yoon, S.R.; Dubeau, L.; Veitch, N.J.; Vonsattel, J.P.; Wexler, N.S.; Group, U.S.-V.C.R.; Arnheim, N.; et al. Triplet repeat mutation length gains correlate with cell-type specific vulnerability in Huntington disease brain. *Hum. Mol. Genet.* **2007**, *16*, 1133–1142. [[CrossRef](#)]
137. Labadorf, A.; Hoss, A.G.; Lagomarsino, V.; Latourelle, J.C.; Hadzi, T.C.; Bregu, J.; MacDonald, M.E.; Gusella, J.F.; Chen, J.F.; Akbarian, S.; et al. RNA Sequence Analysis of Human Huntington Disease Brain Reveals an Extensive Increase in Inflammatory and Developmental Gene Expression. *PLoS ONE* **2015**, *10*, e0143563. [[CrossRef](#)]
138. Madabhushi, R.; Pan, L.; Tsai, L.H. DNA damage and its links to neurodegeneration. *Neuron* **2014**, *83*, 266–282. [[CrossRef](#)] [[PubMed](#)]
139. Lin, Y.; Hubert, L., Jr.; Wilson, J.H. Transcription destabilizes triplet repeats. *Mol. Carcinog.* **2009**, *48*, 350–361. [[CrossRef](#)]
140. Petruska, J.; Hartenstine, M.J.; Goodman, M.F. Analysis of strand slippage in DNA polymerase expansions of CAG/CTG triplet repeats associated with neurodegenerative disease. *J. Biol. Chem.* **1998**, *273*, 5204–5210. [[CrossRef](#)] [[PubMed](#)]
141. Bettencourt, C.; Hensman-Moss, D.; Flower, M.; Wiethoff, S.; Brice, A.; Goizet, C.; Stevanin, G.; Koutsis, G.; Karadima, G.; Panas, M.; et al. DNA repair pathways underlie a common genetic mechanism modulating onset in polyglutamine diseases. *Ann. Neurol.* **2016**, *79*, 983–990. [[CrossRef](#)] [[PubMed](#)]
142. Lee, J.M.; Chao, M.J.; Harold, D.; Abu Elneel, K.; Gillis, T.; Holmans, P.; Jones, L.; Orth, M.; Myers, R.H.; Kwak, S.; et al. A modifier of Huntington’s disease onset at the MLH1 locus. *Hum. Mol. Genet.* **2017**, *26*, 3859–3867. [[CrossRef](#)] [[PubMed](#)]
143. Moss, D.J.H.; Pardini, A.F.; Langbehn, D.; Lo, K.; Leavitt, B.R.; Roos, R.; Durr, A.; Mead, S.; investigators, T.-H.; investigators, R.; et al. Identification of genetic variants associated with Huntington’s disease progression: A genome-wide association study. *Lancet Neurol.* **2017**, *16*, 701–711. [[CrossRef](#)]
144. Tome, S.; Manley, K.; Simard, J.P.; Clark, G.W.; Slean, M.M.; Swami, M.; Shelbourne, P.F.; Tillier, E.R.; Monckton, D.G.; Messer, A.; et al. MSH3 polymorphisms and protein levels affect CAG repeat instability in Huntington’s disease mice. *PLoS Genet.* **2013**, *9*, e1003280. [[CrossRef](#)] [[PubMed](#)]
145. Kovtun, I.V.; Liu, Y.; Bjoras, M.; Klungland, A.; Wilson, S.H.; McMurray, C.T. OGG1 initiates age-dependent CAG trinucleotide expansion in somatic cells. *Nature* **2007**, *447*, 447–452. [[CrossRef](#)]
146. Tejwani, L.; Lim, J. Pathogenic mechanisms underlying spinocerebellar ataxia type 1. *Cell. Mol. Life Sci.* **2020**, *77*, 4015–4029. [[CrossRef](#)] [[PubMed](#)]
147. Kraus-Perrotta, C.; Lagalwar, S. Expansion, mosaicism and interruption: Mechanisms of the CAG repeat mutation in spinocerebellar ataxia type 1. *Cerebellum Ataxias* **2016**, *3*, 20. [[CrossRef](#)]
148. Mouro Pinto, R.; Arning, L.; Giordano, J.V.; Razghandi, P.; Andrew, M.A.; Gillis, T.; Correia, K.; Mysore, J.S.; Grote Urtubey, D.-M.; Parwez, C.R.; et al. Patterns of CAG repeat instability in the central nervous system and periphery in Huntington’s disease and in spinocerebellar ataxia type 1. *Hum. Mol. Genet.* **2020**, *29*. [[CrossRef](#)]
149. van Blitterswijk, M.; DeJesus-Hernandez, M.; Rademakers, R. How do C9ORF72 repeat expansions cause amyotrophic lateral sclerosis and frontotemporal dementia. *Curr. Opin. Neurol.* **2012**, *25*. [[CrossRef](#)] [[PubMed](#)]
150. Gendron, T.F.; Petrucelli, L. Disease Mechanisms of C9ORF72 Repeat Expansions. *Cold Spring Harb. Perspect. Med.* **2018**, *8*, a024224. [[CrossRef](#)]
151. Babić Leko, M.; Župunski, V.; Kirincich, J.; Smilović, D.; Hortobágyi, T.; Hof, P.R.; Šimić, G. Molecular Mechanisms of Neurodegeneration Related to C9orf72 Hexanucleotide Repeat Expansion. *Behav. Neurol.* **2019**, *2019*. [[CrossRef](#)] [[PubMed](#)]
152. Buchman, V.L.; Cooper-Knock, J.; Connor-Robson, N.; Higginbottom, A.; Kirby, J.; Razinskaya, O.D.; Ninkina, N.; Shaw, P.J. Simultaneous and independent detection of C9ORF72 alleles with low and high number of GGGGCC repeats using an optimised protocol of Southern blot hybridisation. *Mol. Neurodegener.* **2013**, *8*, 12. [[CrossRef](#)] [[PubMed](#)]
153. Dols-Icardo, O.; García-Redondo, A.; Rojas-García, R.; Sánchez-Valle, R.; Noguera, A.; Gómez-Tortosa, E.; Pastor, P.; Hernández, I.; Esteban-Pérez, J.; Suárez-Calvet, M.; et al. Characterization of the repeat expansion size in C9orf72 in amyotrophic lateral sclerosis and frontotemporal dementia. *Hum. Mol. Genet.* **2014**, *23*, 749–754. [[CrossRef](#)]
154. Nordin, A.; Akimoto, C.; Wuolikainen, A.; Alstermark, H.; Jonsson, P.; Birve, A.; Marklund, S.L.; Graffmo, K.S.; Forsberg, K.; Brännström, T.; et al. Extensive size variability of the GGGGCC expansion in C9orf72 in both neuronal and non-neuronal tissues in 18 patients with ALS or FTD. *Hum. Mol. Genet.* **2015**, *24*, 3133–3142. [[CrossRef](#)]
155. Ross, J.P.; Leblond, C.S.; Catoire, H.; Volkening, K.; Strong, M.; Zinman, L.; Robertson, J.; Dion, P.A.; Rouleau, G.A. Somatic expansion of the C9orf72 hexanucleotide repeat does not occur in ALS spinal cord tissues. *Neurol. Genet.* **2019**, *5*, e317. [[CrossRef](#)]
156. Lee, M.H.; Siddoway, B.; Kaeser, G.E.; Segota, I.; Rivera, R.; Romanow, W.J.; Liu, C.S.; Park, C.; Kennedy, G.; Long, T.; et al. Somatic APP gene recombination in Alzheimer’s disease and normal neurons. *Nature* **2018**, *563*, 639–645. [[CrossRef](#)]
157. Hozumi, N.; Tonegawa, S. Evidence for somatic rearrangement of immunoglobulin genes coding for variable and constant regions. *Proc. Natl. Acad. Sci. USA* **1976**, *73*, 3628–3632. [[CrossRef](#)] [[PubMed](#)]
158. Lee, M.-H.; Liu, C.S.; Zhu, Y.; Kaeser, G.E.; Rivera, R.; Romanow, W.J.; Kihara, Y.; Chun, J. Reply to: APP gene copy number changes reflect exogenous contamination. *Nature* **2020**, *584*, E29–E33. [[CrossRef](#)] [[PubMed](#)]
159. Pacific Biosciences Data Release: Alzheimer Brain Isoform Sequencing (Iso-Seq) Dataset; Pacific Biosciences: Menlo Park, CA, USA, 2016.
160. Park, J.S.; Lee, J.; Jung, E.S.; Kim, M.H.; Kim, I.B.; Son, H.; Kim, S.; Kim, S.; Park, Y.M.; Mook-Jung, I.; et al. Brain somatic mutations observed in Alzheimer’s disease associated with aging and dysregulation of tau phosphorylation. *Nat. Commun.* **2019**, *10*. [[CrossRef](#)] [[PubMed](#)]

161. Kim, J.; Zhao, B.; Huang, A.Y.; Miller, M.B.; Lodato, M.A.; Walsh, C.A.; Lee, E.A. APP gene copy number changes reflect exogenous contamination. *Nature* **2020**, *584*, E20–E28. [[CrossRef](#)]
162. Kaeser, G.; Chun, J. Brain cell somatic gene recombination and its phylogenetic foundations. *J. Biol. Chem.* **2020**, *295*, 12786–12795. [[CrossRef](#)] [[PubMed](#)]
163. Hooli, B.V.; Kovacs-Vajna, Z.M.; Mullin, K.; Blumenthal, M.A.; Mattheisen, M.; Zhang, C.; Lange, C.; Mohapatra, G.; Bertram, L.; Tanzi, R.E. Rare autosomal copy number variations in early-onset familial Alzheimer's disease. *Mol. Psychiatry* **2014**, *19*, 676–681. [[CrossRef](#)]
164. Kaeser, G.E.; Chun, J. Mosaic Somatic Gene Recombination as a Potentially Unifying Hypothesis for Alzheimer's Disease. *Front. Genet.* **2020**, *11*, 390. [[CrossRef](#)] [[PubMed](#)]
165. Bauer, N.C.; Corbett, A.H.; Doetsch, P.W. The current state of eukaryotic DNA base damage and repair. *Nucleic Acids Res.* **2015**, *43*, 10083–10101. [[CrossRef](#)]
166. Frank, S.A. Somatic evolutionary genomics: Mutations during development cause highly variable genetic mosaicism with risk of cancer and neurodegeneration. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 1725–1730. [[CrossRef](#)] [[PubMed](#)]
167. Bae, T.; Tomasini, L.; Mariani, J.; Zhou, B.; Roychowdhury, T.; Franjic, D.; Pletikos, M.; Pattni, R.; Chen, B.J.; Venturini, E.; et al. Different mutational rates and mechanisms in human cells at pregastrulation and neurogenesis. *Science* **2018**, *359*, 550–555. [[CrossRef](#)]
168. Lodato, M.A.; Woodworth, M.B.; Lee, S.; Evrony, G.D.; Mehta, B.K.; Karger, A.; Lee, S.; Chittenden, T.W.; D'Gama, A.M.; Cai, X.; et al. Somatic mutation in single human neurons tracks developmental and transcriptional history. *Science* **2015**, *350*, 94–98. [[CrossRef](#)] [[PubMed](#)]
169. Wei, W.; Keogh, M.J.; Aryaman, J.; Golder, Z.; Kullar, P.J.; Wilson, I.; Talbot, K.; Turner, M.R.; McKenzie, C.A.; Troakes, C.; et al. Frequency and signature of somatic variants in 1461 human brain exomes. *Genet. Med.* **2019**, *21*, 904–912. [[CrossRef](#)] [[PubMed](#)]
170. García-Nieto, P.E.; Morrison, A.J.; Fraser, H.B. The somatic mutation landscape of the human body. *Genome Biol.* **2019**, *20*, 1–20. [[CrossRef](#)]
171. Hoang, M.L.; Kinde, I.; Tomasetti, C.; McMahon, K.W.; Rosenquist, T.A.; Grollman, A.P.; Kinzler, K.W.; Vogelstein, B.; Papadopoulos, N. Genome-wide quantification of rare somatic mutations in normal human tissues using massively parallel sequencing. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 9846–9851. [[CrossRef](#)]
172. Lodato, M.A.; Rodin, R.E.; Bohrsen, C.L.; Coulter, M.E.; Barton, A.R.; Kwon, M.; Sherman, M.A.; Vitzthum, C.M.; Luquette, L.J.; Yandava, C.N.; et al. Aging and neurodegeneration are associated with increased mutations in single human neurons. *Science* **2018**, *359*, 555–559. [[CrossRef](#)]
173. Yizhak, K.; Aguet, F.; Kim, J.; Hess, J.M.; Kübler, K.; Grimsby, J.; Frazer, R.; Zhang, H.; Haradhvala, N.J.; Rosebrock, D.; et al. RNA sequence analysis reveals macroscopic somatic clonal expansion across normal tissues. *Science* **2019**, *364*. [[CrossRef](#)]
174. Beck, J.A.; Poulter, M.; Campbell, T.A.; Uphill, J.B.; Adamson, G.; Geddes, J.F.; Revesz, T.; Davis, M.B.; Wood, N.W.; Collinge, J.; et al. Somatic and germline mosaicism in sporadic early-onset Alzheimer's disease. *Hum. Mol. Genet.* **2004**, *13*, 1219–1224. [[CrossRef](#)]
175. Hisahara, S.; Nishiyama, A.; Tsuda, E.; Suzuki, S.; Matsumura, A.; Ishikawa, A.; Sakurai, A.; Motoike, I.N.; Aoki, M.; Aoki, Y.; et al. Possible Somatic Mosaicism of Novel FUS Variant in Familial Amyotrophic Lateral Sclerosis. *Neurol. Genet.* **2021**, *7*, e552. [[CrossRef](#)]
176. Parcerisas, A.; Rubio, S.E.; Muhaisen, A.; Gómez-Ramos, A.; Pujadas, L.; Puiggros, M.; Rossi, D.; Ureña, J.; Burgaya, F.; Pascual, M.; et al. Somatic Signature of Brain-Specific Single Nucleotide Variations in Sporadic Alzheimer's Disease. *J. Alzheimer's Dis.* **2014**, *42*, 1357–1382. [[CrossRef](#)]
177. Leija-Salazar, M.; Pittman, A.; Mokretar, K.; Morris, H.; Schapira, A.H.; Proukakis, C. Investigation of Somatic Mutations in Human Brains Targeting Genes Associated With Parkinson's Disease. *Front. Neurol.* **2020**, *11*, 8–12. [[CrossRef](#)]
178. Proukakis, C.; Houlden, H.; Schapira, A.H. Somatic Alpha-synuclein mutations in Parkinson's disease: Hypothesis and preliminary data. *Mov. Disord.* **2013**, *28*, 705–712. [[CrossRef](#)]
179. Proukakis, C.; Shoaee, M.; Morris, J.; Brier, T.; Kara, E.; Sheerin, U.M.; Charlesworth, G.; Tolosa, E.; Houlden, H.; Wood, N.W.; et al. Analysis of Parkinson's disease brain-derived DNA for Alpha-synuclein coding somatic mutations. *Mov. Disord.* **2014**, *29*, 1060–1064. [[CrossRef](#)]
180. Keogh, M.J.; Wei, W.; Aryaman, J.; Walker, L.; van den Amele, J.; Coxhead, J.; Wilson, I.; Bashton, M.; Beck, J.; West, J.; et al. High prevalence of focal and multi-focal somatic genetic variants in the human brain. *Nat. Commun.* **2018**, *9*. [[CrossRef](#)] [[PubMed](#)]
181. Mendez, M.F. Early-Onset Alzheimer Disease. *Neurol. Clin.* **2017**, *35*, 263–281. [[CrossRef](#)] [[PubMed](#)]
182. Nicolas, G.; Acuña-Hidalgo, R.; Keogh, M.J.; Quenez, O.; Stehouwer, M.; Lelieveld, S.; Rousseau, S.; Richard, A.C.; Oud, M.S.; Marguet, F.; et al. Somatic variants in autosomal dominant genes are a rare cause of sporadic Alzheimer's disease. *Alzheimer's Dement.* **2018**, *14*, 1632–1639. [[CrossRef](#)] [[PubMed](#)]
183. Helgadottir, H.T.; Lundin, P.; Wallén Artzt, E.; Lindström, A.K.; Graff, C.; Eriksson, M. Somatic mutation that affects transcription factor binding upstream of CD55 in the temporal cortex of a late-onset Alzheimer disease patient. *Hum. Mol. Genet.* **2019**, *28*, 2675–2685. [[CrossRef](#)] [[PubMed](#)]
184. Shen, Y.; Yang, L.; Li, R. What does complement do in Alzheimer's disease? Old molecules with new insights. *Transl. Neurodegener.* **2013**, *2*, 1. [[CrossRef](#)] [[PubMed](#)]

185. Ivashko-Pachima, Y.; Hadar, A.; Grigg, I.; Korenková, V.; Kapitansky, O.; Karmon, G.; Gershovits, M.; Sayas, C.L.; Kooy, R.F.; Attems, J.; et al. Discovery of autism/intellectual disability somatic mutations in Alzheimer's brains: Mutated ADNP cytoskeletal impairments and repair as a case study. *Mol. Psychiatry* **2019**. [[CrossRef](#)]
186. Rohrbach, S.; Siddoway, B.; Liu, C.S.; Chun, J. Genomic mosaicism in the developing and adult brain. *Dev. Neurobiol.* **2018**, *78*, 1026–1048. [[CrossRef](#)]
187. Lähnemann, D.; Köster, J.; Szczurek, E.; McCarthy, D.J.; Hicks, S.C.; Robinson, M.D.; Vallejos, C.A.; Campbell, K.R.; Beerenwinkel, N.; Mahfouz, A.; et al. Eleven grand challenges in single-cell data science. *Genome Biol.* **2020**, *21*, 31. [[CrossRef](#)] [[PubMed](#)]
188. Loman, N.J.; Misra, R.V.; Dallman, T.J.; Constantinidou, C.; Gharbia, S.E.; Wain, J.; Pallen, M.J. Performance comparison of benchtop high-throughput sequencing platforms. *Nat. Biotechnol.* **2012**, *30*. [[CrossRef](#)] [[PubMed](#)]
189. Dou, Y.; Gold, H.D.; Luquette, L.J.; Park, P.J. Detecting Somatic Mutations in Normal Cells. *Trends Genet.* **2018**, *34*. [[CrossRef](#)]
190. Wang, K.; Lai, S.; Yang, X.; Zhu, T.; Lu, X.; Wu, C.-I.; Ruan, J. Ultrasensitive and high-efficiency screen of de novo low-frequency mutations by o2n-seq. *Nat. Commun.* **2017**, *8*. [[CrossRef](#)] [[PubMed](#)]
191. Kennedy, S.R.; Schmitt, M.W.; Fox, E.J.; Kohrn, B.F.; Salk, J.J.; Ahn, E.H.; Prindle, M.J.; Kuong, K.J.; Shen, J.-C.; Risques, R.-A.; et al. Detecting ultralow-frequency mutations by Duplex Sequencing. *Nat. Protoc.* **2014**, *9*. [[CrossRef](#)] [[PubMed](#)]
192. Ahn, E.H.; Lee, S.H. Detection of Low-Frequency Mutations and Identification of Heat-Induced Artifactual Mutations Using Duplex Sequencing. *Int. J. Mol. Sci.* **2019**, *20*, 199. [[CrossRef](#)] [[PubMed](#)]
193. Kim, J.; Kim, D.; Lim, J.S.; Maeng, J.H.; Son, H.; Kang, H.-C.; Nam, H.; Lee, J.H.; Kim, S. The use of technical replication for detection of low-level somatic mutations in next-generation sequencing. *Nat. Commun.* **2019**, *10*. [[CrossRef](#)]
194. Reiner, B.C.; Doyle, G.A.; Weller, A.E.; Levinson, R.N.; Namoglu, E.; Pigeon, A.; Perea, E.D.; Weickert, C.S.; Turecki, G.; Mash, D.C.; et al. Restriction Enzyme Based Enriched L1Hs Sequencing (REBELseq): A Scalable Technique for Detection of Ta Subfamily L1Hs in the Human Genome. *G3 Genes Genomes Genet.* **2020**, *10*. [[CrossRef](#)]
195. Doan, R.N.; Miller, M.B.; Kim, S.N.; Rodin, R.E.; Ganz, J.; Bizzotto, S.; Morillo, K.S.; Huang, A.Y.; Digumarthy, R.; Zimmel, Z.; et al. MIPP-Seq: Ultra-sensitive rapid detection and validation of low-frequency mosaic mutations. *BMC Med. Genom.* **2021**, *14*. [[CrossRef](#)]
196. Cooke, S.L.; Shlien, A.; Marshall, J.; Pipinikas, C.P.; Martincorena, I.; Tubio, J.M.C.; Li, Y.; Menzies, A.; Mudie, L.; Ramakrishna, M.; et al. Processed pseudogenes acquired somatically during cancer development. *Nat. Commun.* **2014**, *5*, 3644. [[CrossRef](#)]
197. Henssen, A.G.; Koche, R.; Zhuang, J.; Jiang, E.; Reed, C.; Eisenberg, A.; Still, E.; MacArthur, I.C.; Rodríguez-Fos, E.; Gonzalez, S.; et al. PGBD5 promotes site-specific oncogenic mutations in human tumors. *Nat. Genet.* **2017**, *49*, 1005–1014. [[CrossRef](#)]
198. Siudeja, K.; Beek, M.; Riddiford, N.; Boumard, B.; Wurmser, A.; Stefanutti, M.; Lameiras, S.; Bardin, A.J. Unraveling the features of somatic transposition in the *Drosophila* intestine. *EMBO J.* **2021**, *40*. [[CrossRef](#)]
199. Davie, K.; Janssens, J.; Koldere, D.; De Waegeneer, M.; Pech, U.; Kreft, L.; Aibar, S.; Makhzami, S.; Christiaens, V.; Bravo González-Blas, C.; et al. A Single-Cell Transcriptome Atlas of the Aging *Drosophila* Brain. *Cell* **2018**, *174*, 982–998.e20. [[CrossRef](#)]
200. Adams, M.D. The Genome Sequence of *Drosophila melanogaster*. *Science* **2000**, *287*, 2185–2195. [[CrossRef](#)]
201. von Bartheld, C.S.; Bahney, J.; Herculano-Houzel, S. The search for true numbers of neurons and glial cells in the human brain: A review of 150 years of cell counting. *J. Comp. Neurol.* **2016**, *524*, 3865–3895. [[CrossRef](#)]
202. Herculano-Houzel, S. The remarkable, yet not extraordinary, human brain as a scaled-up primate brain and its associated cost. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 10661–10668. [[CrossRef](#)]
203. Wu, S.; Turner, K.M.; Nguyen, N.; Raviram, R.; Erb, M.; Santini, J.; Luebeck, J.; Rajkumar, U.; Diao, Y.; Li, B.; et al. Circular ecDNA promotes accessible chromatin and high oncogene expression. *Nature* **2019**, *575*, 699–703. [[CrossRef](#)] [[PubMed](#)]
204. Møller, H.D.; Mohiyuddin, M.; Prada-Luengo, I.; Sailani, M.R.; Halling, J.F.; Plomgaard, P.; Maretty, L.; Hansen, A.J.; Snyder, M.P.; Pilegaard, H.; et al. Circular DNA elements of chromosomal origin are common in healthy human somatic tissue. *Nat. Commun.* **2018**, *9*, 1069. [[CrossRef](#)]
205. Morton, A.R.; Dogan-Artun, N.; Faber, Z.J.; MacLeod, G.; Bartels, C.F.; Piazza, M.S.; Allan, K.C.; Mack, S.C.; Wang, X.; Gimple, R.C.; et al. Functional Enhancers Shape Extrachromosomal Oncogene Amplifications. *Cell* **2019**, *179*, 1330–1341.e13. [[CrossRef](#)] [[PubMed](#)]
206. Koche, R.P.; Rodriguez-Fos, E.; Helmsauer, K.; Burkert, M.; MacArthur, I.C.; Maag, J.; Chamorro, R.; Munoz-Perez, N.; Puiggròs, M.; Dorado Garcia, H.; et al. Extrachromosomal circular DNA drives oncogenic genome remodeling in neuroblastoma. *Nat. Genet.* **2020**, *52*, 29–34. [[CrossRef](#)] [[PubMed](#)]
207. Shibata, Y.; Kumar, P.; Layer, R.; Willcox, S.; Gagan, J.R.; Griffith, J.D.; Dutta, A. Extrachromosomal MicroDNAs and Chromosomal Microdeletions in Normal Tissues. *Science* **2012**, *336*, 82–86. [[CrossRef](#)]
208. Qiu, G.-H.; Zheng, X.; Fu, M.; Huang, C.; Yang, X. The decreased exclusion of nuclear ecDNA: From molecular and subcellular levels to human aging and age-related diseases. *Ageing Res. Rev.* **2021**, *67*, 101306. [[CrossRef](#)] [[PubMed](#)]
209. Ain, Q.; Schmeer, C.; Wengerodt, D.; Witte, O.W.; Kretz, A. Extrachromosomal Circular DNA: Current Knowledge and Implications for CNS Aging and Neurodegeneration. *Int. J. Mol. Sci.* **2020**, *21*, 2477. [[CrossRef](#)] [[PubMed](#)]
210. Deshpande, V.; Luebeck, J.; Nguyen, N.-P.D.; Bakhtiari, M.; Turner, K.M.; Schwab, R.; Carter, H.; Mischel, P.S.; Bafna, V. Exploring the landscape of focal amplifications in cancer using AmpliconArchitect. *Nat. Commun.* **2019**, *10*, 392. [[CrossRef](#)]
211. Silva, A.C.; Lobo, D.D.; Martins, I.M.; Lopes, S.M.; Henriques, C.; Duarte, S.P.; Dodart, J.-C.; Nobre, R.J.; Pereira de Almeida, L. Antisense oligonucleotide therapeutics in neurodegenerative diseases: The case of polyglutamine disorders. *Brain* **2020**, *143*. [[CrossRef](#)]

-
212. Derbis, M.; Kul, E.; Niewiadomska, D.; Sekrecki, M.; Piasecka, A.; Taylor, K.; Hukema, R.K.; Stork, O.; Sobczak, K. Short antisense oligonucleotides alleviate the pleiotropic toxicity of RNA harboring expanded CGG repeats. *Nat. Commun.* **2021**, *12*. [[CrossRef](#)]
 213. Hornung, S.; Dutta, S.; Bitan, G. CNS-Derived Blood Exosomes as a Promising Source of Biomarkers: Opportunities and Challenges. *Front. Mol. Neurosci.* **2020**, *13*. [[CrossRef](#)]