DR CHRISTOS PROUKAKIS (Orcid ID : 0000-0001-6423-6539)

Article type : Invited Review

Title page

Title: Somatic mutations in neurodegeneration

Running title: somatic mutations

Authors:

Melissa Leija-Salazar, Charlotte Lucie Piette, Christos Proukakis*

Department of Clinical Neuroscience, Rowland Hill Street, University College London Institute of Neurology, London NW3 2PF, UK

*Corresponding author: c.proukakis@ucl.ac.uk

Abstract:

Somatic mutations are post-zygotic mutations which may lead to mosaicism, the presence of cells with genetic differences in an organism. Their role in cancer is well established, but detailed investigation in health and other diseases has only been recently possible. This has been empowered by the improvements of sequencing techniques, including single cell sequencing, which can still be error-prone but is rapidly improving. Mosaicism appears relatively common in the human body, including the normal brain, probably arising in early development, but also potentially during ageing.

In this review, we first discuss theoretical considerations and current evidence relevant to somatic mutations in the brain. We present a framework to explain how they may be integrated with current views on neurodegeneration, focusing mainly on sporadic late onset neurodegenerative diseases (Parkinson's disease, Alzheimer's disease, and amyotrophic lateral sclerosis). We review the relevant studies so far, with the first evidence emerging in Alzheimer's in particular. We also discuss

the role of mosaicism in inherited neurodegenerative disorders, particularly somatic instability of tandem repeats. We summarise existing views and data to present a model whereby the time of origin and spatial distribution of relevant somatic mutations, combined with any additional risk factors, may partly determine the development and onset age of sporadic neurodegenerative diseases.

List of abbreviations (in order encountered)

PD	Parkinson's disease
DLB	Dementia with Lewy bodies
AD	Alzheimer's disease
ALS	amyotrophic lateral sclerosis
FTLD	frontotemporal lobar degeneration
SNV	single nucleotide variant
CNV	copy number variants
SV	structural variant
TE	transposable element
mtDNA	mitochondrial DNA
SSB	single-strand break
DSB	double-strand break
FISH	fluorescence *in situ* hybridisation
CISH	chromogenic *in situ* hybridisation
WGS	Whole genome sequencing
WGA	whole genome amplification
PBL	peripheral blood leucocyte
DCV	DNA content variation

MSA multiple system atrophy

LOY loss of chromosome Y

HD Huntington's disease

GWAS Genome-wide association study

AT ataxia-telangiectasia

CS Cockayne Syndrome

XP Xeroderma Pigmentosum

Background and theoretical considerations

Mutations are termed "somatic" if they occur post-zygotically, and can lead to mosaicism, the presence of genetically different cells within a single organism [1, 2]. Somatic mutations underlie antigenic variation, and have long been linked to cancer, with decades of studies and technological advances supporting and elucidating their role [3]. The latest advances in DNA sequencing have shown mosaicism to be more prevalent than previously thought in humans [2, 4–6], whether arising in development or ageing, with the term "somatic evolutionary genomics" used to describe the study of the accumulation of somatic mutations in the body [7]. Somatic mutations can indeed be used to reconstruct the developmental cell lineage in an organism [7, 8]. Mosaicism can be classified as somatic (affecting the "soma", or body), gonadal (affecting the germline), and gonosomal (affecting both) [5]. Somatic and gonosomal mosaicism could affect the nervous system, but only gonadal and gonosomal are heritable. Indeed, an important recent realisation with profound implications in genetic counselling, is that apparently "*de novo*" mutations in a patient can actually reflect gonadal or gonosomal mosaicism in a parent [9]. Studying this is obviously easier in sperm than ova, with striking sperm mosaicism recently reported [10].

The existence of somatic mutations in apparently healthy nervous system tissues, and a possible role in non-neoplastic neurological disease, have only been explored relatively recently. The evidence for a role of somatic mutations in a wide range of neurodevelopmental and neuropsychiatric disorders is already strong [11]. A role of somatic mutations in neurodegenerative diseases has been repeatedly hypothesised [7, 12–14]. In this review, we discuss the theoretical basis and current evidence for somatic mutations in neurodegeneration, focusing on the common, mainly sporadic,

late onset neurodegenerative diseases such as Parkinson's disease (PD) and dementia with Lewy bodies (DLB), Alzheimer's disease (AD), and amyotrophic lateral sclerosis / frontotemporal lobar degeneration (ALS / FTLD) spectrum. Somatic mutations may represent good candidates to explain some of the main features of these illnesses, such as sporadic occurrence, focal onset, and progressive spread of the pathology [15]. We will also discuss rare and inherited conditions, including mixed neurodevelopmental and neurodegenerative disorders, where relevant.

Genetic variation between individuals can occur due to several broad classes of mutations: single nucleotide variants (SNVs) and indels (small insertions / deletions), copy number variants (CNVs) and other structural variants (SVs), tandem repeats such as trinucleotides, L1 retrotransposons and other transposable elements (TEs), and aneuploidy. All these types of variations can also occur somatically, leading to mosaicism [6], and have now been reported in normal brain [11]. We use the term somatic mutation to refer to any acquired post-zygotic change in the nuclear genome, from single base to whole chromosome, encompassing all the above. Broadly, aneuploidy arises from spindle errors in cell division, TE insertions arise from their retrotransposition activity, while other mutations arise during either DNA replication, or repair of DNA damage (single-strand breaks (SSBs) and double-strand breaks (DSBs)) [16, 17]. DNA damage may lead to transient chemical lesions, but if they are erroneously repaired, a somatic mutation will result [18].

In its broader definition, mosaicism can be considered to result from epigenetic changes rather than mutations [1]. There is also the special case of mitochondrial DNA (mtDNA), which is highly variable between, and even within, cells (heteroplasmy). Although these are both beyond the scope of this review, we should note that a role of mtDNA mutations in neurodegeneration, particularly deletions in PD, has been repeatedly claimed, with the latest evidence suggesting accumulating deletions and relative depletion of normal mtDNA are most relevant [19].

Are somatic mutations likely to be common, and could they confer an early selective advantage?

The "disposable soma" theory postulates that less effort will be invested in maintenance of an organism, than in securing it achieves reproduction [20]. Extrapolating this concept to mutations, one might expect a higher somatic than germline mutation rate. Indeed this was recently demonstrated for SNVs in both humans and mice, with a human somatic mutation rate of 2.66×10^{-9} per base per mitosis in fibroblasts [21]. With a minimum number of cell divisions estimated at 10^{13} to 10^{14} [7], this would lead to each base mutating several thousand times, and each cell could have hundreds of deleterious mutations [22]. As these estimates exclude other mutation types, and have

not been performed in rapidly proliferating neuronal precursors, the total burden of somatic mutations in neurons could be even higher.

Although somatic mutations in a normal brain might appear at first to only have negative consequences, they have been proposed to confer advantageous diversity [23–25]. It was first suggested half a century ago that a form of "chromosomal programming", analogous to the immune system recombination of antibody-coding genes, might underlie the extreme diversity and specialisation of neurons [26], and later that neurons may pass through a stage when they are particularly sensitive to DNA DSB [27]. DSB in mouse neuronal precursors were first reported almost 20 years ago [28], and recently demonstrated conclusively, mostly in genes involved in synaptic function or neural cell adhesion [29]. A specific developmental "window" for sensitivity to DNA damage *in vivo* has now been demonstrated in mouse enteric neuronal cells [30]. Neuronal precursors therefore join lymphocytes as cells with naturally high rates of localised DSB, providing a mechanism for generation of mosaicism in early embryogenesis [31], although distinct from the mechanism of immunoglobulin gene rearrangements [27].

A particular issue in the developing nervous system is competition for survival from apoptotic programmed cell death of both neuroblasts and post-mitotic neurons [27]. It appears unlikely that cells which ultimately survive are selected randomly, as "an advantage must be conferred on them by genetic differences" [32], with somatic mutations a possible substrate for positive and negative selection [27]. Somatic mutations arising exclusively in embryonic neurogenesis should have a stable level in ageing post-mitotic neurons, unless they impact neuronal survival. There are several possible reasons, however, why they might increase with age [33, 34]. Oxidative stress, which increases with age, could cause mutations by errors in DNA repair. Aberrant cell cycle re-entry as a protective mechanism could trigger large-scale DNA gains. Transcription-related DNA damage could also lead to mutations in post-mitotic neurons. A fascinating study showed that neuronal activity can lead to DSB in mice [35]. Such events might be limited to single neurons individually, but common collectively. Finally, mutagenesis in adult brain neurogenesis is unexplored.

Evidence for existence of somatic mutations in healthy brain

The field is evolving so rapidly that we feel an overview of the current state is warranted. Aneuploidy has been traditionally studied using fluorescence *in situ* hybridisation (FISH) or its chromogenic equivalent (CISH), which can visualise single nuclei directly. DNA analysis has been revolutionised by "next generation" sequencing, whether targeted or whole genome. Traditionally, DNA is extracted

from the tissue of interest, and this "bulk" DNA is analysed. One obvious limitation is the level at which a somatic mutation of any type is present, as it may be difficult to distinguish it from background. In addition, DNA extraction can lead to significant GC-dependent bias, which may confound determination of subtle copy number differences [36, 37]. DNA quality and results could also be affected by pre-morbid factors like hypoxia, which leads to brain DNA fragmentation particularly in elderly rats [38], and post-mortem interval. Interpretation of data from fixed paraffin-embedded (rather than frozen) specimens may require additional attention [39]. Single-cell sequencing has revolutionised somatic mutation studies [40]. Whole genome sequencing (WGS) after whole genome amplification (WGA) of single cells is now possible, and the amplification, sequencing and analysis protocol can be geared towards the type of variant of interest. The main issues related to this method, apart from cost limiting the sequencing "depth" or "coverage", and the number of cells, are efficient cell characterisation and isolation, and accurate and even amplification of single-cell DNA, which is influenced by the WGA method chosen [11, 40].

Aneuploidy (the gain or loss of chromosomes) was the first type of mosaicism described in mouse brain, where one third of neuroblasts appear aneuploid [41], due to chromosome segregation defects [42], and later in adult humans [43]. A mean aneusomy (aneuploidy rate per chromosome) of 0.1-0.8% was found in post-mortem adult neurons [44], and >1% in foetal brain [45]. Neuronal aneusomy for chromosome 21 was estimated at ~4% [43] and ~12% [46]. Single cell sequencing studies have confirmed the existence of aneuploidy in healthy human brains, but at apparently lower rates, with overall aneuploidy 0.7%-5% in single frontal neurons [47–49], and 2.2% in unselected cells from frontal lobe grey matter [50]. The higher frequency of aneuploid neuroblasts suggests selection against them, and this was indeed demonstrated in a mouse study where inhibition of programmed cell death led to a marked rise of aneuploid cells, with increased levels of 'extreme' aneuploidies [51].

CNVs are gains or losses of DNA fragments (size cut-off defined variably as 50 bases - 1 kb). Several WGS studies, mostly of single control frontal cortical neurons, showed individual neuronal CNVs as aberrations in read depth at very low coverage WGS (<1x) [47, 48, 52, 53], although one which looked for CNVs as a secondary aim did not find any definitive evidence [24](table 1). Only large (megabase-scale) CNVs were detectable, except for a study using a microfluidic platform, which minimised WGA variability and called smaller CNVs, mostly gains [52]. Importantly, reanalysis of data with other algorithms can lead to fewer CNV calls [53, 54], highlighting the pitfalls of current methods. Some CNVs were shared between cells, suggesting a developmental origin [47, 53, 55]. In addition to simple CNVs, various types of structural variants (SVs) exist, and translocations or

inversions can occur alone, not leading to CNVs, or combined with copy number changes in complex patterns. The complexity of inherited SVs is becoming increasingly evident [56], and their origins can be traced back to processes known as chromothripsis and chromoanasynthesis [57], first described in cancer. A remarkable cloning study generated viable mice from single olfactory neurons, and thus allowed robust calling of SVs and other mutations, free of WGA artefacts [58]. An average of 1.5 SV per neuron were reported, of which one third were complex, suggesting chromothripsis. Most were very small (seven of nine <35 kb), but the cloning protocol may have selected against neurons with larger genomic abnormalities. The SV origin was felt most likely to be post-mitotic.

L1 retrotransposition is the insertion of mobile LINE-1 (L1) elements in a new genomic location. L1 elements comprise one sixth of the human genome, and can "copy-paste" themselves back into new positions in the genome, leading to somatic L1 insertions [59]. Retrotransposition activity was demonstrated in neuronal precursor final divisions [60], and it may also occur in post-mitotic neurons [61]. Somatic insertions were detected in adult human brain [62], and confirmed in several single cell studies, mostly reporting <1 insertions per human neuronal genome [63–65], and a similar frequency in glia [65], and mouse olfactory neurons [58]. Hippocampal neurons may each have 13.7 L1 insertions [24], although reanalysis by another group suggested a true positive rate ~0.2 per neuron [66]. While the frequency may vary between neuronal types, and analysis methods require standardisation, it is clear that L1 brain mosaicism does exist.

DNA content variation (DCV) may occur in individual neurons, with frontal neurons having up to 250 Mb extra DNA compared to cerebellar neurons or lymphocytes from the same healthy individual [67]. This phenomenon affects ~11.5% of neurons, with no difference across cortical regions, and the frequency of neurons with increased DNA content declining with age [68]. This effect was not noted in another study, although the data suggest a possible similar trend in non-diseased prefrontal cortex and cerebellar neurons [69]. This suggests that neurons with higher DNA content may be preferentially prone to loss in ageing. DCV could be due to gains arising through a variety of mechanisms, including aneuploidy, CNVs, and L1 or other TE insertions. The low prevalence of aneuploidy and somatic L1 insertions make them unlikely to be major contributors, and the predominance of gains in one single neuron CNV study may be relevant [52].

SNVs ("point mutations") may be the commonest somatic variants in normal brains [11]. Deep WGS of single control frontal cortical neurons revealed ~1500 somatic SNVs in each (80% C>T transitions) [70]. The subsequent study of cloned mouse olfactory neurons estimated <100 SNVs per neuron (40% C>T)[58]. These SNVs may arise post-mitotically [58, 70], leading to the question of whether they increase with age, although the typical signatures of mutations arising from aberrant repair

after oxidative stress were not seen [58]. Very recent work by the same group showed increasing single neuron SNVs with age in prefrontal cortex and hippocampus [71]. This is consistent with a study using WGS of "bulk" DNA, with modifications of the library preparation to detect low level mutations, which reported >1 somatic SNV per 10 million bases in frontal cortex from children, rising six-fold in older adults [72]. One concern has been the abundance of C>T changes, which can arise by chemical modification of DNA [22, 73]. They were, however, shown to vary with age, suggesting a biological cause [71].

Distribution of brain somatic mutations depends on when they arise (figure 1). Those arising in very early development, before the germ layers split in gastrulation, might be detectable in a wide range of tissues, including blood, albeit at possibly very low levels [14, 15]. Indeed, somatic SNVs present in >5-10% of brain cells were also found outside the brain, including in tissues derived from mesoderm and endoderm [70]. After gastrulation, somatic mutations affecting brain tissues would be restricted to the ectodermal lineage, but even at this stage they could be widely dispersed throughout the nervous system. It has become clear that there is profound mixing of neuronal cells in very early development, with an example of a somatic SNV found throughout the cerebral cortex, cerebellum and spinal cord [70]. Mutations occurring later, but still in dividing neuronal precursors, would reach a lower level and be restricted to a smaller brain region, with examples of such SNVs and L1 insertions also demonstrated [64, 70]. Post-mitotic mutations should be confined to one cell, but could be common collectively.

Rationale for a role of somatic mutations in neurodegeneration

The commonest neurodegenerative diseases (AD, PD, ALS) are generally sporadic, but inherited cases due to mutations in several genes are well described. These genes are prime candidates for somatic mutations leading to sporadic cases, possibly with milder phenotypes [74]. This would be analogous to hypertension caused by primary aldosteronism, where somatic mutations lead to milder phenotypes than inherited [75]. Notably, inherited mutations in several genes do lead to earlier onset than sporadic disease, eg *SNCA* and recessive genes in PD [76, 77], *APP* and particularly *PSEN1* in AD [78], and *FUS*, *SOD1*, and genes of juvenile forms in ALS [79].

It appears likely that an individual somatic mutation could only have a significant effect if it arose early enough to be present in a substantial proportion of cells of the relevant type. The actual level needed in a region or cell type to contribute to neurodegeneration is completely unknown. Accumulating evidence suggest the importance of spread of pathology in neurodegenerative

disorders [80]. If a somatic mutation were responsible, it could be a local trigger that leads to spread from a small number or even a single mutant cell [7, 15, 81]. Alternatively, widespread somatic mutations, even at low levels, could increase individual cell vulnerability to a spreading agent. Conversely, or additionally, other factors, such as intrinsic neuronal type vulnerability, may be important in pathology development, at least in PD [82–84], which may have multifocal onset [85]. Could somatic mutations have a role in such a scenario? Clearly this could not be a localised mutation [86], but there are two non-mutually exclusive possibilities. Somatic mutations arising independently in many locations could affect the "threshold" for disease development. Alternatively, a single mutation in early embryogenesis could "spread" to multiple tissues, even if spatially separated, if their lineage is shared, for example in PD where multiple neural crest-derived structures may be affected early [14]. The pattern of dispersed somatic mutations could determine the clinical and pathological phenotype, eg with *SNCA* mutation distribution determining a predominantly PD or DLB picture, with variable peripheral involvement, or indeed multiple system atrophy (MSA) if the glia and autonomic system were more affected. Somatic mutations may also affect penetrance of inherited mutations, including in discordant monozygotic twins [70], such as a twin pair with *LRRR2* G2019S mutation discordant for PD [87].

A significant role of somatic mutations in sporadic neurodegenerative diseases appears more likely if the following conditions are met.

(1) Increased somatic mutability of genes involved in sporadic neurodegenerative diseases

Fragile sites are genomic regions prone to damage particularly during DNA replication stress, where somatic CNVs may arise [88, 89]. Two key genes causing inherited PD, *SNCA* and *PARK2,* are within common fragile sites [90, 91], although it is not known if they are fragile in neuroblasts. A gain of several megabases across that fragile site, including *SNCA*, was indeed reported in 1/110 frontal neurons from controls [47]. Additionally, *PARK2* is unstable in various cancers, including melanoma [92]. Furthermore, somatic L1 insertions in both these genes were reported in caudate neurons from controls [62]. *GBA*, a major genetic risk factor in PD and DLB, has a highly homologous nearby pseudogene, which results in SV caused by aberrant recombination. Although there are no such known somatic events, we hypothesise that they may occur, and could account for example for the loss of glucocerebebrosidase enzymatic function in PD patients not carrying inherited mutations [93].

(2) Increased tendency of patients with sporadic neurodegenerative diseases to somatic mutations in general

A generalised tendency to develop somatic mutations might be detectable in other tissues, perhaps even blood. Studies of DNA damage in response to UV radiation in lymphoblasts showed that PD and AD-derived lines, but not ALS, were more sensitive than controls [94, 95]. Micronuclei, nuclear structures where chromothripsis may occur [96], were found more often in peripheral blood leucocytes (PBL) from PD and AD than controls [97]. The increased occurrence of melanoma in PD remains unexplained, with melanocytes sharing a neural crest origin. Melanocytes from individuals with inherited risk of PD may have genomic instability predisposing to both, with UV damage triggering melanoma more readily [98].

(3) Conferring a selective advantage in development

Several neurodegenerative disease genes have important roles in early neurodevelopment, eg *SNCA*, which regulates the synaptic vesicle pool [99], and *APP*, which increases axonal growth cone size when overexpressed [100]. One can therefore speculate the existence of somatic mutations beneficial to the developing neuron, but predisposing to protein aggregation and neurodegeneration in the ageing brain [14](“antagonistic pleiotropy” [101]).

Evidence for a role of somatic mutations in neurodegeneration

Synucleinopathies (Parkinson’s disease, dementia with Lewy bodies, multiple system atrophy)

Limited work has been done to investigate mosaicism in synucleinopathies. We analysed *SNCA* coding exons in at least one brain region in over 500 cases of PD and DLB using a PCR-based assay, able to detect SNVs present at levels of at least 5-10% [14, 102] (table 2), but did not find any. Early studies investigated genome-wide CNV mosaicism using microarrays. One found some specific to PD brains, but they were not independently confirmed, and did not involve PD genes [103]. An interesting study of PBL from monozygotic twins discordant for PD / DLB showed different post-zygotic CNVs, but again none involved PD genes, and no brain tissue was analysed [104]. One very large mosaic duplication involving *SNCA* was reported in PBL of a 69-year old case, which appears likely pathogenic, but again no brain tissue was available [105]. An interesting study claimed high level mosaicism for *SNCA* gains in ectodermal cells from the oral mucosa in two early onset cases

based on FISH [106], but another assay (multiple ligation probe amplification, MLPA) was negative, even in the case with apparent 75% mosaicism, and no brain was available.

Mosaicism for large scale genomic changes was investigated in three studies. One group reported mosaic aneuploidy (gains for both chromosomes studied) in PD nigra neuromelanin-positive neurons, and proposed aberrant DNA synthesis leading to whole genome duplication [107]. Another study assessed DCV in neurons selected by NeuN; in Lewy body diseases, several cell types, including nigral pigmented neurons, had higher DNA content than controls [108]. Interestingly, the same group did not find DCV in MSA [109].

Alzheimer's disease

Three different targeted studies of brain DNA are summarised in table 2. A single case of a somatic SNV apparently directly causing AD (a missense mutation in *PSEN1*) has been reported [110]. Interestingly, this was only detected because it was transmitted to the offspring, who was therefore heterozygous for the mutation, and developed a severe phenotype. The level of the mutation in the mosaic parent was 14% in cerebral cortex and 8% in PBL. Although the germline DNA was not available, the parent clearly was a gonosomal mosaic. The phenotype was milder in the mosaic, as discussed earlier [74]. The mutation must have arisen very early in embryogenesis, leading to gonosomal mosaicism (see figure 1).

A later systematic study using very high coverage sequencing of AD genes in entorhinal cortex from sporadic cases and controls, focusing primarily on SNVs, detected three somatic mutations at ~1% level, but none in the most relevant genes in patients [111]. Two of these were novel missense mutations in *MAPT*, likely to be damaging, in patients, while the third was a known *PSEN2* variant in a control. The study had a low validation rate (3/107), illustrating the difficulties in detecting low level SNVs in "bulk" DNA. No evidence of mosaic gains in *APP* or other AD genes was found, but sensitivity was estimated at 10% or more of cells having a gain [111]. A study using multiple methods, including single cell analysis, found somatic gains in *APP* in AD prefrontal cortical neurons, with up to 12 copies per cell [69]. The mean copy number, assessed by single cell qPCR for two different exons, was 3.8 and 3.4 copies for AD cortical neurons, compared to 1.6 and 1.44 for controls (with cerebellar neurons around 2). A recent single neuron WGS study did not detect *APP* gains in AD [49], but it was designed for aneuploidy detection, and the mean number of reads per cell (324,000) may have been inadequate for CNV calling.

DCV with increased DNA content in AD has been reported by three groups. The study showing *APP* gains, performed by the group which had originally described DCV, found a 8% increase in DNA content of AD frontal cortical nuclei compared to controls [69]. Another group showed that neurons with higher DNA content were preferentially lost in AD [112], and the increased DNA content in AD was highest in the most vulnerable cortical areas [113]. In a study of DLB and controls, DNA content increase in the cortex and hippocampus correlated with Alzheimer pathology [108]. Because of the presence of *APP* on chromosome 21, and the increased incidence of AD in trisomy 21, aneusomy of this chromosome has been of particular interest, and could contribute to DCV. One study reported excess chromosome 21 aneusomy in AD frontal cortical neurons against controls (10.7% v 1.7%), although losses were almost as common as gains [114]. Another early study showed no difference in disease and control brains, although a surprisingly high aneusomy level was claimed [46]. More recent studies of single neurons did not find excess chromosome 21 aneusomy in Alzheimer [69], including by single cell WGS [49]. It thus appears that the increase in *APP* copy number, and in DNA content overall, in AD cortex, is not primarily due to chromosome 21 aneusomy. Interestingly, these gains might not arise in embryogenesis. One hypothesis of AD pathogenesis considers aberrant cell cycle re-entry, with DNA synthesis in a non-dividing neuron ultimately leading to apoptosis, as an early event [115]. DNA damage has been reported in early AD [116], with evidence of deficient DSB repair [117], and this could trigger cell cycle re-entry [118].

An intriguing finding, potentially linking AD to mosaicism outside the CNS, is the mosaic loss of chromosome Y (LOY) in PBL throughout lifetime, associated with shorter survival and increased risk of cancer [119]. AD males showed a higher LOY compared to controls, and prospectively men with LOY mosaicism in PBL had greater risk of being diagnosed with AD later [120]. Defects in immune surveillance were proposed to underlie this association.

Other sporadic neurodegenerative diseases

An early study of cerebral cortex DNA, the first looking for somatic mutations in a neurodegenerative disease to our knowledge, did not find any somatic *SOD1* mutations [121]. We are not aware of any somatic mutation in a known ALS gene, on an inherited normal allele, causing sporadic disease, although the special case of *C9ORF72* is discussed in the next section. WGS of PBL DNA in monozygotic twins discordant for ALS did not reveal any relevant somatic mutations [122], although interestingly epigenetic modifications were recently reported [123]. There has been one report of a sporadic case of Creutzfeldt-Jakob disease with a known disease-causing mutation which was absent

in the parents, and appeared to be a somatic event [124]. This was estimated at a level of 97% in PBL and brain, and must have therefore arisen in one of the first post-zygotic cell divisions (and presumably undergone positive selection to reach such a level). The alternative possibility, of a partial sample contamination, with the mutation actually arising in the parental germline, cannot be excluded.

A possible role for glia

Glia have been studied less, despite their likely importance in neurodegeneration. Glial aneuploidy was reported to increase with age in mouse cortex [125]. The existence of neurodegeneration in histiocytosis prompted a mouse study of a somatic mutation in the B-raf proto-oncogene introduced in the microglial lineage. Remarkably, this led to a severe late-onset neurodegenerative disorder [126].

Somatic mutations in genes causing rare genetic neurodegenerative diseases

There are numerous rare familial neurodegenerative conditions, such as hereditary spastic paraplegia (HSP), with some cases with no family history [127]. In addition to recessive inheritance, incomplete penetrance, and *de novo* mutations arising in meiosis, another possibility is that the unaffected parent is a mosaic [5]. This was documented in a family with spastin-related HSP, where the unaffected mother had mosaicism in her PBL, and presumably germline, DNA [128]. This is similar to the AD case discussed, although the mosaic parent there had been affected [110]. A further twist in X-linked dominant disorders, such as a form of neurodegeneration with brain iron accumulation, is that males with severe mutations may only be viable if they are mosaics [129].

Somatic mutations as a mediator or modifier in inherited neurodegenerative diseases

MAPT mutations underlie a subset of FTLD. Peripheral tissues (fibroblasts, PBLs, and lymphoblasts) from patient carrying *MAPT* mutations had somatic chromosomal aberrations [130], and two transgenic mouse models had excess aneuploidy in PBLs, although the brain was not studied [131]. In Drosophila tauopathy models, the 4-repeat tau protein isoform expression affected mitosis, inducing chromosomal missegregation and higher levels of aneuploidy in neuronal tissue [132]. Studies of human brains with *MAPT* mutations will be needed to confirm these intriguing findings.

Expansion of microsatellite repeats underlies many inherited neurodegenerative disorders including Huntington's disease (HD), several spinocerebellar ataxias, and ALS / FTLD caused by *C9ORF72* expansions. The inherent instability of such repeats has prompted a search for somatic instability leading to mosaicism, which has been found in HD, and several other CAG repeat disorders [133]. In HD, the inherited repeat length correlates with the chance of somatic expansion [134]. The age of onset variability is not fully explained by the inherited repeat size, and the expansion size in the prefrontal cortex also correlates with onset age [135]. Somatic instability is tissue-specific, with the cerebellum apparently spared [135], and the striatum particularly affected, which may explain the differential vulnerability [136]. The somatic instability appears to be due to DNA damage and errors in repair in post-mitotic neurons [137]. Remarkably, suppression of somatic instability delayed the phenotype in a mouse model [138]. In GWAS analysis of CAG repeat disorders as a whole, DNA repair genes were associated with onset age, further suggesting the importance of somatic expansion [133]. Somatic instability may influence the rate of progression, with a coding *MSH3* variant having the strongest association in a GWAS [139]. *MSH3* is a neuron-specific gene for DNA mismatch repair, already implicated in the regulation of somatic instability in HD mouse models, and in patients with a different expansion disorder, myotonic dystrophy type 1 [140]. These data point to a mechanism whereby HD is modified and exacerbated by both the inherent somatic instability of the causative mutation, and inherited alleles which influence the somatic mutation burden. It is plausible that reduced penetrance alleles in HD (36-39 repeats), and intermediate alleles (27-35 repeats), which may cause disease very rarely [141], become pathogenic only when expanded in the striatum. This could apply to other CAG repeat disorders. In a case of MSA with additional SCA17 pathology and an intermediate expansion size of 41, however, no evidence of further somatic expansion was found [142].

In ALS, the repeat expansion in *C9ORF72* is somatically unstable, when an expanded allele is present already [143, 144]. The variable phenotype caused by *C9ORF72* expansions could be due to different patterns of somatic expansion [145]. In a study of multiple tissues in a series with and without pathogenic expansions [144], there were two cases with modest expansions in PBL (61-92) and extreme ones in CNS (highly variable, up to 3500). Given their incomplete penetrance, small or borderline expansions may only be pathogenic if they expand somatically in the CNS. Sporadic cases could be due to *de novo* somatic expansions from an allele with a normal repeat length, in which case testing of PBL might not reveal the expansion if it was limited to the brain, or ectoderm. However, repeats sizes up to 15 appear stable [144], consistent with an earlier suggestion that pathogenic mosaicism may not arise from a truly normal allele [146].

Several inherited DNA repair disorders have neurological phenotypes [16, 34]. Although severe DNA damage leads to early cell loss and microcephaly, progressive neurodegeneration also occurs in some of these, with ataxia as the common feature, suggesting that these DNA repair pathways may be most important in the cerebellum. Ataxia-telangiectasia (AT) is caused by mutations in *ATM*, which lead to defective DNA repair, particularly affecting DSBs. Patients usually develop cancer and progressive neurodegeneration, most prominent in the cerebellum. Excess aneuploidy has been reported in AT brains [147]. Disorders of DSB repair dysfunction with neurodegeneration include Xeroderma Pigmentosum (XP) and Cockayne Syndrome (CS). XP has more neurodegeneration in general, except for the cerebellum, where neuronal loss is more severe in CS [148, 149]. The pathogenesis of ataxia with oculomotor apraxia 2, and the related ALS4, may also involve impaired DSB repair and other deficiencies of genome maintenance [150]. In ataxia with oculomotor apraxia 1, SSB repair is affected [34]. Firm conclusions on how the DNA repair deficit leads to neurodegeneration require detailed sequencing studies of affected brains, to determine whether the main driver are mutations resulting from aberrant repair, or apoptosis as a direct result of DNA damage [151]. The former explanation is supported by very recent work showing greater than two-fold age-adjusted increase in single neuron SNVs in CS and XP [71]. Another possibility, at least for AT, is compromise of a different protein function altogether [152]. Interestingly, ALS caused by mutations in *FUS* also belongs here, as the protein is involved in DSB repair, and patients had significantly higher levels of DNA damage in motor cortex compared to controls [153]. Finally, progranulin, a cause of familial FTLD with variable phenotypes, also appears to regulate a DNA repair pathway [154].

Could proteins involved in neurodegeneration cause somatic mutations?

We have focused on the possibility of somatic mutations contributing to neurodegeneration. The reverse question is whether protein aggregation, inflammation, and other processes occurring during neurodegeneration, can cause DNA damage, and somatic mutations in post-mitotic neurons. DNA breaks induced by neuronal activity in mice were higher in *APP* transgenic mice, and potentiated by Aβ oligomers in neuronal cultures [35]. Tau oligomers damaged DNA in hippocampal neurons from transgenic mice [155], and alpha-synuclein caused DNA breaks in human cell models [156]. The idea of accumulating protein oligomers leading to somatic mutations is tantalising, but requires further evidence.

Conclusions, study design considerations, and future directions

The evidence of mosaicism in apparently healthy brain tissue has increased rapidly in recent years. The ultimate cause of sporadic neurodegenerative disease remains partly unexplained. We have discussed the theoretical basis and evidence so far for a role of somatic mutations in neurodegenerative disorders. Somatic expansion can clearly influence the phenotype of inherited expansion repeat disorders. The real challenge is to determine how much of a role somatic mutations may have in predominantly sporadic disorders like PD, AD, and ALS. The evidence so far is stronger in Alzheimer's, with common *APP* somatic CNVs and rare potentially relevant SNVs, but there is still only one case which was proven to be due to a somatic mutation [110], and the *APP* copy number gains in sporadic AD have not been fully defined [69]. If age of onset and / or severity corresponds to the somatic mutation burden [7, 14], younger and more severe patients should be studied, but they may, however, also be the ones most likely to carry inherited risk alleles, confounding the picture. Clear-cut cases, where a particularly detrimental somatic mutation essentially causes the disease, may be relatively rare. If, however, somatic mutations in relevant genes are relatively common, even at low or modest levels, they may act as risk factors, alongside inherited risk alleles, and any epigenetic and environmental influences. The origin and distribution of the mutation(s) could determine the phenotype and age of onset, with late or low level mutations perhaps clinically silent (figure 2).

Study design should depend on the mutation type and level being sought. If somatic mutations are seen as potential initiators of spread, rather than determinants of regional vulnerability, very low levels may be relevant. Targeted detection of low level SNVs can be optimised by use of "molecular barcodes" to minimise PCR artefacts [157]. For post-mitotic changes, single neuron strategies appear mandatory. Although single neuron WGS has revolutionised the field, deep WGS of hundreds of single neurons from multiple cases and controls is beyond the reach of most labs. Large-scale collaborations, such as the Brain Somatic Mosaicism Network for neurodevelopmental and neuropsychiatric disorders [11], are the likely answer. As discussed earlier, the technical and bioinformatics aspects require further development, as orthogonal validation in DNA from a single neuron is essentially impossible once it has been amplified. New WGA methods may minimise artefacts and uneven amplification [73, 158]. Fluorescence-activated sorting has been used to isolate nuclei from specific neuron types, although laser-capture can select neurons in their proper spatial location [40], and was indeed used for targeted *SOD1* sequencing in cortical motor neurons [159]. Sorting methods can also be used to isolate pools of specific cell types [11].

An important "elephant in the room" is the fact that the likely consequence of a detrimental somatic mutation is the death of the neuron carrying it, before the death of the individual, unlike cancer, where a mutation would lead to clonal proliferation. The absence of somatic mutations in an end-stage brain may simply be a result of loss of the cells carrying them. This has been illustrated in AD, where neurons with increased DNA content are preferentially lost [112], and HD, where large somatic gains are present in the cortex, but not the striatum, where somatic instability is pronounced, but neurons are inherently more vulnerable to CAG expansions [136]. Analysing autopsy cases with short disease duration may ensure that some neurons with relevant somatic mutations are still present. Utilising healthy individuals with incidental autopsy findings, such as incidental Lewy body disease which may represent preclinical PD [160], is another strategy, but these individuals might have never developed the disease clinically.

The brain region chosen is also a matter for discussion. It would intuitively make sense to analyse the brain region or cell type affected earliest, or most severely, yet this may be where mutation-carrying vulnerable cells do not survive. Using an alternative region, where the cells are resistant to the effects of mutations in inherited cases (eg cerebellum in familial PD), might help, but only if the mutation(s) occurred in a shared lineage. Tissues accessible during life, which could be used as biomarkers, are worth considering, such as salivary glands in PD [161]. CSF could also be a source of DNA from dying neurons [162]. An often neglected easily accessible source of ectodermal DNA are hair follicles [163] (figure 1). Comparison of DNA from different germ layers is often helpful, and if no blood is available, meninges can be used as a source of mesodermal DNA.

In conclusion, somatic genomic variation in the brain is increasingly recognised. We have discussed the theoretical considerations regarding a role of somatic mutations in neurodegeneration, especially in sporadic cases where the ultimate cause or trigger remains largely unproven, and presented the emerging evidence in Alzheimer's. There is clear need for investment into considering the contribution of somatic mutations to disorders which are becoming more frequent as the population ages. Large studies, coupled with improvements in single-cell sequencing, are required for reliable estimation of the somatic mutation burden in healthy and diseased brains.

Note added: A very recent study of clonally expanded human foetal forebrain cells has shown 200-400 SNVs in each. The mutation rate during neurogenesis was estimated at 5.1 per day per progenitor, and was higher than in the first few cell divisions [164].

Acknowledgements

We would like thank the reviewers for their the extremely helpful comments. Christos Proukakis acknowledges funding from the Michael J Fox Foundation to investigate somatic mutations in Parkinson's disease. Melissa Leija-Salazar is funded by a fellowship from CONACYT.

Author contributions

Melissa Leija-Salazar and Charlotte Piette wrote the initial draft. Christos Proukakis recommended a structure for the review, substantially advanced the draft, and prepared the final version.

Conflicts of interest

None of the authors have any conflicts of interest.

Captions for tables and figures

Table 1. Summary of human single neuron CNV studies. Frontal cortical neurons from healthy individuals, selected by sorting with NeuN, unless specified otherwise.

Table 2. Studies of brain DNA in PD and AD targeted to low level mutations in specific genes.

Figure 1. Distribution of somatic mutations depending on timing of origin. In this highly simplified view of development, a zygote, outer view of a blastocyst (day 4), and the trilaminar disc where the three germ layers separate (day 16), are shown. Squares with different colours represent mutations which may be shared between several tissues. Green squares may be present in cells derived from all three germ layers. Blue squares will be restricted to ectodermal-derived tissues, red squares to mesodermal, and orange squares to endodermal, although they may not be present in all tissues from that germ layer. Triangles are later mutations, restricted to a specific tissue type. A "green square", "blue square" and "purple triangle" mutation can all affect the brain, but they will be potentially detectable in all other tissues, ectodermal tissues only, and nervous system only, respectably. Green squares may lead to gonosomal mosaicism, blue and orange squares and red triangles will lead to somatic mosaicism only, and pink triangles in the germline will lead to germline mosaicism only.

Figure 2. Schematic representation of the effect of somatic mutations on the risk and severity of sporadic neurodegenerative disease. This may depend on the somatic mutation overall burden, which will be influenced by the timing of their occurrence, but also selective pressures. The

likelihood of each outcome, from severe disease to no pathological changes, is represented by the width of each arrow. This may be on a continuum, influenced by the mutation burden, and their functional consequence, but also by inherited risk alleles, epigenetic changes, and any environmental factors. Even if somatic mutations contribute to a disease, it could still arise in cases with none, due to other risk factors.

References

1. Youssoufian H, Pyeritz RE. Mechanisms and consequences of somatic mosaicism in humans. *Nat Rev Genet* 2002; **3**: 748–58

2. Biesecker LG, Spinner NB. A genomic view of mosaicism and human disease. *Nat Rev Genet* 2013; **14**: 307–20

3. Martincorena I, Campbell PJ. Somatic mutation in cancer and normal cells. *Science* 2015; **349**: 1483–9

4. Macosko EZ, McCarroll SA. Exploring the variation within. *Nat Genet* 2012; **44**: 614–6

5. Campbell IM, Shaw CA, Stankiewicz P, Lupski JR. Somatic mosaicism: implications for disease and transmission genetics. *Trends Genet* 2015; **31**: 382-92

6. Forsberg LA, Gisselsson D, Dumanski JP. Mosaicism in health and disease — clones picking up speed. *Nat Rev Genet* 2017; **18**: 128-42.

7. Frank SA. Evolution in health and medicine Sackler colloquium: Somatic evolutionary genomics: mutations during development cause highly variable genetic mosaicism with risk of cancer and neurodegeneration. *Proc Natl Acad Sci USA* 2010; **107** Suppl: 1725–30

8. Woodworth MB, Girskis KM, Walsh CA. Building a lineage from single cells: genetic techniques for cell lineage tracking. *Nat Rev Genet* 2017; **18**: 230–44

9. Campbell IM, Yuan B, Robberecht C, Pfundt R, Szafranski P, McEntagart ME, Nagamani SCS, Erez A, Bartnik M, Wiśniowiecka-Kowalnik B, Plunkett KS, Pursley AN, Kang S-HL, Bi W, Lalani SR, Bacino CA, Vast M, Marks K, Patton M, Olofsson P, Patel A, Veltman JA, Cheung SW, Shaw CA, Vissers LELM, Vermeesch JR, Lupski JR, Stankiewicz P. Parental Somatic Mosaicism Is Underrecognized and Influences Recurrence Risk of Genomic Disorders. *Am J Hum Genet* 2014; **95**: 1-10

10. Breuss M, Kleiber M, George RD, Antaki D, James KN, Ball LL, Hong O, Garcia CAB, Musaev D, Nguyen A, McEvoy-Venneri J, Knox R, Sticca E, Devinsky O, Gymrek M, Sebat J, Gleeson JG. Quantification of autism recurrence risk by direct assessment of paternal sperm mosaicism. bioRxiv **2017**; 208165

11. McConnell MJ, Moran J V., Abyzov A, Akbarian S, Bae T, Cortes-Ciriano I, Erwin JA, Fasching L, Flasch DA, Freed D, Ganz J, Jaffe AE, Kwan KY, Kwon M, Lodato MA, Mills RE, Paquola ACM, Rodin RE, Rosenbluh C, Sestan N, Sherman MA, Shin JH, Song S, Straub RE, Thorpe J, Weinberger DR, Urban AE, Zhou B, Gage FH, Lehner T, Senthil G, Walsh CA, Chess A, Courchesne E, Gleeson JG, Kidd JM, Park PJ, Pevsner J, Vaccarino FM, Brain Somatic Mosaicism Network. Intersection of diverse neuronal genomes and neuropsychiatric disease: The Brain Somatic Mosaicism Network. Science 2017; **356**:

eaal1641

12. Pamphlett R. Somatic mutation: a cause of sporadic neurodegenerative diseases? *Med Hypotheses* 2004; **62**: 679–82

13. Van Broeckhoven C. The future of genetic research on neurodegeneration. *Nat Med* 2010; **16**: 1215–7

14. Proukakis C, Houlden H, Schapira AH. Somatic alpha-synuclein mutations in Parkinson's disease: hypothesis and preliminary data. *Mov Disord* 2013; **28**: 705–12

15. Pamphlett R. The "somatic-spread" hypothesis for sporadic neurodegenerative diseases. *Med Hypotheses* 2011; **77**: 544–7

16. McKinnon PJ. DNA repair deficiency and neurological disease. *Nat Rev Neurosci* 2009; **10**: 100–12

17. Rulten SL, Caldecott KW. DNA strand break repair and neurodegeneration. DNA Repair (Amst) 2013; **12**: 558–67

18. Vijg J, Suh Y. Genome Instability and Aging. *Annu Rev Physiol* 2013; **75**: 645–68

19. Dölle C, Flønes I, Nido GS, Miletic H, Osuagwu N, Kristoffersen S, Lilleng PK, Larsen JP, Tysnes O-B, Haugarvoll K, Bindoff LA, Tzoulis C. Defective mitochondrial DNA homeostasis in the substantia nigra in Parkinson disease. *Nat Commun*; 2016; **7**: 13548

20. Kirkwood TBL, Rose MR. Evolution of Senescence: Late Survival Sacrificed for Reproduction. *Philos Trans R Soc B Biol Sci* 1991; **332**: 15–24

21. Milholland B, Dong X, Zhang L, Hao X, Suh Y, Vijg J. Differences between germline and somatic mutation rates in humans and mice. *Nat Commun* 2017; **8**: 15183

22. Dong X, Zhang L, Milholland B, Lee M, Maslov AY, Wang T, Vijg J. Accurate identification of single-nucleotide variants in whole-genome-amplified single cells. *Nat Methods* 2017; **14**: 491–3

23. Chun J. Selected comparison of immune and nervous system development. *Adv Immunol* 2001; **77**: 297–322

24. Upton KR, Gerhardt DJ, Jesuadian JS, Richardson SR, Sánchez-Luque FJ, Bodea GO, Ewing AD, Salvador-Palomeque C, van der Knaap MS, Brennan PM, Vanderver A, Faulkner GJ. Ubiquitous L1 Mosaicism in Hippocampal Neurons. *Cell* 2015; **161**: 228–39

25. Singer T, McConnell MJ, Marchetto MC, Coufal NG, Gage FH. LINE-1 retrotransposons: mediators of somatic variation in neuronal genomes? *Trends Neurosci* 2010; **33**: 345–54

26. Dreyer WJ, Gray WR, 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**: 353–67

27. Chun J, Schatz DG. Rearranging Views on Neurogenesis: Neuronal Death in the Absence of DNA End-Joining Proteins. *Neuron* 1999; **22**: 7–10

28. Gao Y, Sun Y, Frank KM, Dikkes P, Fujiwara Y, Seidl KJ, Sekiguchi JM, Rathbun GA, Swat W, Wang J, Bronson RT, Malynn BA, Bryans M, Zhu C, Chaudhuri J, Davidson L, Ferrini R, Stamato T, Orkin SH, Greenberg ME, Alt FW. A critical role for DNA end-joining proteins in both lymphogenesis and neurogenesis. *Cell* 1998; 95: 891–902

29. Schwer B, Wei P-C, Chang AN, Kao J, Du Z, Meyers RM, Alt FW. Transcription-associated

processes cause DNA double-strand breaks and translocations in neural stem/progenitor cells. *Proc Natl Acad Sci USA* 2016; **113**: 2258-63

30. Konstantinidou C, Taraviras S, Pachnis V. Geminin prevents DNA damage in vagal neural crest cells to ensure normal enteric neurogenesis. *BMC Biol* ; 2016; **14**: 94

31. Weissman IL, Gage FH. A Mechanism for Somatic Brain Mosaicism. *Cell* 2016; **164**: 593–5

32. Oppenheim RW. Cell death during development of the nervous system. *Annu Rev Neurosci* 1991; 14: 453–501

33. Chow H-M, Herrup K. Genomic integrity and the ageing brain. *Nat Rev Neurosci* 2015; **16**: 672-84.

34. Madabhushi R, Pan L, Tsai LH. DNA damage and its links to neurodegeneration. *Neuron* 2014; **83**: 266–82.

35. Suberbielle E, Sanchez PE, Kravitz A V, Wang X, Ho K, Eilertson K, Devidze N, Kreitzer AC, Mucke L. Physiologic brain activity causes DNA double-strand breaks in neurons, with exacerbation by amyloid-β. *Nat Neurosci* 2013; **16**: 613–21

36. van Heesch S, Mokry M, Boskova V, Junker W, Mehon R, Toonen P, de Bruijn E, Shull JD, Aitman TJ, Cuppen E, Guryev V. Systematic biases in DNA copy number originate from isolation procedures. *Genome Biol* 2013; **14**: R33

37. Nacheva E, Mokretar K, Soenmez A, Pittman AM, Grace C, Valli R, Ejaz A, Vattathil S, Maserati E, Houlden H, Taanman J-W, Schapira AH, Proukakis C. DNA isolation protocol effects on nuclear DNA analysis by microarrays, droplet digital PCR, and whole genome sequencing, and on mitochondrial DNA copy number estimation. *PLoS One* 2017; **12**: e0180467

38. Azhar G, Liu L, Zhang X, Wei JY. Influence of age on hypoxia/reoxygenation-induced DNA fragmentation and bcl-2, bcl-xl, bax and fas in the rat heart and brain. *Mech Ageing Dev* 1999; **112**: 5–25

39. Do H, Dobrovic A. Sequence Artifacts in DNA from Formalin-Fixed Tissues: Causes and Strategies for Minimization. Clin Chem 2015; **61**: 64–71

40. Gawad C, Koh W, Quake SR: Single-cell genome sequencing: current state of the science. *Nat Rev Genet* 2016, **17**: 175-88

41. Rehen SK, McConnell MJ, Kaushal D, Kingsbury MA, Yang AH, Chun J. Chromosomal variation in neurons of the developing and adult mammalian nervous system. *Proc Natl Acad Sci U S A* 2001, **98**: 13361–6

42. Yang AH, Kaushal D, Rehen SK, Kriedt K, Kingsbury MA, McConnell MJ, Chun J. Chromosome segregation defects contribute to aneuploidy in normal neural progenitor cells. *J Neurosci* 2003; **23**: 10454–62

43. Rehen SK, Yung YC, McCreight MP, Kaushal D, Yang AH, Almeida BS V, Kingsbury MA, Cabral KMS, McConnell MJ, Anliker B, Fontanoz M, Chun J. Constitutional aneuploidy in the normal human brain. *J Neurosci* 2005, **25**: 2176–80

44. Yurov YB, Iourov IY, Monakhov V V., Soloviev I V., Vostrikov VM, Vorsanova SG. The Variation of Aneuploidy Frequency in the Developing and Adult Human Brain Revealed by an Interphase FISH Study. *J Histochem Cytochem* 2005, **53**: 385–390

45. Yurov YB, Iourov IY, Vorsanova SG, Liehr T, Kolotii AD, Kutsev SI, Pellestor F, Beresheva AK,

Demidova IA, Kravets VS, Monakhov V V, Soloviev IV. Aneuploidy and confined chromosomal mosaicism in the developing human brain. *PLoS One* 2007, **2**:e558

46. Thomas P, Fenech M. Chromosome 17 and 21 aneuploidy in buccal cells is increased with ageing and in Alzheimer's disease. *Mutagenesis* 2007, **23**:57–65

47. McConnell MJ, Lindberg MR, Brennand KJ, Piper JC, Voet T, Cowing-Zitron C, Shumilina S, Lasken RS, Vermeesch JR, Hall IM, Gage FH. Mosaic copy number variation in human neurons. *Science* 2013, **342**: 632–7

48. Cai X, Evrony GD, Lehmann HS, Elhosary PC, Mehta BK, Poduri A, Walsh CA. Single-Cell, Genome-wide Sequencing Identifies Clonal Somatic Copy-Number Variation in the Human Brain. *Cell Rep* 2014; **8**: 1280–9

49. van den Bos H, Spierings DCJ, Taudt AS, Bakker B, Porubský D, Falconer E, Novoa C, Halsema N, Kazemier HG, Hoekstra-Wakker K, Guryev V, den Dunnen WFA, Foijer F, Tatché MC, Boddeke HWGM, Lansdorp PM. Single-cell whole genome sequencing reveals no evidence for common aneuploidy in normal and Alzheimer's disease neurons. *Genome Biol* 2016; **17**: 116

50. Knouse KA, Wu J, Whittaker CA, Amon A. Single cell sequencing reveals low levels of aneuploidy across mammalian tissues. *Proc Natl Acad Sci U S A* 2014, **111**:13409–14

51. Peterson SE, Yang AH, Bushman DM, Westra JW, Yung YC, Barral S, Mutoh T, Rehen SK, Chun J. Aneuploid cells are differentially susceptible to caspase-mediated death during embryonic cerebral cortical development. *J Neurosci* 2012, **32**: 16213–22

52. 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**: 1126–32

53. Knouse KA, Wu J, Amon A. Assessment of megabase-scale somatic copy number variation using single cell sequencing. *Genome Res* 2016; **26**: 376-84

54. Garvin T, Aboukhalil R, Kendall J, Baslan T, Atwal GS, Hicks J, Wigler M, Schatz MC. Interactive analysis and assessment of single-cell copy-number variations. *Nat Methods* 2015, **12**: 1058–60

55. Cai J, Miao X, Li Y, Smith C, Tsang K, Cheng L, Wang Q-F. Whole-genome sequencing identifies genetic variances in culture-expanded human mesenchymal stem cells. *Stem cell reports* 2014; **3**: 227–33

56. Collins RL, Brand H, Redin CE, Hanscom C, Antolik C, Stone MR, Glessner JT, Mason T, Pregno G, Dorrani N, Mandrile G, Giachino D, Perrin D, Walsh C, Cipicchio M, Costello M, Stortchevoi A, An J-Y, Currall BB, Seabra CM, Ragavendran A, Margolin L, Martinez-Agosto JA, Lucente D, Levy B, Sanders SJ, Wapner RJ, Quintero-Rivera F, Kloosterman W, Talkowski ME. Defining the diverse spectrum of inversions, complex structural variation, and chromothripsis in the morbid human genome. *Genome Biol* 2017; **18**: 36

57. Zhang C-Z, Leibowitz ML, Pellman D. Chromothripsis and beyond: rapid genome evolution from complex chromosomal rearrangements. *Genes Dev* 2013; **27**: 2513–30

58. Hazen JL, Faust GG, Rodriguez AR, Ferguson WC, Shumilina S, Clark RA, Boland MJ, Martin G, Chubukov P, Tsunemoto RK, Torkamani A, Kupriyanov S, Hall IM, Baldwin KK. The Complete Genome Sequences, Unique Mutational Spectra, and Developmental Potency of Adult Neurons Revealed by Cloning. *Neuron* 2016; **89**: 1223–36

59. Kazazian HH. L1 Retrotransposons Shape the Mammalian Genome. *Science* 2000; **289**: 1152-3

60. Muotri AR, Chu VT, Marchetto MC, Deng W, Moran J V, Gage FH. Somatic mosaicism in neuronal precursor cells mediated by L1 retrotransposition. *Nature* 2005; **435**: 903–10

61. Macia A, Widmann TJ, Heras SR, Ayllon V, Sanchez L, Benkaddour-Boumzaouad M, Muñoz-Lopez M, Rubio A, Amador-Cubero S, Blanco-Jimenez E, Garcia-Castro J, Menendez P, Ng P, Muotri AR, Goodier JL, Garcia-Perez JL. Engineered LINE-1 retrotransposition in non-dividing human neurons. *Genome Res* 2016; **27**: 335-48

62. Baillie JK, Barnett MW, Upton KR, Gerhardt DJ, Richmond TA, De Sapio F, Brennan PM, Rizzu P, Smith S, Fell M, Talbot RT, Gustincich S, Freeman TC, Mattick JS, Hume DA, Heutink P, Carninci P, Jeddeloh JA, Faulkner GJ. Somatic retrotransposition alters the genetic landscape of the human brain. *Nature* 2011; **479**: 534–7

63. Evrony GD, Cai X, Lee E, Hills LB, Elhosary PC, Lehmann HS, Parker JJ, Atabay KD, Gilmore EC, Poduri A, Park PJ, Walsh CA. Single-neuron sequencing analysis of L1 retrotransposition and somatic mutation in the human brain. *Cell* 2012; **151**: 483–96

64. Evrony GD, Lee E, Mehta BK, Benjamini Y, Johnson RM, Cai X, Yang L, Haseley P, Lehmann HS, Park PJ, Walsh CA. Cell Lineage Analysis in Human Brain Using Endogenous Retroelements. *Neuron* 2015, **85**: 49–59

65. Erwin JA, Paquola ACM, Singer T, Gallina I, Novotny M, Quayle C, Bedrosian TA, Alves FIA, Butcher CR, Herdy JR, Sarkar A, Lasken RS, Muotri AR, Gage FH. L1-associated genomic regions are deleted in somatic cells of the healthy human brain. *Nat Neurosci* 2016; **19**: 1583-91

66. Evrony GD, Lee E, Park PJ, Walsh CA. Resolving rates of mutation in the brain using single-neuron genomics. *Elife* 2016, **5**: e12966

67. Westra JW, Rivera RR, Bushman DM, Yung YC, Peterson SE, 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

68. Fischer H-G, Morawski M, Brückner MK, Mittag A, Tarnok A, Arendt T. Changes in neuronal DNA content variation in the human brain during aging. *Aging Cell* 2012; **11**: 628–33

69. Bushman DM, Kaeser GE, Siddoway B, Westra JW, Rivera RR, Rehen SK, Yung YC, 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**: e05116

70. Lodato MA, Woodworth MB, Lee S, Evrony GD, Mehta BK, Karger A, Chittenden TW, D'Gama AM, Cai X, Luquette LJ, Lee E, Park PJ, Walsh CA. Somatic mutation in single human neurons tracks developmental and transcriptional history. *Science* 2015; **350**: 94–98

71. Lodato MA, Rodin RE, Bohrson CL, Coulter ME, Barton AR, Kwon M, Sherman MA, Vitzhum CM, Luquette LJ, Yandava C, Yang P, Chittenden TW, Hatem NE, Ryu SC, Woodworth MB, Park PJ, Walsh CA. Aging and neurodegeneration are associated with increased mutations in single human neurons. Science 2017; 10.1126/science.aao4426

72. Hoang ML, Kinde I, Tomasetti C, McMahon KW, Rosenquist TA, Grollman AP, Kinzler KW, 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**: 9846–51

73. Chen C, Xing D, Tan L, Li H, Zhou G, Huang L, Xie XS. Single-cell whole-genome analyses by Linear

Amplification via Transposon Insertion (LIANTI). *Science* 2017; **356**: 189–94

74. Forsberg LA, Absher D, Dumanski JP. Non-heritable genetics of human disease: spotlight on post-zygotic genetic variation acquired during lifetime. *J Med Genet* 2013, **50**: 1–10

75. Fernandes-Rosa FL, Boulkroun S, Zennaro M-C. Somatic and inherited mutations in primary aldosteronism. *J Mol Endocrinol* 2017; **59**: R47–R63

76. Kasten M, Klein C. The many faces of alpha-synuclein mutations. *Mov Disord* 2013; **28**: 697–701

77. Verstraeten A, Theuns J, Van Broeckhoven C. Progress in unraveling the genetic etiology of Parkinson disease in a genomic era. *Trends Genet* 2015; **31**: 140–9

78. Ryan NS, Nicholas JM, Weston PSJ, Liang Y, Lashley T, Guerreiro R, Adamson G, Kenny J, Beck J, Chavez-Gutierrez L, de Strooper B, Revesz T, Holton J, Mead S, Rossor MN, Fox NC. Clinical phenotype and genetic associations in autosomal dominant familial Alzheimer's disease: a case series. *Lancet Neurol* 2016; **15**: 1326–35

79. Sabatelli M, Conte A, Zollino M. Clinical and genetic heterogeneity of amyotrophic lateral sclerosis. *Clin Genet* 2013; **83**: 408–16

80. Brettschneider J, Tredici K Del, Lee VM-Y, Trojanowski JQ. Spreading of pathology in neurodegenerative diseases: a focus on human studies. *Nat Rev Neurosci* 2015; **16**: 109–20

81. Frank SA. Somatic Mosaicism and Disease. *Curr Biol* 2014; **24**: R577–81

82. Engelender S, Isacson O. The Threshold Theory for Parkinson's Disease. *Trends Neurosci* 2017; **40**: 4–14

83. Walsh DM, Selkoe DJ. A critical appraisal of the pathogenic protein spread hypothesis of neurodegeneration. *Nat Rev Neurosci* 2016; **17**: 251–60

84. Surmeier DJ, Obeso JA, Halliday GM. Selective neuronal vulnerability in Parkinson disease. *Nat Rev Neurosci* 2017; **18**: 101–13

85. Dickson DW. Parkinson's disease and parkinsonism: neuropathology. *Cold Spring Harb Perspect Med* 2012; **2**: a009258

86. Frank SA. Puzzles in modern biology. IV. Neurodegeneration, localized origin and widespread decay. *F1000Research* 2016; **5**: 2537

87. Xiromerisiou G, Houlden H, Sailer A, Silveira-Moriyama L, Hardy J, Lees AJ. Identical twins with Leucine rich repeat kinase type 2 mutations discordant for Parkinson's disease. *Mov Disord* 2012; **27**: 1323

88. Hussein SM, Batada NN, Vuoristo S, Ching RW, Autio R, Närvä E, Ng S, Sourour M, Hämäläinen R, Olsson C, Lundin K, Mikkola M, Trokovic R, Peitz M, Brüstle O, Bazett-Jones DP, Alitalo K, Lahesmaa R, Nagy A, Otonkoski T. Copy number variation and selection during reprogramming to pluripotency. *Nature* 2011; **471**: 58–62

89. Arlt MF, Wilson TE, Glover TW. Replication stress and mechanisms of CNV formation. *Curr Opin Genet Dev* 2012; **22**: 204–10

90. Rozier L, El-Achkar E, Apiou F, Debatisse M. Characterization of a conserved aphidicolin-sensitive common fragile site at human 4q22 and mouse 6C1: possible association with an inherited disease and cancer. *Oncogene* 2004; **23**: 6872–80

91. Mitsui J, Takahashi Y, Goto J, Tomiyama H, Ishikawa S, Yoshino H, Minami N, Smith DI, Lesage S, Aburatani H, Nishino I, Brice A, Hattori N, Tsuji S. Mechanisms of genomic instabilities underlying two common fragile-site-associated loci, PARK2 and DMD, in germ cell and cancer cell lines. *Am J Hum Genet* 2010; **87**: 75–89

92. Hu H-H, Kannengiesser C, Lesage S, André J, Mourah S, Michel L, Descamps V, Basset-Seguin N, Bagot M, Bensussan A, Lebbé C, Deschamps L, Saiag P, Leccia M-T, Bressac-de-Paillerets B, Tsalamlal A, Kumar R, Klebe S, Grandchamp B, Andrieu-Abadie N, Thomas L, Brice A, Dumaz N, Soufir N. PARKIN Inactivation Links Parkinson's Disease to Melanoma. *J Natl Cancer Inst* 2015; **108**: djv340

93. Gegg ME, Burke D, Heales SJR, Cooper JM, Hardy J, Wood NW, Schapira AH. Glucocerebrosidase deficiency in substantia nigra of parkinson disease brains. *Ann Neurol* 2012; **72**: 455–63

94. Robbins J, Otsuka F, Tarone R, Polinsky R, Brumback R, Moshell A, Nee L, Ganges M, Cayeux S. Radiosensitivity in Alzheimer Disease and Parkinson Disease. *Lancet* 1983; **321**: 468–69

95. Robbins JH, Otsuka F, Tarone RE, Polinsky RJ, Brumback RA, Nee LE. Parkinson's disease and Alzheimer's disease: hypersensitivity to X rays in cultured cell lines. *J Neurol Neurosurg Psychiatry* 1985; **48**: 916–23

96. Zhang C-Z, Spektor A, Cornils H, Francis JM, Jackson EK, Liu S, Meyerson M, Pellman D. Chromothripsis from DNA damage in micronuclei. *Nature* 2015; **52**: 179-84

97. Migliore L, Petrozzi L, Lucetti C, Gambaccini G, Bernardini S, Scarpato R, Trippi F, Barale R, Frenzilli G, Rodilla V, Bonuccelli U. Oxidative damage and cytogenetic analysis in leukocytes of Parkinson's disease patients. *Neurology* 2002; **58**: 1809–15

98. Inzelberg R, Samuels Y, Azizi E, Qutob N, Inzelberg L, Domany E, Schechtman E, Friedman E. Parkinson disease (PARK) genes are somatically mutated in cutaneous melanoma. *Neurol Genet* 2016; **2**: e70

99. Murphy DD, Rueter SM, Trojanowski JQ, Lee VM. Synucleins are developmentally expressed, and alpha-synuclein regulates the size of the presynaptic vesicular pool in primary hippocampal neurons. *J Neurosci* 2000; **20**: 3214–20

100. Sosa LJ, Bergman J, Estrada-Bernal A, Glorioso TJ, Kittelson JM, Pfenninger KH. Amyloid Precursor Protein Is an Autonomous Growth Cone Adhesion Molecule Engaged in Contact Guidance. *PLoS One* 2013; **8**: e64521

101. Williams GC. Pleiotropy, Natural Selection, and the Evolution of Senescence. Evolution 1957; **11**: 398-411.

102. Proukakis C, Shoaee M, Morris J, Brier T, Kara E, Sheerin U-M, Charlesworth G, Tolosa E, Houlden H, Wood NW, Schapira AH. Analysis of Parkinson's disease brain-derived DNA for alpha-synuclein coding somatic mutations. *Mov Disord* 2014; **29**: 1060–4

103. Pamphlett R, Morahan JM, Luquin N, Yu B. An approach to finding brain-situated mutations in sporadic Parkinson's disease. *Park & Relat Disord* 2012; **18**: 82-5

104. Bruder CE, Piotrowski A, Gijsbers AA, Andersson R, Erickson S, Diaz de Stahl T, Menzel U, Sandgren J, von Tell D, Poplawski A, Crowley M, Crasto C, Partridge EC, Tiwari H, Allison DB, Komorowski J, van Ommen GJ, Boomsma DI, Pedersen NL, den Dunnen JT, Wirdefeldt K, Dumanski JP. Phenotypically concordant and discordant monozygotic twins display different DNA copy-number-variation profiles. *Am J Hum Genet* 2008; **82**: 763–71

105. Pankratz N, Dumitriu A, Hetrick KN, Sun M, Latourelle JC, Wilk JB, et al. Copy Number Variation in Familial Parkinson Disease. *PLoS One* 2011; **6**: e20988

106. Perandones C, Giugni JC, Calvo DS, Raina GB, De Jorge Lopez L, Volpini V, et al. Mosaicism of alpha-synuclein gene rearrangements: Report of two unrelated cases of early-onset parkinsonism. *Parkinsonism Relat Disord* 2014; **20**: 558–61

107. Höglinger GU, Breunig JJ, Depboylu C, Rouaux C, Michel PP, Alvarez-Fischer D, Halter C, Doheny KF, Gusella JF, Nichols WC, Myers RH, Foroud T, DeStefano AL, PSG-PROGENI and GenePD Investigators C and MGL. The pRb/E2F cell-cycle pathway mediates cell death in Parkinson's disease. *Proc Natl Acad Sci USA* 2007; **104**: 3585–90

108. Yang Y, Shepherd C, Halliday G. Aneuploidy in Lewy body diseases. *Neurobiol Aging* 2015; 36: 1253-60

109. Yang Y, Shepherd CE, Halliday GM. Increased aneuploidy is not a universal feature across α-synucleinopathies. *Mov Disord* 2017; **32**: 475–6

110. Beck JA, Poulter M, Campbell TA, Uphill JB, Adamson G, Geddes JF, Revesz T, Davis MB, Wood NW, Collinge J, Tabrizi SJ. Somatic and germline mosaicism in sporadic early-onset Alzheimer's disease. *Hum Mol Genet* 2004; **13**: 1219–24

111. Frigerio CS, 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 disease patients. *Alzheimers Dement* 2015; **11**: 1265-75

112. Arendt T, Bruckner MK, Mosch B, Losche A. Selective cell death of hyperploid neurons in Alzheimer's disease. *Am J Pathol* 2010; **177**: 15–20

113. Arendt T, Bruckner MK, Losche A. Regional mosaic genomic heterogeneity in the elderly and in Alzheimer's disease as a correlate of neuronal vulnerability. *Acta Neuropathol* 2015; **130**: 501–10

114. Iourov IY, Vorsanova SG, Liehr T, Yurov YB. Aneuploidy in the normal, Alzheimer's disease and ataxia-telangiectasia brain: differential expression and pathological meaning. *Neurobiol Dis* 2009; **34**: 212–20

115. Herrup K, Arendt T. Re-expression of cell cycle proteins induces neuronal cell death during Alzheimer's disease. *J Alzheimers Dis* 2002; **4**: 243–7.

116. Simpson JE, Ince PG, Matthews FE, Shaw PJ, Heath PR, Brayne C, Garwood C, Higginbottom A, Wharton SB, MRC Cognitive Function and Ageing Neuropathology Study Group. A neuronal DNA damage response is detected at the earliest stages of Alzheimer's neuropathology and correlates with cognitive impairment in the Medical Research Council's Cognitive Function and Ageing Study ageing brain cohort. *Neuropathol Appl Neurobiol* 2015; **41**: 483–96

117. Shackelford DA. DNA end joining activity is reduced in Alzheimer's disease. *Neurobiol Aging* 2006, **27**: 596–605

118. Kruman II, Wersto RP, Cardozo-Pelaez F, Smilenov L, Chan SL, Chrest FJ, Emokpae R, Gorospe M, Mattson MP. Cell cycle activation linked to neuronal cell death initiated by DNA damage. *Neuron* 2004,; **41**: 549–61

119. Forsberg LA, Rasi C, Malmqvist N, Davies H, Pasupulati S, Pakalapati G, Sandgren J, Diaz de Ståhl T, Zaghlool A, Giedraitis V, Lannfelt L, Score J, Cross NCP, Absher D, Janson ET, Lindgren CM, Morris AP, Ingelsson E, Lind L, Dumanski JP. Mosaic loss of chromosome Y in peripheral blood is associated

with shorter survival and higher risk of cancer. *Nat Genet* 2014; **46**: 624–8

120. Dumanski JP, Lambert J-C, Rasi C, Giedraitis V, Davies H, Grenier-Boley B, Lindgren CM, Campion D, Dufouil C, Pasquier F, Amouyel P, Lannfelt L, Ingelsson M, Kilander L, Lind L, Forsberg LA, Forsberg LA. Mosaic Loss of Chromosome Y in Blood Is Associated with Alzheimer Disease. *Am J Hum Genet* 2016; **98**: 1208–19

121. Shaw PJ, Tomkins J, Slade JY, Usher P, Curtis A, Bushby K, Ince PG. CNS tissue Cu/Zn superoxide dismutase (SOD1) mutations in motor neurone disease (MND). *Neuroreport* 1997; **8**: 3923–7

122. Meltz Steinberg K, Nicholas TJ, Koboldt DC, Yu B, Mardis E, Pamphlett R. Whole genome analyses reveal no pathogenetic single nucleotide or structural differences between monozygotic twins discordant for amyotrophic lateral sclerosis. *Amyotroph Lateral Scler Front Degener* 2015; **16**: 385–39

123. Young PE, Kum Jew S, Buckland ME, Pamphlett R, Suter CM. Epigenetic differences between monozygotic twins discordant for amyotrophic lateral sclerosis (ALS) provide clues to disease pathogenesis. Cereda C, editor. *PLoS One* 2017; **12**: e0182638

124. Alzualde A, Moreno F, Martínez-Lage P, Ferrer I, Gorostidi A, Otaegui D, et al. Somatic mosaicism in a case of apparently sporadic Creutzfeldt-Jakob disease carrying a de novo D178N mutation in the PRNP gene. *Am J Med Genet Part B Neuropsychiatr Genet* 2010; **153B**: 1283–91

125. Faggioli F, Wang T, Vijg J, Montagna C. Chromosome-specific accumulation of aneuploidy in the aging mouse brain. *Hum Mol Genet* 2012; **21**: 5246–53

126. Mass E, Christian E, Jacome-Galarza TB. A somatic mutation in erythro-myeloid progenitors causes neurodegenerative disease. *Nature* 2017; **549**: 389-93

127. Salinas S, Proukakis C, Crosby A, Warner TT. Hereditary spastic paraplegia: clinical features and pathogenetic mechanisms. *Lancet Neurol* 2008; **7**: 1127–38

128. Aulitzky A, Friedrich K, Gläser D, Gastl R, Kubisch C, Ludolph AC, Volk AE. A complex form of hereditary spastic paraplegia in three siblings due to somatic mosaicism for a novel SPAST mutation in the mother. *J Neurol Sci* 2014; **347**: 352–5

129. Haack TB, Hogarth P, Kruer MC, Gregory A, Wieland T, Schwarzmayr T, Graf E, Sanford L, Meyer E, Kara E, Cuno SM, Harik SI, Dandu VH, Nardocci N, Zorzi G, Dunaway T, Tarnopolsky M, Skinner S, Frucht S, Hanspal E, Schrander-Stumpel C, Héron D, Mignot C, Garavaglia B, Bhatia K, Hardy J, Strom TM, Boddaert N, Houlden HH, Kurian M, Meitinger T, Prokisch H, Hayflick S. Exome sequencing reveals de novo WDR45 mutations causing a phenotypically distinct, X-linked dominant form of NBIA. *Am J Hum Genet* 2012; **91**: 1144–9

130. Rossi G, Conconi D, Panzeri E, Redaelli S, Piccoli E, Paoletta L, Dalprà L, Tagliavini F. Mutations in MAPT gene cause chromosome instability and introduce copy number variations widely in the genome. *J Alzheimers Dis* 2013; **33**: 969–82

131. Rossi G, Conconi D, Panzeri E, Paoletta L, Piccoli E, Ferretti MG, Mangieri M, Ruggerone M, Dalprà L, Tagliavini F. Mutations in MAPT give rise to aneuploidy in animal models of tauopathy. *Neurogenetics* 2014; **15**: 31–40

132. Malmanche N, Dourlen P, Gistelinck M, Demiautte F, Link N, Dupont C, Vanden Broeck L, Werkmeister E, Amouyel P, Bongiovanni A, Bauderlique H, Moechars D, Royou A, Bellen HJ, Lafont F, Callaerts P, Lambert J-C, Dermaut B. Developmental Expression of 4-Repeat-Tau Induces Neuronal Aneuploidy in Drosophila Tauopathy Models. *Sci Rep* 2017; **7**: 40764

133. Bettencourt C, Hensman-Moss D, Flower M, Wiethoff S, Brice A, Goizet C, Stevanin G, Koutsis G, Karadima G, Panas M, Yescas-Gómez P, García-Velázquez LE, Alonso-Vilatela ME, Lima M, Raposo M, Traynor B, Sweeney M, Wood N, Giunti P, Durr A, Holmans P, Houlden H, Tabrizi SJ, Jones L, Jones L. DNA repair pathways underlie a common genetic mechanism modulating onset in polyglutamine diseases. *Ann Neurol* 2016; **79**: 983–90

134. Veitch NJ, Ennis M, McAbney JP, US-Venezuela Collaborative Research Project PF, Shelbourne PF, Monckton DG. Inherited CAG.CTG allele length is a major modifier of somatic mutation length variability in Huntington disease. *DNA Repair (Amst)* 2007; **6**: 789–96

135. Swami M, Hendricks AE, Gillis T, Massood T, Mysore J, Myers RH, et al. 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**: 3039–47

136. Shelbourne PF, Keller-McGandy C, Bi WL, Yoon S-R, Dubeau L, Veitch NJ, Wheeler VC. Triplet repeat mutation length gains correlate with cell-type specific vulnerability in Huntington disease brain. *Hum Mol Genet* 2007; **16**: 1133–42

137. Jones L, Houlden H, Tabrizi SJ. DNA repair in the trinucleotide repeat disorders. *Lancet Neurol* 2017; **16**: 88–96

138. Budworth H, Harris FR, Williams P, Lee DY, Holt A, Pahnke J, Szczesny B, Acevedo-Torres K, Ayala-Peña S, McMurray CT. Suppression of Somatic Expansion Delays the Onset of Pathophysiology in a Mouse Model of Huntington's Disease. *PLoS Genet* 2015; **11**: e1005267

139. Moss DJH, Pardiñas AF, Langbehn D, Lo K, Leavitt BR, Roos R, Durr A, Mead S, TRACK-HD investigators, REGISTRY investigators, Holmans P, Jones L, Tabrizi SJ. Identification of genetic variants associated with Huntington's disease progression: a genome-wide association study. *Lancet Neurol* 2017; **16**: 701–11

140. Morales F, Vásquez M, Santamaría C, Cuenca P, Corrales E, Monckton DG. A polymorphism in the MSH3 mismatch repair gene is associated with the levels of somatic instability of the expanded CTG repeat in the blood DNA of myotonic dystrophy type 1 patients. *DNA Repair (Amst)* 2016; **40**: 57–66

141. Oosterloo M, Van Belzen MJ, Bijlsma EK, Roos RAC. Is There Convincing Evidence that Intermediate Repeats in the HTT Gene Cause Huntington's Disease? *J Huntingtons Dis* 2015; **4**: 141–8

142. Doherty KM, De Pablo-Fernandez E, Houlden H, Polke JM, Lees AJ, Warner TT, Holton JL. MSA-C or SCA 17? A clinicopathological case update. *Mov Disord* 2016; **31**: 1582–4

143. van Blitterswijk M, DeJesus-Hernandez M, Niemantsverdriet E, Murray ME, Heckman MG, Diehl NN, Brown PH, Baker MC, Finch NA, Bauer PO, Serrano G, Beach TG, Josephs KA, Knopman DS, Petersen RC, Boeve BF, Graff-Radford NR, Boylan KB, Petrucelli L, Dickson DW, Rademakers R. Association between repeat sizes and clinical and pathological characteristics in carriers of C9ORF72 repeat expansions (Xpansize-72): a cross-sectional cohort study. *Lancet Neurol* 2013; **12**: 978–88

144. Nordin A, Akimoto C, Wuolikainen A, Alstermark H, Jonsson P, Birve A, Marklund SL, Graffmo KS, Forsberg K, Brännström T, Andersen PM. 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**: 3133–42

145. Cooper-Knock J, Shaw PJ, Kirby J. The widening spectrum of C9ORF72-related disease; genotype/phenotype correlations and potential modifiers of clinical phenotype. *Acta Neuropathol* 2014; **127**: 333–45

146. Pamphlett R, Cheong PL, Trent RJ, Yu B. Can ALS-associated C9orf72 repeat expansions be diagnosed on a blood DNA test alone? *PLoS One* 2013; **8** :e70007

147. Iourov IY, Vorsanova SG, Liehr T, Kolotii AD, Yurov YB. Increased chromosome instability dramatically disrupts neural genome integrity and mediates cerebellar degeneration in the ataxia-telangiectasia brain. *Hum Mol Genet* 2009; **18**: 2656–69

148. Itoh M, Hayashi M, Shioda K, Minagawa M, Isa F, Tamagawa K, Morimatsu Y, Oda M. Neurodegeneration in hereditary nucleotide repair disorders. *Brain Dev* 1999; **21**: 326–33

149. Weidenheim KM, Dickson DW, Rapin I. Neuropathology of Cockayne syndrome: Evidence for impaired development, premature aging, and neurodegeneration. *Mech Ageing Dev* 2009; **130**: 619–36

150. Groh M, Albulescu LO, Cristini A, Gromak N. Senataxin: Genome Guardian at the Interface of Transcription and Neurodegeneration. *J Mol Biol* 2019; **429**: 3181-95

151. McKinnon PJ. ATM and the molecular pathogenesis of ataxia telangiectasia. *Annu Rev Pathol* 2012; **7**: 303–21

152. Rothblum-Oviatt C, Wright J, Lefton-Greif MA, McGrath-Morrow SA, Crawford TO, Lederman HM. Ataxia telangiectasia: a review. *Orphanet J Rare Dis* 2016; **11**: 159

153. Wang W-Y, Pan L, Su SC, Quinn EJ, Sasaki M, Jimenez JC, A Mackenzie IR, Huang EJ, Tsai L-H. Interaction of FUS and HDAC1 regulates DNA damage response and repair in neurons. *Nature Neurosci* 2013; **16**: 1383-91

154. Bandey I, Chiou S-H, Huang A-P, Tsai J-C, Tu P. Progranulin promotes Temozolomide resistance of glioblastoma by orchestrating DNA repair and tumor stemness. *Oncogene* 2015; 34: 1853–64

155. Violet M, Chauderlier A, Delattre L, Tardivel M, Chouala MS, Sultan A, Marciniak E, Humez S, Binder L, Kayed R, Lefebvre B, Bonnefoy E, Buée L, Galas M-C. Prefibrillar Tau oligomers alter the nucleic acid protective function of Tau in hippocampal neurons in vivo. *Neurobiol Dis* 2015; **82**: 540–51

156. Vasquez V, Mitra J, Hegde PM, Pandey A, Sengupta S, Mitra S, Rao KS, Hegde ML. Chromatin-Bound Oxidized α-Synuclein Causes Strand Breaks in Neuronal Genomes in in vitro Models of Parkinson's Disease. *J Alzheimer's Dis* 2017; **60**: S133-50

157. Kinde I, Wu J, Papadopoulos N, Kinzler KW, Vogelstein B. Detection and quantification of rare mutations with massively parallel sequencing. *Proc Natl Acad Sci USA* 2011; **108**: 9530–5

158. Vitak SA, Torkenczy KA, Rosenkrantz JL, Fields AJ, Christiansen L, Wong MH, Carbone L, Steemers FJ, Adey A. Sequencing thousands of single-cell genomes with combinatorial indexing. *Nat Methods* 2017; **14**: 302-8

159. Pamphlett R, Heath PR, Holden H, Ince PG, Shaw PJ. Detection of mutations in whole genome-amplified DNA from laser-microdissected neurons. *J Neurosci Methods* 2005; **147**: 65–7

160. Dickson DW, Fujishiro H, DelleDonne A, Menke J, Ahmed Z, Klos KJ, Josephs K, Frigerio R, Burnett M, Parisi JE, Ahlskog JE. Evidence that incidental Lewy body disease is pre-symptomatic Parkinson's disease. *Acta Neuropathol* 2008; **115**: 437–44

161. Beach TG, Adler CH, Serrano G, Sue LI, Walker DG, Dugger BN, Shill HA, Driver-Dunckley E, Caviness JN, Intorcia A, Filon J, Scott S, Garcia A, Hoffman B, Belden CM, Davis KJ, Sabbagh MN.

Prevalence of Submandibular Gland Synucleinopathy in Parkinson's Disease, Dementia with Lewy Bodies and other Lewy Body Disorders. *J Parkinsons Dis* 2016; **6**: 153-63

162. Pentsova EI, Shah RH, Tang J, Boire A, You D, Briggs S, Omuro A, Lin X, Fleisher M, Grommes C, Panageas KS, Meng F, Selcuklu SD, Ogilvie S, Distefano N, Shagabayeva L, Rosenblum M, DeAngelis LM, Viale A, Mellinghoff IK, Berger MF. Evaluating Cancer of the Central Nervous System Through Next-Generation Sequencing of Cerebrospinal Fluid. *J Clin Oncol* 2016; **34**: 2404-15

163. Higuchi R, von Beroldingen CH, Sensabaugh GF, Erlich HA. DNA typing from single hairs. *Nature* 1988; **332**: 543–6

164. Bae T, Tomasini L, Mariani J, Zhou B, Roychowdhury T, Franjic D, Pletikos M, Pattni R, Chen B-J, Venturini E, Riley-Gillis B, Sestan N, Urban AE, Abyzov A, Vaccarino FM. Different mutational rates and mechanisms in human cells at pregastrulation and neurogenesis. *Science* 2017:eaan8690

Number of neurons	**Total losses**	**Total gains**	**Mean CNVs per neuron**	**Ref**	**Comments**
110	99	50	1.35	47	
6	14	22	4.5	52	Microfluidic platform. Included four neurons from a Down syndrome case.
19	64	1	3.4	48	
92	0	0	0	24	Hippocampal neurons. CNV calling secondary aim.
80	15	0	0.19	53	Obtained from dissociated cortical gray matter. High quality cells only included.

Disease	Cases	Controls	Brain region(s) analysed	Technique(s)	Gene(s) targeted	Mutation type(s) targeted	Lowest level detectable	Ref
PD / DLB	567	0	Cerebellum (all), also nigra (in 25)	PCR based (high resolution melt curve analysis)	*SNCA* (coding exons only)	SNVs	5-10%	14, 102
AD	1	0	Cerebral cortex	Allele-specific PCR and oligonucleotide hybridisation, Sanger sequencing	*PSEN1*	Specific SNV	N/A. 14% detected.	110
AD	72	58	Entorhinal cortex	Targeted deep sequencing, qPCR	*APP, PSEN1, PSEN2, MAPT*	SNV CNV	~1% ~10%	111
AD	32	40	Prefrontal cortex (all), cerebellum (in 15 each)	Small pool and single cell qPCR, FISH	*APP*	CNV	Single cell	69