SOMATIC GENOMIC MOSAICISM & HUMAN DISEASE

EDITED BY: Ivan Y. Iourov, Henry H. Heng and Svetlana G. Vorsanova

PUBLISHED IN: Frontiers in Genetics

frontiers Research Topics

frontiers

ISSN 1664-8714
ISBN 978-2-83250-562-5
DOI 10.3389/978-2-83250-562-5

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

SOMATIC GENOMIC MOSAICISM & HUMAN DISEASE

Topic Editors:
Ivan Y. Iourov, Mental Health Research Center of Russian Academy of Medical Sciences, Russia
Henry H. Heng, Wayne State University, United States
Svetlana G. Vorsanova, Veltischev Research and Clinical Institute for Pediatrics of the Pirogov Russian National Research Medical University, Russia

Citation: Iourov, I. Y., Heng, H. H., Vorsanova, S. G., eds. (2022). Somatic Genomic Mosaicism & Human Disease. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-83250-562-5

Table of Contents

frontiers | Frontiers in Genetics

TYPE Editorial
PUBLISHED 07 October 2022
DOI 10.3389/fgene.2022.1045559

OPEN ACCESS

EDITED AND REVIEWED BY
Jared C. Roach,
Institute for Systems Biology (ISB),
United States

*CORRESPONDENCE
Ivan Y. Iourov,
ivan.iourov@gmail.com

SPECIALTY SECTION
This article was submitted to Human and Medical Genomics, a section of the journal Frontiers in Genetics

RECEIVED 15 September 2022
ACCEPTED 21 September 2022
PUBLISHED 07 October 2022

CITATION
Iourov IY and Heng HH (2022), Editorial: Somatic genomic mosaicism & human disease.
Front. Genet. 13:1045559.
doi: 10.3389/fgene.2022.1045559

Editorial: Somatic genomic mosaicism & human disease

Ivan Y. Iourov[1,2,3]* and Henry H. Heng[4,5]

[1]Yurov's Laboratory of Molecular Genetics and Cytogenomics of the Brain, Mental Health Research Center, Moscow, Russia, [2]Vorsanova's Laboratory of Molecular Cytogenetics of Neuropsychiatric Diseases, Veltischev Research and Clinical Institute for Pediatrics and Pediatric Surgery of the Pirogov Russian National Research Medical University of the Russian Ministry of Health, Moscow, Russia, [3]Department of Medical Biological Disciplines, Belgorod State University, Belgorod, Russia, [4]Center for Molecular Medicine and Genetics, Wayne State University School of Medicine, Detroit, United States, [5]Department of Pathology, Wayne State University School of Medicine, Detroit, United States

KEYWORDS

somatic mosaicism, aging, brain disease, cancer, Alzheimer's disease, chromosome, fuzzy inheritance, karyotype coding

Editorial on the Research Topic
Somatic genomic mosaicism & human disease

Somatic genomic mosaicism has become a major focus of genetic research during the last decade (Campbell et al., 2015; D'Gama and Walsh, 2018). Considering the number of cellular divisions required to produce ~10^{14} of cells in an average human being, it is highly unlikely that all these cells share identical genomes. Thus, all humans are apparently genetic mosaics (Iourov et al., 2012). This viewpoint is endorsed by the new genomic concept of "Fuzzy Inheritance" (Heng, 2019). With the introduction of new genomic technologies, somatic mosaicism has been found to be a mechanism for human morbidity. Additionally, somatic (chromosomal and single-gene) mosaicism appears to be a mechanism for human interindividual diversity, development and aging (Campbell et al., 2015; D'Gama and Walsh, 2018; Iourov et al., 2012; Heng, 2019; Vijg, 2014). More precisely, monogenic and chromosomal diseases, neurodevelopmental/neurobehavioral and neuropsychiatric disorders, neurodegeneration, cancer and healthy/unhealthy aging are associated with a wide spectrum of somatic genomic mosaicism types (D'Gama and Walsh, 2018; Iourov et al., 2012; Heng, 2019; Vijg, 2014; Iourov et al., 2019; Yurov et al., 2019; Vorsanova et al., 2020; Ye et al., 2020; Miller et al., 2021; Iourov et al., 2021a; Iourov et al., 2021b). According to the Genome Architecture Theory, somatic mosaicism-mediated heterogeneity is essential for cellular adaptation, and at the same time, as an evolutionary trade-off, somatic mosaicism may be a disease mechanism, as well (Heng, 2019; Ye et al., 2019; Iourov et al., 2020; Iourov et al., 2021b; Heng and Heng, 2021). Timely recognition of the importance of somatic mosaicism is required to understand genetic mechanisms of human morbidity and physiological changes during the ontogeny for improving life quality and span.

This Research Topic presents the knowledge about somatic genomic mosaicism acquired by molecular genetic and cytogenetic/cytogenomic studies. Moreover, original hypotheses about the role of somatic mosaicism in human diseases and innovative approaches to the detection are described.

The role of chromosome instability and mosaic aneuploidy in the pathogenesis of neurodegenerative and neurodevelopmental disorders is generally overlooked. The paper by Potter et al. fills this gap in current biomedical literature and describes a hypothesis suggesting the involvement of aneuploidy in the cognitive deficits that characterize the neurological symptoms of these disorders by promoting apoptosis in the diseased brain. The analysis of somatic chromosomal mosaicism is continued by Liehr and Al-Rikabi, who provided a timely systematic review of mosaic small supernumerary marker chromosomes detected in unaffected individuals.

Since Alzheimer's disease is repeatedly associated with a variety of types of somatic mosaicism (Iourov et al., 2012; Yurov et al., 2019; Ye et al., 2020; Miller et al., 2021), it is not surprising that this common and devastative disease is a focus of four articles of this Research Topic. The description of somatic mosaicism in Alzheimer's disease is started by a review by Bajic et al., who described the role of X chromosome-specific mosaicism and instability in the pathogenesis. Barrio-Alonso et al. hypothesize that neuronal hyperploidization is a highly probable mechanism of Alzheimer's disease. Ueberham and Arendt review genomic indexing by somatic gene recombination of mRNA/ncRNA and suggest that related processes underlie several symptoms of Alzheimer's disease. Still, this process probably has both advantageous and deleterious consequences. Finally, Alzheimer's disease-associated somatic mosaicism is addressed by Kaeser and Chun. The authors present their original potentially unifying hypothesis suggesting that mosaic somatic gene recombination is a novel mechanism to explain the pathogenesis of this currently untreatable disease.

Recently, somatic mosaicism has been found to be involved in cancer pathogenesis (Iourov et al., 2021a; Iourov et al., 2021b; Heng and Heng, 2021). This involvement is highlighted by Ye et al., who used multiple myeloma as a model for describing cancerous aspects of somatic genomic mosaicism.

The Research Topic is finalized by two articles describing approaches to study somatic mosaicism. Kuroki et al. present a study performed for establishing quantitative PCR assays for active Long Interspersed Nuclear Element-1 (LINE1) subfamilies, which may be applied to the analysis of aging-associated retrotransposition, which is a common cause of somatic mosaicism. Dong et al. describe their original and freely available software tool (SCCNV), which may be used for identifying mosaic copy number variation by analysing single-cell whole-genome sequencing data.

Recently, a series of publications have further highlighted the importance of somatic chromosomal mosaicism in cancer and aging. Because karyotype codes the "system information" that organizes gene interactive networks, altered karyotypes represent newly formed information packages. The somatic chromosomal mosaicism should certainly alter genetic-environmental interactions offering therapeutic opportunities in disease and pathological aging. We regret that some of these papers are not included in this Research Topic, but readers are able to read them elsewhere (Ye et al., 2019; Iourov et al., 2020; Vorsanova et al., 2020; Ye et al., 2020; Iourov et al., 2021a; Iourov et al., 2021b; Heng and Heng, 2021; Miller et al., 2021) for complementing their views on somatic genomic mosaicism in humans.

To this end, we have to inform the readers that our co-editor, Svetlana G, Vorsanova, has tragically passed away during the finalization of this topic (for more information, please see (Iourov, 2022)). Accordingly, we dedicate our editorial and Research Topic to her memory.

Author contributions

II and HH have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Campbell, I. M., Shaw, C. A., Stankiewicz, P., and Lupski, J. R. (2015). Somatic mosaicism: Implications for disease and transmission genetics. *Trends Genet.* 31 (7), 382–392. doi:10.1016/j.tig.2015.03.013

D'Gama, A. M., and Walsh, C. A. (2018). Somatic mosaicism and neurodevelopmental disease. *Nat. Neurosci.* 21 (11), 1504–1514. doi:10.1038/s41593-018-0257-3

Heng, H. H. (2019). *Genome chaos: Rethinking genetics, evolution, and molecular medicine*. Cambridge, MA, USA: Academic Press Elsevier.

Heng, J., and Heng, H. H. (2021). Genome chaos, information creation, and cancer emergence: Searching for new frameworks on the 50th anniversary of the "war on cancer. *Genes* 13 (1), 101. doi:10.3390/genes13010101

Iourov, I. Y., Vorsanova, S. G., Kurinnaia, O. S., Zelenova, M. A., Vasin, K. S., and Yurov, Y. B. (2021). Causes and consequences of genome instability in psychiatric and neurodegenerative diseases. *Mol. Biol.* 55 (1), 42–53. doi:10.31857/S0026898421010158

Iourov, I. Y., Vorsanova, S. G., Yurov, Y. B., and Kutsev, S. I. (2019). Ontogenetic and pathogenetic views on somatic chromosomal mosaicism. *Genes* 10 (5), 379. doi:10.3390/genes10050379

Iourov, I. Y., Vorsanova, S. G., and Yurov, Y. B. (2012). Single cell genomics of the brain: Focus on neuronal diversity and neuropsychiatric diseases. *Curr. Genomics* 13 (6), 477–488. doi:10.2174/138920212802510439

Iourov, I. Y., Vorsanova, S. G., Yurov, Y. B., Zelenova, M. A., Kurinnaia, O. S., Vasin, K. S., et al. (2020). The cytogenomic "theory of everything": Chromohelkosis may underlie chromosomal instability and mosaicism in disease and aging. *Int. J. Mol. Sci.* 21 (21), 8328. doi:10.3390/ijms21218328

Iourov, I. Y., Yurov, Y. B., Vorsanova, S. G., and Kutsev, S. I. (2021). Chromosome instability, aging and brain diseases. *Cells* 10 (5), 1256. doi:10.3390/cells10051256

Iourov, I. Y. (2022). Svetlana G. Vorsanova (1945-2021). *Mol. Cytogenet.* 15 (1), 35. doi:10.1186/s13039-022-00613-1

Miller, M. B., Reed, H. C., and Walsh, C. A. (2021). Brain somatic mutation in aging and Alzheimer's disease. *Annu. Rev. Genomics Hum. Genet.* 22, 239–256. doi:10.1146/annurev-genom-121520-081242

Vijg, J. (2014). Somatic mutations, genome mosaicism, cancer and aging. *Curr. Opin. Genet. Dev.* 26, 141–149. doi:10.1016/j.gde.2014.04.002

Vorsanova, S. G., Yurov, Y. B., and Iourov, I. Y. (2020). Dynamic nature of somatic chromosomal mosaicism, genetic-environmental interactions and therapeutic opportunities in disease and aging. *Mol. Cytogenet.* 13, 16. doi:10.1186/s13039-020-00488-0

Ye, C. J., Sharpe, Z., and Heng, H. H. (2020). Origins and consequences of chromosomal instability: From cellular adaptation to genome chaos-mediated system survival. *Genes (Basel)* 11 (10), 1162. doi:10.3390/genes11101162

Ye, C. J., Stilgenbauer, L., Moy, A., Liu, G., and Heng, H. H. (2019). What is karyotype coding and why is genomic topology important for cancer and evolution? *Front. Genet.* 10, 1082. doi:10.3389/fgene.2019.01082

Yurov, Y. B., Vorsanova, S. G., and Iourov, I. Y. (2019). Chromosome instability in the neurodegenerating brain. *Front. Genet.* 10, 892. doi:10.3389/fgene.2019.00892

frontiers in Genetics

HYPOTHESIS AND THEORY
published: 07 November 2019
doi: 10.3389/fgene.2019.01092

Chromosome Instability and Mosaic Aneuploidy in Neurodegenerative and Neurodevelopmental Disorders

Huntington Potter[1,2†], Heidi J. Chial[1,2†], Julbert Caneus[3], Mihret Elos[1,2], Nina Elder[1,2], Sergiy Borysov[4] and Antoneta Granic[5,6,7]*

[1] Department of Neurology, Rocky Mountain Alzheimer's Disease Center, University of Colorado, Aurora, CO, United States, [2] Linda Crnic Institute for Down Syndrome, University of Colorado, Aurora, CO, United States, [3] NanoScience Technology Center, University of Central Florida, Orlando, FL, United States, [4] Department of Math and Science, Saint Leo University, Saint Leo, FL, United States, [5] AGE Research Group, Institute of Neuroscience, Newcastle University, Newcastle upon Tyne, United Kingdom, [6] Newcastle University Institute for Ageing, NIHR Newcastle Biomedical Research Centre, Newcastle upon Tyne, United Kingdom, [7] Newcastle upon Tyne Hospitals, NHS Foundation Trust, Newcastle upon Tyne, United Kingdom

OPEN ACCESS

Edited by:
Svetlana G. Vorsanova, Russian National Research Medical University, Russia

Reviewed by:
Vladan Bajic, University of Belgrade, Serbia
Thomas Liehr, Friedrich Schiller University Germany
Thomas Arendt, Leipzig University, Germany

***Correspondence:**
Huntington Potter
huntington.potter@cuanschutz.edu

†These authors have contributed equally to this work

Specialty section:
This article was submitted to Human Genomics, a section of the journal Frontiers in Genetics

Received: *12 September 2019*
Accepted: *09 October 2019*
Published: *07 November 2019*

Citation:
Potter H, Chial HJ, Caneus J, Elos M, Elder N, Borysov S and Granic A (2019) Chromosome Instability and Mosaic Aneuploidy in Neurodegenerative and Neurodevelopmental Disorders. Front. Genet. 10:1092. doi: 10.3389/fgene.2019.01092

Evidence from multiple laboratories has accumulated to show that mosaic neuronal aneuploidy and consequent apoptosis characterizes and may underlie neuronal loss in many neurodegenerative diseases, particularly Alzheimer's disease and frontotemporal dementia. Furthermore, several neurodevelopmental disorders, including Seckel syndrome, ataxia telangiectasia, Nijmegen breakage syndrome, Niemann–Pick type C, and Down syndrome, have been shown to also exhibit mosaic aneuploidy in neurons in the brain and in other cells throughout the body. Together, these results indicate that both neurodegenerative and neurodevelopmental disorders with apparently different pathogenic causes share a cell cycle defect that leads to mosaic aneuploidy in many cell types. When such mosaic aneuploidy arises in neurons in the brain, it promotes apoptosis and may at least partly underlie the cognitive deficits that characterize the neurological symptoms of these disorders. These findings have implications for both diagnosis and treatment/prevention.

Keywords: Alzheimer's disease, Mosaic aneuploidy, Frontotemporal lobar degeneration, Neuronal apoptosis, Huntington's disease (HD)

INTRODUCTION

Age-associated neurodegenerative diseases exhibit different brain pathologies and different clinical features, and all are associated with reduced neuronal numbers in specific brain regions. Furthermore, when caused by a mutation, each disorder evidently involves a unique pathogenic pathway because the mutant proteins are usually involved in very different physiological processes. Indeed, the normal function of the associated mutant gene should provide insights into that specific disease's pathogenic pathway. For example, the mutations that cause autosomal dominant familial Alzheimer's disease (FAD) arise in only three genes: the amyloid precursor protein (*APP*) gene, the presenilin 1 (*PSEN1*) gene, and the presenilin 2 (*PSEN2*) gene, all three of which encode proteins involved in the production of the Aβ peptide, which is the main pathogenic molecule of AD (Goate and Hardy, 2012; Hardy, 2017). Specifically, PSEN1 and PSEN2 are components of the γ-secretase complex, which, together with the β-secretase

enzyme, cleaves APP to release Aβ. In contrast, patients with frontotemporal dementia (also called frontotemporal lobar degeneration, FTLD) exhibit clinical and pathological characteristics that differ from AD, and most of the autosomal dominantly inherited familial forms of FTLD are caused by a mutation in the *MAPT/Tau* gene, by a mutation in the progranulin (*PRGN*) gene, or by a triplet repeat expansion in the *C9ORF72* gene, all three of which carry out vastly different normal functions in the cell and are also unrelated to the genes mutated in FAD (Rademakers et al., 2012). Interestingly, in some families, instead of or in addition to FTLD, the triplet repeat expansion in the *C9ORF72* gene can cause familial amyotrophic lateral sclerosis, a motor neuron degenerative disease that is associated with cognitive decline only during later stages of the disease. Huntington's disease similarly exhibits a unique pathology and clinical course and is caused by a triplet repeat expansion in the huntingtin (*HTT*) gene, whose normal function is also apparently different from any of the genes associated with familial forms of AD or FTLD (Podvin et al., 2019). Thus, there appears to be no common feature of neurodegenerative disorders beyond the fact that they all result in neuronal loss.

Neurodevelopmental disorders are associated with different pathologies and clinical phenotypes, although they again usually include microcephaly or another indication of a defect in neurogenesis or neuronal survival. For example, ataxia telangiectasia and the related disorder Seckel syndrome are caused by mutations in the *ATM* and *ATR* genes, respectively, which encode two related protein kinases. Because both ataxia telangiectasia and Seckel syndrome appear to involve the loss of neurons, they might be considered neurodegenerative disorders without an essential aging component associated with their underlying mechanisms. Down syndrome also results in reduced neuronal numbers in key brain regions during development, and, interestingly, every person with Down syndrome develops AD brain neuropathology by age 30–40, which usually leads to dementia by age 50–60 (Hartley et al., 2015; Hithersay et al., 2019).

Because these multiple disorders have different pathologies and different clinical symptoms and they involve different pathogenic pathways, as evidenced by the different genes in which causal mutations can arise, it is reasonable to hypothesize that each disorder is distinct and would require different approaches to therapy and prevention. However, if it were possible to identify a key step that is shared among the pathogenic pathways of many neurodegenerative and/or neurodevelopmental disorders, then it would also be reasonable to hypothesize that insights into the causal mechanism might be gained and the potential for a common approach to the development of new therapies might be recognized. Of course, any shared mechanistic features that are identified must also be considered as potentially a mere correlate of the degenerative process rather than as an *essential* step in the pathogenic pathways. To distinguish between these two hypotheses, the strength of the genetics can be exploited because it is self-evident that a direct effect of a mutant gene is likely to be part of the causal mechanism. If multiple neurodevelopmental and neurodegenerative disease-causing mutations impact a common pathogenic step, then that step should be considered a potential key to preventing the neurodegeneration and should thus serve as a prime target for the development of therapeutic interventions that could be applied to multiple disorders.

Over the past decade, we and others have identified a potential common step in the pathogenic pathways that lead to multiple distinct neurodegenerative and neurodevelopmental disorders. Specifically, mutations in genes linked to many of these disorders have been shown to lead to chromosome segregation defects and mosaic aneuploidy in cell types throughout the body, including in brain neurons, which likely contributes to the neuronal cell loss/apoptosis that underlies their neurological features.

Mosaic Aneuploidy in Alzheimer's Disease

Mosaic aneuploidy/hyperploidy, including trisomy 21, was first hypothesized (Potter, 1991) and has been most thoroughly investigated in AD (Geller and Potter, 1999; Yang et al., 2001; Kingsbury et al., 2006; Mosch et al., 2007; Thomas and Fenech, 2008; Iourov et al., 2009b; Arendt et al., 2010; Iourov et al., 2011). Arendt and colleagues' extensive study of brains from AD patients showed that 20–30% of brain neurons are aneuploid during the early preclinical stages of AD and that their specific loss in later stages of the disease can account for 90% of the neuronal atrophy observed at autopsy (Arendt et al., 2010). Somatic mosaic aneuploidy can also be detected in cells from peripheral tissues of AD patients, including fibroblasts, peripheral blood mononuclear cells, and buccal cells (Potter et al., 1995; Migliore et al., 1997; Geller and Potter, 1999; Migliore et al., 1999; Trippi et al., 2001; Thomas and Fenech, 2008), providing an opportunity for early detection.

The specific finding that trisomy 21 mosaicism occurs in many tissues in AD patients, including in the brain, is highly relevant because people with Down syndrome are usually fully trisomic for chromosome 21 due to chromosome mis-segregation during meiosis, every person with Down syndrome develops AD brain neuropathology by 35 years of age, the majority of people with Down syndrome develop AD dementia by age 60, and nearly all people with Down syndrome who die after age 35 have dementia (Glenner and Wong, 1984; Wisniewski et al., 1985; Epstein, 1990; Hartley et al., 2015; Hithersay et al., 2019). The *APP* gene is located on chromosome 21, and its presence in three copies in Down syndrome is presumed to underlie the development of early-onset AD in this population (Hartley et al., 2015). Support for this conclusion comes from the observation that individuals who carry an FAD mutation in the *APP* gene or who have three copies of the *APP* gene due to a local duplication on one chromosome (Rovelet-Lecrux et al., 2006; Sleegers et al., 2006) will develop autosomal-dominant early-onset AD. Furthermore, individuals who only have 1–10% trisomy 21 cells and show no features of Down syndrome also develop early-onset, sporadic AD dementia, suggesting that even low levels of trisomy 21 mosaicism can lead to AD (reviewed in Potter, 1991). Evidence that three copies of the *APP* gene are not only *sufficient* but also *necessary* to cause AD is provided by the fact that rare individuals

who have full trisomy 21 and all of the clinical and physiological features of Down syndrome but have only two copies of the *APP* gene due to a localized deletion on chromosome 21 fail to develop AD symptoms or AD pathology even at an old age (Prasher et al., 1998; Doran et al., 2017).

Evidence that an underlying susceptibility to chromosome segregation defects may be associated with an increased risk of AD was first provided by a study showing that women who were 35 or younger when they gave birth to a child with Down syndrome have a fivefold increased risk of developing AD later in life compared to control mothers or compared to mothers who gave birth to a child with Down syndrome after age 35 (Schupf et al., 1994). More direct support for the idea that both trisomy 21 and AD are associated with an underlying predisposition for chromosome mis-segregation comes from a study of cultured peripheral blood lymphocytes from mothers under the age of 35 who gave birth to a child with Down syndrome (Migliore et al., 2006). In that study, they used fluorescence in situ hybridization probes for chromosomes 21 and 13 and observed significantly higher levels of chromosome non-disjunction involving both chromosomes within the first cell cycle in the lymphocytes from mothers who gave birth to a child with Down syndrome compared to control mothers who had not had a miscarriage and whose children did not have genetic disorders (Migliore et al., 2006). Taken together, these findings suggest that an underlying predisposition for chromosome instability may lead to increased AD risk.

The association of chromosome instability and aneuploidy with AD has been reinforced by mechanistic studies. The key proteins whose mutant genes cause the majority of FAD—the presenilin proteins (PSEN1 and PSEN2) and the amyloid precursor protein (APP)—localize to centromeres/kinetochores, centrosomes, and/or the nuclear envelope (Li et al., 1997; Annaert et al., 1999; Honda et al., 2000; Jeong et al., 2000; Kimura et al., 2001; Nizzari et al., 2007a; Nizzari et al., 2007b; Young-Pearse et al., 2010; Judge et al., 2011). Furthermore, FAD mutations in *PSEN1* or *APP* cause mitotic spindle abnormalities and aneuploidy in transgenic mice and in transfected cells (Boeras et al., 2008; Granic et al., 2010). Treatment of karyotypically normal cells with oligomeric Aβ peptide, which is the product of PSEN1- and PSEN2-dependent cleavage of APP, also disrupts the mitotic spindle and induces chromosome mis-segregation and aneuploidy by competitively inhibiting certain microtubule motors, particularly Kinesin-5/KIF11/Eg5, MCAK/KIF2C, and KIF4A, in both cell culture experiments and in *Xenopus* egg extracts (Boeras et al., 2008; Borysov et al., 2011). These mechanistic studies established that cell cycle defects and the resultant mosaic aneuploidy are a direct effect of FAD mutant genes and are thus likely to be part of the AD pathogenic pathway and not merely a correlate of neurodegeneration in the brain.

The role of chromosome aneuploidy in AD suggests that studying mechanisms that regulate mitosis may lead to novel insights into AD. For example, Shugoshin-1 (SGO1) encodes a protein that is involved in chromosome cohesion and is needed for normal chromosome segregation, and SGO1 haploinsufficiency leads to chromosome missegregation and tumorigenesis (Yamada et al., 2012). Building on the role of aneuploidy in AD, Rao and Yamada and colleagues hypothesized that *SGO1* heterozygous knockout mice may serve as a potential model of sporadic late-onset AD, and they indeed discovered some AD-related pathology as the mice aged, which was associated with prolonged mitosis and spindle checkpoint activation (Rao et al., 2018a; Rao et al., 2018b).

Mosaic Aneuploidy in Frontotemporal Lobar Degeneration

In addition to AD, mosaic aneuploidy has been observed in FTLD (Rossi et al., 2013; Rossi et al., 2014; Caneus et al., 2018). In earlier studies, mosaic aneuploidy was reported in skin fibroblasts and peripheral blood lymphocytes from FTLD patients and in splenic lymphocytes from mouse models of FTLD-MAPT (Rossi et al., 2008; Rossi et al., 2013; Rossi et al., 2014). In the subsequent study, we found mosaic aneuploidy and associated apoptosis in both neuronal and non-neuronal brain cells from patients with familial FTLD who carry a mutation in the *MAPT/Tau* gene (Caneus et al., 2018). Expression of FTLD-causing mutant MAPT induced mitotic spindle abnormalities, chromosome mis-segregation, aneuploidy, and apoptosis in neurons and other cells in the brains of transgenic mice and in transfected cells (Caneus et al., 2018). Furthermore, we showed in our FTLD study that apoptosis occurs in the same brain neurons that are aneuploid and that, in cultured cells expressing FTLD-causing mutant forms of human MAPT, apoptosis follows and depends upon aneuploidy-generating cell cycle defects (Caneus et al., 2018). If the cell cycle is blocked by inhibiting the interaction between MDM2 and p53 by treatment with low doses of Nutlin-3 at 24 h after expression of FTLD-causing mutant MAPT, no aneuploid cells arise (Caneus et al., 2018). Importantly, treatment with Nutlin-3 also blocks apoptosis, indicating that the toxicity of mutant MAPT depends on cells aberrantly proceeding through the cell cycle and becoming aneuploid.

Evidence linking neuronal aneuploidy, neurodegeneration, and MAPT was reported recently by two other groups in *Drosophila* models of FTLD. Specifically, a study by Bougé and Parmentier showed that excess Tau causes mitotic spindle defects, aneuploidy, and apoptosis in neurons by inhibiting the microtubule-dependent motor protein Kinesin-5 (Bouge and Parmentier, 2016). Similar results have been reported by Malmanche et al. who examined photoreceptors and brain neurons in *Drosophila* and found that adult-onset neurodegeneration mediated by MAPT overexpression included the generation of aneuploid cells (Malmanche et al., 2017). The former result is of particular interest in view of our previous finding that Aβ induces chromosome mis-segregation and aneuploidy by competitively inhibiting the activity of Kinesin-5/KIF11/Eg5 (Borysov et al., 2011). Thus, causal mutations leading to AD and FTLD-MAPT appear to lead to chromosome mis-segregation, aneuploidy, and apoptosis through inhibition of the same target enzyme: Kinesin-5/KIF11/Eg5.

In addition to MAPT-FTLD, we have found that mosaic neuronal aneuploidy and dependent apoptosis are also characteristic of brains of individuals with sporadic FTLD or with FTLD caused by mutations in *PRGN* or by triplet repeat expansions in *C9ORF72* (Elos and Caneus et al., unpublished results, manuscript in preparation).

It is likely that other neurodegenerative diseases are also associated with mosaic aneuploidy in the brain. For example, autism spectrum disorder (Yurov et al., 2007; Iourov et al., 2008), ataxia telangiectasia (McConnell et al., 2004; Iourov et al., 2009a; Iourov et al., 2009b), and Lewy body disease, which includes Parkinson's disease (Yang et al., 2015), have all been reported to exhibit either general hyperploidy or mosaic aneuploidy for numerous chromosomes in brain and/or peripheral tissues. Our laboratory also has preliminary evidence for mosaic aneuploidy in both brain cells and fibroblasts from Huntington's disease patients (Elos and Caneus et al., unpublished results, manuscript in preparation).

Mosaic Aneuploidy in Neurodevelopmental Disorders

Mosaic aneuploidy in neurons and other types of cells also characterizes neurodevelopmental disorders. For example, loss-of-function mutations in the ataxia telangiectasia mutated and Rad3-related (ATR) encoded kinase cause Seckel syndrome, a rare autosomal recessive disorder characterized by pre- and postnatal growth delays, microcephaly, and intellectual disability. Loss of ATR function and of the related kinase ataxia telangiectasia mutated (ATM) have been linked to defective DNA repair, which has been assumed to cause the genomic instability, including aneuploidy, observed in these disorders and to make ataxia telangiectasia patients prone to cancer (Wright et al., 1998; Spring et al., 2002; Shen et al., 2005; Murga et al., 2009; Lang et al., 2016; Yazinski and Zou, 2016; Blackford and Jackson, 2017; Quek et al., 2017). Previous studies showed that ATR localizes to centrosomes (Zhang et al., 2007) and that loss of ATR function causes centrosome overduplication (Alderton et al., 2004; Collis et al., 2008; Stiff et al., 2016) and genomic instability (Casper et al., 2004; Mokrani-Benhelli et al., 2013). In a recent study, Kabeche and colleagues reported a mechanism by which loss of ATR function leads to chromosome mis-segregation and aneuploidy (Kabeche et al., 2018; Saldivar and Cimprich, 2018). Specifically, they elegantly demonstrated that ATR localizes to centromeres and is required for proper chromosome segregation, in addition to and *independent of* its roles in DNA damage repair and replication stress responses (Kabeche et al., 2018). Although not discussed by Kabeche and colleagues or in previous publications, the links between ATR and mitosis provide an explanation for how reduced ATR function and subsequent aneuploidy may underlie the neuronal cell loss during development that leads to microcephaly and cognitive dysfunction, the major clinical, pathological, and disabling features of Seckel syndrome: reduced ATR function results in aneuploidy that leads to neuronal apoptosis.

In addition to Seckel syndrome, mosaic aneuploidy has been observed in brain neurons in ataxia telangiectasia itself (Iourov et al., 2007; Iourov et al., 2009a; Iourov et al., 2009b) and in Niemann–Pick type C disease (Granic and Potter, 2013), and in peripheral cells in Nijmegen breakage syndrome (Vessey et al., 1999; Shimada et al., 2009; Shimada et al., 2010; Hou et al., 2012), Fanconi anemia (Nalepa et al., 2013), and xeroderma pigmentosum (Amiel et al., 2004). All of these developmental disorders are characterized by microcephaly or other evidence of poor neurogenesis and/or of neuronal loss, and all are associated with cognitive disfunction.

Mechanisms by Which Neuronal Aneuploidy and Apoptosis Can Arise

Because neurons have been traditionally considered to be post-mitotic (Bhardwaj et al., 2006), it has been unclear how extensive mosaic aneuploidy can arise in neurodegenerative or neurodevelopmental disorders. More recently, it has become appreciated that neurogenesis is more widespread than previously thought and that the capacity for neurogenesis continues into old age, even if not normally utilized (Zhao et al., 2008; Spalding et al., 2013; Boldrini et al., 2018; Sorrells et al., 2018). In the adult brain, three processes have been identified that may generate the neuronal aneuploidy observed at autopsy in patients with AD, FTLD-MAPT, and other neurodegenerative and neurodevelopmental disorders. In principle, the generation and accumulation of aneuploidy in dividing or regenerating cell populations might arise by both genetic and environmental stressors at any time in life (discussed in Potter, 1991; Oromendia and Amon, 2014). Indeed, there is strong evidence that neurogenesis can occur throughout life in several regions of the brain (Zhao et al., 2008; Mu and Gage, 2011; Spalding et al., 2013; Ernst et al., 2014). Furthermore, data from many studies provide evidence that neurogenesis can be induced in many brain regions in adult mice and rats in response to brain damage and attempted self-repair by the brain (Zhou et al., 2004; Zhao et al., 2008; Spalding et al., 2013; Zheng et al., 2013; Ibrahim et al., 2016), or as part of an ongoing process in the sub-ventricular/granular zone of the brain (Eriksson et al., 1998; Hallbergson et al., 2003; Sakamoto et al., 2014). Thus, neuronal damage and the mitotic defects evident in AD, FTLD-MAPT, and other neurodegenerative and neurodevelopmental disorders could result in the production of new aneuploid neurons, which would not be fully functional and would be particularly prone to apoptosis and degeneration. Indeed, aneuploidy has been shown to promote cell death, including neurodegeneration, in many experimental systems (Rajendran et al., 2008; Kai et al., 2009; Arendt et al., 2010; Oromendia and Amon, 2014).

The second potential mechanism for the generation of neuronal aneuploidy in neurodegenerative disease is cell cycle reentry. Neurons in the AD brain express phospho-proteins usually detected only during mitosis, such as cyclin B1, cyclin D1, cdc2, and Ki67 (Vincent et al., 1996; McShea et al., 1997; Vincent et al., 1997; Yang et al., 2001; Arendt, 2012). In AD mice, the loss of preexisting neurons induces the remaining neurons to reenter the cell cycle (Lopes et al., 2009). Indeed, Aβ has been shown to induce the expression of mitotic proteins and cell cycle

reentry in mature neurons in culture (Majd et al., 2008; Absalon et al., 2013; Seward et al., 2013), which we have confirmed (Nina Elder, unpublished observation).

The third potential mechanism for generating aneuploid neurons is based on the recent discovery that striatal astrocytes can transdifferentiate into new neurons capable of forming functional neuronal circuits with preexisting neurons following ischemic brain injury (Magnusson et al., 2014; Duan et al., 2015). This finding suggests that at least some of the aneuploid neurons in AD and FTLD-MAPT brains may be derived from the aneuploid glia that we have shown are present in our preliminary and published studies. In additional preliminary studies, we have found that low numbers of primary astrocytes exposed to Aβ in culture can begin to express the neuronal marker NeuN (Nina Elder, unpublished observation). Taken together, these findings provide evidence that aneuploidy can arise *de novo* in mature neurons by cell cycle reactivation or can be carried over from previously dividing cells that generate new neurons. It is reasonable that age may exacerbate all of these processes because neuronal and non-neuronal aneuploidy have been shown to increase with age (Arendt et al., 2009; Yurov et al., 2009; Yurov et al., 2010; Fischer et al., 2012; Fantin et al., 2019). Aging is also associated with increasing total exposure to environmental stressors, some of which can promote chromosome missegregation and aneuploidy (for reviews, see Potter, 1991; Iourov et al., 2013).

In addition to the close and mechanistic association between aneuploidy and induced apoptosis discussed above, multiple reports in different systems have shown that aneuploid or other copy number variant cells are prone to degeneration/apoptosis (Oromendia and Amon, 2014; Ohashi et al., 2015; Potter et al., 2016; Andriani et al., 2017; Chronister et al., 2019). As mentioned earlier, Arendt and colleagues conducted a pathological study of AD patients' brains across the disease spectrum and showed that neuronal aneuploidy arises before neurodegeneration or clinical symptoms are evident (Arendt et al., 2010). Specifically, they found that the number of aneuploid neurons increases steadily from around 10% in normal controls to around 30% during the early preclinical stages of AD and then declines back to around 10% during the transition from preclinical AD to severe AD when neuronal loss occurs. In addition, they calculated that the loss of aneuploid, but not diploid, neurons accounted for 90% of the neuronal atrophy observed at autopsy of late-stage AD brains (Arendt et al., 2010). Based on their findings, it can be concluded that: 1) aneuploidy in neurons arises in the AD brain before extensive neuronal cell loss occurs and thus the aneuploidy is not likely to be caused by neurodegeneration/neuronal apoptosis, and 2) the vast majority of later neuronal cell loss selectively affects aneuploid neurons, indicating that the neurodegeneration is likely caused by a cell-autonomous cell cycle defect in the neurons themselves rather than by a tissue-wide mechanism (such as an unidentified, diffusible toxic insult released from damaged cells). Possible cell-autonomous effects of aneuploidy that could contribute to cell death include DNA replication stress (Yurov et al., 2011) and proteotoxic stress (Oromendia et al., 2012).

Linking Development and Aging: A Role for Catalysts in Age-Associated Proteinopathies

In view of these considerations, we note that developmental disorders, such as Seckel syndrome, ataxia telangiectasia, Niemann–Pick type C, Nijmegen breakage syndrome, Fanconi anemia, and xeroderma pigmentosum, all of which lead to neuronal apoptosis, degeneration, and microcephaly, result from mutations in genes whose products impact mitosis, directly or indirectly. In contrast, aging-associated neurodegenerative diseases, such as AD, FTLD, Lewy body disease (Yang et al., 2015), and potentially Huntington's disease (Sathasivam et al., 2001; Elos and Caneus et al., unpublished results) and prion disease (Basu et al., 1998; Borchsenius et al., 2000; Nieznanska et al., 2012) all develop abnormal protein deposits in the brain in addition to aneuploidy. The formation of these deposits apparently involves not only the seminal protein itself but often requires inflammation or other aging-associated catalysts. For example, inheritance of the ε4 allele of the apolipoprotein E (*APOE*) gene is the strongest risk factor for the development of AD besides age itself. Interestingly, the *APOE4* allele and an AD-linked *PSEN1* polymorphism have each been shown to increase the risk of meiosis II chromosome segregation errors, leading to Down syndrome, and a mother carrying both the *APOE4* allele and the *PSEN1* polymorphism has an even higher risk of a trisomy 21 conception (Avramopoulos et al., 1996; Petersen et al., 2000; Rodriguez-Manotas et al., 2007; Bhaumik et al., 2017). Indeed, a recent study of older adults with Down syndrome reported that those who were *APOE4* carriers were at increased risk of both dementia and death (Hithersay et al., 2019). Notably, ApoE, particularly ApoE4, catalyzes the conversion of Aβ into the toxic oligomers that directly disrupt the mitotic spindle and chromosome segregation and also leads to amyloid deposition (Potter and Wisniewski, 2012). A similar co-pathological protein likely exists in prion disease too, although this exacerbating protein has been shown not to be ApoE (Tatzelt et al., 1996). This two-hit mechanism involving a mutant aggregation-prone protein plus an amyloid catalyst may underlie the fact that, in AD, amyloid deposits, symptoms, and aneuploidy all arise with aging. Similar two-hit mechanisms may underlie other aging-associated neurodegenerative diseases and neurodevelopmental disorders. Furthermore, the region-specific expression of the second hit (such as with ApoE in AD) may underlie the region-specific pathology and neuronal loss in different disorders.

Constitutional Aneuploidy in the Normal Brain

In addition to its association with neurodegenerative and neurodevelopmental disorders, aneuploidy and possibly copy number variations on a smaller scale are considered potential contributors to diversity in brain function (Iourov et al.,

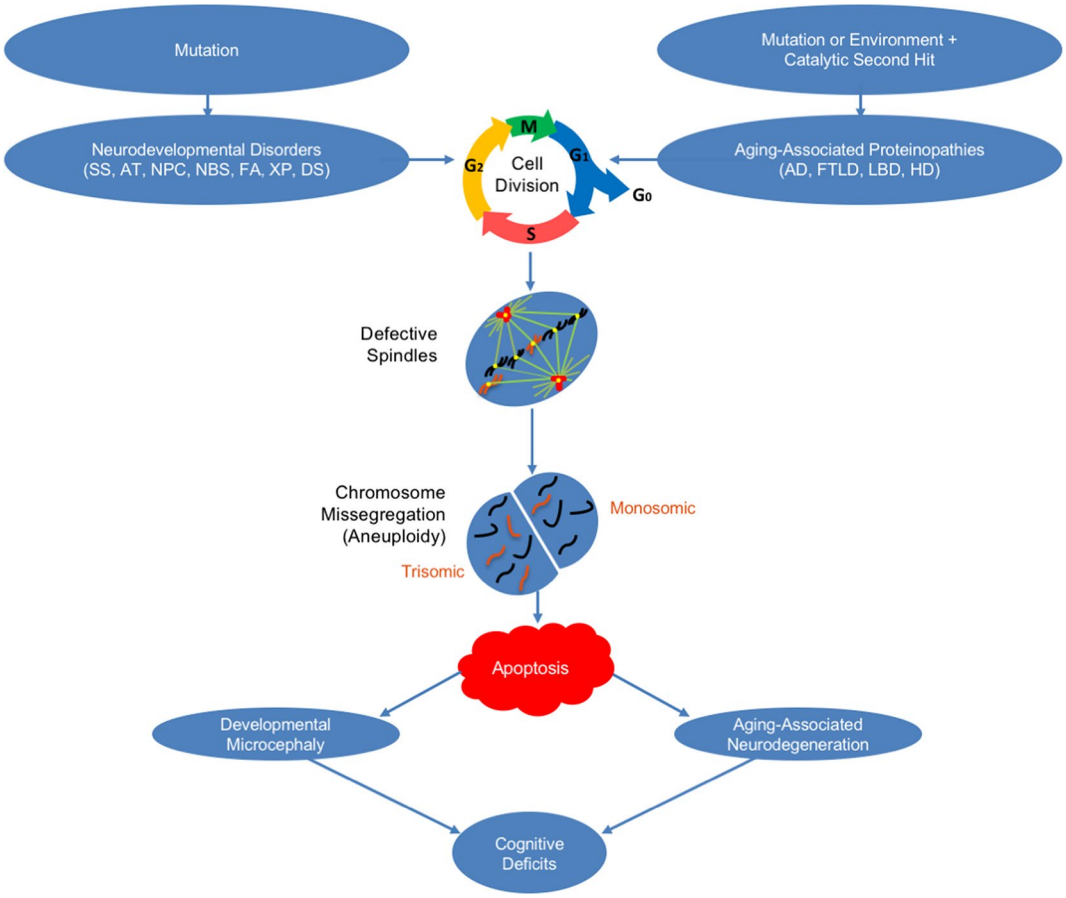

FIGURE 1 | How mosaic aneuploidy may lead to neuronal loss in neurodegenerative and neurodevelopmental disorders. Shown is a schematic of how mosaic aneuploidy may occur in neurons leading to apoptosis and cognitive deficits in neurodevelopmental and aging-associated neurodegenerative disorders. Seckel syndrome (*SS*), ataxia telangiectasia (*AT*), Niemann–Pick type C disease (*NPC*), Nijmegen breakage syndrome (*NBS*), Fanconi anemia (*FA*), xeroderma pigmentosum (*XP*), Down syndrome (*DS*), Alzheimer's disease (*AD*), frontotemporal lobar degeneration (*FTLD*), Lewy body disease (*LBD*), and Huntington's disease (*HD*).

2009b; Mkrtchyan et al., 2010; Bushman and Chun, 2013; Rohrback et al., 2018). Although extensive whole chromosome aneuploidy has not been found by all investigators (Knouse et al., 2014), it is likely that new methods will reveal more aneuploid cells in both normal aged and diseased brains (Caneus et al., 2017).

Summary

In sum, recent work reinforces our emerging understanding of the important role that chromosome mis-segregation and mosaic aneuploidy in neurons may play in an ever-growing list of both neurodevelopmental disorders and aging-associated neurodegenerative disorders (**Figure 1**). These findings have potentially important implications for the development of: 1) novel diagnoses because, as discussed, in addition to neurons in the brain, peripheral cells also exhibit mosaic aneuploidy in these disorders, and 2) innovative preventions/treatments because interventions can now be sought that specifically promote correct chromosome segregation in the presence of aneugenic mutations and/or aneugenic protein structures that lead to neuropathogenesis.

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

AUTHOR CONTRIBUTIONS

All authors contributed to the design and/or execution of the experiments. HP and HC wrote the manuscript. JC primarily designed the figure.

ACKNOWLEDGMENTS

Relevant research was supported by the Alzheimer's Association, the NIH (AG026711; AG037942; NS072076; AG21886), and private donors.

REFERENCES

Absalon, S., Kochanek, D. M., Raghavan, V., and Krichevsky, A. M. (2013). MiR-26b, upregulated in Alzheimer's disease, activates cell cycle entry, tau-phosphorylation, and apoptosis in postmitotic neurons. *J. Neurosci.* 33 (37), 14645–14659. doi: 10.1523/JNEUROSCI.1327-13.2013

Alderton, G. K., Joenje, H., Varon, R., Borglum, A. D., Jeggo, P. A., and O'Driscoll, M. (2004). Seckel syndrome exhibits cellular features demonstrating defects in the ATR-signalling pathway. *Hum. Mol. Genet.* 13 (24), 3127–3138. doi: 10.1093/hmg/ddh335

Amiel, A., Peretz, G., Slor, H., Weinstein, G., and Fejgin, M. D. (2004). Molecular cytogenetic parameters in fibroblasts from patients and carriers of xeroderma pigmentosum. *Cancer Genet. Cytogenet.* 149 (2), 154–160. doi: 10.1016/j.cancergencyto.2003.07.004

Andriani, G. A., Vijg, J., and Montagna, C. (2017). Mechanisms and consequences of aneuploidy and chromosome instability in the aging brain. *Mech. Ageing Dev.* 161 (Pt A), 19–36. doi: 10.1016/j.mad.2016.03.007

Annaert, W. G., Levesque, L., Craessaerts, K., Dierinck, I., Snellings, G., Westaway, D., et al. (1999). Presenilin 1 controls gamma-secretase processing of amyloid precursor protein in pre-golgi compartments of hippocampal neurons. *J. Cell Biol.* 147 (2), 277–294. doi: 10.1083/jcb.147.2.277

Arendt, T. (2012). Cell cycle activation and aneuploid neurons in Alzheimer's disease. *Mol. Neurobiol.* 46 (1), 125–135. doi: 10.1007/s12035-012-8262-0

Arendt, T., Bruckner, M. K., Mosch, B., and Losche, A. (2010). Selective cell death of hyperploid neurons in Alzheimer's disease. *Am. J. Pathol.* 177 (1), 15–20. doi: 10.2353/ajpath.2010.090955

Arendt, T., Mosch, B., and Morawski, M. (2009). Neuronal aneuploidy in health and disease: a cytomic approach to understand the molecular individuality of neurons. *Int. J. Mol. Sci.* 10 (4), 1609–1627. doi: 10.3390/ijms10041609

Avramopoulos, D., Mikkelsen, M., Vassilopoulos, D., Grigoriadou, M., and Petersen, M. B. (1996). Apolipoprotein E allele distribution in parents of Down's syndrome children. *Lancet* 347 (9005), 862–865. doi: 10.1016/s0140-6736(96)91346-x

Basu, J., Williams, B. C., Li, Z., Williams, E. V., and Goldberg, M. L. (1998). Depletion of a Drosophila homolog of yeast Sup35p disrupts spindle assembly, chromosome segregation, and cytokinesis during male meiosis. *Cell Motil. Cytoskeleton* 39 (4), 286–302. doi: 10.1002/(SICI)1097-0169(1998)39:4<286::AID-CM4>3.0.CO;2-1

Bhardwaj, R. D., Curtis, M. A., Spalding, K. L., Buchholz, B. A., Fink, D., Bjork-Eriksson, T., et al. (2006). Neocortical neurogenesis in humans is restricted to development. *Proc. Natl. Acad. Sci. U.S.A.* 103 (33), 12564–12568. doi: 10.1073/pnas.0605177103

Bhaumik, P., Ghosh, P., Ghosh, S., Feingold, E., Ozbek, U., Sarkar, B., et al. (2017). Combined association of Presenilin-1 and Apolipoprotein E polymorphisms with maternal meiosis II error in Down syndrome births. *Genet. Mol. Biol.* 40 (3), 577–585. doi: 10.1590/1678-4685-GMB-2016-0138

Blackford, A. N., and Jackson, S. P. (2017). ATM, ATR, and DNA-PK: The Trinity at the Heart of the DNA Damage Response. *Mol. Cell* 66 (6), 801–817. doi: 10.1016/j.molcel.2017.05.015

Boeras, D. I., Granic, A., Padmanabhan, J., Crespo, N. C., Rojiani, A. M., and Potter, H. (2008). Alzheimer's presenilin 1 causes chromosome missegregation and aneuploidy. *Neurobiol Aging* 29 (3), 319–328. doi: 10.1016/j.neurobiolaging.2006.10.027

Boldrini, M., Fulmore, C. A., Tartt, A. N., Simeon, L. R., Pavlova, I., Poposka, V., et al. (2018). Human Hippocampal Neurogenesis Persists throughout Aging. *Cell Stem Cell* 22 (4), 589–599 e585. doi: 10.1016/j.stem.2018.03.015

Borchsenius, A. S., Tchourikova, A. A., and Inge-Vechtomov, S. G. (2000). Recessive mutations in SUP35 and SUP45 genes coding for translation release factors affect chromosome stability in Saccharomyces cerevisiae. *Curr. Genet.* 37 (5), 285–291. doi: 10.1007/s002940050529

Borysov, S. I., Granic, A., Padmanabhan, J., Walczak, C. E., and Potter, H. (2011). Alzheimer Abeta disrupts the mitotic spindle and directly inhibits mitotic microtubule motors. *Cell Cycle* 10 (9), 1397–1410. doi: 10.4161/cc.10.9.15478

Bouge, A. L., and Parmentier, M. L. (2016). Tau excess impairs mitosis and kinesin-5 function, leading to aneuploidy and cell death. *Dis. Model Mech.* 9 (3), 307–319. doi: 10.1242/dmm.022558

Bushman, D. M., and Chun, J. (2013). The genomically mosaic brain: aneuploidy and more in neural diversity and disease. *Semin Cell Dev. Biol.* 24 (4), 357–369. doi: 10.1016/j.semcdb.2013.02.003

Caneus, J., Granic, A., Chial, H. J., and Potter, H., (2017). "Using Fluorescence In Situ Hybridization (FISH) Analysis to Measure Chromosome Instability and Mosaic Aneuploidy in Neurodegenerative Diseases," in *Genomic Mosaicism in Neurons and Other Cell Types.* Eds. Frade, J. M., and Gage, F. H. (New York, NY: Humana Press), vol. 329-359. doi: 10.1007/978-1-4939-7280-7

Caneus, J., Granic, A., Rademakers, R., Dickson, D. W., Coughlan, C. M., Chial, H. J., et al. (2018). Mitotic defects lead to neuronal aneuploidy and apoptosis in frontotemporal lobar degeneration caused by MAPT mutations. *Mol. Biol. Cell* 29 (5), 575–586. doi: 10.1091/mbc.E17-01-0031

Casper, A. M., Durkin, S. G., Arlt, M. F., and Glover, T. W. (2004). Chromosomal instability at common fragile sites in Seckel syndrome. *Am. J. Hum Genet.* 75 (4), 654–660. doi: 10.1086/422701

Chronister, W. D., Burbulis, I. E., Wierman, M. B., Wolpert, M. J., Haakenson, M. F., Smith, A. C. B., et al. (2019). Neurons with Complex Karyotypes Are Rare in Aged Human Neocortex. *Cell Rep.* 26825-835 (4), e827. doi: 10.1016/j.celrep.2018.12.107

Collis, S. J., Ciccia, A., Deans, A. J., Horejsi, Z., Martin, J. S., Maslen, S. L., et al. (2008). FANCM and FAAP24 function in ATR-mediated checkpoint signaling independently of the Fanconi anemia core complex. *Mol. Cell* 32 (3), 313–324. doi: 10.1016/j.molcel.2008.10.014

Doran, E., Keator, D., Head, E., Phelan, M. J., Kim, R., Totoiu, M., et al. (2017). Down Syndrome, Partial Trisomy 21, and Absence of Alzheimer's Disease: The Role of APP. *J. Alzheimers Dis.* 56 (2), 459–470. doi: 10.3233/JAD-160836

Duan, C. L., Liu, C. W., Shen, S. W., Yu, Z., Mo, J. L., Chen, X. H., et al. (2015). Striatal astrocytes transdifferentiate into functional mature neurons following ischemic brain injury. *Glia* 63 (9), 1660–1670. doi: 10.1002/glia.22837

Epstein, C. J. (1990). The consequences of chromosome imbalance. *Am. J. Med. Genet. Suppl.* 7, 31–37. doi: 10.1002/ajmg.1320370706

Eriksson, P. S., Perfilieva, E., Bjork-Eriksson, T., Alborn, A. M., Nordborg, C., Peterson, D. A., et al. (1998). Neurogenesis in the adult human hippocampus. *Nat. Med.* 4 (11), 1313–1317. doi: 10.1038/3305

Ernst, A., Alkass, K., Bernard, S., Salehpour, M., Perl, S., Tisdale, J., et al. (2014). Neurogenesis in the striatum of the adult human brain. *Cell* 156 (5), 1072–1083. doi: 10.1016/j.cell.2014.01.044

Fantin, C., Moraes Nunes, K., Brito, D. V., Moura Carvalho, N. D., and Benzaquem, D. C. (2019). Chromosomal Alterations in Patients with Alzheimer Disease in Manaus, Amazonas, Brazil. *J. Pharmacy Pharmacol.* 7, 451–458. doi: 10.17265/2328-2150/2019.08.001

Fischer, H. G., Morawski, M., Bruckner, M. K., Mittag, A., Tarnok, A., and Arendt, T. (2012). Changes in neuronal DNA content variation in the human brain during aging. *Aging Cell* 11 (4), 628–633. doi: 10.1111/j.1474-9726.2012.00826.x

Geller, L. N., and Potter, H. (1999). Chromosome missegregation and trisomy 21 mosaicism in Alzheimer's disease. *Neurobiol Dis.* 6 (3), 167–179. doi: 10.1006/nbdi.1999.0236

Glenner, G. G., and Wong, C. W. (1984). Alzheimer's disease and Down's syndrome: sharing of a unique cerebrovascular amyloid fibril protein. *Biochem. Biophys Res. Commun.* 122 (3), 1131–1135. doi: 10.1016/0006-291x(84)91209-9

Goate, A., and Hardy, J. (2012). Twenty years of Alzheimer's disease-causing mutations. *J. Neurochem* 120 Suppl 1, 3–8. doi: 10.1111/j.1471-4159.2011.07575.x

Granic, A., Padmanabhan, J., Norden, M., and Potter, H. (2010). Alzheimer Abeta peptide induces chromosome mis-segregation and aneuploidy, including trisomy 21: requirement for tau and APP. *Mol. Biol. Cell* 21 (4), 511–520. doi: 10.1091/mbc.E09-10-0850

Granic, A., and Potter, H. (2013). Mitotic spindle defects and chromosome mis-segregation induced by LDL/cholesterol-implications for Niemann-Pick C1, Alzheimer's disease, and atherosclerosis. *PloS One* 8 (4), e60718. doi: 10.1371/journal.pone.0060718

Hallbergson, A. F., Gnatenco, C., and Peterson, D. A. (2003). Neurogenesis and brain injury: managing a renewable resource for repair. *J. Clin. Invest* 112 (8), 1128–1133. doi: 10.1172/JCI20098

Hardy, J. (2017). The discovery of Alzheimer-causing mutations in the APP gene and the formulation of the "amyloid cascade hypothesis". *FEBS J.* 284 (7), 1040–1044. doi: 10.1111/febs.14004

Hartley, D., Blumenthal, T., Carrillo, M., DiPaolo, G., Esralew, L., Gardiner, K., et al. (2015). Down syndrome and Alzheimer's disease: Common pathways, common goals. *Alzheimers Dement.* 11 (6), 700–709. doi: 10.1016/j.jalz.2014.10.007

Hithersay, R., Startin, C. M., Hamburg, S., Mok, K. Y., Hardy, J., Fisher, E. M. C., et al. (2019). Association of Dementia With Mortality Among Adults With Down Syndrome Older Than 35 Years. *JAMA Neurol.* 76 (2), 152–160. doi: 10.1001/jamaneurol.2018.3616

Honda, T., Nihonmatsu, N., Yasutake, K., Ohtake, A., Sato, K., Tanaka, S., et al. (2000). Familial Alzheimer's disease-associated mutations block translocation of full-length presenilin 1 to the nuclear envelope. *Neurosci. Res.* 37 (2), 101–111. doi: 10.1016/s0168-0102(00)00106-1

Hou, Y. Y., Toh, M. T., and Wang, X. (2012). NBS1 deficiency promotes genome instability by affecting DNA damage signaling pathway and impairing telomere integrity. *Cell Biochem. Funct.* 30 (3), 233–242. doi: 10.1002/cbf.1840

Ibrahim, S., Hu, W., Wang, X., Gao, X., He, C., and Chen, J. (2016). Traumatic Brain Injury Causes Aberrant Migration of Adult-Born Neurons in the Hippocampus. *Sci. Rep.* 6, 21793. doi: 10.1038/srep21793

Iourov, I. Y., Vorsanova, S. G., Liehr, T., Kolotii, A. D., and Yurov, Y. B. (2009a). Increased chromosome instability dramatically disrupts neural genome integrity and mediates cerebellar degeneration in the ataxia-telangiectasia brain. *Hum. Mol. Genet.* 18 (14), 2656–2669. doi: 10.1093/hmg/ddp207

Iourov, I. Y., Vorsanova, S. G., Liehr, T., and Yurov, Y. B. (2009b). Aneuploidy in the normal, Alzheimer's disease and ataxia-telangiectasia brain: differential expression and pathological meaning. *Neurobiol. Dis.* 34 (2), 212–220. doi: 10.1016/j.nbd.2009.01.003

Iourov, I. Y., Vorsanova, S. G., and Yurov, Y. B. (2007). Ataxia telangiectasia paradox can be explained by chromosome instability at the subtissue level. *Med. Hypotheses* 68 (3), 716. doi: 10.1016/j.mehy.2006.09.021

Iourov, I. Y., Vorsanova, S. G., and Yurov, Y. B. (2011). Genomic landscape of the Alzheimer's disease brain: chromosome instability–aneuploidy, but not tetraploidy–mediates neurodegeneration. *Neurodegener. Dis.* 8 (1-2), 35–37. doi: 10.1159/000315398

Iourov, I. Y., Vorsanova, S. G., and Yurov, Y. B. (2013). Somatic cell genomics of brain disorders: a new opportunity to clarify genetic-environmental interactions. *Cytogenet Genome Res.* 139 (3), 181–188. doi: 10.1159/000347053

Iourov, I. Y., Yurov, Y. B., and Vorsanova, S. G. (2008). Mosaic X chromosome aneuploidy can help to explain the male-to-female ratio in autism. *Med. Hypotheses* 70 (2), 456. doi: 10.1016/j.mehy.2007.05.037

Jeong, S. J., Kim, H. S., Chang, K. A., Geum, D. H., Park, C. H., Seo, J. H., et al. (2000). Subcellular localization of presenilins during mouse preimplantation development. *FASEB J.* 14 (14), 2171–2176. doi: 10.1096/fj.99-1068com

Judge, M., Hornbeck, L., Potter, H., and Padmanabhan, J. (2011). Mitosis-specific phosphorylation of amyloid precursor protein at threonine 668 leads to its altered processing and association with centrosomes. *Mol. Neurodegener* 6, 80. doi: 10.1186/1750-1326-6-80

Kabeche, L., Nguyen, H. D., Buisson, R., and Zou, L. (2018). A mitosis-specific and R loop-driven ATR pathway promotes faithful chromosome segregation. *Science* 359 (6371), 108–114. doi: 10.1126/science.aan6490

Kai, Y., Wang, C. C., Kishigami, S., Kazuki, Y., Abe, S., Takiguchi, M., et al. (2009). Enhanced apoptosis during early neuronal differentiation in mouse ES cells with autosomal imbalance. *Cell Res.* 19 (2), 247–258. doi: 10.1038/cr.2008.305

Kimura, N., Nakamura, S. I., Honda, T., Takashima, A., Nakayama, H., et al., (2001). Age-related changes in the localization of presenilin-1 in cynomolgus monkey brain. *Brain Res.* 922 (1), 30–41. doi: 10.1016/s0006-8993(01)03146-8

Kingsbury, M. A., Yung, Y. C., Peterson, S. E., Westra, J. W., and Chun, J. (2006). Aneuploidy in the normal and diseased brain. *Cell Mol. Life Sci.* 63 (22), 2626–2641. doi: 10.1007/s00018-006-6169-5

Knouse, K. A., Wu, J., Whittaker, C. A., and Amon, A. (2014). Single cell sequencing reveals low levels of aneuploidy across mammalian tissues. *Proc. Natl. Acad. Sci. U.S.A.* 111 (37), 13409–13414. doi: 10.1073/pnas.1415287111

Lang, P. Y., Nanjangud, G. J., Sokolsky-Papkov, M., Shaw, C., Hwang, D., Parker, J. S., et al. (2016). ATR maintains chromosomal integrity during postnatal cerebellar neurogenesis and is required for medulloblastoma formation. *Development* 143 (21), 4038–4052. doi: 10.1242/dev.139022

Li, J., Xu, M., Zhou, H., Ma, J., and Potter, H. (1997). Alzheimer presenilins in the nuclear membrane, interphase kinetochores, and centrosomes suggest a role in chromosome segregation. *Cell* 90 (5), 917–927. doi: 10.1016/s0092-8674(00)80356-6

Lopes, J. P., Blurton-Jones, M., Yamasaki, T. R., Agostinho, P., and LaFerla, F. M. (2009). Activation of cell cycle proteins in transgenic mice in response to neuronal loss but not amyloid-beta and tau pathology. *J. Alzheimers Dis.* 16 (3), 541–549. doi: 10.3233/JAD-2009-0993

Magnusson, J. P., Goritz, C., Tatarishvili, J., Dias, D. O., Smith, E. M., Lindvall, O., et al. (2014). A latent neurogenic program in astrocytes regulated by Notch signaling in the mouse. *Science* 346 (6206), 237–241. doi: 10.1126/science.346.6206.237

Majd, S., Zarifkar, A., Rastegar, K., and Takhshid, M. A. (2008). Different fibrillar Abeta 1-42 concentrations induce adult hippocampal neurons to reenter various phases of the cell cycle. *Brain Res.* 1218, 224–229. doi: 10.1016/j.brainres.2008.04.050

Malmanche, N., Dourlen, P., Gistelinck, M., Demiautte, F., Link, N., Dupont, C., et al. (2017). Developmental Expression of 4-Repeat-Tau Induces Neuronal Aneuploidy in Drosophila Tauopathy Models. *Sci. Rep.* 7, 40764. doi: 10.1038/srep40764

McConnell, M. J., Kaushal, D., Yang, A. H., Kingsbury, M. A., Rehen, S. K., Treuner, K., et al. (2004). 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.* 24 (37), 8090–8096. doi: 10.1523/JNEUROSCI.2263-04.2004

McShea, A., Harris, P. L., Webster, K. R., Wahl, A. F., and Smith, M. A. (1997). Abnormal expression of the cell cycle regulators P16 and CDK4 in Alzheimer's disease. *Am. J. Pathol.* 150 (6), 1933–1939.

Migliore, L., Boni, G., Bernardini, R., Trippi, F., Colognato, R., Fontana, I., et al. (2006). Susceptibility to chromosome malsegregation in lymphocytes of women who had a Down syndrome child in young age. *Neurobiol Aging* 27 (5), 710–716. doi: 10.1016/j.neurobiolaging.2005.03.025

Migliore, L., Botto, N., Scarpato, R., Petrozzi, L., Cipriani, G., and Bonuccelli, U. (1999). Preferential occurrence of chromosome 21 malsegregation in peripheral blood lymphocytes of Alzheimer disease patients. *Cytogenet Cell Genet.* 87 (1-2), 41–46. doi: 15389

Migliore, L., Testa, A., Scarpato, R., Pavese, N., Petrozzi, L., and Bonuccelli, U. (1997). Spontaneous and induced aneuploidy in peripheral blood lymphocytes of patients with Alzheimer's disease. *Hum. Genet.* 101 (3), 299–305. doi: 10.1007/s004390050632

Mkrtchyan, H., Gross, M., Hinreiner, S., Polytiko, A., Manvelyan, M., Mrasek, K., et al. (2010). The human genome puzzle - the role of copy number variation in somatic mosaicism. *Curr. Genomics* 11 (6), 426–431. doi: 10.2174/138920210793176047

Mokrani-Benhelli, H., Gaillard, L., Biasutto, P., Le Guen, T., Touzot, F., Vasquez, N., et al. (2013). Primary microcephaly, impaired DNA replication, and genomic instability caused by compound heterozygous ATR mutations. *Hum Mutat* 34 (2), 374–384. doi: 10.1002/humu.22245

Mosch, B., Morawski, M., Mittag, A., Lenz, D., Tarnok, A., and Arendt, T. (2007). Aneuploidy and DNA replication in the normal human brain and Alzheimer's disease. *J. Neurosci.* 27 (26), 6859–6867. doi: 10.1523/JNEUROSCI.0379-07.2007

Mu, Y., and Gage, F. H. (2011). Adult hippocampal neurogenesis and its role in Alzheimer's disease. *Mol. Neurodegener* 6, 85. doi: 10.1186/1750-1326-6-85

Murga, M., Bunting, S., Montana, M. F., Soria, R., Mulero, F., Canamero, M., et al. (2009). A mouse model of ATR-Seckel shows embryonic replicative stress and accelerated aging. *Nat. Genet.* 41 (8), 891–898. doi: 10.1038/ng.420

Nalepa, G., Enzor, R., Sun, Z., Marchal, C., Park, S. J., Yang, Y., et al. (2013). Fanconi anemia signaling network regulates the spindle assembly checkpoint. *J. Clin. Invest.* 123 (9), 3839–3847. doi: 10.1172/JCI67364

Nieznanska, H., Dudek, E., Zajkowski, T., Szczesna, E., Kasprzak, A. A., and Nieznanski, K. (2012). Prion protein impairs kinesin-driven transport. *Biochem. Biophys Res. Commun.* 425 (4), 788–793. doi: 10.1016/j.bbrc.2012.07.153

Nizzari, M., Venezia, V., Bianchini, P., Caorsi, V., Diaspro, A., Repetto, E., et al. (2007a). Amyloid precursor protein and Presenilin 1 interaction studied by FRET in human H4 cells. *Ann N Y Acad. Sci.* 1096, 249–257. doi: 10.1196/annals.1397.091

Nizzari, M., Venezia, V., Repetto, E., Caorsi, V., Magrassi, R., Gagliani, M. C., et al. (2007b). Amyloid precursor protein and Presenilin1 interact with the adaptor GRB2 and modulate ERK 1,2 signaling. *J. Biol. Chem.* 282 (18), 13833–13844. doi: 10.1074/jbc.M610146200

Ohashi, A., Ohori, M., Iwai, K., Nakayama, Y., Nambu, T., Morishita, D., et al. (2015). Aneuploidy generates proteotoxic stress and DNA damage concurrently with p53-mediated post-mitotic apoptosis in SAC-impaired cells. *Nat. Commun.* 6, 7668. doi: 10.1038/ncomms8668

Oromendia, A. B., and Amon, A. (2014). Aneuploidy: implications for protein homeostasis and disease. *Dis. Model Mech.* 7 (1), 15–20. doi: 10.1242/dmm.013391

Oromendia, A. B., Dodgson, S. E., and Amon, A. (2012). Aneuploidy causes proteotoxic stress in yeast. *Genes Dev.* 26 (24), 2696–2708. doi: 10.1101/gad.207407.112

Petersen, M. B., Karadima, G., Samaritaki, M., Avramopoulos, D., Vassilopoulos, D., et al. (2000). Association between presenilin-1 polymorphism and maternal meiosis II errors in Down syndrome. *Am. J. Med. Genet.* 93 (5), 366–372. doi: 10.1002/1096-8628(20000828)93:5<366::aid-ajmg5>3.0.co;2-g

Podvin, S., Reardon, H. T., Yin, K., Mosier, C., and Hook, V. (2019). Multiple clinical features of Huntington's disease correlate with mutant HTT gene CAG repeat lengths and neurodegeneration. *J. Neurol.* 266 (3), 551–564. doi: 10.1007/s00415-018-8940-6

Potter, H. (1991). Review and hypothesis: Alzheimer disease and Down syndrome–chromosome 21 nondisjunction may underlie both disorders. *Am. J. Hum. Genet.* 48 (6), 1192–1200.

Potter, H., Granic, A., and Caneus, J. (2016). Role of Trisomy 21 Mosaicism in Sporadic and Familial Alzheimer's Disease. *Curr. Alzheimer Res.* 13 (1), 7–17. doi: 10.2174/1567205013011512071006l6

Potter, H., Ma, J., Das, S., Geller, L. N., Benjamin, M., Kayyali, U. S., et al., (1995). "Beyond β-protein: New steps in the pathogenic pathway to Alzheimer's disease," in *Research Advances in Alzheimer's Disease and Related Disorders*. Eds. Iqbal, K., Mortimer, J. A., Winblad, B., and Wisniewski, H. M. (New York: John Wiley and Sons Ltd.), 643–654.

Potter, H., and Wisniewski, T. (2012). Apolipoprotein e: essential catalyst of the Alzheimer amyloid cascade. *Int. J. Alzheimers Dis.* 2012, 489428. doi: 10.1155/2012/489428

Prasher, V. P., Farrer, M. J., Kessling, A. M., Fisher, E. M., West, R. J., Barber, P. C., et al. (1998). Molecular mapping of Alzheimer-type dementia in Down's syndrome. *Ann Neurol* 43 (3), 380–383. doi: 10.1002/ana.410430316

Quek, H., Luff, J., Cheung, K., Kozlov, S., Gatei, M., Lee, C. S., et al. (2017). A rat model of ataxia-telangiectasia: evidence for a neurodegenerative phenotype. *Hum. Mol. Genet.* 26 (1), 109–123. doi: 10.1093/hmg/ddw371

Rademakers, R., Neumann, M., and Mackenzie, I. R. (2012). Advances in understanding the molecular basis of frontotemporal dementia. *Nat. Rev. Neurol.* 8 (8), 423–434. doi: 10.1038/nrneurol.2012.117

Rajendran, R. S., Wellbrock, U. M., and Zupanc, G. K. (2008). Apoptotic cell death, long-term persistence, and neuronal differentiation of aneuploid cells generated in the adult brain of teleost fish. *Dev Neurobiol* 68 (10), 1257–1268. doi: 10.1002/dneu.20656

Rao, C. V., Farooqui, M., Asch, A. S., and Yamada, H. Y. (2018a). Critical role of mitosis in spontaneous late-onset Alzheimer's disease; from a Shugoshin 1 cohesinopathy mouse model. *Cell Cycle* 17 (19-20), 2321–2334. doi: 10.1080/15384101.2018.1515554

Rao, C. V., Farooqui, M., Zhang, Y., Asch, A. S., and Yamada, H. Y. (2018b). Spontaneous development of Alzheimer's disease-associated brain pathology in a Shugoshin-1 mouse cohesinopathy model. *Aging Cell* 17 (4), e12797. doi: 10.1111/acel.12797

Rodriguez-Manotas, M., Amorin-Diaz, M., Canizares-Hernandez, F., Ruiz-Espejo, F., Martinez-Vidal, S., Gonzalez-Sarmiento, R., et al. (2007). Association study and meta-analysis of Alzheimer's disease risk and presenilin-1 intronic polymorphism. *Brain Res.* 1170, 119–128. doi: 10.1016/j.brainres.2007.07.032

Rohrback, S., Siddoway, B., Liu, C. S., and Chun, J. (2018). Genomic mosaicism in the developing and adult brain. *Dev. Neurobiol.* 78 (11), 1026–1048. doi: 10.1002/dneu.22626

Rossi, G., Conconi, D., Panzeri, E., Paoletta, L., Piccoli, E., Ferretti, M. G., et al. (2014). Mutations in MAPT give rise to aneuploidy in animal models of tauopathy. *Neurogenetics* 15 (1), 31–40. doi: 10.1007/s10048-013-0380-y

Rossi, G., Conconi, D., Panzeri, E., Redaelli, S., Piccoli, E., Paoletta, L., et al. (2013). Mutations in MAPT gene cause chromosome instability and introduce copy number variations widely in the genome. *J. Alzheimers Dis.* 33 (4), 969–982. doi: 10.3233/JAD-2012-121633

Rossi, G., Dalpra, L., Crosti, F., Lissoni, S., Sciacca, F. L., et al., (2008). A new function of microtubule-associated protein tau: involvement in chromosome stability. *Cell Cycle* 7 (12), 1788–1794. doi: 10.4161/cc.7.12.6012

Rovelet-Lecrux, A., Hannequin, D., Raux, G., Le Meur, N., Laquerriere, A., Vital, A., et al. (2006). APP locus duplication causes autosomal dominant early-onset Alzheimer disease with cerebral amyloid angiopathy. *Nat. Genet.* 38 (1), 24–26. doi: 10.1038/ng1718

Sakamoto, M., Ieki, N., Miyoshi, G., Mochimaru, D., Miyachi, H., Imura, T., et al. (2014). Continuous postnatal neurogenesis contributes to formation of the olfactory bulb neural circuits and flexible olfactory associative learning. *J. Neurosci.* 34 (17), 5788–5799. doi: 10.1523/JNEUROSCI.0674-14.2014

Saldivar, J. C., and Cimprich, K. A. (2018). A new mitotic activity comes into focus. *Science* 359 (6371), 30–31. doi: 10.1126/science.aar4799

Sathasivam, K., Woodman, B., Mahal, A., Bertaux, F., Wanker, E. E., Shima, D. T., et al. (2001). Centrosome disorganization in fibroblast cultures derived from R6/2 Huntington's disease (HD) transgenic mice and HD patients. *Hum. Mol. Genet.* 10 (21), 2425–2435. doi: 10.1093/hmg/10.21.2425

Schupf, N., Kapell, D., Lee, J. H., Ottman, R., and Mayeux, R. (1994). Increased risk of Alzheimer's disease in mothers of adults with Down's syndrome. *Lancet* 344 (8919), 353–356. doi: 10.1016/s0140-6736(94)91398-6

Seward, M. E., Swanson, E., Norambuena, A., Reimann, A., Cochran, J. N., Li, R., et al. (2013). Amyloid-beta signals through tau to drive ectopic neuronal cell cycle re-entry in Alzheimer's disease. *J. Cell Sci.* 126 (Pt 5), 1278–1286. doi: 10.1242/jcs.1125880

Shen, K. C., Heng, H., Wang, Y., Lu, S., Liu, G., Deng, C. X., et al. (2005). ATM and p21 cooperate to suppress aneuploidy and subsequent tumor development. *Cancer Res.* 65 (19), 8747–8753. doi: 10.1158/0008-5472.CAN-05-1471

Shimada, M., Kobayashi, J., Hirayama, R., and Komatsu, K. (2010). Differential role of repair proteins, BRCA1/NBS1 and Ku70/DNA-PKcs, in radiation-induced centrosome overduplication. *Cancer Sci.* 101 (12), 2531–2537. doi: 10.1111/j.1349-7006.2010.01702.x

Shimada, M., Sagae, R., Kobayashi, J., Habu, T., and Komatsu, K. (2009). Inactivation of the Nijmegen breakage syndrome gene leads to excess centrosome duplication *via the* ATR/BRCA1 pathway. *Cancer Res.* 69 (5), 1768–1775. doi: 10.1158/0008-5472.CAN-08-3016

Sleegers, K., Brouwers, N., Gijselinck, I., Theuns, J., Goossens, D., Wauters, J., et al. (2006). APP duplication is sufficient to cause early onset Alzheimer's dementia with cerebral amyloid angiopathy. *Brain* 129 (Pt 11), 2977–2983. doi: 10.1093/brain/awl203

Sorrells, S. F., Paredes, M. F., Cebrian-Silla, A., Sandoval, K., Qi, D., Kelley, K. W., et al. (2018). Human hippocampal neurogenesis drops sharply in children to undetectable levels in adults. *Nature* 555 (7696), 377–381. doi: 10.1038/nature25975

Spalding, K. L., Bergmann, O., Alkass, K., Bernard, S., Salehpour, M., Huttner, H. B., et al. (2013). Dynamics of hippocampal neurogenesis in adult humans. *Cell* 153 (6), 1219–1227. doi: 10.1016/j.cell.2013.05.002

Spring, K., Ahangari, F., Scott, S. P., Waring, P., Purdie, D. M., Chen, P. C., et al. (2002). Mice heterozygous for mutation in Atm, the gene involved in ataxia-telangiectasia, have heightened susceptibility to cancer. *Nat. Genet.* 32 (1), 185–190. doi: 10.1038/ng958

Stiff, T., Casar Tena, T., O'Driscoll, M., Jeggo, P. A., and Philipp, M. (2016). ATR promotes cilia signalling: links to developmental impacts. *Hum Mol. Genet.* 25 (8), 1574–1587. doi: 10.1093/hmg/ddw034

Tatzelt, J., Maeda, N., Pekny, M., Yang, S. L., Betsholtz, C., Eliasson, C., et al. (1996). Scrapie in mice deficient in apolipoprotein E or glial fibrillary acidic protein. *Neurology* 47 (2), 449–453. doi: 10.1212/wnl.47.2.449

Thomas, P., and Fenech, M. (2008). Chromosome 17 and 21 aneuploidy in buccal cells is increased with ageing and in Alzheimer's disease. *Mutagenesis* 23 (1), 57–65. doi: 10.1093/mutage/gem044

Trippi, F., Botto, N., Scarpato, R., Petrozzi, L., Bonuccelli, U., Latorraca, S., et al. (2001). Spontaneous and induced chromosome damage in somatic cells of sporadic and familial Alzheimer's disease patients. *Mutagenesis* 16 (4), 323–327. doi: 10.1093/mutage/16.4.323

Vessey, C. J., Norbury, C. J., and Hickson, I. D. (1999). Genetic disorders associated with cancer predisposition and genomic instability. *Prog. Nucleic Acid Res. Mol. Biol.* 63, 189–221.

Vincent, I., Jicha, G., Rosado, M., and Dickson, D. W. (1997). Aberrant expression of mitotic cdc2/cyclin B1 kinase in degenerating neurons of Alzheimer's disease brain. *J. Neurosci.* 17 (10), 3588–3598. doi: 10.1083/jcb.132.3.413

Vincent, I., Rosado, M., and Davies, P. (1996). Mitotic mechanisms in Alzheimer's disease? *J. Cell Biol.* 132 (3), 413–425. doi: 10.1083/jcb.132.3.413

Wisniewski, K. E., Wisniewski, H. M., and Wen, G. Y. (1985). Occurrence of neuropathological changes and dementia of Alzheimer's disease in Down's syndrome. *Ann. Neurol.* 17 (3), 278–282. doi: 10.1002/ana.410170310

Wright, J. A., Keegan, K. S., Herendeen, D. R., Bentley, N. J., Carr, A. M., et al., (1998). Protein kinase mutants of human ATR increase sensitivity to UV and

ionizing radiation and abrogate cell cycle checkpoint control. *Proc. Natl. Acad. Sci. U.S.A.* 95 (13), 7445–7450. doi: 10.1073/pnas.95.13.7445

Yamada, H. Y., Yao, Y., Wang, X., Zhang, Y., Huang, Y., Dai, W., et al. (2012). Haploinsufficiency of SGO1 results in deregulated centrosome dynamics, enhanced chromosomal instability and colon tumorigenesis. *Cell Cycle* 11 (3), 479–488. doi: 10.4161/cc.11.3.18994

Yang, Y., Geldmacher, D. S., and Herrup, K. (2001). DNA replication precedes neuronal cell death in Alzheimer's disease. *J. Neurosci.* 21 (8), 2661–2668. doi: 10.1523/JNEUROSCI.21-08-02661.2001

Yang, Y., Shepherd, C., and Halliday, G. (2015). Aneuploidy in Lewy body diseases. *Neurobiol Aging* 36 (3), 1253–1260. doi: 10.1016/j.neurobiolaging.2014.12.016

Yazinski, S. A., and Zou, L. (2016). Functions, Regulation, and Therapeutic Implications of the ATR Checkpoint Pathway. *Annu. Rev. Genet.* 50, 155–173. doi: 10.1146/annurev-genet-121415-121658

Young-Pearse, T. L., Suth, S., Luth, E. S., Sawa, A., and Selkoe, D. J. (2010). Biochemical and functional interaction of disrupted-in-schizophrenia 1 and amyloid precursor protein regulates neuronal migration during mammalian cortical development. *J. Neurosci.* 30 (31), 10431–10440. doi: 10.1523/JNEUROSCI.1445-10.2010

Yurov, Y. B., Vorsanova, S. G., and Iourov, I. Y. (2009). GIN'n'CIN hypothesis of brain aging: deciphering the role of somatic genetic instabilities and neural aneuploidy during ontogeny. *Mol. Cytogenet* 2, 23. doi: 10.1186/1755-8166-2-23

Yurov, Y. B., Vorsanova, S. G., and Iourov, I. Y. (2010). Ontogenetic variation of the human genome. *Curr. Genomics* 11 (6), 420–425. doi: 10.2174/138920210793175958

Yurov, Y. B., Vorsanova, S. G., and Iourov, I. Y. (2011). The DNA replication stress hypothesis of Alzheimer's disease. *ScientificWorldJournal* 11, 2602–2612. doi: 10.1100/2011/625690

Yurov, Y. B., Vorsanova, S. G., Iourov, I. Y., Demidova, I. A., Beresheva, A. K., Kravetz, V. S., et al. (2007). Unexplained autism is frequently associated with low-level mosaic aneuploidy. *J. Med. Genet.* 44 (8), 521–525. doi: 10.1136/jmg.2007.049312

Zhang, S., Hemmerich, P., and Grosse, F. (2007). Centrosomal localization of DNA damage checkpoint proteins. *J. Cell Biochem.* 101 (2), 451–465. doi: 10.1002/jcb.21195

Zhao, C., Deng, W., and Gage, F. H. (2008). Mechanisms and functional implications of adult neurogenesis. *Cell* 132 (4), 645–660. doi: 10.1016/j.cell.2008.01.033

Zheng, W., ZhuGe, Q., Zhong, M., Chen, G., Shao, B., Wang, H., et al. (2013). Neurogenesis in adult human brain after traumatic brain injury. *J. Neurotrauma* 30 (22), 1872–1880. doi: 10.1089/neu.2010.1579

Zhou, L., Del Villar, K., Dong, Z., and Miller, C. A. (2004). Neurogenesis response to hypoxia-induced cell death: map kinase signal transduction mechanisms. *Brain Res.* 1021 (1), 8–19. doi: 10.1016/j.brainres.2004.05.115

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Genetics

SYSTEMATIC REVIEW
published: 11 November 2019
doi: 10.3389/fgene.2019.01131

Mosaicism: Reason for Normal Phenotypes in Carriers of Small Supernumerary Marker Chromosomes With Known Adverse Outcome. A Systematic Review

Thomas Liehr and Ahmed Al-Rikabi*

Institute of Human Genetics, Jena University Hospital, Friedrich Schiller University, Jena, Germany

Small supernumerary marker chromosomes (sSMCs) are present in ~3.3 million of presently living human beings. The majority of these sSMC carriers (i.e. ~2.1 million) will never know about their condition, as they are perfectly healthy and just may learn by chance about it, e.g. if chromosomal analysis is done for some reason during their life time. The remainder ~1.2 million of sSMC carriers are clinically affected either due to adverse effects of gained genetic material being present on the sSMC and/or by uniparental disomy of the sSMC's sister chromosomes. Influence of mosaicism being present in 50% of sSMC carriers is controversy discussed in the literature. Even though genotype–phenotype correlation for sSMCs progressed during last years, still there are only eight sSMC-associated syndromes characterized yet, which may go together with mosaicism. Here we summarize presently available data for carriers of sSMCs normally leading to these well-defined syndromes, however, showing (almost) no clinical signs. This can be observed in ~1 to 30% of the corresponding sSMC-carriers, thus, a high impact for counselling in corresponding prenatal *de novo* cases is not to be neglected.

Keywords: small supernumerary marker chromosomes, genotype–phenotype correlation, Pallister–Killian syndrome, tetrasomy 9p, cat-eye syndrome, proximal tetrasomy 15q, isochromosome 18p

OPEN ACCESS

Edited by:
Ivan Y. Iourov, Mental Health Research Center of Russian Academy of Medical Sciences, Russia

Reviewed by:
Maria Zelenova, Mental Health Research Center of Russian Academy of Medical Sciences, Russia
Marine Manvelyan, Yerevan State University, Armenia

****Correspondence:***
Thomas Liehr
Thomas.Liehr@med.uni-jena.de

Specialty section:
This article was submitted to Human Genomics, a section of the journal Frontiers in Genetics

Received: *21 August 2019*
Accepted: *17 October 2019*
Published: *11 November 2019*

Citation:
Liehr T and Al-Rikabi A (2019) Mosaicism: Reason for Normal Phenotypes in Carriers of Small Supernumerary Marker Chromosomes With Known Adverse Outcome. A Systematic Review. Front. Genet. 10:1131. doi: 10.3389/fgene.2019.01131

INTRODUCTION

Small supernumerary marker chromosomes (sSMCs) are at the same time structural chromosomal abnormalities as well as numerical ones (Liehr, 2012; Liehr, 2019). Thirty percent of sSMCs are inherited from a parent, while 70% are *de novo*. It is suggested that *de novo* sSMCs are products of trisomic rescue. The latter may be conveyed by different mechanisms, like U-type formation, ring chromosome-formation (Liehr, 2012; Liehr, 2019), or chromothripsis (Liehr, 2018; Kurtas et al., 2019). sSMCs are preferentially detected in three groups of patients: (i) infertile, (ii) patients with some kind of physical or mental impairment, and (iii) prenatally, in children with and without sonographic abnormalities. It can be estimated that in a world population of almost eight billion people, 3.3 million sSMC carriers should be present. Approximately 30% of sSMC carriers (1.2 million) are clinically impaired and may get the diagnoses to have an sSMC during life-time. Still most about 2.1 million of these extra chromosome carriers (70%) will never or only by chance learn about their condition (Liehr and Weise, 2007; Liehr, 2012).

An sSMC can derive from each of the 24 human chromosomes, can consist of continuous stretches of DNA from one or more chromosomes, can also be constituted from discontinuous parts of the same or different chromosomes, and contain hetero- and/or euchromatic DNA. Thus, especially for genetic counselling of prenatal *de novo* sSMC genotype–phenotype correlations are urgently needed. Research during last decades showed that there are two major players influencing clinical outcome: (a) gene content of the sSMC and (b) epigenetic influences mediated by imprinting (Liehr, 2012; Al-Rikabi et al., 2018). For (a) it is important to understand that only genes being dosage sensitive can have an impact on the sSMC carrier's phenotype. Accordingly, sSMCs with euchromatin not necessarily are connected with adverse effects for its carrier, and it was already possible to characterize pericentric dosage-insensitive regions for each human chromosome (Al-Rikabi et al., 2018; Liehr, 2019). Presently, there are eight sSMC-related syndromes defined, which are due to adverse gene dosage effects, particularly partial tri- or tetrasomies; these are: isochromosome 5p- [Online Mendelian Inheritance in Man (OMIM) # n.a.], isochromosome 8p- (OMIM # n.a.), tetrasomy 9p- (OMIM # n.a.), proximal tetrasomy 15q- (OMIM # n.a.), Pallister–Killian (OMIM # 601803), isochromosome 18p- (OMIM #614290), isochromosome 20p- (OMIM # n.a.), and cat-eye-syndrome (OMIM #115470) (Liehr, 2019). Concerning (b) one must remember that *de novo* sSMCs normally derive from (incomplete) trisomic rescue. In most cases of any trisomy there are two copies of one maternal chromosome and one copy of a paternal one. In >95% of such cases where e.g. trisomic rescue is initiated one of the two maternal derived chromosomes is degraded. In the remainder cases the paternal copy is lost, which leads to a maternal uniparental disomy. The same may happen vice versa starting with two paternal chromosome copies in a trisomy. Especially if chromosomes 6, 7, 11, 14, 15, or 20 are concerned sSMC-presence may be a hint on an imprinting disease (Liehr et al., 2011; Liehr, 2012).

Another feature for sSMC carriers is that in 50% of the cases a mosaicism of cells with and without sSMC can be observed (Liehr et al., 2010). However, a human being comprises literally hundreds of different tissues while in diagnostics of a living person it is routine to study one, two maximally five different tissues only, for sSMC presence. Normally mosaicism as being observed in one tissue is suggested to be the approximate rate being present in all other tissues of this studied person (Liehr, 2012). Still, singular studies in aborted fetuses showed, that there is/maybe at least a substantial degree of variance in different tissues, and more important, that there is no obvious scheme behind the observable patterns. Particularly it is absolutely impossible to predict reliably the percentage of cells carrying an sSMC in the brain of a prenatally detected carrier by studying amnion-, chorion-, or even blood-cells (Fickelscher et al., 2007). Even though in the majority of the cases the presence of an sSMC known to be deleterious will lead to the expected adverse clinical outcome, during the last decade there were single case reports showing a normal or much less severe than to be expected outcome, especially in case of mosaicism (**Table 1**). Here we summarize these reports and estimate the frequency of clinically normal/only minor affected sSMC carriers in the eight sSMC-related syndromes listed in **Table 1**. Besides those eight sSMC-associated syndromes mentioned before there are three further syndromes being associated with so-called "complex sSMC" (Liehr et al., 2013); as these have a different mode of formation, never show mosaicism and also show no complete absence of phenotypes in sSMC carriers, Emanuel- (OMIM #609029), derivative chromosome 8 and 22- (OMIM #613700), and derivative chromosome 13/21 and 18-syndrome (OMIM # n.a.) were not included in this review.

MATERIALS AND METHODS

Literature Search

All reported sSMC cases are collected in the database: "Small supernumerary marker chromosomes" (accessible *via* http://ssmc-tl.com/sSMC.html, http://molbiol.sci.am/ssmc/ssmc-tl.com/sSMC.html or http://markerchromosomes.ag.vu/= Liehr, 2019). Cases reported with eight sSMC-related syndromes presenting with and without clinical symptoms were identified in this database and summarized in **Table 1**.

RESULTS

Overall, 48 cases out of 2,331 reported cases with sSMC-related syndromes (~2%) showed (almost) normal outcomes, most likely due to mosaicism, reducing the normally adverse clinical signs and symptoms in parts to zero.

In seven out of eight sSMC-related syndromes cases without or only minor clinical symptoms were reported (**Table 1**). In isochromosome 20p-syndrome no clinically healthy sSMC carriers were identified, yet. For the remainder syndromes following percentages of clinically not or less affected than to be expected sSMC carriers were found (**Table 1**): isochromosome 5p-syndrome 30% (out of 27 cases), tetrasomy 9p-syndrome 8.4% (out of 107 cases), isochromosome 8p-syndrome 4% (out of 23 cases), cat-eye syndrome 3.2% (out of 242 cases), isochromosome 18p-syndrome 1.6% (out of 320 cases), Pallister–Killian syndrome 0.8% (out of 608 cases), and proximal tetrasomy 15q-syndrome <0.7% (out of >1,000 cases).

sSMC were found in different percentages of studied tissues of the tested persons listed in **Table 1**. Interestingly, there were several sSMC carriers without symptoms but 100% of cells with sSMC in peripheral blood lymphocytes, as observed in three cases with isochromosome 9p normally associated with tetrasomy 9p-syndrome, one case with isochromosome 18p-syndrome associated sSMC (plus 1 such case in amnion), and two cases with cat-eye syndrome-like sSMC.

DISCUSSION

sSMCs are a challenge especially for prenatal diagnostics and counselling. Here a yet underscored factor for predicting clinical outcome is reviewed, highlighted and discussed: the influence of mosaicism in cases with sSMC. As shown in a previous study

TABLE 1 | sSMC-associated syndromes, number of reported cases are given together with details on cases with no or minor phenotypical signs irrespective of deleterious sSMC and mosaicism with normal cells detected in studied tissues.

Case #	Tissue studied	sSMC %	Phenotype/Frequency
Chromosome 5: isochromosome 5p-syndrome		**27 cases reported**	**30%**
05-W-iso/1-13	CVS; AF; PBL	10/0/0	None
05-W-iso/1-14	CVS; AF; PBL	10/0/0	None
05-W-iso/1-15	CVS; AF; PBL	10/0/0	None
05-W-iso/1-16	CVS; AF; PBL	10/0/0	None
05-W-iso/1-17	CVS; AF; PBL	10/2/0	None
05-W-iso/1-18	CVS; AF; PBL	0/2/0	None
05-W-iso/1-19	PBL; skin; urine	16/0/0	INF
05-W-iso/1-23	AF/PBL; skin (normal); skin (hyperpigm.); urine; buccal mucosa	7/0/13/85/7/70	None
Chromosome 8: isochromosome 8p-syndrome		**23 cases reported**	**4%**
08-W-iso/2-1	PBL	70	None but dwarphism
Chromosome 9: tetrasomy 9p-syndrome		**107cases reported**	**8.4%**
09-W-iso/2-1	PBL; skin	16/0	None
09-W-iso/2-2	PBL; buccal mucosa	100/65	RAB
09-W-iso/2-3	PBL	47	INF
09-W-iso/2-4	PBL	n.a.	None
09-W-iso/2-5	PBL	72	INF
09-W-iso/3-1	PBL	100	Klinefelter like
09-W-iso/4-1	PBL; buccal mucosa	6/5	Klinefelter like
09-W-iso/4-2	PBL; buccal mucosa	100/85	None but dwarphism
09-W-iso/4-3	PBL; skin	30/0	None but dwarphism and Blashko lines
Chromosome 12: Pallister–Killian syndrome		**608 cases reported**	**0.8%**
12-Wpks-1	PBL; skin	0/37	Much less severe than normal PKS
12-Wpks-1a	PBL; skin	0/mosaic	Much less severe than normal PKS
12-Wpks-328	PBL; skin	0/mosaic	Much less severe than normal PKS
12-Wpks-329	Skin; buccal mucosa	mosaic/36%	Much less severe than normal PKS
12-Wpks-357b	PBL; buccal mucosa; hair root cells	50/0/0	None
Chromosome 15: proximal tetrasomy 15q-syndrome		**>1,000 cases reported**	**<0.7%**
15-O-q13/1-1	PBL	56	None
15-O-q13/1-2	AF (1); AF (2); PBL (birth); PBL (2y), PBL (4y)	23/6/26/46/36	None
15-O-q13/2-1	PBL	30	INF
15-O-q13/3-1	AF; PBL; skin; buccal mucosa	6/45/25/8	None
15-O-q13.1/1-1	AF/PBL derived cell line	79/61	None
Mother of 15-O-q13.1/1-1	PBL	10	None
15-O-q13.1/2-1	PBL	93	None
Chromosome 18: isochromosome 18p-syndrome		**320 cases reported**	**1.6%**
18-Wi-158	AF; PBL	35/0	None
18-Wi-158a	PBL	100	None
18-Wi-158b	AF	100	None
18-Wi-158c	AF (1); AF (2); PBL	21/14/0	None
18-Wi-272	PBL	11	Slight DD Much less severe than in i(18p) syndrome
Chromosome 20: isochromosome 20p-syndrome		**4 cases reported**	**0%**
n.a.	n.a.	n.a.	n.a.
Chromosome 22: Cat-eye-syndrome		**242 cases reported**	**3.2%**
22-Wces-5-168; father	PBL; buccal mucosa; spermatozoa	2.8/5.4/49.6	None
22-Wces-5-168; daughter 1	PBL; buccal mucosa	20/32	None
22-Wces-5-168; daughter "	PBL; buccal mucosa	29/63	Mild CES symptoms
22-Wces-5-168; son 1	PBL; buccal mucosa	27/47	Very minor CES symptoms
22-Wces-5-175	PBL	100	(None) No typical CES signs at all
22-Wces-5-192	PBL	100	(None) No typical CES signs at all
22-Wces-5-200	PBL	20	None
22-Wces-5-201	PBL	4.5	None

AF, amnion fluid; case #, identifier of the case acc. to Liehr (2019); CES, cat-eye-syndrome; CVS, chorion villi sampling; i(18p), isochromosome 18p-syndrome; INF, infertile; PBL, peripheral blood lymphocytes; PKS, Pallister–Killian syndrome; RAB, repeated abortions; sSMC %, percentage of cells with sSMC per tissue mentioned in column before. The frequency for (almost) normal phenotype for each of the 8 sSMC-associated syndromes is given in the column "Phenotype/ Frequency".

for different tissues of an aborted sSMC carrier (Fickelscher et al., 2007) also the here summarized cases did not show any tendencies for defined mosaicism rates in different tissues (e.g. case 05-W-iso/1-23 with low rates of sSMC presence in amnion and blood but high rates in skin and urine, or case 22-Wces-5-168 with low sSMC rates in blood and buccal mucosa, but high rates in spermatozoa). Thus, general conclusions or clear predictions about grade of mosaicism in not studied tissues are not possible.

Frequencies of cases (almost) without clinical symptoms are different between the eight sSMC-associated syndromes (**Table 1**) and this has different reasons. No normal sSMC carriers were detected in isochromosome 20p-syndrome, which is most likely due to the small number of (i.e. only four) reported cases. For isochromosome 5p-syndrome 8 of 27 cases (30%) show normal outcomes. Here it must be considered also the small number of available reports, as well as the fact that six cases are prenatal ones from one single study with low rates of cells with isochromosome 5p in chorion or amnion (Venci and Bettio, 2009). Such cases may be more frequent for each numerical chromosomal abnormality, but normally are not reported in scientific papers (Yurov et al., 2018). Still there remain 2/21 cases (9.5%) with iso-chromosome 5p in adult without clinical symptoms, apart from infertility in one of the two cases. In isochromosome 8p-syndrome there is also necessary to consider the small number of reported cases, still 1 in 23 cases without symptoms gives a rate of 4%.

For remainder five other syndromes discussed here >100 case reports, and rates of <0.7 to 3.2% for proximal tetrasomy 15q-syndrome, cat-eye-syndrome, Pallister–Killian syndrome and isochromosome 18p-syndrome were determined, which are close to the overall 2% rate for normal outcomes in otherwise sSMC-related syndromes found here. Still the 8.4% rate for clinically (almost) healthy tetrasomy 9p-syndrome cases is remarkable, especially as this is the largest existent sSMC with overall 94.6 megabases of DNA being present as extra copy to the normal genetic content of a cell. For Pallister–Killian syndrome it must be admitted that for this condition mosaicism is rather the rule than exception, as the disease causing sSMC(12) is known to be lost in fast dividing tissues, regularly. However, patients still show the typical syndrome-associated clinical features. Accordingly, case 12-Wpks-357b with 50% of Pallister–Killian syndrome-typical sSMC in peripheral blood, but completely healthy, got the sSMC(12) restricted in him to peripheral blood most likely by fetal-fetal blood transfusion from his affected twin-sib, who had the sSMC in all body tissues. All other Pallister–Killian syndrome cases included in **Table 1** just show reduced but not completely absent symptoms.

The fact that 6 or 1/48 cases included here showed the sSMC in 100% of their peripheral blood cells or in 100% of amnion cells, is alarming. This means that among prenatal cases identified to be carriers of an sSMC known to be normally deleterious, there are ~2% (for isochromosome 5p-, 8p-, and 9p-syndromes maybe much more) of such fetuses which have a normal clinical outcome.

Overall this review shows that somatic mosaicism being present in at least 50% of sSMC carriers is the third player besides genetic content and uniparental disomy influencing the clinical outcome. Even though overall only 2% of cases may be unexpectedly influenced positively by low mosaicism, e.g. in brain, this needs to be discussed in prenatal genetic counselling. Especially in case of isochromsomes 9, 8, and 5 this possibility could be even more important. Finally, in present times when main stream of human genetics promotes shifting all diagnostic efforts to high throughput approaches, it must be stressed here that (low-level) mosaicism like present in sSMC can only reliably be detected by single cell oriented approaches like banding and/or molecular cytogenetics.

DATA AVAILABILITY STATEMENT

The underlying datasets are available on http://ssmc-tl.com/sSMC.html, http://molbiol.sci.am/ssmc/ssmc-tl.com/sSMC.html and http://markerchromosomes.wg.am/.

AUTHOR CONTRIBUTIONS

TL drafted the paper, and did the literature search and development of discussion part together with AA-R.

REFERENCES

Al-Rikabi, A. B. H., Pekova, S., Fan, X., Jančušková, T., and Liehr, T. (2018). Small supernumerary marker chromosomes may provide information on dosage-insensitive pericentric regions in human. *Curr. Genomics* 19, 192–199. doi: 10.2174/1389202918666170717163830

Fickelscher, I., Starke, H., Schulze, E., Ernst, G., Kosyakova, N., Mkrtchyan, H., et al. (2007). A further case with a small supernumerary marker chromosome (sSMC) derived from chromosome 1–evidence for high variability in mosaicism in different tissues of sSMC carriers. *Prenat. Diagn.* 27, 783–785. doi: 10.1002/pd.1776

Kurtas, N. E., Xumerle, L., Leonardelli, L., Delledonne, M., Brusco, A., Chrzanowska, K., et al. (2019). Small supernumerary marker chromosomes: A legacy of trisomy rescue? *Hum. Mutat.* 40, 193–200. doi: 10.1002/humu.23683

Liehr, T. (2012). *Small Supernumerary Marker Chromosomes (sSMC) - A Guide for Human Geneticists and Clinicians*. Berlin: Springer New York. doi: 10.1007/978-3-642-20766-2

Liehr, T. (2018). Chromothripsis detectable in small supernumerary marker chromosomes (sSMC) using fluorescence in situ hybridization (FISH). *Methods Mol. Biol.* 1769, 79–84. doi: 10.1007/978-1-4939-7780-2_6

Liehr, T. (2019). Small supernumerary marker chromosomes Accessible via http://ssmc-tl.com/sSMC.html, http://molbiol.sci.am/ssmc/ssmc-tl.com/sSMC.html or http://markerchromosomes.ag.vu/.

Liehr, T., and Weise, A. (2007). Frequency of small supernumerary marker chromosomes in prenatal, newborn, developmentally retarded and infertility diagnostics. *Int. J. Mol. Med.* 19, 719–731. doi: 10.3892/ijmm.19.5.719

Liehr, T., Ewers, E., Hamid, A. B., Kosyakova, N., Voigt, M., Weise, A., et al. (2011). Small supernumerary marker chromosomes and uniparental disomy have a story to tell. *J. Histochem. Cytochem.* 59, 842–848. doi: 10.1369/0022155411412780

Liehr, T., Karamysheva, T., Merkas, M., Brecevic, L., Hamid, A. B., Ewers, E., et al. (2010). Somatic mosaicism in cases with small supernumerary marker chromosomes. *Curr. Genomics* 11, 432–439. doi: 10.2174/138920210793176029

Liehr, T., Cirkovic, S., Lalic, T., Guc-Scekic, M., de Almeida, C., Weimer, J., et al. (2013). Complex small supernumerary marker chromosomes - an update. *Mol. Cytogenet.* 6, 46. doi: 10.1186/1755-8166-6-46

Venci, A., and Bettio, D. (2009). Tetrasomy 5p mosaicism due to an additional isochromosome 5p in a man with normal phenotype. *Am. J. Med. Genet. A.* 149A, 2889–2891. doi: 10.1002/ajmg.a.33131

Yurov, Y. B., Vorsanova, S. G., Demidova, I. A., Kolotii, A. D., Soloviev, I. V., and Iourov, I. Y. (2018). Mosaic brain aneuploidy in mental illnesses: An association of low-level post-zygotic aneuploidy with schizophrenia and comorbid psychiatric disorders. *Curr. Genomics* 19, 163–172. doi: 10.2174/1389202918666170717154340

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Genetics

REVIEW
published: 28 January 2020
doi: 10.3389/fgene.2019.01368

The X Files: "The Mystery of X Chromosome Instability in Alzheimer's Disease"

Vladan P. Bajic[1†], Magbubah Essack[2†], Lada Zivkovic[3], Alan Stewart[4], Sonja Zafirovic[1], Vladimir B. Bajic[2‡], Takashi Gojobori[2,5], Esma Isenovic[1] and Biljana Spremo-Potparevic[3]*

[1] Laboratory for Radiobiology and Molecular Genetics, Vinca Institute of Nuclear Sciences, University of Belgrade, Belgrade, Serbia, [2] Computational Bioscience Research Center (CBRC), Computer, Electrical and Mathematical Sciences and Engineering (CEMSE) Division, King Abdullah University of Science and Technology (KAUST), Thuwal, Saudi Arabia, [3] Department of Physiology, Faculty of Pharmacy, University of Belgrade, Belgrade, Serbia, [4] School of Medicine, University of St Andrews, St Andrews, United Kingdom, [5] Biological and Environmental Sciences and Engineering Division (BESE), King Abdullah University of Science and Technology (KAUST), Thuwal, Saudi Arabia

OPEN ACCESS

Edited by:
Ivan Y. Iourov, Mental Health Research Center of Russian Academy of Medical Sciences, Russia

Reviewed by:
Huntington Potter, University of Colorado Denver, United States
Anja Weise, University Hospital Jena, Germany

***Correspondence:**
Vladan P. Bajic
vladanbajic@yahoo.com

†These authors have contributed equally in this work and share first authorship

‡Deceased

Specialty section:
This article was submitted to Human Genomics, a section of the journal Frontiers in Genetics

Received: *25 September 2019*
Accepted: *13 December 2019*
Published: *28 January 2020*

Citation:
Bajic VP, Essack M, Zivkovic L, Stewart A, Zafirovic S, Bajic VB, Gojobori T, Isenovic E and Spremo-Potparevic B (2020) The X Files: "The Mystery of X Chromosome Instability in Alzheimer's Disease". Front. Genet. 10:1368. doi: 10.3389/fgene.2019.01368

Alzheimer's disease (AD) is a neurodegenerative disease that affects millions of individuals worldwide and can occur relatively early or later in life. It is well known that genetic components, such as the amyloid precursor protein gene on chromosome 21, are fundamental in early-onset AD (EOAD). To date, however, only the apolipoprotein E4 (ApoE4) gene has been proved to be a genetic risk factor for late-onset AD (LOAD). In recent years, despite the hypothesis that many additional unidentified genes are likely to play a role in AD development, it is surprising that additional gene polymorphisms associated with LOAD have failed to come to light. In this review, we examine the role of X chromosome epigenetics and, based upon GWAS studies, the PCDHX11 gene. Furthermore, we explore other genetic risk factors of AD that involve X-chromosome epigenetics.

Keywords: X chromosome, Alzheimer's disease, sex chromosome dosage, protocadherin 11, centromere instability

INTRODUCTION

In the first two decades of the 21st century, the proportion of individuals living with Alzheimer's disease (AD) [AD (MIM: 104300)] has been on the rise with an increasingly aging population. Today, two basic forms of AD exist, early-onset AD (EOAD) and late-onset AD (LOAD). EOAD correlates with the occurrence of mutations on specific genes that have given rise to inherited forms of the disease, whilst LOAD - which occurs later in life - has no specified etiology (Smith, 1998; Selkoe, 2001). Familial studies have identified a point mutation associated with EOAD on chromosome 21. This mutation is located in a gene called amyloid precursor protein (APP), and all members of these families show signs of the Alzheimer's phenotype at a relatively early stage of

Abbreviations: AD, Alzheimer's disease; EOAD, Early onset AD; GWAS, genome-wide association studies; LOAD, Late onset AD; NRC/MASC, N-methyl-D-aspartate receptor complex/membrane-associated guanylate kinase-associated signaling complex; PCDH11X, protocadherin 11; PSD, postsynaptic density; SCD, Sex Chromosome Dosage; Xist, X-inactive-specific transcript gene; Xi, inactive X chromosome.

life (Wiseman et al., 2018). Novel mutations located on chromosomes 14 and 1 in genes encoding presenilin-1 and presenilin-2 have also been identified in EOAD (Guven et al., 2019). Unfortunately, specific genetic determinants that can explain the high prevalence of LOAD have yet to be identified.

Today we are aware that EOAD comprises only 1-3% of all AD cases (Smith, 1998; Selkoe, 2001; Bekris et al., 2010). In LOAD subjects, disease prevalence changes with age; 5% after 65 years of age, 20% after 75 years of age, 30% after 80 years of age (Bekris et al., 2010). Also women are twice as likely to suffer from AD than men (Pike, 2017). This prevalence is suggested to be due to differences in the life expectancy between males and females and to hormonal status (Vest and Pike, 2013; Pike, 2017). Studies with twins clearly signal that a strong genetic component is present in LOAD cases (Gatz et al., 2006; Seripa et al., 2009). Various genes have been implicated in AD and identified by using genetic approaches, such as Genome-Wide Association Studies (GWAS). However, the only "single gene" risk factor for LOAD without opposition in the research community concerns the gene encoding apolipoprotein E4 (Giri et al., 2017). In LOAD, the percentage of individuals carrying the at-risk allele of the ApoE4 gene was found to be between 20% and 70%, suggesting that there are additional genetic, and perhaps also epigenetic, factors that underlie the development of LOAD (Slooter et al., 1998; Giri et al., 2017). Carrasquillo et al. (Carrasquillo et al., 2009) found that an alteration of a single-nucleotide polymorphism (SNP; rs5984894) on the Xq21.3 in a gene called protocadherin 11 (PCDH11X) in a cohort of women, was significantly associated with LOAD (**Figure 1**). Other GWAS, however, have been unable to confirm the existence of these connections (Beecham et al., 2010; Wu et al., 2010; Miar et al., 2011; Chung et al., 2013). We hypothesize that one of the possible answers to these observed genetic discrepancies is based on the epigenetics of the X chromosome.

We have previously identified centromere impairment or premature centromere separation (PCS) of the X chromosome in neuronal nuclei of the cerebral cortex in AD women (Spremo-Potparevic et al., 2008). In addition, Yurov et al. discovered X chromosome aneuploidy in AD-affected neurons (Yurov et al., 2014), which suggests that premature centromere separation is a mechanism of X chromosome instability (Spremo-Potparevic et al., 2008). Epigenetically, chromosome X can be affected by skewed X chromosome inactivation, asynchronous replication patterns of the inactive X chromosome (Xi), X-inactivation escape, aneuploidy, and premature centromere separation. All these epigenetic X chromosome changes could potentially affect X chromosome genes through changes in sex chromosome dosage (SCD), and consequently promote AD pathogenesis (Amiel et al., 1998; Gribnau et al., 2005; Ahn and Lee, 2008; Spremo-Potparevic et al., 2008; Hong and Reiss, 2014; Mugford et al., 2014; Yurov et al., 2014; Bajic et al., 2015a; Balaton and Brown, 2016; Le Gall et al., 2017; Graham et al., 2019). Raznahan et al. found recently that sex chromosome dosage not only influenced the adjacent sex chromosomes X and Y, but also autosomal gene expression (Raznahan et al., 2018).

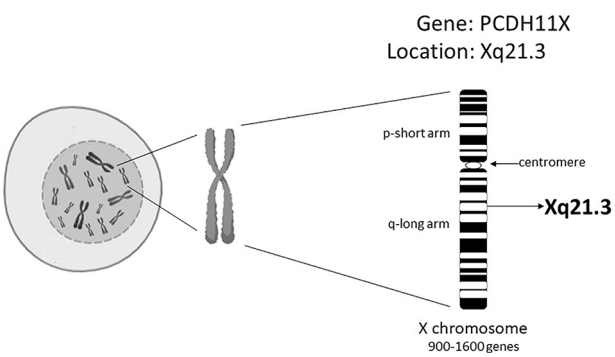

FIGURE 1 | The genetic position of the PCDH11X on the X chromosome.

The X Chromosome Is Unique

In women, there is a systematic demand to compensate for SCD by silencing one of the copies of the X chromosome. With two X-chromosomes, women are more prone to inheriting potentially deleterious mutations in X-encoded genes, which, because of Xi, may all be expressed in different cells. The first finding of inactivation of the X chromosome was reported by Lyon, (1961). It was found that one of the X chromosomes, paternal or maternal, was always inactivated, suggesting that an inactivation mechanism only allows active transcription at one X chromosome (Splinter et al., 2011). This process of X-chromosome inactivation (XCI) evolved as a mechanism to regulate gene dosage. As a compensation mechanism, it does not affect all genes equally, and those genes that are not affected are known to escape XCI [termed escapees; (Pessia et al., 2012)].

Human embryos initially have non-random imprinted XCI, where the X-chromosome from the mother remains active, and XCI applies only to the X-chromosome inherited from the father. The imprint is not constant; XCI resets at the embryonic implantation stage. At this point the XCI reset leaves the maternal and paternal X open to random inactivation (Sun and Lee, 2006). Because XCI at this stage is random it causes most women to be mosaic for two cell lines, one harboring the active chromosome, the paternal X, and the other the maternal X. The randomness of this process causes an XCI ratio of approximately 50%:50% to be associated with the two cell lines in the female population. However, on rare occasions, in approximately 9% of the female population, a bias towards one of the two X chromosomes produces a skewed ratio (> 80%:20%; (Amos-Landgraf et al., 2006). In this regard, Renault et al., analyzed the distribution of X-inactivation patterns (the relative abundance of the two cell populations) in a large cohort study of normal females, and reported that human XCI distribution pattern is more genetically influenced in comparison to the Xi model, which suggests a completely random selection of XCI (Renault et al., 2013).

The genetically influenced selection of XCI may be indicative of mutations in genes (Orstavik, 2009; Shvetsova et al., 2019), suggesting the inactive X chromosome often harbors the mutated allele of an X-linked gene. This would mean that with a 50%:50% XCI ratio, wild type cells generally ameliorate disease

phenotypes. Changes in the XCI ratio towards an increased expression of mutated genes can increase disease phenotype severity, as it is in the case of female hemophilia A (Renault et al., 2007), and sideroblastic anemia (Cazzola et al., 2000), where the majority of cells express the mutated allele. Changes in the XCI ratio where expression of the mutated allele is increased to exhibit the disease phenotype can also occur, as in Rett syndrome. In this case, the hemizygous mutation of the methyl-CpG–binding protein 2 (MeCP2) gene in males causes lethality, while the MeCP2 heterozygous mutation in females weakens such phenotypic consequences. It seems that the loss of MeCP2 function contributes to Rett syndrome, while the gain in MeCP2 dosage does not necessarily ameliorate the disease phenotype but may manifest as a less aggressive form in other neurological diseases. Increased expression of MeCP2 was found to be associated with other neurological diseases, such as AD and Huntington's disease (Amir et al., 1999; Ausio et al., 2014; McFarland et al., 2014; Maphis et al., 2017).

Xi acquires several features of heterochromatin, such as hypermethylation, hypercondensation, altered replication patterns (late *vs.* early), and depletion of acetylated histones (Chow and Brown, 2003; Ng et al., 2007). Methylation patterns have been extensively used to determine the inactive chromosome (Shvetsova et al., 2019), enabling an analysis of non-random inactivation processes in diseases that are X chromosome-linked (Yuan et al., 2015). In our published study we suggest that changes in the inactivation patterns of the X-chromosome could have an impact on AD pathogenesis (Bajic et al., 2015a).

BRAIN AND THE X CHROMOSOME

The X chromosome harbors 3-5% of all the genes in a genome (Skuse, 2005). There has been a debate on how many genes reside on the X chromosome and how many genes are expressed in the brain alone, compared to genes that are X-linked and expressed in the placenta, testes, muscles, and ovary. It is estimated that between 1,100 and 1,500 genes are present on the X chromosome (Skuse, 2005; Laumonnier et al., 2007). By using the Mart View software it was found that 1,500 X-linked genes are expressed in the brain, which represent numerous candidate genes that could be responsible for X-linked brain diseases (Laumonnier et al., 2007). Many of the proteins expressed from the genes linked to the X chromosome represent channels, receptors, repair, transcription factors, and DNA/RNA binding proteins. Most of these proteins are located in the postsynaptic cleft and postsynaptic density (PSD) and are regulated through signaling complexes (Nguyen and Disteche, 2006; Laumonnier et al., 2007). It is intriguing that even if the X chromosome harbors 3-5% of all the genes, it is responsible for 10% of all diseases with Mendelian inheritance (Germain, 2006).

Another aspect that makes the X chromosome unique is that it harbors a higher proportion of brain-expressed miRNAs than would be expected (Goncalves et al., 2019), with 20% of these related to autoimmune diseases such as rheumatoid arthritis and systemic lupus erythematosus (Khalifa et al., 2016). Most of these miRNA are clustered, for example, miR532/188, miR-221/222, miR-98/Let7f, and miR-363/106a/20b/92a (DeMarco et al., 2019; Goncalves et al., 2019). Many of these are also intronic and it is believed that they are co-transcribed and co-expressed with other genes linked to chromosome X (deleted X-linked genes) and may be susceptible to SCD, skewing, and Xi escape processes. It is important to point out that inflammation and altered immunity are features of AD (Forloni et al., 1992; Hauss-Wegrzyniak et al., 1998; Eikelenboom et al., 2000; McGeer and McGeer, 2002; Castellani et al., 2008; Krstic et al., 2012; Bajic et al., 2015c; Regen et al., 2017).

How these X-linked genes interact with genes controlling the immune system in AD is still unknown. For individual genes involved in diseases of the brain, a more complex hypothesis is that interplay occurs in disease genes embedded in multiprotein neuronal complexes. Many of the most important components of neuronal complexes are encoded on the X chromosome (Laumonnier et al., 2007). Such complexes, which are essential for neuronal plasticity, cognitive processes, and cell signaling, are thought to be in the PSD cleft (Muddashetty et al., 2011; Yudowski et al., 2013). Taking N-methyl-D-aspartate receptor/membrane-associated guanylate kinase-associated signaling complex/(NRC/MASC) as an example; combining its 185 proteins and with the other proteins in PSD gives a total of 1100 proteins. The X chromosome plays an essential role, and the percentage of genes related to synaptic plasticity, some 86% of all the genes in NRC/MASC are genes linked to chromosome X (Grant et al., 2005; Laumonnier et al., 2007). It is interesting that these genes are also presented or expressed in human cognitive disorders (Grant et al., 2005; Pocklington et al., 2006; Fernandez et al., 2009; Tam et al., 2009). An analysis of the number of altered proteins in X-linked mental retardation disorders shows that from 69 genes currently known, 19 (or 28%) of these genes belong to postsynaptic proteins (Laumonnier et al., 2007). The same pattern is conserved in the mouse X chromosome, and this suggests a network of multiprotein complexes functioning as integrated entities or complex molecular machines. If one component of this complex machinery is disrupted, the whole complex/network fails thus impairing the overall role of the multiprotein complex in processes of cognition (Grant et al., 2005; Nguyen and Disteche, 2006; Pocklington et al., 2006; Laumonnier et al., 2007; Fernandez et al., 2009; Tam et al., 2009).

X-Linked miRNA and the Brain

The X chromosome is enriched in ncRNAs and harbors several miRNAs essential to brain function (Goncalves et al., 2019). It is important to note that miRNAs not only affect mRNA through translation repression but also work through other ncRNAs, such as lncRNAs and circRNAs, affecting downstream genes (Khalifa et al., 2016; DeMarco et al., 2019; Goncalves et al., 2019). Bian et al. revealed that a miRNA located on the X chromosome, a miR-374 family member, plays a role in cell growth and differentiation not only in various cancers, but also in AD. This miR-374 member is located at the X chromosome

inactivation center and targets the VEGF, PTEN, Wnt, and Fas signaling pathways (Bian et al., 2019). Importantly, the PTEN pathway is of importance to the progression of AD through a mechanism that includes altered autophagy (Wani et al., 2019), mitophagy (Fang et al., 2019), and apoptosis (Cui et al., 2017). A report by Manzine et al. suggested that miR-374 directly targets the beta-secretase 1 to regulate the progress of AD, as the levels of miR-374 were significantly decreased in comparison to controls (Manzine et al., 2018). In addition to miR-374, several miRNAs have also been found to correlate to X chromosome-linked intellectual disability syndrome, and among them are miR-223-3p, miR-362-5p, miR-504-5p.1, miR-361-5p, miR-505-3p.1 and miR-505-3p.2. All these miRNAs act as key regulators of genes linked to chromosome X but also of many autosomal intellectual disability genes that are connected in a complex network (Goncalves et al., 2019).

In the future, it is hoped that further work will reveal the extent to which genes on the X chromosome and miRNAs expressed in the brain, that together regulate processes including nervous system development, cell proliferation and transcription regulation, are altered by X chromosome skewing and asynchronous replication, which lead to aneuploidy and deregulation of cohesion dynamics in AD. Also, RNA genes that are linked to the X chromosome are prone to escape inactivation of the X chromosome (Peeters et al., 2019). These epigenetic processes may prove to be gender-associated as research shows that expression of an X-linked miRNA in rheumatoid arthritis is more prevalent in women than in men (Khalifa et al., 2016).

PCDH11X

Carrasquillo et al. previously identified an SNP (rs 5984894) on the X chromosome (Xq21,3) in a gene called PCDH11X (Carrasquillo et al., 2009). This locus is associated with LOAD in women of European origin from the USA. The PCDH11X gene encodes the protein, protocadherin 11. Women who are homozygous for this SNP have a greater risk of developing AD, not only when compared to women without the SNP, but also when compared to women that are heterozygotes, and male hemizygotes (Carrasquillo et al., 2009). Zubenko et al. reported that the DXS1047 genotype is correlated with AD (Zubenko et al., 1999) and that this genotype is associated with the PCDH11X gene (Zubenko et al., 1998). The results from the same authors indicate an association between the variation in the PCDH11X gene and the risk of acquiring AD, but these results have not been confirmed in other GWAS (Beecham et al., 2010; Wu et al., 2010; Miar et al., 2011). Our suggestion is that these discrepancies in GWAS results may well be due to the changes in the epigenetics of the X chromosome.

Does PCDH11X Escape X Inactivation?

Pseudoautosomal genes and functional Y chromosome orthologues (X-linked genes with Y homology) tend to escape X inactivation (Disteche et al., 2002; Brown and Greally, 2003). Sudbrak et al. reported that PCDH11X expression might also escape X inactivation, and this assumption was verified by using an X chromosome-specific cDNA microarray where elevated expression of PCDH11X was identified in cells expressing multiple X chromosomes (Sudbrak et al., 2001). Lopes et al. indirectly found that PCDH11X expression was higher in women than in men by looking at CpG islands and their methylation patterns. By using bisulfite sequencing analysis, the same authors found the absence of CpG island methylation on both the active and the Xi chromosomes and that these processes coincide with possible PCDH11X escape from X inactivation (Lopes et al., 2006). Another study found that PCDH11X can undergo asynchronous replication, and that PCDHX11 is also prone to escape the inactivation process (Wilson et al., 2007). Replication asynchrony of the X pseudoautosomal locus has been identified (Vorsanova et al., 2001), and suggests that other genes that replicate asynchronously are also prone to escape inactivation (Anderson and Brown, 2005; Carrel and Willard, 2005; Escamilla-Del-Arenal et al., 2011).

PCDH11X Asynchronous Replication

Xi is associated with a sequence of epigenetic modifications (Chow and Brown, 2003), and goes through a phase of changes involving DNA methylation and histone modification resulting in Xi condensation in a body called the Barr body. This results in changes in DNA replication – more specifically, the Xi in the S phase replicates later than its active counterpart. Imperfect chromosome replication can be a consequence of "escapees" (genes that escape the inactivation process). Such genes include hypoxanthine-guanine phosphoribosyltransferase and Fragile X-chromosome genes that display asynchronicity. The X-inactive-specific transcript (Xist) gene (important for inactivation) that is expressed from the Xi also replicates asynchronously (Boggs and Chinault, 1994; Aladjem and Fu, 2014). Wilson et al. reported that PCDH11X displays replication asynchrony in both female and male cells (Wilson et al., 2007). The data from these authors, together with those from others (Orstavik, 2009), show that a complex relationship exists between X-inactivation, replication asynchrony, and the status of expression of individual genes on chromosome X (Bajic et al., 2008; Bajic et al., 2009).

It thus appears that synchronous replication occurs more frequently than previously thought, and is found not only through imprinting, but also through randomized monoallelic expression, pathologies, and tandem duplications (Wilson et al., 2007). Clinically, an increase in asynchronous replication increases the risk in women for aneuploidy (Amiel et al., 2000). The relationship between centromere instability, control of replication, and nondisjunction are best exemplified by the fact that young women that have children with Down's syndrome have twice the risk of developing AD (Hardy et al., 1989; Goate et al., 1990; Fidani et al., 1992; Schellenberg et al., 1992; Schupf et al., 1994; Petersen et al., 2000; Schupf et al., 2001; Migliore et al., 2006; Migliore et al., 2009; Iourov et al., 2010; Goate and Hardy, 2012).

Chromosomes 21, X, and 18 were primarily affected, showing repeated non-disjunction and centromere impairment (Potter and Geller, 1996; Geller and Potter, 1999; Petersen et al., 2000;

Migliore et al., 2006; Migliore et al., 2009; Iourov et al., 2010; Potter, 2016). We suggest that X chromosome replication asynchrony is likely to lead to accelerated instability of chromosome X in AD (Bajic et al., 2009).

SEX CHROMOSOME DOSAGE (SCD): AN ENGINE OF STABILITY

The crosstalk that exists between X chromosomes and autologous genes is a relatively new paradigm that has emerged as a result of the biology of sex differences, and gives rise to the question of how SCD shapes the genome function. To explore this, human sex aneuploidies were analyzed from a genome-wide expression dataset by Raznahan et al. where they found a dosage sensitivity of the X-Y chromosome pair resulting in increased expression of genes that decrease X/Y chromosomal dosage (Raznahan et al., 2018). The most interesting finding was that X-linked genes were found to regulate co-expression of networks of autosomal genes that are SCD-sensitive and, in addition to these findings, suggest that the autosomal genes and their corresponding networks are crucial for cellular functions. This highlights the potential of SCD to affect the occurrence of disease.

The most common aneuploidy in AD is XO mosaicism (Spremo-Potparevic et al., 2004; Spremo-Potparevic et al., 2008; Yurov et al., 2014; Spremo-Potparevic et al., 2015). In respect to SCD and the XO status, Raznahan et al. have demonstrated up-regulation of the protein networks, noncoding RNA metabolism, suppression of the cell cycle, changes in regulation of DNA/chromatin organization, glycolysis, and response to stress (Raznahan et al., 2018). Changes in these collective networks through XO and supernumerary XXY, and XXYY syndromes may enhance the risk of AD (Raznahan et al., 2018; Graham et al., 2019).

There is a small but constant number of neuronal cells that express a different number of chromosomes, such as aneuploidy (Iourov et al., 2006; Yurov et al., 2007; Iourov et al., 2008; Iourov et al., 2009; Yurov et al., 2014), but also copy number variation on chromosome 21, which is crucial in AD (Cai et al., 2014), DNA content variation (Madrigal et al., 2007; Westra et al., 2010), and LINE elements (Evrony et al., 2012).

Mosaic aneuploidy in the brain revealed that not only was chromosome 21 affected in AD, but also that the X chromosome was found to be supernumerary and presumed to be affected through a mechanism that involves altered cohesion/cohesin dynamics (Spremo-Potparevic et al., 2004; Spremo-Potparevic et al., 2008; Bajic et al., 2009; Zivković et al., 2010; Zivkovic et al., 2013; Yurov et al., 2014; Bajic et al., 2015b; Spremo-Potparevic et al., 2015; Yurov et al., 2019).

Yurov et al. (2014) suggested that chromosome 21 might not be the only chromosome to influence changes in genome stability of a neuron, which leads to a cascade of processes that result in neuronal loss. The finding that affected brains show a two-fold increase in X chromosome aneuploidy in the hippocampus and cerebrum - areas of the brain most affected by AD - suggesting that altered sex chromosome dosage plays a role in the large scale genomic variation in neuronal cells in AD compared to controls. These results have been recently corroborated by the finding that the sex chromosomes were distinct from autosomes in the dorsolateral prefrontal cortex and that X chromosome aneuploidy was associated with a faster rate of cognitive decline which is a hallmark of AD (Graham et al., 2019). Therefore, X chromosome aneuploidy may contribute to aging, but also to processes leading to pathological changes in brains affected by AD.

Previously we proposed the "post-mitotic state-maintained protein hypothesis" where we distinguished aneuploization in the brain as constitutional aneuploidy with non-pathological diversification of the neurons (Bajic et al., 2015b). These aneuploidogenic processes are balanced with cohesin and cohesion-related proteins. Alteration of this balance develops as a link between neuronal development and chromosomal instability, intracellular diversity and human brain diseases including AD (Hong and Reiss, 2014). Looking closely at the overall somatic mosaicism found in the brain, we, together with others, suggest that micro aneuploidy or segmental aneuploidy is a more proper measure of changes in gene dosage leading to AD (Dierssen et al., 2009). These processes are heavily realized when looking at SCD effects on gene expression in humans (Raznahan et al., 2018).

An additional complexity of genome mosaicism in the brain relates to findings concerning DNA and gene copy number variations. Regional variations of DNA content has been identified with higher DNA content found in the frontal cortex and cerebellum compared with other brain regions (Westra et al., 2010). Copy number variations may be considered as an independent genetic factor not related to other genomic changes, suggesting its plays a role in neurodevelopmental disorders in patients with sex chromosome aneuploidies (Haack et al., 2013; Le Gall et al., 2017). It has been reported that 11% of neurons in the brain cortex exhibit a DNA content that is above the diploid level (Fischer et al., 2012), and similar findings have also been reported in the AD brain (Ueberham and Arendt, 2005; Arendt et al., 2010; Yurov, 2017; Barrio-Alonso et al., 2018). These somatic gene variations in neurons appear to be generated by chromosome segregation defects. Some of these cells are expelled by apoptosis, but several cells are introduced as a pool of variability of the neuronal genome. These cell populations are thus vulnerable in the sense that they are more prone to genome instability and thus may contribute to age-related mental disorders, such as AD. Gómez-Ramos et al. presented distinct X-chromosome single nucleotide variants from some sporadic AD samples (Gómez-Ramos et al., 2015). In samples from LOAD patients, a higher number of single nucleotide variants in genes present at the X chromosome were identified using exome sequencing compared to age-matched controls. Two genes that were not previously described as risk factors, UBE2NL and ATXN3L, were found to have variants important for the ubiquitin pathway in LOAD (Gómez-Ramos et al., 2015).

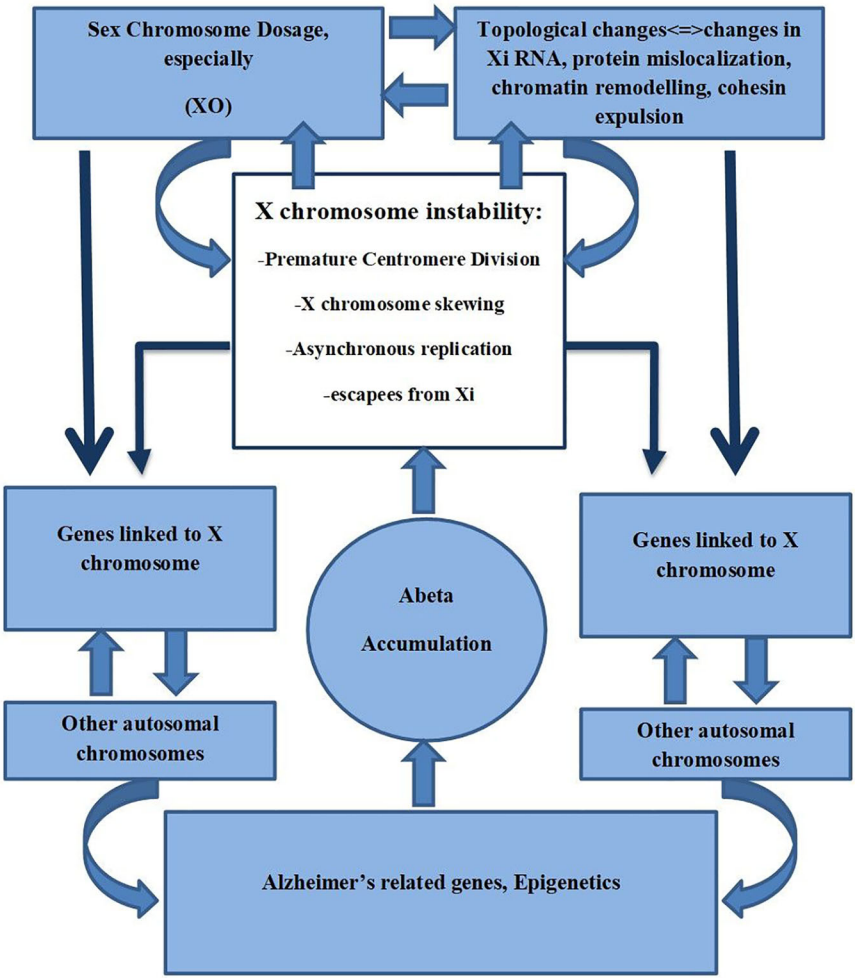

FIGURE 2 | X chromosome instability, Sex Chromosome Dosage, Topological changes of Chromosomes, and its possible role in AD.

Maintenance of the interphase state in neurons is an active process. The 3D organization of the genome is correlated to gene expression in the interphase. In the 3D domain, chromosomes occupy preferential positions by self-organizing into topologically-associated domains, which may change due to the cell lineage or stage of the organism (Laskowski et al., 2019). There is a possible exchange between the inactive and active chromosome in gene regulatory information. Cohesin is indirectly associated with the Xi 3D position in the genome. Minajigil et al. reported that a reaction between Xist and cohesins results in the repulsion of the latter from Xi, thus changing its 3D shape (Minajigi et al., 2015). The Xi is much more complex, and it also represents a reservoir of genes that could replace mutated genes from the active X chromosome. At present, this untapped potential known as the X interactome requires further investigation (Minajigi et al., 2015). Progress in understanding the Xist interactome requires more understanding of how it is used and how epigenetically-regulated long ncRNAs potentially influence disease. By utilizing a specific technique named iDRiP, some 200 proteins in the Xist interactome were identified (Minajigi et al., 2015). Most of the proteins are from several categories, such as cohesins, condesins, topoisomerases, RNA helicase, histone modifiers, methyltransferases, nuclear matrix proteins, and nucleoskeletal factors. Cohesin may play a more important role in the complex relationship between the Xi and active X chromosome (Minajigi et al., 2015). Even though these processes are an important mechanism of diversity, alterations may lead to an increased structural and topological variation of the genome in the brain, enhancing the susceptibility of affected neurons to genome instability that may lead to AD (Bajic et al., 2015b; Graham et al., 2019; Yurov et al., 2019).

A number of publications have reported mislocalization of some critical proteins responsible for chromatin organization and epigenetic modifications in brain diseases including AD (Gill et al., 2007; Lu et al., 2014; Luperchio et al., 2014; Quinodoz and Guttman, 2014; Guo et al., 2015; Mastroeni et al., 2015; Pombo and Dillon, 2015; Sen et al., 2015; Winick-Ng and Rylett, 2018).

All these data suggest, that in AD chromatin, organizers are deregulated and chromatin topology is changed in a manner that alters gene expression leading to synaptic dysfunction, a major

pathological change in AD, and consequently neurodegeneration (Gill et al., 2007; Lu et al., 2014; Luperchio et al., 2014; Quinodoz and Guttman, 2014; Guo et al., 2015; Mastroeni et al., 2015; Pombo and Dillon, 2015; Sen et al., 2015; Winick-Ng and Rylett, 2018).

Xist RNA can act as a scaffold for proteins required to maintain the inactive state of neurons. It has been shown that it can act as a repulsion mechanism that expels architectural factors such as cohesins in order to avoid unwanted chromatin conformation that could increase unfavorable transcription (Raznahan et al., 2018). Minajigi et al. suggest that Xi RNA plays an important role in the organization of how chromosomes are regulated into chromosome territories and that Xi inactivation is fundamentally important in these processes (Minajigi et al., 2015). It could be suggested that X chromosome instability found in AD may result in changes in the Xi pattern, Xi escapees, SCD, and consequently changes in the topological organization, thus altering chromatin organization that may affect already other genes related to AD (**Figure 2**).

The cohesin-associated protein, shugoshin-1, seems to be fundamental in repressing the accumulation of amyloid-β and Tau phosphorylation in shugoshin-1 gene (Sgo1) haploinsufficient mice (Rao et al., 2018).

SUMMARY

Conflicting results from studies of the PCDH11X gene in AD could be explained by cohort size, ethnicity, and environmental factors *per se* but also by the influence of X chromosome epigenetics. Thus, GWAS of sex chromosomes should take into account any alterations of the epigenetic processes in the X chromosome (Schurz et al., 2019).

The findings that chromosome X expresses all of the somatic genomic neuronal variability properties and can *de novo* express several epigenetic mechanisms suggest that the X chromosome instability phenotype may be viewed as an important risk factor in AD pathogenesis.

AUTHOR CONTRIBUTIONS

All authors (VPB, ME, LZ, AS, SZ, VBB, TG, EI, and BS-P) contributed to the design and writing of the manuscript. VPB and SZ designed the figures.

FUNDING

This work is part of the collaboration between the Laboratory of Radiobiology and Molecular Genetics, Vinca Institute of Nuclear Sciences, University of Belgrade, Belgrade, Serbia and King Abdullah University of Science and Technology (KAUST), Computational Bioscience Research Center (CBRC), Thuwal, Saudi Arabia. This work has been supported by grants No. 173033 (EI) and No. 173034 (VPB) from the Ministry of Education, Science and Technological Development, Republic of Serbia and by the KAUST grant OSR#4129 (to EI and VBB), which also supported SZ and VPB. VBB has been supported by the KAUST Base Research Fund (BAS/1/1606-01-01), while VBB and ME have been supported by KAUST Office of Sponsored Research (OSR) grant no. FCC/1/1976-17-01. TG has been supported by the King Abdullah University of Science and Technology (KAUST) Base Research Fund (BAS/1/1059-01-01).

REFERENCES

Ahn, J., and Lee, J. (2008). X chromosome: X inactivation. *Nature Education* 1, 24.

Aladjem, M. I., and Fu, H. (2014). A new light on DNA replication from the inactive X chromosome. *Bioessays* 36, 591–597. doi: 10.1002/bies.201400021

Amiel, A., Avivi, L., Gaber, E., and Fejgin, M. D. (1998). Asynchronous replication of allelic loci in down syndrome. *Eur. J. Hum. Genet.* 6, 359–364. doi: 10.1038/sj.ejhg.5200199

Amiel, A., Reish, O., Gaber, E., Kedar, I., Diukman, R., and Fejgin, M. (2000). Replication asynchrony increases in women at risk for aneuploid offspring. *Chromosome Res.* 8, 141–150. doi: 10.1023/a:1009246603868

Amir, R. E., Van den Veyver, I. B., Wan, M., Tran, C. Q., Francke, U., and Zoghbi, H. Y. (1999). Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nat. Genet.* 23, 185–188. doi: 10.1038/13810

Amos-Landgraf, J. M., Cottle, A., Plenge, R. M., Friez, M., Schwartz, C. E., Longshore, J., et al. (2006). X chromosome-inactivation patterns of 1,005 phenotypically unaffected females. *Am. J. Hum. Genet.* 79, 493–499. doi: 10.1086/507565

Anderson, C. L., and Brown, C. J. (2005). Epigenetic predisposition to expression of TIMP1 from the human inactive X chromosome. *BMC Genet.* 6, 48. doi: 10.1186/1471-2156-6-48

Arendt, T., Bruckner, M. K., Mosch, B., and Losche, A. (2010). Selective cell death of hyperploid neurons in Alzheimer's disease. *Am. J. Pathol.* 177, 15–20. doi: 10.2353/ajpath.2010.090955

Ausio, J., Martinez de Paz, A., and Esteller, M. (2014). MeCP2: the long trip from a chromatin protein to neurological disorders. *Trends Mol. Med.* 20, 487–498. doi: 10.1016/j.molmed.2014.03.004

Bajic, V. P., Spremo-Potparevic, B., Zivkovic, L., Djelic, N., and Smith, M. A. (2008). Is the time dimension of the cell cycle re-entry in AD regulated by centromere cohesion dynamics? *Biosci. Hypotheses* 1, 156–161. doi: 10.1016/j.bihy.2008.03.006

Bajic, V. P., Spremo-Potparevic, B., Zivkovic, L., Bonda, D. J., Siedlak, S. L., Casadesus, G., et al. (2009). The X-chromosome instability phenotype in Alzheimer's disease: a clinical sign of accelerating aging? *Med. Hypotheses* 73, 917–920. doi: 10.1016/j.mehy.2009.06.046

Bajic, V., Mandusic, V., Stefanova, E., Bozovic, A., Davidovic, R., Zivkovic, L., et al. (2015a). Skewed X-chromosome inactivation in women affected by Alzheimer's disease. *J. Alzheimers Dis.* 43, 1251–1259. doi: 10.3233/JAD-141674

Bajic, V., Spremo-Potparevic, B., Zivkovic, L., Isenovic, E. R., and Arendt, T. (2015b). Cohesion and the aneuploid phenotype in Alzheimer's disease: A tale of genome instability. *Neurosci. Biobehav. Rev.* 55, 365–374. doi: 10.1016/j.neubiorev.2015.05.010

Bajic, V., Stanojevic, B., Zivkovic, L., Cabarkapa, A., Perry, G., Arendt, T., et al. (2015c). Cyclin dependent kinase 11, neuroinflammation and alzheimer's disease: a review. *J. Clin. Cell Immunol.* 6, 305. doi: 10.4172/2155-9899.1000305

Balaton, B. P., and Brown, C. J. (2016). Escape Artists of the X Chromosome. *Trends Genet.* 32, 348–359. doi: 10.1016/j.tig.2016.03.007

Barrio-Alonso, E., Hernández-Vivanco, A., Walton, C. C., Perea, G., and Frade, J. M. (2018). Cell cycle reentry triggers hyperploidization and synaptic

dysfunction followed by delayed cell death in differentiated cortical neurons. *Sci. Rep.* 8, 14316–14316. doi: 10.1038/s41598-018-32708-4

Beecham, G. W., Naj, A. C., Gilbert, J. R., Haines, J. L., Buxbaum, J. D., and Pericak-Vance, M. A. (2010). PCDH11X variation is not associated with late-onset Alzheimer disease susceptibility. *Psychiatr. Genet.* 20, 321–324. doi: 10.1097/YPG.0b013e32833b635d

Bekris, L. M., Yu, C. E., Bird, T. D., and Tsuang, D. W. (2010). Genetics of Alzheimer disease. *J. Geriatr. Psychiatry Neurol.* 23, 213–227. doi: 10.1177/0891988710383571

Bian, H., Zhou, Y., Zhou, D., Zhang, Y., Shang, D., and Qi, J. (2019). The latest progress on miR-374 and its functional implications in physiological and pathological processes. *J. Cell Mol. Med.* 23, 3063–3076. doi: 10.1111/jcmm.14219

Boggs, B. A., and Chinault, A. C. (1994). Analysis of replication timing properties of human X-chromosomal loci by fluorescence *in situ* hybridization. *Proc. Natl. Acad. Sci. U.S.A.* 91, 6083–6087. doi: 10.1073/pnas.91.13.6083

Brown, C. J., and Greally, J. M. (2003). A stain upon the silence: genes escaping X inactivation. *Trends Genet.* 19, 432–438. doi: 10.1016/S0168-9525(03)00177-X

Cai, X., Evrony, G. D., Lehmann, H. S., Elhosary, P. C., Mehta, B. K., Poduri, A., et al. (2014). Single-cell, genome-wide sequencing identifies clonal somatic copy-number variation in the human brain. *Cell Rep.* 8, 1280–1289. doi: 10.1016/j.celrep.2014.07.043

Carrasquillo, M. M., Zou, F., Pankratz, V. S., Wilcox, S. L., Ma, L., Walker, L. P., et al. (2009). Genetic variation in PCDH11X is associated with susceptibility to late-onset Alzheimer's disease. *Nat. Genet.* 41, 192–198. doi: 10.1038/ng.305

Carrel, L., and Willard, H. F. (2005). X-inactivation profile reveals extensive variability in X-linked gene expression in females. *Nature* 434, 400–404. doi: 10.1038/nature03479

Castellani, R. J., Lee, H. G., Zhu, X., Perry, G., and Smith, M. A. (2008). Alzheimer disease pathology as a host response. *J. Neuropathol. Exp. Neurol.* 67, 523–531. doi: 10.1097/NEN.0b013e318177eaf4

Cazzola, M., May, A., Bergamaschi, G., Cerani, P., Rosti, V., and Bishop, D. F. (2000). Familial-skewed X-chromosome inactivation as a predisposing factor for late-onset X-linked sideroblastic anemia in carrier females. *Blood* 96, 4363–4365. doi: 10.1182/blood.V96.13.4363

Chow, J. C., and Brown, C. J. (2003). Forming facultative heterochromatin: silencing of an X chromosome in mammalian females. *Cell Mol. Life Sci.* 60, 2586–2603. doi: 10.1007/s00018-003-3121-9

Chung, S. J., Lee, J. H., Kim, S. Y., You, S., Kim, M. J., Lee, J. Y., et al. (2013). Association of GWAS top hits with late-onset Alzheimer disease in Korean population. *Alzheimer Dis. Assoc. Disord.* 27, 250–257. doi: 10.1097/WAD.0b013e31826d7281

Cui, W., Wang, S., Wang, Z., Wang, Z., Sun, C., and Zhang, Y. (2017). Inhibition of PTEN attenuates endoplasmic reticulum stress and apoptosis *via* activation of PI3K/AKT pathway in Alzheimer's disease. *Neurochem. Res.* 42, 3052–3060. doi: 10.1007/s11064-017-2338-1

DeMarco, B., Stefanovic, S., Williams, A., Moss, K. R., Anderson, B. R., Bassell, G. J., et al. (2019). FMRP - G-quadruplex mRNA - miR-125a interactions: implications for miR-125a mediated translation regulation of PSD-95 mRNA. *PLoS One* 14, e0217275. doi: 10.1371/journal.pone.0217275

Dierssen, M., Herault, Y., and Estivill, X. (2009). Aneuploidy: from a physiological mechanism of variance to Down syndrome. *Physiol. Rev.* 89, 887–920. doi: 10.1152/physrev.00032.2007

Disteche, C. M., Filippova, G. N., and Tsuchiya, K. D. (2002). Escape from X inactivation. *Cytogenet. Genome Res.* 99, 36–43. doi: 10.1159/000071572

Eikelenboom, P., Rozemuller, A. J., Hoozemans, J. J., Veerhuis, R., and van Gool, W. A. (2000). Neuroinflammation and Alzheimer disease: clinical and therapeutic implications. *Alzheimer Dis. Assoc. Disord.* 14 Suppl 1, S54–S61. doi: 10.1097/00002093-200000001-00009

Escamilla-Del-Arenal, M., da Rocha, S. T., and Heard, E. (2011). Evolutionary diversity and developmental regulation of X-chromosome inactivation. *Hum. Genet.* 130, 307–327. doi: 10.1007/s00439-011-1029-2

Evrony, G. D., Cai, X., Lee, E., Hills, L. B., Elhosary, P. C., Lehmann, H. S., et al. (2012). Single-neuron sequencing analysis of L1 retrotransposition and somatic mutation in the human brain. *Cell* 151, 483–496. doi: 10.1016/j.cell.2012.09.035

Fang, E. F., Hou, Y., Palikaras, K., Adriaanse, B. A., Kerr, J. S., Yang, B., et al. (2019). Mitophagy inhibits amyloid-β and tau pathology and reverses cognitive deficits in models of Alzheimer's disease. *Nat. Neurosci.* 22, 401–412. doi: 10.1038/s41593-018-0332-9

Fernandez, E., Collins, M. O., Uren, R. T., Kopanitsa, M. V., Komiyama, N. H., Croning, M. D., et al. (2009). Targeted tandem affinity purification of PSD-95 recovers core postsynaptic complexes and schizophrenia susceptibility proteins. *Mol. Syst. Biol.* 5, 269. doi: 10.1038/msb.2009.27

Fidani, L., Rooke, K., Chartier-Harlin, M. C., Hughes, D., Tanzi, R., Mullan, M., et al. (1992). Screening for mutations in the open reading frame and promoter of the beta-amyloid precursor protein gene in familial Alzheimer's disease: identification of a further family with APP717 Val–> Ile. *Hum. Mol. Genet.* 1, 165–168. doi: 10.1093/hmg/1.3.165

Fischer, H. G., Morawski, M., Bruckner, M. K., Mittag, A., Tarnok, A., and Arendt, T. (2012). Changes in neuronal DNA content variation in the human brain during aging. *Aging Cell* 11, 628–633. doi: 10.1111/j.1474-9726.2012.00826.x

Forloni, G., Demicheli, F., Giorgi, S., Bendotti, C., and Angeretti, N. (1992). Expression of amyloid precursor protein mRNAs in endothelial, neuronal and glial cells: modulation by interleukin-1. *Brain Res. Mol. Brain Res.* 16, 128–134. doi: 10.1016/0169-328x(92)90202-m

Gómez-Ramos, A., Podlesniy, P., Soriano, E., and Avila, J. (2015). Distinct X-chromosome SNVs from some sporadic AD samples. *Sci. Rep.* 5, 18012. doi: 10.1038/srep18012

Gatz, M., Reynolds, C. A., Fratiglioni, L., Johansson, B., Mortimer, J. A., Berg, S., et al. (2006). Role of genes and environments for explaining Alzheimer disease. *Arch. Gen. Psychiatry* 63, 168–174. doi: 10.1001/archpsyc.63.2.168

Geller, L. N., and Potter, H. (1999). Chromosome missegregation and trisomy 21 mosaicism in Alzheimer's disease. *Neurobiol. Dis.* 6, 167–179. doi: 10.1006/nbdi.1999.0236

Germain, D. P. (2006). "General aspects of X-linked diseases," in *Fabry Disease: Perspectives from 5 Years of FOS*. Eds. A. Mehta, M. Beck and G. Sunder-Plassmann (Oxford: Oxford PharmaGenesis).

Gill, S. K., Ishak, M., Dobransky, T., Haroutunian, V., Davis, K. L., and Rylett, R. J. (2007). 82-kDa choline acetyltransferase is in nuclei of cholinergic neurons in human CNS and altered in aging and Alzheimer disease. *Neurobiol. Aging* 28, 1028–1040. doi: 10.1016/j.neurobiolaging.2006.05.011

Giri, M., Shah, A., Upreti, B., and Rai, J. C. (2017). Unraveling the genes implicated in Alzheimer's disease. *Biomed. Rep.* 7, 105–114. doi: 10.3892/br.2017.927

Goate, A., and Hardy, J. (2012). Twenty years of Alzheimer's disease-causing mutations. *J. Neurochem.* 120 Suppl 1, 3–8. doi: 10.1111/j.1471-4159.2011.07575.x

Goate, A. M., Hardy, J. A., Owen, M. J., Haynes, A., James, L., Farrall, M., et al. (1990). Genetics of Alzheimer's disease. *Adv. Neurol.* 51, 197–198.

Goncalves, T. F., Piergiorge, R. M., Dos Santos, J. M., Gusmao, J., Pimentel, M. M. G., and Santos-Reboucas, C. B. (2019). Network profiling of brain-expressed X-chromosomal microRNA genes implicates shared key microRNAs in intellectual disability. *J. Mol. Neurosci.* 67, 295–304. doi: 10.1007/s12031-018-1235-7

Graham, E. J., Vermeulen, M., Vardarajan, B., Bennett, D., De Jager, P., Pearse, R. V., 2nd, et al. (2019). Somatic mosaicism of sex chromosomes in the blood and brain. *Brain. Res.* 1721, 146345. doi: 10.1016/j.brainres.2019.146345

Grant, S. G. N., Marshall, M. C., Page, K.-L., Cumiskey, M. A., and Armstrong, J. D. (2005). Synapse proteomics of multiprotein complexes: en route from genes to nervous system diseases. *Hum. Mol. Genet.* 14, R225–R234. doi: 10.1093/hmg/ddi330

Gribnau, J., Luikenhuis, S., Hochedlinger, K., Monkhorst, K., and Jaenisch, R. (2005). X chromosome choice occurs independently of asynchronous replication timing. *J. Cell Biol.* 168, 365–373. doi: 10.1083/jcb.200405117

Guo, Y., Xu, Q., Canzio, D., Shou, J., Li, J., Gorkin, D. U., et al. (2015). CRISPR Inversion of CTCF sites alters genome topology and enhancer/promoter function. *Cell* 162, 900–910. doi: 10.1016/j.cell.2015.07.038

Guven, G., Erginel-Unaltuna, N., Samanci, B., Gulec, C., Hanagasi, H., and Bilgic, B. (2019). A patient with early-onset Alzheimer's disease with a novel PSEN1 p.Leu424Pro mutation. *Neurobiol. Aging.* 84, 238.e1–238.e4. doi: 10.1016/j.neurobiolaging.2019.05.014

Haack, T. B., Hogarth, P., Gregory, A., Prokisch, H., and Hayflick, S. J. (2013). BPAN: the only X-linked dominant NBIA disorder. *Int. Rev. Neurobiol.* 110, 85–90. doi: 10.1016/b978-0-12-410502-7.00005-3

Hardy, J., Goate, A., Owen, M., and Rossor, M. (1989). Presenile dementia associated with mosaic trisomy 21 in a patient with a Down syndrome child. *Lancet* 2, 743. doi: 10.1016/s0140-6736(89)90805-2

Hauss-Wegrzyniak, B., Dobrzanski, P., Stoehr, J. D., and Wenk, G. L. (1998). Chronic neuroinflammation in rats reproduces components of the neurobiology of Alzheimer's disease. *Brain Res.* 780, 294–303. doi: 10.1016/s0006-8993(97)01215-8

Hong, D. S., and Reiss, A. L. (2014). Cognitive and neurological aspects of sex chromosome aneuploidies. *Lancet Neurol.* 13, 306–318. doi: 10.1016/s1474-4422(13)70302-8

Iourov, I. Y., Vorsanova, S. G., and Yurov, Y. B. (2006). Chromosomal variation in mammalian neuronal cells: known facts and attractive hypotheses. *Int. Rev. Cytol.* 249, 143–191. doi: 10.1016/s0074-7696(06)49003-3

Iourov, I. Y., Vorsanova, S. G., and Yurov, Y. B. (2008). Molecular cytogenetics and cytogenomics of brain diseases. *Curr. Genomics* 9, 452–465. doi: 10.2174/138920208786241216

Iourov, I. Y., Vorsanova, S. G., Liehr, T., and Yurov, Y. B. (2009). Aneuploidy in the normal, Alzheimer's disease and ataxia-telangiectasia brain: differential expression and pathological meaning. *Neurobiol. Dis.* 34, 212–220. doi: 10.1016/j.nbd.2009.01.003

Iourov, I. Y., Vorsanova, S. G., and Yurov, Y. B. (2010). Somatic genome variations in health and disease. *Curr. Genomics* 11, 387–396. doi: 10.2174/138920210793176065

Khalifa, O., Pers, Y. M., Ferreira, R., Senechal, A., Jorgensen, C., Apparailly, F., et al. (2016). X-Linked miRNAs associated with gender differences in rheumatoid arthritis. *Int. J. Mol. Sci.* 17, 1852–1863. doi: 10.3390/ijms17111852

Krstic, D., Madhusudan, A., Doehner, J., Vogel, P., Notter, T., Imhof, C., et al. (2012). Systemic immune challenges trigger and drive Alzheimer-like neuropathology in mice. *J. Neuroinflammation* 9, 151. doi: 10.1186/1742-2094-9-151

Laskowski, A. I., Neems, D. S., Laster, K., Strojny-Okyere, C., Rice, E. L., Konieczna, I. M., et al. (2019). Varying levels of X chromosome coalescence in female somatic cells alters the balance of X-linked dosage compensation and is implicated in female-dominant systemic lupus erythematosus. *Sci. Rep.* 9, 8011. doi: 10.1038/s41598-019-44229-9

Laumonnier, F., Cuthbert, P. C., and Grant, S. G. (2007). The role of neuronal complexes in human X-linked brain diseases. *Am. J. Hum. Genet.* 80, 205–220. doi: 10.1086/511441

Le Gall, J., Nizon, M., Pichon, O., Andrieux, J., Audebert-Bellanger, S., Baron, S., et al. (2017). Sex chromosome aneuploidies and copy-number variants: a further explanation for neurodevelopmental prognosis variability? *Eur. J. Hum. Genet.* 25, 930–934. doi: 10.1038/ejhg.2017.93

Lopes, A. M., Ross, N., Close, J., Dagnall, A., Amorim, A., and Crow, T. J. (2006). Inactivation status of PCDH11X: sexual dimorphisms in gene expression levels in brain. *Hum. Genet.* 119, 267–275. doi: 10.1007/s00439-006-0134-0

Lu, T., Aron, L., Zullo, J., Pan, Y., Kim, H., Chen, Y., et al. (2014). REST and stress resistance in ageing and Alzheimer's disease. *Nature* 507, 448–454. doi: 10.1038/nature13163

Luperchio, T. R., Wong, X., and Reddy, K. L. (2014). Genome regulation at the peripheral zone: lamina associated domains in development and disease. *Curr. Opin. Genet. Dev.* 25, 50–61. doi: 10.1016/j.gde.2013.11.021

Lyon, M. F. (1961). Gene action in the X-chromosome of the mouse (Mus musculus L.). *Nature* 190, 372–373. doi: 10.1038/190372a0

Madrigal, I., Rodríguez-Revenga, L., Armengol, L., González, E., Rodriguez, B., Badenas, C., et al. (2007). X-chromosome tiling path array detection of copy number variants in patients with chromosome X-linked mental retardation. *BMC Genomics* 8, 443. doi: 10.1186/1471-2164-8-443

Manzine, P. R., Pelucchi, S., Horst, M. A., Vale, F. A. C., Pavarini, S. C. I., Audano, M., et al. (2018). microRNA 221 targets ADAM10 mRNA and is downregulated in Alzheimer's disease. *J. Alzheimers Dis.* 61, 113–123. doi: 10.3233/jad-170592

Maphis, N. M., Jiang, S., Binder, J., Wright, C., Gopalan, B., Lamb, B. T., et al. (2017). Whole Genome Expression Analysis in a Mouse Model of Tauopathy Identifies MECP2 as a Possible Regulator of Tau Pathology. *Front. Mol. Neurosci.* 10, 69. doi: 10.3389/fnmol.2017.00069

Mastroeni, D., Delvaux, E., Nolz, J., Tan, Y., Grover, A., Oddo, S., et al. (2015). Aberrant intracellular localization of H3k4me3 demonstrates an early epigenetic phenomenon in Alzheimer's disease. *Neurobiol. Aging* 36, 3121–3129. doi: 10.1016/j.neurobiolaging.2015.08.017

McFarland, K. N., Huizenga, M. N., Darnell, S. B., Sangrey, G. R., Berezovska, O., Cha, J. H., et al. (2014). MeCP2: a novel Huntingtin interactor. *Hum. Mol. Genet.* 23, 1036–1044. doi: 10.1093/hmg/ddt499

McGeer, P. L., and McGeer, E. G. (2002). Local neuroinflammation and the progression of Alzheimer's disease. *J. Neurovirol.* 8, 529–538. doi: 10.1080/13550280290100969

Miar, A., Alvarez, V., Corao, A. I., Alonso, B., Diaz, M., Menendez, M., et al. (2011). Lack of association between protocadherin 11-X/Y (PCDH11X and PCDH11Y) polymorphisms and late onset Alzheimer's disease. *Brain Res.* 1383, 252–256. doi: 10.1016/j.brainres.2011.01.054

Migliore, L., Boni, G., Bernardini, R., Trippi, F., Colognato, R., Fontana, I., et al. (2006). Susceptibility to chromosome malsegregation in lymphocytes of women who had a down syndrome child in young age. *Neurobiol. Aging* 27, 710–716. doi: 10.1016/j.neurobiolaging.2005.03.025

Migliore, L., Migheli, F., and Coppede, F. (2009). Susceptibility to aneuploidy in young mothers of down syndrome children. *Sci. World J.* 9, 1052–1060. doi: 10.1100/tsw.2009.122

Minajigi, A., Froberg, J. E., Wei, C., Sunwoo, H., Kesner, B., Colognori, D., et al. (2015). A comprehensive Xist interactome reveals cohesin repulsion and an RNA-directed chromosome conformation. *Science* 349, aab2276. doi: 10.1126/science.aab2276

Muddashetty, R. S., Nalavadi, V. C., Gross, C., Yao, X., Xing, L., Laur, O., et al. (2011). Reversible inhibition of PSD-95 mRNA translation by miR-125a, FMRP phosphorylation, and mGluR signaling. *Mol. Cell* 42, 673–688. doi: 10.1016/j.molcel.2011.05.006

Mugford, J. W., Starmer, J., Williams, R. L., Jr., Calabrese, J. M., Mieczkowski, P., Yee, D., et al. (2014). Evidence for local regulatory control of escape from imprinted X chromosome inactivation. *Genetics* 197, 715–723. doi: 10.1534/genetics.114.162800

Ng, K., Pullirsch, D., Leeb, M., and Wutz, A. (2007). Xist and the order of silencing. *EMBO Rep.* 8, 34–39. doi: 10.1038/sj.embor.7400871

Nguyen, D. K., and Disteche, C. M. (2006). Dosage compensation of the active X chromosome in mammals. *Nat. Genet.* 38, 47–53. doi: 10.1038/ng1705

Orstavik, K. H. (2009). X chromosome inactivation in clinical practice. *Hum. Genet.* 126, 363–373. doi: 10.1007/s00439-009-0670-5

Peeters, S. B., Korecki, A. J., Baldry, S. E. L., Yang, C., Tosefsky, K., Balaton, B. P., et al. (2019). How do genes that escape from X-chromosome inactivation contribute to Turner syndrome? *Am. J. Med. Genet.* 181, 28–35. doi: 10.1002/ajmg.c.31672

Pessia, E., Makino, T., Bailly-Bechet, M., McLysaght, A., and Marais, G. A. B. (2012). Mammalian X chromosome inactivation evolved as a dosage-compensation mechanism for dosage-sensitive genes on the X chromosome. *Proc. Natl. Acad. Sci.* 109, 5346. doi: 10.1073/pnas.1116763109

Petersen, M. B., Karadima, G., Samaritaki, M., Avramopoulos, D., Vassilopoulos, D., and Mikkelsen, M. (2000). Association between presenilin-1 polymorphism and maternal meiosis II errors in down syndrome. *Am. J. Med. Genet.* 93, 366–372. doi: 10.1002/1096-8628(20000828)93:5<366::AID-AJMG5>3.0.CO;2-G

Pike, C. J. (2017). Sex and the development of Alzheimer's disease. *J. Neurosci. Res.* 95, 671–680. doi: 10.1002/jnr.23827

Pocklington, A. J., Cumiskey, M., Armstrong, J. D., and Grant, S. G. (2006). The proteomes of neurotransmitter receptor complexes form modular networks with distributed functionality underlying plasticity and behaviour. *Mol. Syst. Biol.* 2, 2006 0023. doi: 10.1038/msb4100041

Pombo, A., and Dillon, N. (2015). Three-dimensional genome architecture: players and mechanisms. *Nat. Rev. Mol. Cell Biol.* 16, 245. doi: 10.1038/nrm3965

Potter, H., and Geller, L. N. (1996). Alzheimer's disease, Down's syndrome, and chromosome segregation. *Lancet* 348. doi: 10.1016/s0140-6736(05)64399-1

Potter, H. (2016). Beyond trisomy 21: phenotypic variability in people with down syndrome explained by further chromosome mis-segregation and mosaic aneuploidy. *J. Down Syndr. Chromosom. Abnorm.* 2, 1–6. doi: 10.4172/2472-1115.1000109

Quinodoz, S., and Guttman, M. (2014). Long noncoding RNAs: an emerging link between gene regulation and nuclear organization. *Trends Cell Biol.* 24, 651–663. doi: 10.1016/j.tcb.2014.08.009

Rao, C. V., Farooqui, M., Zhang, Y., Asch, A. S., and Yamada, H. Y. (2018). Spontaneous development of Alzheimer's disease-associated brain pathology in

a Shugoshin-1 mouse cohesinopathy model. *Aging Cell* 17, e12797. doi: 10.1111/acel.12797

Raznahan, A., Parikshak, N. N., Chandran, V., Blumenthal, J. D., Clasen, L. S., Alexander-Bloch, A. F., et al. (2018). Sex-chromosome dosage effects on gene expression in humans. *Proc. Natl. Acad. Sci. U.S.A.* 115, 7398–7403. doi: 10.1073/pnas.1802889115

Regen, F., Hellmann-Regen, J., Costantini, E., and Reale, M. (2017). Neuroinflammation and Alzheimer's disease: implications for microglial activation. *Curr. Alzheimer Res.* 14, 1140–1148. doi: 10.2174/1567205014666170203141717

Renault, N. K., Dyack, S., Dobson, M. J., Costa, T., Lam, W. L., and Greer, W. L. (2007). Heritable skewed X-chromosome inactivation leads to haemophilia a expression in heterozygous females. *Eur. J. Hum. Genet.* 15, 628–637. doi: 10.1038/sj.ejhg.5201799

Renault, N. K., Pritchett, S. M., Howell, R. E., Greer, W. L., Sapienza, C., Orstavik, K. H., et al. (2013). Human X-chromosome inactivation pattern distributions fit a model of genetically influenced choice better than models of completely random choice. *Eur. J. Hum. Genet.* 21, 1396–1402. doi: 10.1038/ejhg.2013.84

Schellenberg, G. D., Bird, T. D., Wijsman, E. M., Orr, H. T., Anderson, L., Nemens, E., et al. (1992). Genetic linkage evidence for a familial Alzheimer's disease locus on chromosome 14. *Science* 258, 668–671. doi: 10.1126/science.1411576

Schupf, N., Kapell, D., Lee, J. H., Ottman, R., and Mayeux, R. (1994). Increased risk of Alzheimer's disease in mothers of adults with down's syndrome. *Lancet* 344, 353–356. doi: 10.1016/s0140-6736(94)91398-6

Schupf, N., Kapell, D., Nightingale, B., Lee, J. H., Mohlenhoff, J., Bewley, S., et al. (2001). Specificity of the fivefold increase in AD in mothers of adults with down syndrome. *Neurology* 57, 979–984. doi: 10.1212/wnl.57.6.979

Schurz, H., Salie, M., Tromp, G., Hoal, E. G., Kinnear, C. J., and Möller, M. (2019). The X chromosome and sex-specific effects in infectious disease susceptibility. *Hum. Genomics* 13, 2. doi: 10.1186/s40246-018-0185-z

Selkoe, D. J. (2001). Alzheimer's disease results from the cerebral accumulation and cytotoxicity of amyloid beta-protein. *J. Alzheimers Dis.* 3, 75–80. doi: 10.3233/jad-2001-3111

Sen, A., Nelson, T. J., and Alkon, D. L. (2015). ApoE4 and abeta oligomers reduce BDNF expression *via* HDAC nuclear translocation. *J. Neurosci.* 35, 7538–7551. doi: 10.1523/jneurosci.0260-15.2015

Seripa, D., Panza, F., Franceschi, M., D'Onofrio, G., Solfrizzi, V., Dallapiccola, B., et al. (2009). Non-apolipoprotein E and apolipoprotein E genetics of sporadic Alzheimer's disease. *Ageing Res. Rev.* 8, 214–236. doi: 10.1016/j.arr.2008.12.003

Shvetsova, E., Sofronova, A., Monajemi, R., Gagalova, K., Draisma, H. H. M., White, S. J., et al. (2019). Skewed X-inactivation is common in the general female population. *Eur. J. Hum. Genet.* 27, 455–465. doi: 10.1038/s41431-018-0291-3

Skuse, D. H. (2005). X-linked genes and mental functioning. *Hum. Mol. Genet.* 14 Spec No 1, R27–R32. doi: 10.1093/hmg/ddi112

Slooter, A. J., de Knijff, P., Hofman, A., Cruts, M., Breteler, M. M., Van Broeckhoven, C., et al. (1998). Serum apolipoprotein E level is not increased in Alzheimer's disease: the rotterdam study. *Neurosci. Lett.* 248, 21–24. doi: 10.1016/s0304-3940(98)00339-5

Smith, M. A. (1998). Alzheimer disease. *Int. Rev. Neurobiol.* 42, 1–54. doi: 10.1016/s0074-7742(08)60607-8

Splinter, E., de Wit, E., Nora, E. P., Klous, P., van de Werken, H. J., Zhu, Y., et al. (2011). The inactive X chromosome adopts a unique three-dimensional conformation that is dependent on Xist RNA. *Genes Dev.* 25, 1371–1383. doi: 10.1101/gad.633311

Spremo-Potparevic, B., Zivkovic, L., Djelic, N., and Bajic, V. (2004). Analysis of premature centromere division (PCD) of the X chromosome in Alzheimer patients through the cell cycle. *Exp. Gerontol.* 39, 849–854. doi: 10.1016/j.exger.2004.01.012

Spremo-Potparevic, B., Zivkovic, L., Djelic, N., Plecas-Solarovic, B., Smith, M. A., and Bajic, V. (2008). Premature centromere division of the X chromosome in neurons in Alzheimer's disease. *J. Neurochem.* 106, 2218–2223. doi: 10.1111/j.1471-4159.2008.05555.x

Spremo-Potparevic, B., Bajic, V., Perry, G., and Zivkovic, L. (2015). Alterations of the X chromosome in lymphocytes of alzheimer's disease patients. *Curr. Alzheimer Res.* 12, 990–996. doi: 10.2174/1567205012666151027124154

Sudbrak, R., Wieczorek, G., Nuber, U. A., Mann, W., Kirchner, R., Erdogan, F., et al. (2001). X chromosome-specific cDNA arrays: identification of genes that escape from X-inactivation and other applications. *Hum. Mol. Genet.* 10, 77–83. doi: 10.1093/hmg/10.1.77

Sun, B. K., and Lee, J. T. (2006). "X-Chromosome Inactivation," in *Encyclopedic Reference of Genomics and Proteomics in Molecular Medicine* (Berlin, Heidelberg: Springer Berlin Heidelberg), 2013–2019.

Tam, G. W., Redon, R., Carter, N. P., and Grant, S. G. (2009). The role of DNA copy number variation in schizophrenia. *Biol. Psychiatry* 66, 1005–1012. doi: 10.1016/j.biopsych.2009.07.027

Ueberham, U., and Arendt, T. (2005). The expression of cell cycle proteins in neurons and its relevance for Alzheimer's disease. *Curr. Drug Targets CNS Neurol. Disord.* 4, 293–306. doi: 10.2174/1568007054038175

Vest, R. S., and Pike, C. J. (2013). Gender, sex steroid hormones, and Alzheimer's disease. *Horm. Behav.* 63, 301–307. doi: 10.1016/j.yhbeh.2012.04.006

Vorsanova, S. G., Yurov, Y. B., Kolotii, A. D., and Soloviev, I. V. (2001). FISH analysis of replication and transcription of chromosome X loci: new approach for genetic analysis of Rett syndrome. *Brain Dev.* 23 Suppl 1, S191–S195. doi: 10.1016/s0387-7604(01)00364-3

Wani, A., Gupta, M., Ahmad, M., Shah, A. M., Ahsan, A. U., Qazi, P. H., et al. (2019). Alborixin clears amyloid-β by inducing autophagy through PTEN-mediated inhibition of the AKT pathway. *Autophagy* 15, 1810–1828. doi: 10.1080/15548627.2019.1596476

Westra, J. W., Rivera, R. R., Bushman, D. M., Yung, Y. C., Peterson, S. E., Barral, S., et al. (2010). Neuronal DNA content variation (DCV) with regional and individual differences in the human brain. *J. Comp. Neurol.* 518, 3981–4000. doi: 10.1002/cne.22436

Wilson, N. D., Ross, L. J. N., Close, J., Mott, R., Crow, T. J., and Volpi, E. V. (2007). Replication profile of PCDH11X and PCDH11Y, a gene pair located in the non-pseudoautosomal homologous region Xq21.3/Yp11.2. *Chromosome Res. Int. J. Mol. Supramol. Evol. Aspects Chromosome Biol.* 15, 485–498. doi: 10.1007/s10577-007-1153-y

Winick-Ng, W., and Rylett, R. J. (2018). Into the fourth dimension: dysregulation of genome architecture in aging and Alzheimer's disease. *Front. Mol. Neurosci.* 11, 60. doi: 10.3389/fnmol.2018.00060

Wiseman, F. K., Pulford, L. J., Barkus, C., Liao, F., Portelius, E., Webb, R., et al. (2018). Trisomy of human chromosome 21 enhances amyloid-beta deposition independently of an extra copy of APP. *Brain* 141, 2457–2474. doi: 10.1093/brain/awy159

Wu, Z. C., Yu, J. T., Wang, N. D., Yu, N. N., Zhang, Q., Chen, W., et al. (2010). Lack of association between PCDH11X genetic variation and late-onset Alzheimer's disease in a Han Chinese population. *Brain Res.* 1357, 152–156. doi: 10.1016/j.brainres.2010.08.008

Yuan, D., XiuJuan, W., Yan, Z., JunQin, L., Fang, X., Shirong, Y., et al. (2015). Use of X-chromosome inactivation pattern to analyze the clonality of 14 female cases of kaposi sarcoma. *Med. Sci. Monitor Basic Res.* 21, 116–122. doi: 10.12659/MSMBR.894089

Yudowski, G. A., Olsen, O., Adesnik, H., Marek, K. W., and Bredt, D. S. (2013). Acute inactivation of PSD-95 destabilizes AMPA receptors at hippocampal synapses. *PLoS One* 8, e53965. doi: 10.1371/journal.pone.0053965

Yurov, Y. B., Iourov, I. Y., Vorsanova, S. G., Liehr, T., Kolotii, A. D., Kutsev, S. I., et al. (2007). Aneuploidy and confined chromosomal mosaicism in the developing human brain. *PLoS One* 2, e558. doi: 10.1371/journal.pone.0000558

Yurov, Y. B., Vorsanova, S. G., Liehr, T., Kolotii, A. D., and Iourov, I. Y. (2014). X chromosome aneuploidy in the Alzheimer's disease brain. *Mol. Cytogenet.* 7, 20. doi: 10.1186/1755-8166-7-20

Yurov, Y. B., Vorsanova, S. G., and Iourov, I. Y. (2019). Chromosome instability in the neurodegenerating brain. *Front. Genet.* 10, 892. doi: 10.3389/fgene.2019.00892

Yurov, Y. B. (2017). "FISH-Based Assays for Detecting Genomic (Chromosomal) Mosaicism in Human Brain Cells," in *Genomic Mosaicism in Neurons and Other Cell Types*. Eds. J. M. Frade and F. Gage (London, UK: Springer Nature).

Zivković, L., Spremo-Potparević, B., Plecas-Solarović, B., Djelić, N., Ocić, G., Smiljković, P., et al. (2010). Premature centromere division of metaphase chromosomes in peripheral blood lymphocytes of Alzheimer's disease patients: relation to gender and age. *J. Gerontol. Ser. A Biol. Sci. Med. Sci.* 65, 1269–1274. doi: 10.1093/gerona/glq148

Zivkovic, L., Spremo-Potparevic, B., Siedlak, S. L., Perry, G., Plecas-Solarovic, B., Milicevic, Z., et al. (2013). DNA damage in Alzheimer disease lymphocytes and its relation to premature centromere division. *Neurodegener. Dis.* 12, 156–163. doi: 10.1159/000346114

Zubenko, G. S., Stiffler, J. S., Hughes, H. B., Hurtt, M. R., and Kaplan, B. B. (1998). Initial results of a genome survey for novel Alzheimer's disease risk genes: association with a locus on the X chromosome. *Am. J. Med. Genet.* 81, 196–205. doi: 10.1002/(sici)1096-8628(19980207)81:1<98::aid-ajmg17>3.0.co;2-r

Zubenko, G. S., Hughes, H. B., and Stiffler, J. S. (1999). Clinical and neurobiological correlates of DXS1047 genotype in Alzheimer's disease. *Biol. Psychiatry* 46, 173–181. doi: 10.1016/s0006-3223(99)00035-9

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Genetics

PERSPECTIVE
published: 27 March 2020
doi: 10.3389/fgene.2020.00287

Pathological Aspects of Neuronal Hyperploidization in Alzheimer's Disease Evidenced by Computer Simulation

*Estíbaliz Barrio-Alonso[1], Bérénice Fontana[1], Manuel Valero[2] and José M. Frade[1]**

[1] Department of Molecular, Cellular, and Developmental Neurobiology, Cajal Institute, CSIC, Madrid, Spain, [2] Neuroscience Institute, New York University, New York, NY, United States

When subjected to stress, terminally differentiated neurons are susceptible to reactivate the cell cycle and become hyperploid. This process is well documented in Alzheimer's disease (AD), where it may participate in the etiology of the disease. However, despite its potential importance, the effects of neuronal hyperploidy (NH) on brain function and its relationship with AD remains obscure. An important step forward in our understanding of the pathological effect of NH has been the development of transgenic mice with neuronal expression of oncogenes as model systems of AD. The analysis of these mice has demonstrated that forced cell cycle reentry in neurons results in most hallmarks of AD, including neurofibrillary tangles, Aβ peptide deposits, gliosis, cognitive loss, and neuronal death. Nevertheless, in contrast to the pathological situation, where a relatively small proportion of neurons become hyperploid, neuronal cell cycle reentry in these mice is generalized. We have recently developed an *in vitro* system in which cell cycle is induced in a reduced proportion of differentiated neurons, mimicking the *in vivo* situation. This manipulation reveals that NH correlates with synaptic dysfunction and morphological changes in the affected neurons, and that membrane depolarization facilitates the survival of hyperploid neurons. This suggests that the integration of synaptically silent, hyperploid neurons in electrically active neural networks allows their survival while perturbing the normal functioning of the network itself, a hypothesis that we have tested *in silico*. In this perspective, we will discuss on these aspects trying to convince the reader that NH represents a relevant process in AD.

Keywords: neuronal cell cycle reentry, SV40 large T antigen, neuron hypertrophy, neurite retraction, synaptic dysfunction, neural network modeling, synaptic firing rate, oscillatory patterns

OPEN ACCESS

Edited by:
Svetlana G. Vorsanova, Veltischev Research and Clinical Institute for Pediatrics of the Pirogov Russian National Research Medical University, Russia

Reviewed by:
Biljana Spremo Potparevic, University of Belgrade, Serbia
Xiao Dong, Albert Einstein College of Medicine, United States

***Correspondence:**
José M. Frade
frade@cajal.csic.es

Specialty section:
This article was submitted to Human Genomics, a section of the journal Frontiers in Genetics

Received: *14 September 2019*
Accepted: *09 March 2020*
Published: *27 March 2020*

Citation:
Barrio-Alonso E, Fontana B, Valero M and Frade JM (2020) Pathological Aspects of Neuronal Hyperploidization in Alzheimer's Disease Evidenced by Computer Simulation. Front. Genet. 11:287. doi: 10.3389/fgene.2020.00287

INTRODUCTION

As the nervous system ages, it undergoes functional alterations that diminish its performance and, as these changes increase, brain homeostasis becomes compromised resulting in neurodegenerative conditions including Alzheimer's disease (AD). A plethora of neuroanatomical and functional alterations in the nervous system accompanying the process of aging and leading to AD-associated neurodegeneration has so far been described. Among these changes, DNA level variation and aneuploidy (Cuccaro et al., 2017; Shepherd et al., 2018) as well as cell cycle reentry in neurons

leading to increased DNA levels [i.e., neuronal hyperploidy (NH)] (Frade and Ovejero-Benito, 2015) are known to precede and recapitulate the classical neuropathological signs of AD (Yang et al., 2001; Arendt et al., 2010; Frade and López-Sánchez, 2017). In some cases, NH results in full DNA duplication (i.e., neuronal tetraploidy). This latter condition affects around 2–3% of neurons in AD (Mosch et al., 2007; López-Sánchez et al., 2017), a proportion that increases to around 8% when specific neuronal subtypes are evaluated (López-Sánchez et al., 2017). Once chromosomes have been fully replicated in these neurons, the latter may remain as 2N cells with 4C DNA content, as observed in G2, or as 4N cells, if they undergo premature chromosomal separation (Spremo-Potparević et al., 2008; Bajić et al., 2009). Moreover, above 30% of neurons become hyperploid in the middle stages of AD (Arendt et al., 2010). Since the fate of hyperploid neurons is delayed cell death (Yang et al., 2001; Arendt et al., 2010) these numbers likely underestimate the actual proportion of AD-affected neurons undergoing NH.

The involvement of NH in the etiology of AD has been directly proven by forcing neuronal cell cycle reentry in transgenic mice expressing oncogenes such as SV40 T large antigen (TAg) (Park et al., 2007) or c-Myc (Lee et al., 2009) under the control of the neuron-specific CAMKII promoter. This manipulation results in neuropathological hallmarks of AD, including tau protein hyperphosphorylation and neurofibrillary tangles, extracellular deposits of Aβ peptide, neuronal cell death, gliosis, and cognitive deficits. McShea et al. (2007) have also shown that c-Myc/Ras-induced cell cycle reentry in primary cortical neurons triggers tau phosphorylation that result in conformational changes similar to that seen in AD.

NH might also lead to other alterations compromising normal brain function, thus participating in several aspects of the etiology of AD (Frade and López-Sánchez, 2010). In this regard, the increase of ploidy levels is associated with nuclear and cellular hypertrophy (Orr-Weaver, 2015), and several lines of evidence suggest that these changes can be detected in AD (Frade and López-Sánchez, 2010). In this article, we will explore the morphological changes observed in cortical neurons forced to reactivate the cell cycle in response to TAg expression, a procedure recently used by our laboratory to induce hyperploidy in a small proportion of cortical neurons, thus mimicking the *in vivo* situation (Barrio-Alonso et al., 2018). By using this model, we demonstrated that neuronal hyperploidization correlates with synaptic dysfunction (Barrio-Alonso et al., 2018), a known alteration occurring at early stages of AD (Scheff et al., 2006), and that membrane depolarization with high K^+ facilitates the survival of hyperploid neurons without reversing synaptic dysfunction in these cells (Barrio-Alonso et al., 2018). This suggests that AD-associated hyperploid neurons can be sustained *in vivo* if integrated in active neuronal circuits while remaining synaptically silent (i.e., without capacity to fire action potentials). Given that each cortical neuron can establish connections with hundreds other neuronal cells (Markram et al., 2015), it is conceivable that a relatively small fraction of silent hyperploid neurons could disrupt the normal functioning of the circuits in which they are inserted (Lusch et al., 2018). If this were the case, NH might contribute to cognitive impairment at early stages of AD due to synaptic dysfunction, while triggering neuronal cell death at later stages (Yang et al., 2001; Arendt et al., 2010; Barrio-Alonso et al., 2018). On this basis, we have also explored whether the presence of hyperploid neurons could disrupt the normal functioning of the circuits in which they are inserted. As a first approximation to the problem, this analysis has been performed *in silico*, by simulating the outcome of a neural network that contains different proportions of silent, hyperploid neurons.

RESULTS AND DISCUSSION

Morphological Changes Induced by Cell Cycle Reentry in Cortical Neurons

We exploited the capacity of TAg to induce cell cycle reentry in cortical neurons (Barrio-Alonso et al., 2018) to explore the effects of hyperploidy on neuronal morphology. This analysis demonstrated that, 48 h after lipofection, the soma of cortical neurons expressing TAg was significantly larger than that of control neurons expressing either LacZ or TAg K1 (K1), an E107K TAg variant that lacks binding capacity to the pRb family members and therefore cannot induce cell cycle reactivation (Zalvide and DeCaprio, 1995; **Figures 1A,B**). Since no significant differences were observed between cell somas of neurons lipofected with LacZ or K1 (**Figure 1B**), we concluded that the effect of TAg on cell soma size is specific on its capacity to induce cell cycle reentry/hyperploidy (Barrio-Alonso et al., 2018).

We also found that, at this time point, TAg-induced cell cycle reentry specifically triggered a significant reduction in the length of the dendritic tree of cortical neurons, as compared with neurons lipofected with LacZ or K1 (**Figure 1C**; **Supplementary Figure 1**). The observed length reduction correlated with the degree of dendrite branching, evaluated through Sholl analysis, which was significantly reduced in TAg-lipofected neurons (**Figure 1D**). Again, this effect derives from the capacity of TAg to induce cell cycle re-entry as the K1 construct did not modified the branching profile (**Figure 1D**). The reduction of dendritic length and branching observed in TAg-lipofected neurons is consistent with studies carried out with mouse models of AD and postmortem material from AD patients in which a reduction in the total dendritic area was evident (Moolman et al., 2004). Interestingly, this reduction of dendritic length and branching mimics what has been observed in mitotic neurons induced to reactivate the cell cycle with a truncated form of cyclin E/Cdk2 (Walton et al., 2019).

Simulation of Neural Networks Containing Silent Hyperploid Neurons

The morphological changes observed in neurons that reactivate the cell cycle, along with the capacity of cell cycle reentry to trigger synaptic dysfunction in neurons (Barrio-Alonso et al., 2018) suggest that NH participates in the etiology of Alzheimer by affecting neurons' capacity to fire action potentials and therefore altering the neuronal circuits in which hyperploid neurons are inserted.

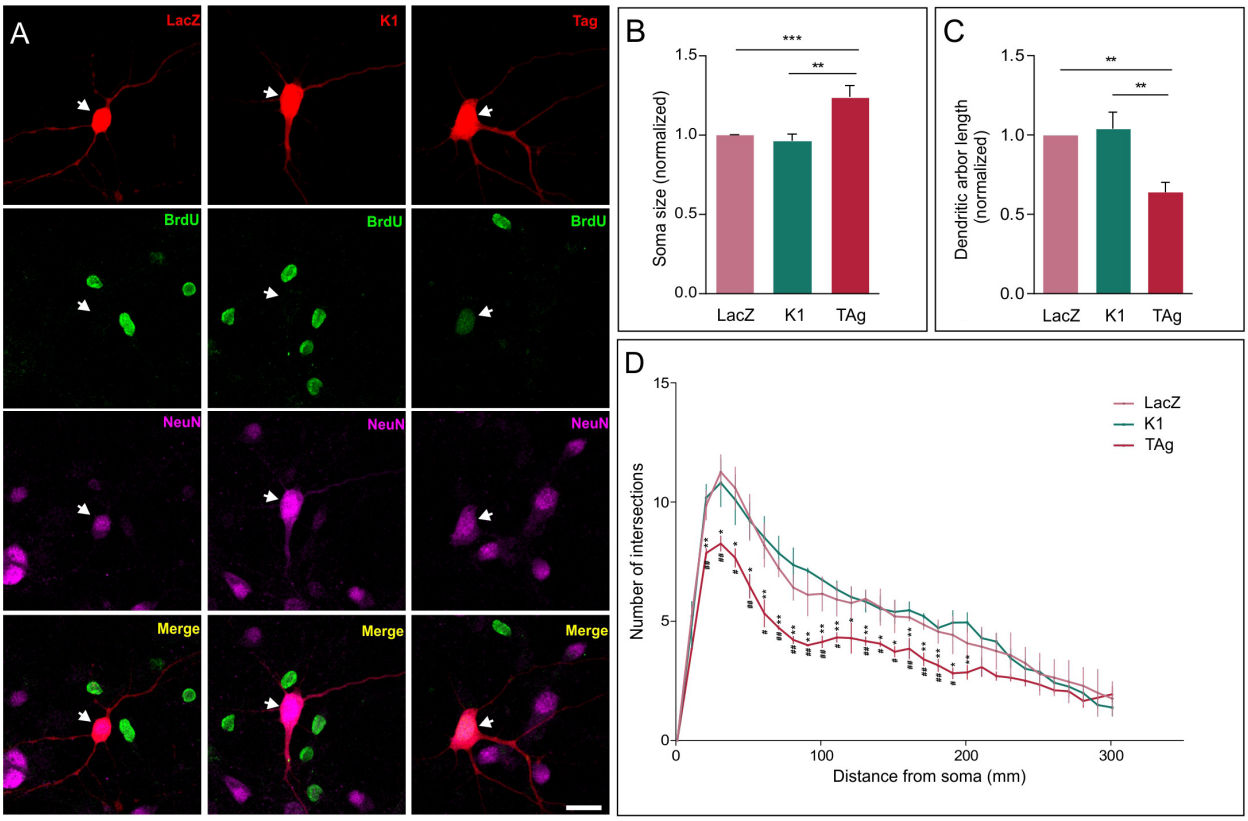

FIGURE 1 | Effect of cell cycle reentry on the neuronal soma and dendritic tree. **(A)** TAg-lipofected neurons incorporating BrdU exhibit a larger soma than both control (LacZ) and K1-lipofected neurons (representative examples are shown). Arrow: neuron lipofected for 48 h with RFP and the indicated protein. Scale bar: 20 μm. **(B)** Soma size quantification in control (LacZ), K1-lipofected, and TAg-lipofected neurons, normalized to control ($n = 3$). **(C)** Normalized length of the dendritic tree in neurons lipofected for 48 h with LacZ, K1, or TAg ($n = 3$). **(D)** Average number of intersections for each distance from the soma in lipofected neurons with LacZ, Tag, or K1 ($n = 3$). Hashtags: statistical significance between LacZ and TAg; asterisks: statistical significance between K1 and TAg. $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$, $^{\#\#}p < 0.01$, $^{\#}p < 0.05$ (One-way ANOVA followed by Tukey's *post hoc* test in **(B,C)**; two-way ANOVA, followed by Tukey's *post hoc* test in **(D)**.

As a first attempt to verify this hypothesis, we employed an *in silico* approach. We simulated the impact that the presence of silent hyperploid neurons may have on the functional connectivity of a neural network through an "integrate-and-fire" model (Knight, 1972) constituted by 4,000 neurons. In this model, the membrane potential of each neuron at any simulation time point ($dt = 1$ ms) depends on two factors: (i) an exponential function, governed by a time constant, which pushes the voltage to its resting membrane potential; and (ii) the amount of excitation and synaptic inhibition received from the partner cells (**Figure 2A**). The local field potential (LFP) was estimated as the average of all transmembrane currents.

Different fractions of silent neurons (0.01, 0.1, 0.2, 0.3, 0.5, or 0.8) were included in (i) the whole neuronal population, (ii) the excitatory neuronal subpopulation, or (iii) the leading neuronal subpopulation, defined as those neurons with higher firing frequency and thus representing functional circuit hubs. In contrast, the control condition included no silent neurons. In a different cohort of simulations, we tested the effect of silencing the inhibitory neuron subpopulation (interneurons). As expected, this manipulation leads to an epileptic-like network with hyper-synchronous activity patterns (**Supplementary Figure 2**), maybe resembling the comorbidity between AD and epilepsy seen in some transgenic mouse models (Palop et al., 2007).

Our results indicate that an increase in the percentage of synaptic silencing in any of the analyzed neuronal types has defined effects on the firing rate of all simulated subpopulations present in the neural network (see **Figure 2B** for the subpopulation of excitatory neurons, and **Supplementary Figures 3A–C** for other neuronal subpopulations). In all cases, a significant inverse correlation between the fraction of silenced leading or excitatory neurons, and the firing frequency was observed (**Figure 2C** and **Supplementary Figures 3D–F**). In contrast, the disruption in the firing ability over the entire network led to an increase in the firing frequency of the excitatory subpopulation and the whole population (**Figure 2C** and **Supplementary Figure 3F**), likely due to silencing of a portion of inhibitory neurons. Therefore, we concluded that the presence of hyperploid neurons with synaptic deficits (silent neurons) affects the firing frequency of the neural network in which they are integrated. This effect is cell-type dependent and

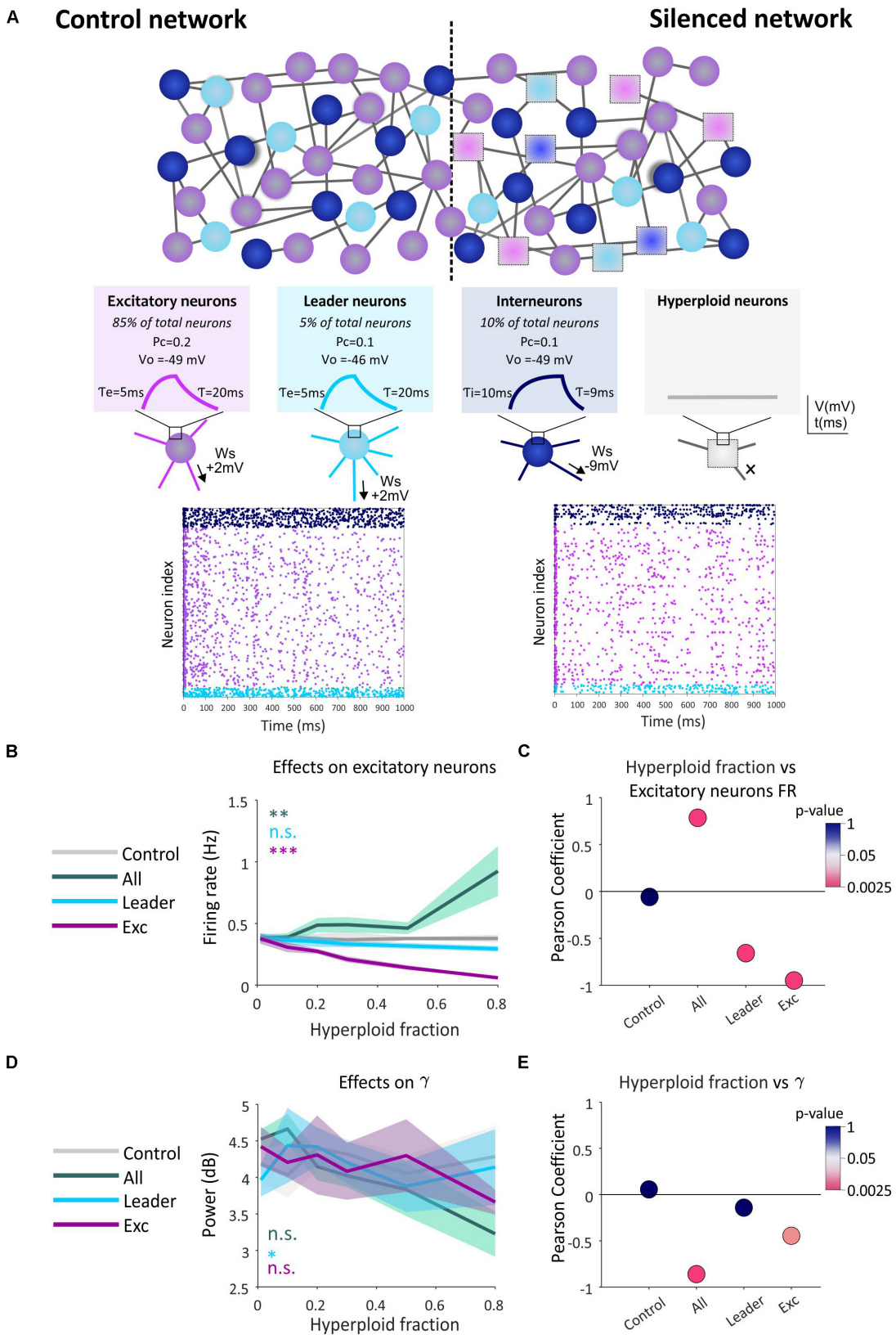

FIGURE 2 | Continued

FIGURE 2 | Effect of the presence of hyperploid neurons on the firing frequency of the excitatory subpopulation and in gamma type oscillations.
(A) Hyperploidization neural network model. An example of the control network (without the presence of hyperploid neurons) is shown in comparison with a network with hyperploid neurons distributed randomly throughout the network (Silenced network). Excitatory neurons are shown in purple, interneurons in dark blue, leader neurons in light blue, and silent hyperploid neurons are shown as squares. Parameters defining the different neuronal subpopulations (*Ws*, P_C, V_0, τ, τ_e, and τ_i) are described in the methodological section. In plots, each dot indicates an action potential emitted by each neuron. The decrease in the number of dots is noticeable in the affected network with a 50% of random neuronal silencing compared to the control network. **(B)** The graph shows how the firing frequency of the excitatory neuron subpopulation is affected by the indicated fraction of silent hyperploid neurons in the whole neuron population (green), leader neurons (blue), or excitatory neurons (purple). Each line shows the average trigger frequency of each population when the corresponding type of neurons have been affected. The gray line shows the average frequency of the population when there is no silencing of neurons. 95% confidence intervals for each line are shown in shading. **(C)** Each point represents the Pearson correlation value. The color code shows the *p*-value (in logarithmic scale) of the correlation. **(D)** The graph shows how the activity of gamma rhythm is affected at the indicated fraction of silent hyperploid neurons in the whole neuron population (green), leader neurons (blue), or excitatory neurons (purple). The gray line shows the average power of this population oscillation when there is no silencing of neurons. 95% confidence intervals for each line are shown in shading. **(E)** Each point represents the Pearson correlation value. The color code shows the *p*-value (in logarithmic scale) of the correlation. $^{**}p < 0.01$; $^{***}p < 0.001$ (two-way ANOVA followed by Tukey's *post hoc* test), n.s.: non-significant, in **(B, C)**.

correlated to the fraction of affected neurons, causing an increase or decrease of firing rate depending on the tested cell type.

We also explored the effect of silencing on the oscillatory patterns generated spontaneously by the network. Brain oscillations, similar to those observed in our model, are an emergent property of the system associated with the coordination of the circuit and the transmission of information between its elements (Buzsáki and Draguhn, 2004; Buzsáki et al., 2013). In this regard, our model presents peaks of synaptic activity in the spectral bands gamma, beta, theta and delta (Buzsáki and Draguhn, 2004; **Supplementary Figure 4**). We observed a significant effect of synaptic silencing on gamma activity, this being dependent on both the proportion of silent neurons and the specific population in which they are found (**Figure 2D**). These effects are mainly appreciated as a negative correlation between silencing in total and excitatory neurons and gamma power (**Figure 2E**), reminiscent of the gamma oscillations impairment observed in AD (Mably and Colgin, 2018). In the rest of the spectral bands we found minor or insignificant changes except for the whole neuronal population and the excitatory subpopulation when the beta and theta rhythms, respectively, were analyzed (**Supplementary Figure 5**).

CONCLUSION

Cumulative evidence indicates that NH may participate in the classical neuropathology observed in AD. In addition, our results suggest that NH can also lead to alterations in neuronal circuit functioning due to the morphological and synaptic changes observed in hyperploid neurons. We believe these alterations, together with any other perturbation underlying the synaptic deficits found in AD (Scheff et al., 2006), could account for the etiology of AD as well. Our simulation study indicates that NH may trigger alterations in the firing frequency of the neural network, an effect that increases as the proportion of hyperploid neurons raises. Therefore, the presence of a high proportion of hyperploid neurons in specific local circuits could lead to major effects in AD. This conclusion should be experimentally tested in the future. In areas such as the entorhinal cortex, where above 30% of neurons become hyperploid in AD patients (Arendt et al., 2010), this condition could have an important impact not only on the firing frequency but also on the oscillations observed in the neural networks (Kitchigina, 2018), which according to our *in silico* model requires a high proportion of silent neurons to be relevant.

MATERIALS AND METHODS

Neuronal Soma and Dendritic Tree Analysis

Primary cortical cultures, co-lipofection with red fluorescent protein (RFP) and plasmids expressing LacZ, TAg, or K1, and immunocytochemistry were performed as described by Barrio-Alonso et al. (2018). RFP-positive neurons were randomly chosen. Image analysis was performed with ImageJ (National Institutes of Health). Sholl analysis (Binley et al., 2014) was carried out with the Sholl analysis module (Fiji) using digital tracings generated with the NeuronJ plugin (Fiji) from confocal projection images of neurons co-lipofected with RFP (n = 30 per condition). Analysis parameters were: starting radius = 1 μm, ending radius = NaN, and radius step size = 10 μm. Linear Sholl plots were generated, representing the average number of intersections with radii in each condition. Total length of neurites was also evaluated. At least 25 lipofected neurons/culture from 3 independent cultures were analyzed for each experimental condition.

Neural Network Simulation

"Integrate-and-fire" simulation (Knight, 1972) of neural networks containing hyperploid neurons was implemented using the Python-based Brian 2 simulator (Goodman and Brette, 2008; Stimberg et al., 2017). In the simulation model (**Figure 2A**), each neuron has a membrane potential (V) governed by the following differential equation:

$$\tau \frac{dV}{dt} = -(V - V_0)$$

where τ is the membrane time constant, which parameterizes the time it takes for the neuron to reach its resting membrane potential (V_0). In turn, V can be disturbed by depolarizing and hyperpolarizing synaptic currents. The excitatory and inhibitory

conductances (ge and gi) of such synaptic currents follow the following differential equations:

$$\frac{dge}{dt} = -\frac{ge}{\tau e}$$

$$\frac{dgi}{dt} = -\frac{gi}{\tau i}$$

The value of ge and gi depends, respectively, on the number of active excitatory and inhibitory synapses according to their synaptic weight (*Ws*). The *Ws* of each excitatory synapse is +2 mV while each inhibitory synapse has a *Ws* of −9 mV. An action potential (AP) from synapse *s* at time *t* induces the following change in V in neuron *j*: $V \rightarrow V + Ws$. Thus, if *s* is an excitatory synapse, the neuron *j* is depolarized by +2 mV at a rate defined by τe. An inhibitory synapse induces a change of −9 mV governed by τi.

Each neuron initializes V in a randomly chosen value between −50 and −60 mV and begins to receive excitatory and inhibitory synapses with their corresponding weights. If V reaches −50 mV, the trigger AP threshold (Vt), the neuron fires, which induces a synaptic current in those neurons to which it is connected. At this point, there is a refractory period of 5 ms in which the neuron cannot fire again.

The network is composed of 4,000 neurons of three types, with different proportion, electrical properties, and probabilities of connection to other neurons (*Pc*): (i) *excitatory neurons* (85% of all neurons), τe = 20 ms, V_0 = −49 mV, and *Pc* = 0.1; (ii) *leading neurons* (5% of all neurons), a subtype of excitatory neurons whose membrane potential is 3 mV more depolarized, and therefore have a higher trigger frequency that simulate neurons constituting relevant hubs of the circuit; and (iii) *interneurons* (10% of all neurons), τi = 9 ms (i.e., high trigger frequency), V_0 = −49 mV, and *Pc* = 0.2. Five repetitions of each condition were simulated.

Parameters established in the model reasonably mimic the physiological situation (Markram et al., 2015) according to the firing frequency: most excitatory neurons fires at a frequency lower than 1 Hz, leading neurons fire at a frequency of 1–5 Hz, and the inhibitory neurons (or interneurons) show a firing frequency of around 2 Hz.

For the simulation of hyperploid neurons (i.e., silent neurons unable to fire APs), a "damage" parameter (*dam*) was incorporated in the differential equation of the model. *dam* has a value equal to −30 mV, which hyperpolarizes the membrane potential, setting its *V* away from *Vt*. The percentages of hyperploid neurons in this study were 1, 10, 20, 30, 50, and 80%. This provides a complete picture of the effects of this variable on the outcome of the neural network.

As an internal control, we found that the simulated neural network has a strong synaptic dependence on its activity patterns. If all synaptic connections are removed, the network presents a synchronous firing rate, reflecting the exponential component that repolarizes V (**Supplementary Figure 6**).

Oscillatory patterns and their power were estimated by the sum of all membrane potentials from the network (**Supplementary Figure 4**).

Statistical Analysis

Statistical analyses were performed with ANOVA, followed by the Tukey's *post hoc* test. Pearson correlation test was also applied in simulation experiments.

DATA AVAILABILITY STATEMENT

The complete code for the simulation is in the **Supplementary Material**. Data were analyzed with MATLAB (MathWorks). The data and the MATLAB scripts used to generate the analyses and representations can be downloaded from http://dx.doi.org/10.20350/digitalCSIC/10541.

ETHICS STATEMENT

This study was carried out in accordance with the principles of the Basel Declaration and the EU guidelines for the care and use of laboratory animals. All of the procedures for handling and sacrificing animals were approved by the CSIC Ethics Committee.

AUTHOR CONTRIBUTIONS

JF conceived the study and wrote the manuscript. EB-A and MV designed the neural network, performed the simulations and analyzed the data. EB-A also contributed to the manuscript writing. BF carried out the neuronal soma and dendritic tree analysis, EB-A and BF prepared figures. All authors approved the final manuscript.

FUNDING

This study was supported by Ministerio de Economía y Competitividad grant SAF2015-68488-R, and Ministerio de Ciencia, Innovación y Universidades grant RTI2018-095030-B-I00.

ACKNOWLEDGMENTS

We acknowledge support of the publication fee by the CSIC Open Access Publication Support Initiative through its Unit of Information Resources for Research (URICI).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2020.00287/full#supplementary-material

REFERENCES

Arendt, T., Brückner, M. K., Mosch, B., and Lösche, A. (2010). Selective cell death of hyperploid neurons in Alzheimer's disease. *Am. J. Pathol.* 177, 15–20. doi: 10.2353/ajpath.2010.090955

Bajić, V. P., Spremo-Potparević, B., Zivković, L., Bonda, D. J., Siedlak, S. L., Casadesus, G., et al. (2009). The X-chromosome instability phenotype in Alzheimer's disease: a clinical sign of accelerating aging? *Med. Hypotheses* 73, 917–920. doi: 10.1016/j.mehy.2009.06.046

Barrio-Alonso, E., Hernández-Vivanco, A., Walton, C. C., Perea, G., and Frade, J. M. (2018). Cell cycle reentry triggers hyperploidization and synaptic dysfunction followed by delayed cell death in differentiated cortical neurons. *Sci. Rep.* 8:14316. doi: 10.1038/s41598-018-32708-4

Binley, K. E., Ng, W. S., Tribble, J. R., Song, B., and Morgan, J. E. (2014). Sholl analysis: a quantitative comparison of semi-automated methods. *J. Neurosci. Methods* 225, 65–70. doi: 10.1016/j.jneumeth.2014.01.017

Buzsáki, G., and Draguhn, A. (2004). Neuronal oscillations in cortical networks. *Science* 304, 1926–1929. doi: 10.1126/science.1099745

Buzsáki, G., Logothetis, N., and Singer, W. (2013). Scaling brain size, keeping timing: evolutionary preservation of brain rhythms. *Neuron* 80, 751–764. doi: 10.1016/j.neuron.2013.10.002

Cuccaro, D., De Marco, E. V., Cittadella, R., and Cavallaro, S. (2017). Copy number variants in Alzheimer's disease. *J. Alzheimers Dis.* 55, 37–52. doi: 10.3233/JAD-160469

Frade, J. M., and López-Sánchez, N. (2010). A novel hypothesis for Alzheimer disease based on neuronal tetraploidy induced by p75NTR. *Cell Cycle* 9, 1934–1941. doi: 10.4161/cc.9.10.11582

Frade, J. M., and López-Sánchez, N. (2017). Neuronal tetraploidy in Alzheimer and aging. *Aging* 9, 2014–2015. doi: 10.18632/aging.101312

Frade, J. M., and Ovejero-Benito, M. C. (2015). Neuronal cell cycle: the neuron itself and its circumstances. *Cell Cycle* 14, 712–720. doi: 10.1080/15384101.2015.1004937

Goodman, D., and Brette, R. (2008). Brian: a simulator for spiking neural networks in Python. *Front. Neuroinform.* 2:5. doi: 10.3389/neuro.11.005.2008

Kitchigina, V. F. (2018). Alterations of coherent theta and gamma network oscillations as an early biomarker of temporal lobe epilepsy and Alzheimer's disease. *Front. Integr. Neurosci.* 12:36. doi: 10.3389/fnint.2018.00036

Knight, B. W. (1972). Dynamics of encoding in a population of neurons. *J. Gen. Physiol.* 59, 734–766. doi: 10.1085/jgp.59.6.734

Lee, H. G., Casadesus, G., Nunomura, A., Zhu, X., Castellani, R. J., Richardson, S. L., et al. (2009). The neuronal expression of MYC causes a neurodegenerative phenotype in a novel transgenic mouse. *Am. J. Pathol.* 174, 891–897. doi: 10.2353/ajpath.2009.080583

López-Sánchez, N., Fontán-Lozano, Á, Pallé, A., González-Álvarez, V., Rábano, A., Trejo, J. L., et al. (2017). Neuronal tetraploidization in the cerebral cortex correlates with reduced cognition in mice and precedes and recapitulates Alzheimer's-associated neuropathology. *Neurobiol. Aging* 56, 50–66. doi: 10.1016/j.neurobiolaging.2017.04.008

Lusch, B., Weholt, J., Maia, P. D., and Kutz, J. N. (2018). Modeling cognitive deficits following neurodegenerative diseases and traumatic brain injuries with deep convolutional neural networks. *Brain Cogn.* 123, 154–164. doi: 10.1016/j.bandc.2018.02.012

Mably, A. J., and Colgin, L. L. (2018). Gamma oscillations in cognitive disorders. *Curr. Opin. Neurobiol.* 52, 182–187. doi: 10.1016/j.conb.2018.07.009

Markram, H., Muller, E., Ramaswamy, S., Reimann, M. W., Abdellah, M., Sanchez, C. A., et al. (2015). Reconstruction and simulation of neocortical microcircuitry. *Cell* 163, 456–492. doi: 10.1016/j.cell.2015.09.029

McShea, A., Lee, H. G., Petersen, R. B., Casadesus, G., Vincent, I., Linford, N. J., et al. (2007). Neuronal cell cycle re-entry mediates Alzheimer disease-type changes. *Biochim. Biophys. Acta.* 1772, 467–472. doi: 10.1016/j.bbadis.2006.09.010

Moolman, D. L., Vitolo, O. V., Vonsattle, J. P., and Shelanski, M. L. (2004). Dendrite and dendritic spine alterations in Alzheimer models. *J. Neurocytol.* 33, 377–387. doi: 10.1023/B:NEUR.0000044197.83514.64

Mosch, B., Morawski, M., Mittag, A., Lenz, D., Tarnok, A., and Arendt, T. (2007). Aneuploidy and DNA replication in the normal human brain and Alzheimer's disease. *J. Neurosci.* 27, 6859–6867. doi: 10.1523/JNEUROSCI.0379-07.2007

Orr-Weaver, T. L. (2015). When bigger is better: the role of polyploidy in organogenesis. *Trends Genet.* 31, 307–315. doi: 10.1016/j.tig.2015.03.011

Palop, J. J., Chin, J., Roberson, E. D., Wang, J., Thwin, M. T., Bien-Ly, N., et al. (2007). Aberrant excitatory neuronal activity and compensatory remodeling of inhibitory hippocampal circuits in mouse models of Alzheimer's disease. *Neuron* 55, 697–711. doi: 10.1016/j.neuron.2007.07.025

Park, K. H., Hallows, J. L., Chakrabarty, P., Davies, P., and Vincent, I. (2007). Conditional neuronal simian virus 40 T antigen expression induces Alzheimer-like tau and amyloid pathology in mice. *J. Neurosci.* 27, 2969–2978. doi: 10.1523/JNEUROSCI.0186-07.2007

Scheff, S. W., Price, D. A., Schmitt, F. A., and Mufson, E. J. (2006). Hippocampal synaptic loss in early Alzheimer's disease and mild cognitive impairment. *Neurobiol. Aging* 27, 1372–1384. doi: 10.1016/j.neurobiolaging.2005.09.012

Shepherd, C. E., Yang, Y., and Halliday, G. M. (2018). Region- and cell-specific aneuploidy in brain aging and neurodegeneration. *Neuroscience* 374, 326–334. doi: 10.1016/j.neuroscience.2018.01.050

Spremo-Potparević, B., Zivković, L., Djelić, N., Plećas-Solarović, B., Smith, M. A., and Bajić, V. (2008). Premature centromere division of the X chromosome in neurons in Alzheimer's disease. *J. Neurochem.* 106, 2218–2223. doi: 10.1111/j.1471-4159.2008.05555.x

Stimberg, M., Goodman, D. F. M., Brette, R., and De Pittà, M. (2017). Modeling neuron-glia interactions with the brian 2 simulator. *bioRxiv* [preprint] doi: 10.1101/198366

Walton, C. C., Zhang, W., Patiño-Parrado, I., Barrio-Alonso, E., Garrido, J. J., and Frade, J. M. (2019). Primary neurons can enter M-phase. *Sci. Rep.* 9:4594. doi: 10.1038/s41598-019-40462-4

Yang, Y., Geldmacher, D. S., and Herrup, K. (2001). DNA replication precedes neuronal cell death in Alzheimer's disease. *J. Neurosci.* 21, 2661–2668. doi: 10.1523/jneurosci.21-08-02661.2001

Zalvide, J., and DeCaprio, J. A. (1995). Role of pRb-related proteins in simian virus 40 large-T-antigen-mediated transformation. *Mol. Cell Biol.* 15, 5800–5810. doi: 10.1128/MCB.15.10.5800

Conflict of Interest: JF is shareholder (ten percent equity ownership) of Tetraneuron, S.L., a biotech company exploiting his patent on the blockade of neuronal tetraploidy as a therapeutic approach against AD and funding his work on therapies against AD based on neuronal tetraploidy blockade.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Genetics

OPINION
published: 22 April 2020
doi: 10.3389/fgene.2020.00388

Somatic Genomic Mosaicism in Multiple Myeloma

Christine J. Ye[1]*, *Jason Chen*[1], *Guo Liu*[2] and *Henry H. Heng*[2,3]*

[1] The Division of Hematology/Oncology, Department of Internal Medicine, University of Michigan, Ann Arbor, MI, United States, [2] Center for Molecular Medicine and Genomics, Wayne State University School of Medicine, Detroit, MI, United States, [3] Department of Pathology, Wayne State University School of Medicine, Detroit, MI, United States

Keywords: cellular heterogeneity, fuzzy inheritance, genome chaos, genome theory, macro-cellular evolution, two phases of cancer evolution, karyotype coding, system inheritance

OPEN ACCESS

Edited by:
Obul Reddy Bandapalli, Hopp Children's Cancer Center Heidelberg (KiTZ), Germany

Reviewed by:
Ezhilarasi Chendamarai, Washington University in St. Louis, United States
Hamid Mahmoudpour, Johannes Gutenberg University Mainz, Germany

***Correspondence:**
Christine J. Ye
jchrisye@med.umich.edu
Henry H. Heng
hheng@med.wayne.edu

Specialty section:
This article was submitted to Human Genomics, a section of the journal Frontiers in Genetics

Received: *11 February 2020*
Accepted: *27 March 2020*
Published: *22 April 2020*

Citation:
Ye CJ, Chen J, Liu G and Heng HH (2020) Somatic Genomic Mosaicism in Multiple Myeloma. Front. Genet. 11:388. doi: 10.3389/fgene.2020.00388

BACKGROUND

Somatic genomic mosaicism occurs when somatic cells of the body display different genotypes (**Table 1**), it has recently received increased attention because of its implications in disease, including neurodegenerative diseases and Down syndrome (Iourov et al., 2008, 2010, 2019; Biesecker and Spinner, 2013; Hultén et al., 2013; Vijg, 2014; Campbell et al., 2015; Fernández et al., 2016). Genomic mosaicism also contributes to high levels of cellular heterogeneity in pathological conditions, which is a distinguishing feature of cancer (Heng, 2015, 2019). In fact, high heterogeneity in cancer represents an extreme example of genomic mosaicism.

The genomic basis of somatic genomic mosaicism, however, remains to be elucidated. Traditional explanations have focused on defective cellular processes, including imperfect DNA replication and repair, abnormal chromosomal machinery, and a faulty stress response to environmental challenges. As illustrated by the evolutionary mechanism of cancer (Ye et al., 2009), nearly all molecular pathways/mechanisms can contribute to variations in cellular systems. The conventional wisdom is that biosystems are not perfect and that error-generating opportunities exist. Thus, the major goals of molecular medicine have been to detect and fix these errors.

Nevertheless, bioerrors (or imperfect-biosystems) do not explain the high degree of genomic mosaicism revealed by large-scale -omics technologies (Vattathil and Scheet, 2016), and plausible mechanisms are not yet revealed (Heng et al., 2016). These novel mechanisms should address (a) both the positive and negative contributions of cellular heterogeneity in normal and disease conditions and (b) the survival strategy of cancer cells to drastically elevate the level of heterogeneity in crisis conditions. Using multiple myeloma (MM) as an example, these mechanisms will be examined in the context of bio-information, adaptive systems (**Table 1**), and emergent behavior during cancer evolution.

A HIGH DEGREE OF SOMATIC GENOMIC MOSAICISM, A NECESSARY AND SUFFICIENT CONDITION FOR EVOLUTION, IS COMMON IN MM

MM patients display a high level of karyotype heterogeneity. Different patient genotypes can involve poly-aneuploidy, hyperdiploidy, hypodiploidy, chromosomal translocation, chaotic genomes (such as chromothripsis) (**Table 1**), and/or a combination of other gene mutations and chromosomal aberrations (Garcia-Sanz et al., 1995; Avet-Loiseau et al., 2007; Klein et al., 2011; Magrangeas et al., 2011; Keats et al., 2012; Bolli et al., 2014; Lee et al., 2017; Kaur et al., 2018; Smetana et al., 2018; Ashby et al., 2019; Maura et al., 2019).

Four key realizations from the Genome Theory (**Table 1**) can explain why such karyotype heterogeneity is observed in MM patients:

TABLE 1 | Definitions/terminologies.

Genome theory
The Genome Theory is a genomic theory of inheritance. The main concept is that traits are passed from parents to offspring through genome package transmission. This departs from the gene theory where genes, representing independent informational units, determine the individual's characteristics. The Genome Theory considers genomic topology as the context for gene interactions, and genomic inheritance defines genomic network structure through karyotype coding. Importantly, under stress, the genomic topology can be altered by re-organizing the genome, leading to the emergence of new systems. Such mechanism is responsible for macroevolution both in somatic cell and organismal evolution (Heng, 2015, 2019; Shapiro, 2017).

Somatic mosaicism vs. genomic heterogeneity
These two terms can refer to the same phenomenon when there are distinctive genetic or genomic cell populations within an organism. Traditionally, mosaicism is thought to occur during early development (both mitotically and meiotically). With increased observations of a high degree of mosaicism in adult tissues associated with normal and disease conditions, and the realization that mosaicism is a stress response needed for somatic evolution at all stages, the usage of somatic mosaicism starts overlapping with genetic/genomic heterogeneity. Here, "somatic mosaicism" rather than "genomic heterogeneity" is used to promote the exchangeable use of these two terms in cancer research.

Karyotype coding vs. gene coding
Karyotype coding is responsible for passing system inheritance, while gene coding determines parts inheritance (Ye et al., 2019b). System inheritance is inherited by the order of genes/DNA sequences along/among chromosomes. In contrast, parts inheritance is stored by the order of base pairs within genes. System inheritance is species-specific, but parts inheritance can be shared among different species. The function of sexual reproduction preserves the karyotype coding through meiosis by checking the order of genes along paired chromosomes (Gorelick and Heng, 2011). In many diseases, somatic mosaicism at the karyotype level is common, suggesting the importance of altered genomic information in cellular populations. However, they have often been ignored due to the popularity of gene-centric concepts. Changing the karyotype coding is a hallmark of somatic and organismal macroevolution (Heng, 2019; Ye et al., 2019a).

Macrocellular evolution vs. microcellular evolution
Macrocellular evolution refers to the punctuated cellular evolution often mediated by karyotype changes, while microcellular evolution refers to the stepwise cellular evolution mediated by gene mutations and epigenetic variations. The two phases of cancer evolution were initially documented by experiments of karyotype evolution in action and then confirmed by cancer genome sequencing (Heng et al., 2006; Heng, 2015). Note that studying punctuated clonal evolution should focus on karyotype profiles as karyotype change-mediated macroevolution differs from gene-mediated microevolution. The relationship between macro- and microevolution also illustrates the interactions among individual molecular mechanisms, genome heterogeneity, system stresses, and evolutionary phase transitions. For example, extremely high stress can change the evolutionary phase. Evolutionary tipping points are often detected within the stress-induced crisis stage, leading to phase transition events such as transformation, metastasis, or drug resistance. Immediately following the event of transition, the degree of heterogeneity falls to the lowest level, after which the growth of a more homogenous population dominates (Ye et al., 2018). The two-phased cancer evolution pattern also challenges the general assumption that the accumulation of microevolution over time leads to macroevolution (Heng, 2015, 2019).

Genome chaos vs. chromothripsis
Genome chaos or karyotype chaos refers to a phenomenon of rapid and massive genome re-organization. Initially described in karyotype studies by watching evolution in action (Heng et al., 2006), this mechanism was confirmed by cancer genome sequencing, albeit mainly illustrated by identifying gene mutations or copy number variations. Many names have been introduced to describe these genome re-organization events, including "chromothripsis," which is a subtype of genome chaos (Heng, 2019). High levels of stress during crises can trigger genome chaos, and the rapid and massive genome re-organization can lead to new survivable genomes essential for macroevolution. Overall, stress response-induced emergent systems and their adaptation is a key component of somatic cell evolution, which provides a unifying framework for understanding diverse molecular mechanisms.

Adaptive systems
Complex systems, which are integrated by a set of interacting or interdependent parts or entities. Such whole systems are able to respond to environmental changes or changes in its own interacting parts (including the parts' topology), often in a non-linear fashion. The key features of adaptive biosystems include feedback loops, part heterogeneity, dynamic emergence, multiple levels of fuzzy inheritance, evolutionary capability, and uncertainty between part alteration and whole system behavior. Biological systems are typical adaptive systems which are much more difficult to predict than non-biological systems. The understanding of lower level parts usually does not lead to the understanding of a whole bio-system, especially its emergent behavior under crises (Heng, 2015, 2019).

(1) Karyotype changes lead to new genomic information packages. According to the Genome Theory, the karyotype codes "system inheritance" (the genomic blueprint), while the genes code for "parts inheritance" (**Table 1**) (Ye et al., 2019b). Specifically, karyotype coding ensures the order of genes and other DNA sequences along and among chromosomes for a given species.

Karyotype coding changes can replace the function of a specific gene (Rancati et al., 2008) and impact global gene interaction, leading to new genome systems (Stevens et al., 2013, 2014). In MM, unique gene expression patterns are associated with recurrent chromosomal translocation and ploidy (Zhou et al., 2009). A recent cancer genome analysis has illustrated that the profile of chromosome aberrations is much more useful than gene mutation profiles when correlated with clinical outcomes either as prognostic or predictive markers (Davoli et al., 2017; Jamal-Hanjani et al., 2017). This result was also confirmed in MM, as karyotypic events have a stronger impact on prognosis than mutations (Bolli et al., 2018). In fact, chromosomal profiles have extensively been associated with prognosis in MM, based on specific translocation, hyperdiploidy, chromosomal amplification/deletion, and chromosomal copy number abnormalities (Garcia-Sanz et al., 1995; Avet-Loiseau et al., 2007, 2009; Walker et al., 2010; Shah et al., 2018). By converting DNA sequence data into aneuploidy data, we showed that the status of aneuploidy can suggest clinical MM outcomes (Ye et al., 2019a).

(2) Cancer often represents an evolutionary trade-off of cellular variation-mediated function. Since genomic variations are needed for cellular adaptation, and many essential bioprocesses often can generate harmful byproducts, genomic variations seem unavoidable. For example, normal B-cell development (affinity maturation in the germinal center) and antibody generation require somatic hypermutation and class-switch recombination. However, these key processes also generate DNA breaks and chromosomal translocations, which are central

characteristics of MM (Manier et al., 2017). This represents an immune system trade-off: performing immune functions comes with the risk of malignant transformation [via translocation of cancer genes into immunoglobulin (Ig) loci and/or new karyotype formation] (Gonzalez et al., 2007).

(3) Even though heterogeneity has growth disadvantages (including in cancer), being highly heterogeneous is the winning strategy for most cancers. Genome chaos is essential for population survival under crises, even though it is extremely expensive due to the massive death and often slow growth of the cell population. The key is to create new survivable genomes (through macro-cellular-evolution) (**Table 1**), after which relatively homogenous growth will soon follow with the help of oncogenes in a stochastic fashion (through micro-cellular-evolution) (Ye et al., 2018; Heng, 2019). This principle is used to develop an MM model by synthesizing new patterns of clonal evolution as well as sequencing data (Manier et al., 2017; Maura et al., 2019; Ye et al., 2019d).

(4) The only way for a new system to emerge is to break the constraints above that system (e.g., cellular competition, tissue organization, immuno-systems, and chemo-drugs). In general, different genome systems are required to break different types of constraints (e.g., different karyotypes are involved during different stages of cancer evolution). It is also difficult for any new genome to become dominant. This high level of aberrated genomes therefore become a sufficient condition for cancer evolution.

In addition to the karyotypic level of mosaicism discussed, different types of somatic mosaicism include copy number variations (CNVs) (Walker et al., 2010, 2015; Lohr et al., 2014; Bolli et al., 2018; Aktas Samur et al., 2019), gene mutations (both driver and passenger) (Chapman et al., 2011; Egan et al., 2012; Keats et al., 2012; Bolli et al., 2014, 2018; Lohr et al., 2014; Walker et al., 2015), and non-genetic variations (e.g., epigenetic variations) (Huang, 2009; Heng, 2019). Together, the multiple levels of genetic variation represent the high degree of somatic genomic mosaicism in MM.

THE MAIN MECHANISM OF SOMATIC GENOMIC MOSAICISM IS "FUZZY INHERITANCE" WHICH IS CODED BY LIVING SYSTEMS TO ADAPT TO MICROENVIRONMENTAL DYNAMICS

Cellular heterogeneity has biological significance and genomic basis. Essential cellular heterogeneity is ensured by fuzzy inheritance, a key component of the self-regulating features in bio-adaptive systems. Specifically, heterogeneity is encoded by the genome and realized by genotype-environment interaction (even though bio-errors can also contribute).

Under classical inheritance theory, the gene codes for a fixed or defined genotype, while the environment can influence the real phenotype. For complex polygenic traits, many individuals are needed to illustrate the mode of inheritance. Unfortunately, as shown by the effort of the genome-wide association studies, the multiple genes that contribute to a polygenic trait are hard to identify despite huge sample sizes used. Many loci are involved, and each only contributes to a tiny portion of the phenotype.

To solve this confusion, the new concept of fuzzy inheritance was proposed: genes and chromosomes code for a potential range or spectrum of phenotypes, and the environment serves as a selective "scanner" to "choose" a specific phenotype among the many defined by the genotype (Heng, 2015, 2019). Although the environment plays an important role in phenotypic selection, it is limited by the range established by the inherited genotype: the ultimate phenotype can only be selected from that range. Since diseases are variable phenotypes defined by the interaction between genomic information and environment (Heng et al., 2016), a normal gene can produce a disease phenotype, and disease-associated gene mutations can display a normal phenotype, depending on the environment.

Interestingly, fuzzy inheritance and dynamic environmental interaction will likely be responsible for the majority of phenotypic plasticity. Given the importance of the microenvironment in MM, the role of fuzzy inheritance in cancer evolution should be a top research priority.

THE IMPORTANCE OF SOMATIC GENOMIC MOSAICISM FOR NEW EMERGENT GENOMES

Cellular heterogeneity can alter emergent properties, and cells that diverge from the average population—outliers—often define the direction of cancer evolution (Heng, 2015, 2019). However, cancer researchers have traditionally ignored the contribution of outliers and focused solely on average profiles or dominant clones. Under normal developmental or physiological conditions, this approach may work (although one must note that, even under normal conditions, the 80/20 principle where about 80% of the effects come from 20% of the causes can still play a role). However, under pathological conditions, especially under cellular crisis conditions, some outliers, such as cells with extremely different phenotypes, often become the dominant population. The general conditions for tipping the balance include new altered genomes that favor survival, environmental constraint, and status of the mosaicism. Interestingly, under the right conditions, even a slight change can trigger the tipping point. For example, when the proportion of outliers in the cellular population changes, even in the range of a few percent, an evolutionary phase transition can occur. Such tipping-point system behavior significantly increases the success of cancer evolution when high heterogeneity exists in the cellular population (Maura et al., 2019). When combined with the difference in initial conditions, cellular heterogeneity makes it very hard to predict the outcomes for most cancer cases.

Equally important, since different subpopulations can be molecularly profiled, especially after becoming dominant clones, a huge number of molecular mechanisms can be characterized. Data from recent studies illustrate diverse genetic variations in MM disease evolution (Egan et al., 2012; Keats et al., 2012; Bolli et al., 2014, 2018; Pawlyn and Morgan, 2017; Aktas Samur et al., 2019; Maura et al., 2019). A better way to understand MM is to study the evolutionary mechanism of cancer (Ye et al., 2009), rather than continue identifying

individual molecular mechanisms: when there are so many, the clinical prediction of any single mechanism is low due to highly dynamic evolutionary processes.

THE CLINICAL IMPLICATIONS OF GENOMIC SOMATIC MOSAICISM AND SYSTEM CONSTRAINT

First, it is important to identify the phase of evolution before initiating or changing treatment. Since different types of inheritance are directly related to micro- and macro-somatic evolution, and all cancer phase transitions are defined by macrocellular evolution, the selection of new systems is significantly different from selection on individual genes, especially since the function of any individual gene is influenced by its genomic context. The relationship between disease progression (from MGUS, smoldering MM to active MM) and evolutionary pattern (micro-and macro-somatic evolution) of MM remains to be determined. This will guide when and how to intervene at different stages of the disease in different subpopulations of patients (**Table 1**).

Applying somatic mosaicism in the clinic represents a new approach. On the surface, it is challenging to directly target mosaicism compared to a molecular pathway. However, this seeming disadvantage is actually an advantage when dealing with adaptive systems in which many pathways are involved (e.g., when the causative role for any pathogenic effect is difficult to elucidate and therapies can lead to toxicity and/or secondary malignancies).

In the case of MM: it is worthwhile to investigate whether asymptomatic patients at the stage of smoldering MM can be distinguished by mosaicism. Of course, it is also possible that this clinical challenge will remain even after analyzing evolutionary profiles. Only future investigations will tell.

Second, the stability of higher systems above cancer cells, i.e., the broader microenvironment, organ system, and immune system, can be applied to constrain cancer evolution by slowing down or stabilizing the specific phase of evolution. As all medical treatment can function as cellular stress that may alter the system's evolutionary dynamics (Kultz, 2005; Horne et al., 2014), caution is crucial when weighing the impact of treatment in the context of evolution. For example, within the stable micro-evolutionary phase, moderately treating cells is a better approach than maximal killing, as an over-killing strategy will trigger genome chaos, leading to rapid drug resistance (Heng, 2015, 2019). MM resistance is frequently associated with chromothripsis (Lee et al., 2017) and likely involves treatment-induced genome chaos. Thus, therapies using an adaptive strategy might confer better long-term benefits (Gatenby et al., 2009; Lohr et al., 2014). So far, clinical trials using adaptive strategies in MM treatment (moderate dosage and treatment schedule) have been explored and likely to yield better clinical outcomes (Ye et al., 2019c). On the other hand, instead of putting stress or therapeutic pressure directly on cancer cells, using immunotherapy to modulate the cancer microenvironment (to enhance immune cytotoxic effects and system constraint) is an attractive strategy.

AUTHOR CONTRIBUTIONS

CY and HH drafted the manuscript. JC and GL participated in the discussion, literature search, and editing of the manuscript.

FUNDING

This work was partially supported by the start-up fund for CY from the University of Michigan's Department of Internal Medicine, Hematology/Oncology Division.

ACKNOWLEDGMENTS

We thank Julie Heng and Jessica Mercer for editing the manuscript.

REFERENCES

Aktas Samur, A., Minvielle, S., Shammas, M., Fulciniti, M., Magrangeas, F., Richardson, P. G., et al. (2019). Deciphering the chronology of copy number alterations in Multiple Myeloma. *Blood Cancer J.* 9:39. doi: 10.1038/s41408-019-0199-3

Ashby, C., Boyle, E. M., Walker, B. A., Bauer, M. A., Ryan, K. R., Dent, J., et al. (2019). Chromoplexy and chromothripsis are important prognostically in myeloma and deregulate gene function by a range of mechanisms. *Blood* 134, 3767–3767. doi: 10.1182/blood-2019-130335

Avet-Loiseau, H., Attal, M., Moreau, P., Charbonnel, C., Garban, F., Hulin, C., et al. (2007). Genetic abnormalities and survival in multiple myeloma: the experience of the Intergroupe Francophone du Myélome. *Blood* 109, 3489–3495. doi: 10.1182/blood-2006-08-040410

Avet-Loiseau, H., Li, C., Magrangeas, F., Gouraud, W., Charbonnel, C., Harousseau, J. L., et al. (2009). Prognostic significance of copy-number alterations in multiple myeloma. *J. Clin. Oncol.* 27, 4585–4590. doi: 10.1200/jco.2008.20.6136

Biesecker, L. G., and Spinner, N. B. (2013). A genomic view of mosaicism and human disease. *Nat. Rev. Genet.* 14, 307–320. doi: 10.1038/nrg3424

Bolli, N., Avet-Loiseau, H., Wedge, D. C., Van Loo, P., Alexandrov, L. B., Martincorena, I., et al. (2014). Heterogeneity of genomic evolution and mutational profiles in multiple myeloma. *Nat. Commun.* 5:2997. doi: 10.1038/ncomms3997

Bolli, N., Biancon, G., Moarii, M., Gimondi, S., Li, Y., de Philippis, C., et al. (2018). Analysis of the genomic landscape of multiple myeloma highlights novel prognostic markers and disease subgroups. *Leukemia* 32, 2604–2616. doi: 10.1038/s41375-018-0037-9

Campbell, I. M., Shaw, C. A., Stankiewicz, P., and Lupski, J. R. (2015). Somatic mosaicism: implications for disease and transmission genetics. *Trends Genet.* 31, 382–392. doi: 10.1016/j.tig.2015.03.013

Chapman, M. A., Lawrence, M. S., Keats, J. J., Cibulskis, K., Sougnez, C., Schinzel, A. C., et al. (2011). Initial genome sequencing and analysis of multiple myeloma. *Nature* 471, 467–472. doi: 10.1038/nature09837

Davoli, T., Uno, H., Wooten, E. C., and Elledge, S. J. (2017). Tumor aneuploidy correlates with markers of immune evasion and with reduced response to immunotherapy. *Science*. 355:eaaf8399. doi: 10.1126/science.aaf8399

Egan, J. B., Shi, C. X., Tembe, W., Christoforides, A., Kurdoglu, A., Sinari, S., et al. (2012). Whole-genome sequencing of multiple myeloma from diagnosis to plasma cell leukemia reveals genomic initiating events, evolution, and clonal tides. *Blood* 120, 1060–1066. doi: 10.1182/blood-2012-01-405977

Fernández, L. C., Torres, M., and Real, F. X. (2016). Somatic mosaicism: on the road to cancer. *Nat. Rev. Cancer* 16, 43–55. doi: 10.1038/nrc.2015.1

Garcia-Sanz, R., Orfao, A., Gonzalez, M., Moro, M. J., Hernandez, J. M., Ortega, F., et al. (1995). Prognostic implications of DNA aneuploidy in 156 untreated multiple myeloma patients. Castelano-Leones (Spain) Cooperative Group for the Study of Monoclonal Gammopathies. *Br. J. Haematol.* 90, 106–112. doi: 10.1111/j.1365-2141.1995.tb03387.x

Gatenby, R. A., Silva, A. S., Gillies, R. J., and Frieden, B. R. (2009). Adaptive therapy. *Cancer Res.* 69, 4894–4903. doi: 10.1158/0008-5472.CAN-08-3658

Gonzalez, D., van der Burg, M., Garcia-Sanz, R., Fenton, J. A., Langerak, A. W., Gonzalez, M., et al. (2007). Immunoglobulin gene rearrangements and the pathogenesis of multiple myeloma. *Blood* 110, 3112–3121. doi: 10.1182/blood-2007-02-069625

Gorelick, R., and Heng, H. H. (2011). Sex reduces genetic variation: a multidisciplinary review. *Evolution* 65, 1088–1098. doi: 10.1111/j.1558-5646.2010.01173.x

Heng, H., Regan, S., and Ye, C. (2016). Genotype, environment, and evolutionary mechanism of diseases. *Environ. Dis.* 1, 14–23. Available online at: http://www.environmentmed.org/text.asp?2016/1/1/14/180332

Heng, H. H. (2015). *Debating Cancer: The Paradox in Cancer Research*. Singapore: World Scientific Publishing Co. Pte Ltd. Available online at: http://www.worldscientific.com/worldscibooks/10.1142/8879#t=toc

Heng, H. H. (2019). *Genome Chaos: Rethinking Genetics, Evolution, and Molecular Medicine*. San Diego, CA: Academic Press. doi: 10.1016/C2016-0-05291-9

Heng, H. H., Stevens, J. B., Liu, G., Bremer, S. W., Ye, K. J., Reddy, P. V., et al. (2006). Stochastic cancer progression driven by non-clonal chromosome aberrations. *J. Cell. Physiol.* 208, 461–472. doi: 10.1002/jcp.20685

Horne, S. D., Chowdhury, S. K., and Heng, H. H. (2014). Stress, genomic adaptation, and the evolutionary trade-off. *Front. Genet.* 5:92. doi: 10.3389/fgene.2014.00092

Huang, S. (2009). Non-genetic heterogeneity of cells in development: more than just noise. *Development* 136, 3853–3862. doi: 10.1242/dev.035139

Hultén, M. A., Jonasson, J., Iwarsson, E., Uppal, P., Vorsanova, S. G., Yurov, Y. B., et al. (2013). Trisomy 21 mosaicism: we may all have a touch of Down syndrome. *Cytogenet. Genome Res.* 139, 189–192. doi: 10.1159/000346028

Iourov, I. Y., Vorsanova, S. G., and Yurov, Y. B. (2008). Chromosomal mosaicism goes global. *Mol. Cytogenet.* 1:26. doi: 10.1186/1755-8166-1-26

Iourov, I. Y., Vorsanova, S. G., and Yurov, Y. B. (2010). Somatic genome variations in health and disease. *Curr. Genomics* 11, 387–396. doi: 10.2174/138920210793176065

Iourov, I. Y., Vorsanova, S. G., Yurov, Y. B., and Kutsev, S. I. (2019). Ontogenetic and pathogenetic views on somatic chromosomal mosaicism. *Genes* 10:379. doi: 10.3390/genes10050379

Jamal-Hanjani, M., Wilson, G. A., McGranahan, N., Birkbak, N. J., Watkins, T. B. K., Veeriah, S., et al. (2017). Tracking the evolution of non-small-cell lung cancer. *N. Engl. J. Med.* 376, 2109–2121. doi: 10.1056/NEJMoa1616288

Kaur, G., Gupta, R., Mathur, N., Rani, L., Kumar, L., Sharma, A., et al. (2018). Clinical impact of chromothriptic complex chromosomal rearrangements in newly diagnosed multiple myeloma. *Leuk. Res.* 76, 58–64. doi: 10.1016/j.leukres.2018.12.005

Keats, J. J., Chesi, M., Egan, J. B., Garbitt, V. M., Palmer, S. E., Braggio, E., et al. (2012). Clonal competition with alternating dominance in multiple myeloma. *Blood* 120, 1067–1076. doi: 10.1182/blood-2012-01-405985

Klein, B., Seckinger, A., Moehler, T., and Hose, D. (2011). Molecular pathogenesis of multiple myeloma: chromosomal aberrations, changes in gene expression, cytokine networks, and the bone marrow microenvironment. *Recent Results Cancer Res.* 183, 39–86. doi: 10.1007/978-3-540-85772-3_3

Kultz, D. (2005). Molecular and evolutionary basis of the cellular stress response. *Annu. Rev. Physiol.* 67, 225–257. doi: 10.1146/annurev.physiol.67.040403.103635

Lee, K. J., Lee, K. H., Yoon, K. A., Sohn, J. Y., Lee, E., Lee, H., et al. (2017). Chromothripsis in treatment resistance in multiple myeloma. *Genomics Inform.* 15, 87–97. doi: 10.5808/GI.2017.15.3.87

Lohr, J. G., Stojanov, P., Carter, S. L., Cruz-Gordillo, P., Lawrence, M. S., Auclair, D., et al. (2014). Widespread genetic heterogeneity in multiple myeloma: implications for targeted therapy. *Cancer Cell* 25, 91–101. doi: 10.1016/j.ccr.2013.12.015

Magrangeas, F., Avet-Loiseau, H., Munshi, N. C., and Minvielle, S. (2011). Chromothripsis identifies a rare and aggressive entity among newly diagnosed multiple myeloma patients. *Blood* 118, 675–678. doi: 10.1182/blood-2011-03-344069

Manier, S., Salem, K. Z., Park, J., Landau, D. A., Getz, G., and Ghobrial, I. M. (2017). Genomic complexity of multiple myeloma and its clinical implications. *Nat. Rev. Clin. Oncol.* 14, 100–113. doi: 10.1038/nrclinonc.2016.122

Maura, F., Bolli, N., Angelopoulos, N., Dawson, K. J., Leongamornlert, D., Martincorena, I., et al. (2019). Genomic landscape and chronological reconstruction of driver events in multiple myeloma. *Nat. Commun.* 10:3835. doi: 10.1038/s41467-019-11680-1

Pawlyn, C., and Morgan, G. J. (2017). Evolutionary biology of high-risk multiple myeloma. *Nat. Rev. Cancer* 17, 543–556. doi: 10.1038/nrc.2017.63

Rancati, G., Pavelka, N., Fleharty, B., Noll, A., Trimble, R., Walton, K., et al. (2008). Aneuploidy underlies rapid adaptive evolution of yeast cells deprived of a conserved cytokinesis motor. *Cell* 135, 879–893. doi: 10.1016/j.cell.2008.09.039

Shah, V., Sherborne, A. L., Walker, B. A., Johnson, D. C., Boyle, E. M., Ellis, S., et al. (2018). Prediction of outcome in newly diagnosed myeloma: a meta-analysis of the molecular profiles of 1905 trial patients. *Leukemia* 32, 102–110. doi: 10.1038/leu.2017.179

Shapiro, J. A. (2017). Living organisms author their read-write genomes in evolution. *Biology* 6:E42. doi: 10.3390/biology6040042

Smetana, J., Oppelt, J., Štork, M., Pour, L., and Kuglík, P. (2018). Chromothripsis 18 in multiple myeloma patient with rapid extramedullary relapse. *Mol. Cytogenet.* 11:7. doi: 10.1186/s13039-018-0357-5

Stevens, J. B., Horne, S. D., Abdallah, B. Y., Ye, C. J., and Heng, H. H. (2013). Chromosomal instability and transcriptome dynamics in cancer. *Cancer Metastasis Rev.* 32, 391–402. doi: 10.1007/s10555-013-9428-6

Stevens, J. B., Liu, G., Abdallah, B. Y., Horne, S. D., Ye, K. J., Bremer, S. W., et al. (2014). Unstable genomes elevate transcriptome dynamics. *Int. J. Cancer* 134, 2074–2087. doi: 10.1002/ijc.28531

Vattathil, S., and Scheet, P. (2016). Extensive hidden genomic mosaicism revealed in normal tissue. *Am. J. Hum. Genet.* 98, 571–578. doi: 10.1016/j.ajhg.2016.02.003

Vijg, J. (2014). Somatic mutations, genome mosaicism, cancer and aging. *Curr. Opin. Genet. Dev.* 26, 141–149. doi: 10.1016/j.gde.2014.04.002

Walker, B. A., Boyle, E. M., Wardell, C. P., Murison, A., Begum, D. B., Dahir, N. M., et al. (2015). Mutational spectrum, copy number changes, and outcome: results of a sequencing study of patients with newly diagnosed myeloma. *J. Clin. Oncol.* 33, 3911–3920. doi: 10.1200/jco.2014.59.1503

Walker, B. A., Leone, P. E., Chiecchio, L., Dickens, N. J., Jenner, M. W., Boyd, K. D., et al. (2010). A compendium of myeloma-associated chromosomal copy number abnormalities and their prognostic value. *Blood* 116, e56–e65. doi: 10.1182/blood-2010-04-279596

Ye, C. J., Regan, S., Liu, G., Alemara, S., and Heng, H. H. (2018). Understanding aneuploidy in cancer through the lens of system inheritance, fuzzy inheritance and emergence of new genome systems. *Mol. Cytogenet.* 11:31. doi: 10.1186/s13039-018-0376-2

Ye, C. J., Sharpe, Z., Alemara, S., Mackenzie, S., Liu, G., Abdallah, B., et al. (2019a). Micronuclei and genome chaos: changing the system inheritance. *Genes* 10:366. doi: 10.3390/genes10050366

Ye, C. J., Stevens, J. B., Liu, G., Bremer, S. W., Jaiswal, A. S., Ye, K. J., et al. (2009). Genome based cell population heterogeneity promotes tumorigenicity: the evolutionary mechanism of cancer. *J. Cell. Physiol.* 219, 288–300. doi: 10.1002/jcp.21663

Ye, C. J., Stilgenbauer, L., Moy, A., Liu, G., and Heng, H. H. (2019b). What is karyotype coding and why is genomic topology important for cancer and evolution? *Front. Genet.* 10:1082. doi: 10.3389/fgene.2019.01082

Ye, J. C., Boonstra, P. S., Boyer, D. F., Anderson, L. D., Lipe, B. C., Kin, A., et al. (2019c). A phase 2 study with Minimal Residual Disease (MRD) driven adaptive strategy in treatment for newly diagnosed multiple myeloma with upfront daratumumab-based therapy. *Blood* 134, 3191–3191. doi: 10.1182/blood-2019-132204

Ye, J. C., Chen, L., Chen, J., Parkin, B., Polk, A., Kandarpa, M., et al. (2019d). Aneuploidy is associated with inferior survival in relapsed refractory multiple myeloma patients. *Blood* 134, 4360–4360. doi: 10.1182/blood-2019-124135

Zhou, Y., Barlogie, B., and Shaughnessy, J. D. J. (2009). The molecular characterization and clinical management of multiple myeloma in the post-genome era. *Leukemia* 23, 1941–1956. doi: 10.1038/leu.2009.160

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Genetics

REVIEW
published: 29 April 2020
doi: 10.3389/fgene.2020.00370

Genomic Indexing by Somatic Gene Recombination of mRNA/ncRNA – Does It Play a Role in Genomic Mosaicism, Memory Formation, and Alzheimer's Disease?

Uwe Ueberham and Thomas Arendt**

Paul Flechsig Institute for Brain Research, University of Leipzig, Leipzig, Germany

Recent evidence indicates that genomic individuality of neurons, characterized by DNA-content variation, is a common if not universal phenomenon in the human brain that occurs naturally but can also show aberrancies that have been linked to the pathomechanism of Alzheimer's disease and related neurodegenerative disorders. Etiologically, this genomic mosaic has been suggested to arise from defects of cell cycle regulation that may occur either during brain development or in the mature brain after terminal differentiation of neurons. Here, we aim to draw attention towards another mechanism that can give rise to genomic individuality of neurons, with far-reaching consequences. This mechanism has its origin in the transcriptome rather than in replication defects of the genome, i.e., somatic gene recombination of RNA. We continue to develop the concept that somatic gene recombination of RNA provides a physiological process that, through integration of intronless mRNA/ncRNA into the genome, allows a particular functional state at the level of the individual neuron to be indexed. By insertion of defined RNAs in a somatic recombination process, the presence of specific mRNA transcripts within a definite temporal context can be "frozen" and can serve as an index that can be recalled at any later point in time. This allows information related to a specific neuronal state of differentiation and/or activity relevant to a memory trace to be fixed. We suggest that this process is used throughout the lifetime of each neuron and might have both advantageous and deleterious consequences.

Keywords: Alzheimer's disease, LINE-1 (L1), mosaicism, neuronal individuality, human brain, amyloid precursor protein (APP), retrotransposition, somatic gene recombination

OPEN ACCESS

Edited by:
Ivan Y. Iourov, Mental Health Research Center of Russian Academy of Medical Sciences, Russia

Reviewed by:
José María Frade, Cajal Institute (CSIC), Spain
Guifeng Wei, University of Oxford, United Kingdom

****Correspondence:***
Uwe Ueberham
uwe.ueberham@medizin.uni-leipzig.de
Thomas Arendt
thomas.arendt@medizin.uni-leipzig.de

Specialty section:
This article was submitted to Human Genomics, a section of the journal Frontiers in Genetics

Received: *06 January 2020*
Accepted: *25 March 2020*
Published: *29 April 2020*

Citation:
Ueberham U and Arendt T (2020) Genomic Indexing by Somatic Gene Recombination of mRNA/ncRNA – Does It Play a Role in Genomic Mosaicism, Memory Formation, and Alzheimer's Disease? Front. Genet. 11:370. doi: 10.3389/fgene.2020.00370

THE INDIVIDUALITY OF NEURONS

The individuality of neurons provides an accepted paradigmatic framework for the nervous systems of invertebrates such as Caenorhabditis elegans or Drosophila melanogaster, where the uniqueness of each single neuron is reflected by its particular spatial and functional position (Alicea, 2018; Davie et al., 2018; Hammarlund et al., 2018). With current technical developments that allow for high-throughput analyses of various cellular markers, it is becoming more and more obvious, however, that in vertebrates, too, and even in primates, neurons are much more heterogeneous than

previously thought (Lake et al., 2016; Lim et al., 2018). Thus, depending on how many parameters can be assessed in depth simultaneously, each neuron might be unique with respect to its functional, structural, and molecular signatures. This cellular diversity and individuality might result from a complex process where different determinants interacting at different levels, such as developmental trajectories, relationships with neighboring cells, functional integration in neuronal networks, and others, may shape and re-shape cellular signatures.

One particular aspect that is crucial to our understanding of the individuality of somatic cells is the individual genetic equipment giving rise to genetic mosaicism. Commonly, mosaicism is defined as the presence of genetically different lineages of cells derived from a single zygote, with additional variations arising in the soma of each cell that are usually not inherited by the next generation (Forsberg et al., 2017). Genomic mosaicism in the human brain has been explored for about 20 years (Muotri et al., 2005; Renthal et al., 2018; Saleh et al., 2019) and is currently a subject of intensive research (see this special issue). While the phenomenon of genomic mosaicism now seems to be an established fact (Rohrback et al., 2018b), there is much less consensus on its extent and distribution (Paquola et al., 2017; Rohrback et al., 2018a), and hardly anything is known about its physiological and potentially pathophysiological meaning.

Research on single-neuronal DNA content in human brain over the last 20 years or so has collected a huge but, to some extent, inconsistent pool of data. Searching for cellular signatures of neurodegenerative diseases such as Alzheimer's disease (AD) was a particular driving force for early studies (Potter, 1991; Potter and Geller, 1996; Geller and Potter, 1999; Rehen et al., 2001; Mosch et al., 2006, 2007; Arendt et al., 2009, 2015, 2017; Arendt, 2009; Fischer et al., 2012).

First reports, based on analyses of bulk DNA, derived from a mixture containing neuronal and non-neuronal cells of the human brain, showed alterations of DNA content (Rehen et al., 2005). Subsequent studies, applying more sophisticated techniques of single-cell analyses based on single-cell isolation by high-throughput cell sorting or laser capture microdissection in combination with single-cell sequencing, identified chromosomal aneuploidy, small and larger copy number variations (CNVs), single nucleotide variations (SNVs), and DNA content variation (DCV), all contributing to the genomic heterogeneity and individuality of each single neuron (Mosch et al., 2006, 2007; Arendt et al., 2009; Iourov et al., 2009; Westra et al., 2010; Fischer et al., 2012; Abdallah et al., 2013; McConnell et al., 2013). Excellent reviews on this phenomenon and the underlying mechanisms are provided in this special issue and elsewhere (McConnell et al., 2017; Rohrback et al., 2018b).

A VERY LARGE PART OF THE HUMAN GENOME MIGHT BE OF RETRO-TRANSPOSABLE ORIGIN

In a comparison of the predicted number of protein-coding genes in a wide range of phylogenetically related vertebrates, only 16,000–26,000 hits are proposed (Holland et al., 2017). Most of them comprise complex exon-intron structures allowing the splicing machinery to generate transcripts in a cell-type- and time-dependent context. However, for the human genome, protein-coding transcripts cover only about 2%, whereas 75% of the human genome can be transcribed (Djebali et al., 2012) and are currently attributed to the ncRNA fraction. In this context, it is important to note that more than 40% of DNA sequences are assumed to be of retro-transposable origin (Cordaux and Batzer, 2009). The function of these sequences is still mostly unknown but is receiving increasing amounts of attention, especially with respect to unveiling the heterogeneity of single cells in selected tissues, particularly in the brain.

Most approaches to assess a potential function of somatic genomic mosaicisms in both health and disease largely ignore the role of RNA or, at the most, attribute to RNA only a canonical function within the context of transcription and translation of genetic information. There are, however, a few most intriguing studies suggesting a role for RNA in DNA sequence modulation, e.g., transcription-associated mutagenesis or transcription-associated recombination resulting from events like RNA collision with replication machinery or co-transcriptional R-loop formation (Green et al., 2003; Majewski, 2003; Polak and Arndt, 2008; Kim and Jinks-Robertson, 2012; Garcia-Muse and Aguilera, 2019; Rondon and Aguilera, 2019). While immunoglobulin class-switch recombination, which generates diverse antibodies, is a beneficial example of R-loop formation (Yu et al., 2003), in several repeat-associated neurological diseases, such RNA-DNA-hybrids produce deleterious DNA sequence modifications (e.g., RNAs from HTT, FXN, or ATXN1) (Richard and Manley, 2017; Neil et al., 2018).

Recently, a study by Lee et al. (2018) reported on the discovery of somatic gene recombination in terminally differentiated human neurons. They identified thousands of variant genomic cDNAs (gencDNA) of the amyloid precursor protein (APP) gene in neurons of Alzheimer's disease brains. These gencDNAs contained no introns but showed a wide range of sequence pattern comprising full-length copies of brain-specific splice variants and many smaller forms with insertions, deletions, single-nucleotide variations, or intra-exonic junctions. According to their data, a "retro-insertion" of RNA is a likely source of these gencDNAs. Though highly enriched in the neurons of AD brains, where several known and some unknown APP mutations could be identified, control brains also showed gencDNA loci of recombined APP. Apparently, genomic recombination is a common process in terminally differentiated neurons in both normal and diseased brain, contributing to mosaicism, individuality, and pathology.

INTRONLESS GENES

A striking observation by Lee et al. (2018) is the detection of thousands of intronless APP-derived sequences in the DNA of single neurons. Though data on further intronless genes of a comparable extent have not yet been reported, the questions arise: what are the possible reasons for the usage of intronless transcripts, and do they fulfill a physiological function?

Firstly, introns are a characteristic feature of eukaryotic genomes. They are genetic elements that can monitor their own gene transcription or the transcription of functionally clustered genes (Hube and Francastel, 2015; Shaul, 2017). Following this idea, a feedback control could avoid unnecessary accumulation of toxic metabolites or proteins to protect cells and to avoid energy/substrate wastage. The presence of introns can thus contribute to better regulation of the genome and increases its coding potential (Heyn et al., 2015). They also provide a mechanism to increase the proteome diversity by alternative splicing (Nilsen and Graveley, 2010). Introns can protect eukaryotic genomes from transcription-associated genetic instability, for example by preventing R-loop formation and DNA damage accumulation (Bonnet et al., 2017). However, a remarkable fraction of constitutively spliced transcripts using the intronic gene structure that might not contribute to substantial regulation has also been identified (Ryu et al., 2015). Accordingly, constitutive exons are evolutionarily older, and their replacement by alternative exons has only restricted functional relevance (Xiong et al., 2018), suggesting a possible role for basic cellular functions. Thus, constitutive exons behave at least partly like intronless transcripts. Of note, several housekeeping genes such as GAPDH or ACTH, the expression of which is assumed to be relatively stable within cells, possess a high number of mostly intronless pseudogenes, comparable in size to their authentic RNA (Sun et al., 2012).

INTRONLESS GENES CAN CONTRIBUTE TO THE GENOMIC DIVERSITY OF CELLS

Up to 10% of sequences that appear as pseudogenes in the human genome seem to be transcribed (Djebali et al., 2012) and could participate in gene expression as a competing endogenous RNA (ceRNA) (Poliseno, 2012; Zhong et al., 2018) or might even code for translated protein (Ingolia et al., 2014; Ji et al., 2015).

Intronless genes, which represent less than 5% of the human genome, lack intron-dependent transcription control, leading to a more constant expression level. Such features have been reported for genes that preferably encode metabolically passive proteins (Hill and Sorscher, 2006).

Most intronless genes are evolutionarily relatively young, are expressed at lower levels compared to intron-containing genes, show a higher tissue specificity, and evolve faster than spliced genes (Shabalina et al., 2010; Louhichi et al., 2011). It seems that intronlessness is a more recent form of evolution to develop tissue-specific functions (Brosius and Gould, 1992; Shabalina et al., 2010) that might be actively involved in brain development and aging.

An unusually high number of intronless genes have neuron-specific expression (Grzybowska, 2012) or at least play a major role in the brain, such as several serotonin receptors, HTR1A, HTR1B, or HTR1D or beta1- and beta2-adrenergic receptors (ADRB1, ADRB2) (see the IGD database[1]; Louhichi et al., 2011).

[1]http://www.bioinfo-cbs.org/igd

Intronless transcripts circumvent the complex splicing process, thereby saving energy and time and allowing for replication of much shorter genes. Splicing mechanisms always pose a definite risk of inaccurate execution. Thus, a globally impaired exon exclusion and selective loss of splicing factors have been shown for AD brains (Berson et al., 2012). Moreover, destruction of cholinergic neurons in mice, a critical feature of AD pathology, leads to disturbances in RNA splicing, dendritic loss, and memory impairment (Kolisnyk et al., 2016). During aging, the number of splicing errors increases in the brain. Recently, an integrative transcriptome analysis of the aging brain provided evidence that deregulated mRNA splicing is a feature in AD, where hundreds of aberrant pre-splicing events could be detected (Raj et al., 2018).

The number of somatic mutations in the human brain increases over the lifetime due to various types of stress and an age-related loss of DNA repair efficiency, which itself is comprised of mutations contributing to this genomic instability (Chow and Herrup, 2015; Verheijen et al., 2018). Usage of such compromised DNA could be prevented by the availability of alternatively saved/stored intronless variants. RNA molecules could serve as templates to repair DNA double-strand breaks leading to intronless genes (Catania, 2017).

Taken together, several lines of evidence suggest that intronless genes, which, to some extent, may appear as pseudogenes, could substantially contribute to the genetic diversity of cells (Kovalenko and Patrushev, 2018).

INTRONLESS GENES ARE A LIKELY CONSEQUENCE OF SOMATICALLY RECOMBINED TRANSCRIPT INCORPORATION INTO THE GENOME AND ARE POTENTIALLY GENERATED BY LINE-1 RETROTRANSPOSITION

A likely source of intronless genes in eukaryotic genomes is the retroposition of cellular mRNAs by retrotransposable elements (Callinan and Batzer, 2006; Baertsch et al., 2008), though Lee et al. (2018) could not confirm this for APP-derived intronless transcripts. However, during the evolution of the primate lineage, there was a burst of retropositions that reached its peak about 38–50 million years ago, when many intronless genes emerged in the genome (Marques et al., 2005).

Retrotransposons are mobile elements that account for more than 40% of the human genome (Lander et al., 2001). They have been identified as an important source for genetic variations during the evolution of the human genome. However, only a limited number of these elements retain full function and are still active in the genome. The vast majority of retrotransposons are silenced at multiple levels, including transcriptional repression, epigenetic modification such as DNA methylation of CpG-rich promoters (Greenberg and Bourc'his, 2019), or other post-transcriptional gene-regulation mechanisms.

The evolutionary bursts of retrotransposable elements in the human genome gave rise to about 700 human intronless

genes, which stabilized biological processes critically required for the survival of the species (Louhichi et al., 2011). Generation of intronless genes through retrotransposable elements can apparently take place both in the germline and somatic cells. Thus, many intronless genes are inherited and, accordingly, show testis-specific expression (Grzybowska, 2012). In addition, during the lifetime of individual organisms, retrotransposable elements might give rise to intronless genes in somatic cells such as neurons, where they could contribute to the genomic individuality of neurons as well as to the individuality of the carrier organism.

The long interspersed element (L1, LINE-1) is the only known active autonomous retrotransposon in human cells (Moran et al., 1996) and covers up to 17% of the human genome. About 100 retrotransposition-competent L1-elements are detected in each individual, while more than 500,000 copies are silent due to truncations, deletions, or other alterations (Myers et al., 2002; Brouha et al., 2003; Salvador-Palomeque et al., 2019). The view of L1 elements has changed over time from being regarded as "selfish" or "parasitic" towards representing functionally critical elements (Paco et al., 2015) that fulfill essential roles in the regulation of gene expression. However, addressing the function of LINE-1 elements has been restricted by technical difficulties in detecting their specific location in the human genome. Their high copy number often gives rise to unreliable data in PCR amplifications or hybridization-based assays, and new methods for mapping active transposable element insertion sites in genomic DNA have been developed only recently (Steranka et al., 2019).

Members of the LINE-1 retrotransposon family typically use target-primed reverse transcription (TPRT) to generate *de novo* insertions into genomic locations of germline and somatic cells. TPRT is catalyzed in cis by ORF1p and ORF2p, two proteins translated from a bicistronic 6 kb L1 mRNA (**Figure 1A**). The L1 ORF2p comprises both endonuclease activity (EN) and reverse transcriptase (RT) activities, which are essential components for successful L1 retrotransposition (Mathias et al., 1991; Feng et al., 1996). Retrotransposition is started by an internal promoter located in the L1 5′-untranslated region (Swergold, 1990). Synthesized L1 mRNA is subsequently transported to the cytoplasm (**Figure 1B**), where ORF1p and ORF2p proteins are translated and bind their own mRNA to form a ribonucleoprotein particle (Wei et al., 2001). After entering the nucleus, TPRT activity catalyzes the retrotransposition (Upton et al., 2015). Intragenic insertions of LINEs can disrupt gene expression, which is often connected to severe diseases (Schwahn et al., 1998; Meischl et al., 2000). Recently, LINE elements have been inferred to participate in recruiting RNA-binding proteins to mammalian introns and to influence the splicing and evolution of tissue-specific exons (Attig et al., 2018). The ability of evolutionarily young LINEs to attract splice-repressive RNA binding proteins (e.g., MATR3, PTBP1) contrasts with evolutionarily old LINEs, which possess less repressive motifs but rather allow for the binding of splice-promoting RNA-binding proteins. These latter LINEs support lineage-specific splicing (Attig et al., 2018) and play an important role in the development of neurons, making the brain a hotspot of somatic mosaicism. Apparently, L1 mobilization operates during the entire life-span of neurons, starting during neurogenesis in neuronal precursor cells (NPC) (Muotri et al., 2005, 2009; Coufal et al., 2011; Upton et al., 2011; Kurnosov et al., 2015; Macia et al., 2017) and persisting into terminally differentiated states (Baillie et al., 2011; Evrony et al., 2012, 2015; Erwin et al., 2016).

Relevant to the above-mentioned generation of somatically recombined transcripts is the ability of LINE-1 transcripts to retrotranspose cellular mRNA in trans (Wei et al., 2001; **Figure 1B**). To this end, both intact ORF1p- and ORF2p-encoded proteins are necessary. Different data on the frequency of pseudogene formation, between 0.01 and 0.05% of the rate of L1 retrotransposition (Wei et al., 2001) and about 10% (Esnault et al., 2000), were reported and suggested different integration mechanisms with respect to L1-endonuclease (Wei et al., 2001). However, *in vitro* mature neurons express detectable L1 mRNA and ORF1p levels and exert efficiently engineered L1 retrotransposition (Macia et al., 2017).

WHAT IS NECESSARY FOR SOMATIC RECOMBINATION OF RNA-BASED TRANSCRIPTS IN TRANS?

The mechanism of retro-insertion of RNA-based transcripts in trans has a number of prerequisites such as reverse transcriptase activity, poly-adenylation, and DNA double-strand break. The availability of these factors will determine the frequency and efficacy of retro-insertion.

Reverse transcriptase (RT) activity has recently been detected in normal human brain extracts (Lee et al., 2018) and blood samples (Steele et al., 2005; MacGowan et al., 2007). Reverse transcriptase activity seems phylogenetically of different origins, with non-LTR reverse transcriptase, including group II intron IEPs, telomerase, and human L1 reverse transcriptase, differing from LTR enzymes, which include the HIV enzymes (Zhao and Pyle, 2017). However, retrotransposons are potential sources for reverse transcriptase activity in human neurons. The family of human-specific LINE-1 retrotransposons is the only family known so far that can actively and autonomously transpose into the human genome, thereby using its own encoded protein activities necessary for retrotransposition (e.g., endonuclease and reverse transcriptase) (Kazazian and Moran, 2017).

Replacement studies demonstrated that a poly(A) sequence is required for LINE-1 directed retrotransposition (Doucet et al., 2015), where not only cis encoding L1 retrotransposons, which end in a 3′ poly(A) sequence, are mobilized, but also cellular mRNA in trans can be a target (Doucet et al., 2015).

The poly(A) tail of mRNA facilitates its export from the nucleus, enhances protein synthesis, and stabilizes mRNA by interacting with poly(A)-binding proteins to prevent exonucleolytic degradation.

Interestingly, it has been reported that non-conserved poly(A) sites are associated with transposable elements to a much greater extent than conserved ones (Lee et al.,

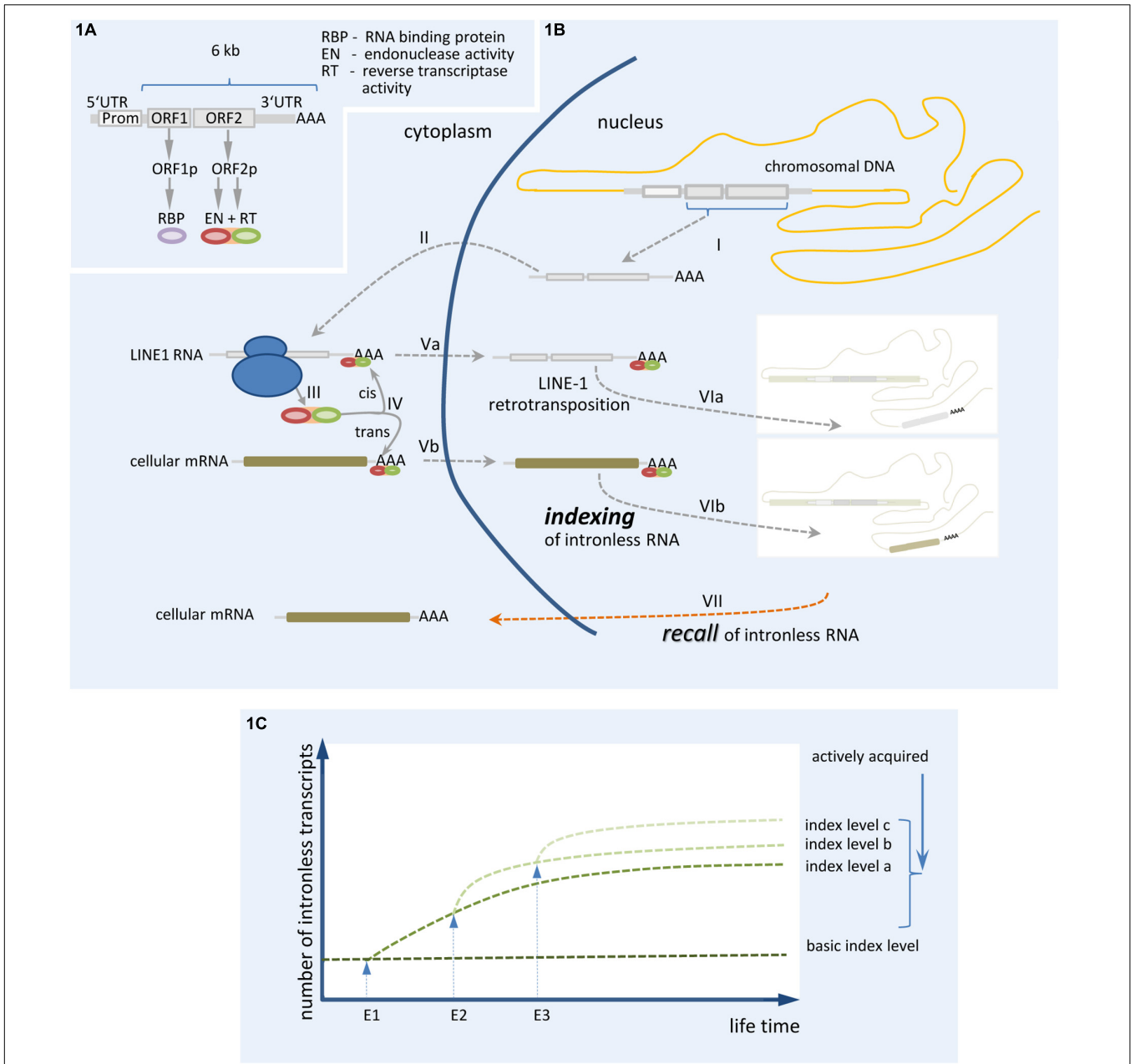

FIGURE 1 | Synopsis of the proposed mechanism of genomic indexing by somatic gene recombination of mRNA/ncRNA. **(A)** The restrotransposition competent (RC) LINE-1 RNA and the encoded proteins are shown. **(B)** The process of LINE-1 directed retrotransposition and genomic indexing by somatic gene recombination of mRNA is depicted: (I) transcription of retrotransposition competent (RC) LINE-1 controlled by endogenous promoter, (II) transport of RC-LINE-1 transcript to cytoplasm, (III) translation of ORF1 and ORF2 proteins, (IV) binding of ORF2 protein (and ORF1 protein, not shown) to their own mRNA (cis) or a cellular mRNA (trans) (potentially representing a specific cellular context) by forming a ribonucleoprotein complex, (Va/Vb) transport of cis- or trans-generated ribonucleoprotein complex into the nucleus, (VIa/b) retrotransposition is controlled by Target Primed Reverse Transcription (TPRT) in "VIb," leading to indexing of a specific cellular context, and (VII) recall of intronless RNA. **(C)** Proposed operational sequence leading to an increasing genomic index or memory trace by somatic gene recombination. E1, E2, and E3 represent events leading to increased index levels due to the insertion of RNA transcripts (generated within a definite temporal context) by somatic recombination. Whether single events finally provide advantageous or deleterious indices depends both on the spatial/temporal context and whether the RNA transcripts used for genomic recombination correspond to a correct or a mutated sequence.

2008). This opens the opportunity for LINE-1 elements to differently use alternative polyadenylation (APA) sites of individual mRNA transcripts, which influence mRNA stability, mRNA localization, or the amount and localization of encoded proteins (Tian and Manley, 2017). Especially for the brain, a wide variety of APA is known, which is typically associated with a particular expression pattern specific to a cell-type or even subcellular compartment

(Miura et al., 2013; Taliaferro et al., 2016). For example, for brain-derived growth factor (BDNF), a short isoform of the mRNA is restricted to the cell body, whereas the long isoform localizes to the dendrites, where it is translated (An et al., 2008).

APAs could thus provide a broad basis for the incorporation of selected transcripts into the genome of each single neuron by somatic recombination according to their individual profile. Additionally, a potential LINE-1 insertion candidate RNA can possess poly(A) tails of different lengths, obtained by somatic mutation, which are finally reverse-transcribed into the genome (Evrony et al., 2015). Data from the same study indicated the existence of highly polymorphic poly(A) tails of varying length, leading to many different somatic mutations, which can contribute to manifestations of local and functional clones. This might also contribute to the highly diverse mosaicism observed in neurons.

While polyadenylation of RNA is required for labeling RNAs to prevent degradation processes, the primordial role of oligoadenylation is RNA tagging for subsequent destabilization, which blurs the boundary between stabilization and destabilization by adenylations (Tudek et al., 2018). It might thus be tempting to speculate that truncated mRNA transcripts, which are intended to be degraded and so are oligoadenylated, might be at risk of being accidently included in "normal polyadenylation processes." This kind of potential RNA "mislabeling" might trigger accidental translation and protein synthesis or even lead to interaction with transposable elements such as LINE-1, which in turn allow the integration into the genome of individual cells and contribute to mosaicism. TENT2, also known as GLD2, a non-canonical poly(A) polymerase, is such a possible candidate, which; performs both polyadenylation and oligoadenylation on many RNAs (e.g., GluN2A RNA), is expressed in the hippocampus, can co-localize with proteins relevant for synaptic plasticity, and may be necessary for long-term potentiation (Rouhana et al., 2005; Swanger et al., 2013). Other non-canonical poly(A) polymerases, e.g., TENT4A/B are involved both in RNA decay and in the stabilization of mRNAs (Gagliardi and Dziembowski, 2018; Warkocki et al., 2018).

A further requirement for the generation of somatically recombined transcripts and their DNA integration are DNA double-strand breaks. Among others, DNA double-strand breaks have been linked to tumorigenesis and genetic instabilities (Aparicio et al., 2014; Mladenov et al., 2016). In addition, disturbances of the underlying repair mechanisms, which involves a coordinated action of TDP2 (tyrosyl DNA phosphodiesterase 2) with enzymes of the NHEJ repair pathway, can lead to neurological diseases associated with intellectual impairment or ataxia (Gomez-Herreros et al., 2014). Moreover, corruption of epigenetically modified DNA in the germline followed by errors in the subsequent repair process could even lead to epigenetic regulatory effects transmissible over generations as an epigenetic memory of repair of DNA double-strand breaks (Orlowski et al., 2011). While any insertion following DNA double-strand breaks can be mutagenic through disrupting coding sequences, it can also influence the expression of adjacent genes by reorganizing the gene structure, providing completely new features, and could therefore also be physiologically relevant.

LINE-1 INTEGRATION IS INVOLVED IN MEMORY FORMATION

Recently, it has been reported that DNA double-strand breaks linked to neuronal activity are a common, basic, and physiological phenomenon. Exploratory activity in mouse, for example, which is associated with increased neuronal activity, has been shown to cause a significant increase in neuronal DNA double-strand breaks (Suberbielle et al., 2013). Moreover, a variety of early-response genes, such as Fos, FosB, and Egr1, other transcription factors, such as Olig2, and ncRNAs, such as Malat1, are targets of activity-induced DNA double-strand breaks in neurons (Madabhushi et al., 2015).

LINE-1 mobilization in brain uses functionally active DNA double-strand breaks to jump into the genomic DNA. The linkage of DNA double-strand breaks to neuronal activity (Suberbielle et al., 2013) might thus provide a mechanism to index the specific activity state of the neurons.

L1 insertions in neurons were proposed to be a mechanism of "genomic plasticity" some years ago (Singer et al., 2010). Accordingly, L1 elements alter the neuronal transcriptome by their genomic integration, which eventually contributes to a modified behavior of the affected individual (Singer et al., 2010). Moreover, the involvement of LINE-1 activation in memory formation has recently been reported (Bachiller et al., 2017). Immediately after a novel place exploration session in mice, a short and temporarily limited increase of LINE-1 orf1- and orf2- mRNA expression was observed in the hippocampus. Remarkably, just 1 h after the exploratory session, a permanently elevated copy number of orf2- insertions in the hippocampal genome was measured, while the content of orf1 sequences did not change. The orf2 copy number increase in genomic DNA could be blocked by the administration of lamivudine, a retrotranscriptase inhibitor. Accordingly, lamivudine application within a time window of 6 h after the training session also impaired long-term memory formation. Both memory formation and orf2 insertion in DNA was also prevented by orf1 antisense or orf2 antisense RNA infusion into hippocampus (Bachiller et al., 2017; Wang et al., 2018). Another study (Kokaeva et al., 2002) has shown that inhibition of the expression of the LINE-1 reverse-transcriptase gene in rats by antisense oligonucleotides disturbed the formation of long-term memory, while short-term memory was not altered.

Taken together, there is thus strong evidence that L1 retrotransposon insertions might be involved in the process of long-term memory formation.

The majority of studies on the *de novo* genomic LINE-1 insertion has mainly focused on the alteration of mRNA/ncRNA expression levels or relocation of splicing-variant-ratios. However, the detection of hundreds of somatically recombined

APP sequences in neurons of both the healthy and diseased human brain (Lee et al., 2018) together with accumulating evidence linking LINE-1 to memory formation, opens up a completely new perspective on the role of transcribed mRNA/ncRNA in indexing functional states of neurons.

We would thus like to further develop the hypothesis that somatic recombination of intronless mRNA/ncRNAs provides a mechanism to index a particular functional state at the level of the individual neuron, a suggestion that was similarly proposed by Lee et al. (2018). By insertion of defined RNAs in a somatic recombination process, the presence of specific mRNA transcripts within a definite temporal context could be "frozen" and serve as an index that can be recalled. This allows the fixing of information related to a specific neuronal state of differentiation and/or activity relevant to a memory trace. This process might take place throughout the lifetime of each neuron and will potentially have both advantageous and deleterious consequences (**Figure 1C**).

In conclusion, it might thus be probable that retrotransposition by LINE-1, which allows the use of defined RNA to index a particular cellular state, represents a powerful and versatile toolbox in somatic cells that can modify the DNA sequence without affecting original gene structures.

ARE LINE-1 ACTIVITY AND gencDNA GENERATION INVOLVED IN THE PATHOMECHANISM OF AD?

While a definite involvement of LINE-1 and gencDNA in the AD pathomechanism remains to be shown, several lines of evidence clearly point in this direction. A recent analysis of more than 600 human cortical transcriptomes indeed revealed evidence for a global transcriptional activation of LINE-1 in AD (Guo et al., 2018). Still, another study analyzing only a small number of AD samples by target PCR failed to detect any differences in L1 genomic copy numbers (Protasova et al., 2017).

Global hypomethylation of DNA, accompanied by a downregulation of neuronal DNA methyltransferase 1 (DNMT1), appears to be a characteristic feature of AD (Mastroeni et al., 2010). At least in human neural progenitor cells, a global DNA hypomethylation by deletion of DNMT1 leads to activation of evolutionarily young hominoid-specific LINE-1 elements while the older L1s remain silent. Accordingly, activated L1s provide alternative promoter activity for many protein-coding genes that are relevant for neuronal functions. This shows that evolutionarily young L1-specific elements are controlled by a DNA methylation pattern (Jonsson et al., 2019). This situation could be provoked in AD brain by dysregulation of LINE-1 elements.

A recent study showed upregulation of the histone demethylase KDM4B in AD brains (Park et al., 2019). This histone demethylase had previously been identified to promote LINE-1 expression and enhance LINE-1 copy number and retrotransposition efficacy, while its depletion reduces LINE-1 expression (Xiang et al., 2019). In addition, SIRT6, a histone deacetylase and powerful repressor of L1-activity by ribosylating KAP1 (van Meter et al., 2014), a nuclear co-repressor protein of LINE-1 (Rowe et al., 2010; Castro-Diaz et al., 2014), is reduced in AD (Kaluski et al., 2017), which could further contribute to the activation of LINE-1.

Many more mutated APP-RNA variants were detected in single neuronal nuclei derived from prefrontal cortices of sporadic AD brains than in control brains (Lee et al., 2018). Some of these APP sequences showed intra-exonic junctions, and some even retained coding potential. Their presence in gencDNA might contribute to manifest sporadic AD cases. Since neurons are able to deliver linear and circular RNAs through exosome-dependent mechanisms (Liu et al., 2019), the propagation of mutated APP RNA transcripts to neighboring cells with the potential to be inserted as gencDNA by LINE-1 elements should be considered as a potential basis for pathology-spreading. Moreover, released extracellular vesicles can also mediate the horizontal transfer of active LINE-1 retrotransposons from one cell to another (Kawamura et al., 2019). In summary, a complete exchange of LINE-1 elements and trans RNAs like a tool kit seems possible.

THERAPEUTIC IMPLICATIONS

If the proposed mechanism is indeed instrumental to the AD pathomechanism, a moderate influence on LINE-1 activity might ameliorate the deleterious insertion of AD-related and mutated APP transcripts. The inhibition of reverse transcriptase activity, as already proposed elsewhere (Lee and Chun, 2019), might thus be a promising approach. Still, RT inhibitors would not prevent the generation of mutated APP mRNAs or their fragments. Moreover, since LINE-1 activity is necessary for memory formation, inhibition of RT would have potentially serious side effects that need to be considered.

Thus, it is necessary to accumulate more detailed knowledge of these mechanisms before any interferences regarding this mechanism can be envisaged.

AUTHOR CONTRIBUTIONS

UU and TA wrote the manuscript.

FUNDING

This work was supported by grants from the Alzheimer Forschungs Initiative (AFI; #17021) to UU and DFG (SPP1738: AR 200/17-1) and SAB (#100327691) to TA.

ACKNOWLEDGMENTS

We acknowledge support from Leipzig University for Open Access Publishing.

REFERENCES

Abdallah, B. Y., Horne, S. D., Stevens, J. B., Liu, G., Ying, A. Y., Vanderhyden, B., et al. (2013). Single cell heterogeneity: why unstable genomes are incompatible with average profiles. *Cell Cycle* 12, 3640–3649. doi: 10.4161/cc.26580

Alicea, B. (2018). The emergent connectome in Caenorhabditis elegans embryogenesis. *Biol. Syst.* 173, 247–255. doi: 10.1016/j.biosystems.2018.09.016

An, J. J., Gharami, K., Liao, G.-Y., Woo, N. H., Lau, A. G., Vanevski, F., et al. (2008). Distinct role of long 3' UTR BDNF mRNA in spine morphology and synaptic plasticity in hippocampal neurons. *Cell* 134, 175–187. doi: 10.1016/j.cell.2008.05.045

Aparicio, T., Baer, R., and Gautier, J. (2014). DNA double-strand break repair pathway choice and cancer. *DNA Repair* 19, 169–175.

Arendt, T. (2009). Synaptic degeneration in Alzheimer's disease. *Acta Neuropathol.* 118, 167–179. doi: 10.1007/s00401-009-0536-x

Arendt, T., Belter, B., Brückner, M. K., Ueberham, U., Morawski, M., and Tarnok, A. (2017). "A Cytomic approach towards genomic individuality of neurons," in *Genomic Mosaicism in Neurons and Other Cell Types. Neuromethods*, eds J. Frade, and F. Gage Vol 131. (New York, NY: Humana Press), 81–106.

Arendt, T., Brückner, M. K., and Lösche, A. (2015). Regional mosaic genomic heterogeneity in the elderly and in Alzheimer's disease as a correlate of neuronal vulnerability. *Acta Neuropathol.* 130, 501–510.

Arendt, T., Mosch, B., and Morawski, M. (2009). Neuronal aneuploidy in health and disease: a cytomic approach to understand the molecular individuality of neurons. *Int. J. Mol. Sci.* 10, 1609–1627. doi: 10.3390/ijms10041609

Attig, J., Agostini, F., Gooding, C., Chakrabarti, A. M., Singh, A., Haberman, N., et al. (2018). Heteromeric RNP Assembly at LINEs Controls Lineage-Specific RNA Processing. *Cell* 174, 1067-1081.e6. doi: 10.1016/j.cell.2018.07.001

Bachiller, S., Del-Pozo-Martín, Y., and Carrión, A. M. (2017). L1 retrotransposition alters the hippocampal genomic landscape enabling memory formation. *Brain Behav. Immun.* 64, 65–70. doi: 10.1016/j.bbi.2016.12.018

Baertsch, R., Diekhans, M., Kent, W. J., Haussler, D., and Brosius, J. (2008). Retrocopy contributions to the evolution of the human genome. *BMC Genomics* 9:466. doi: 10.1186/1471-2164-9-466

Baillie, J. K., Barnett, M. W., Upton, K. R., Gerhardt, D. J., Richmond, T. A., Sapio, F., et al. (2011). Somatic retrotransposition alters the genetic landscape of the human brain. *Nature* 479, 534–537. doi: 10.1038/nature10531

Berson, A., Barbash, S., Shaltiel, G., Goll, Y., Hanin, G., Greenberg, D. S., et al. (2012). Cholinergic-associated loss of hnRNP-A/B in Alzheimer's disease impairs cortical splicing and cognitive function in mice. *EMBO Mol. Med.* 4, 730–742. doi: 10.1002/emmm.201100995

Bonnet, A., Grosso, A. R., Elkaoutari, A., Coleno, E., Presle, A., Sridhara, S. C., et al. (2017). Introns protect eukaryotic genomes from transcription-associated genetic instability. *Mol. Cell.* 67:608-621.e11. doi: 10.1016/j.molcel.2017.07.002

Brosius, J., and Gould, S. J. (1992). On "genomenclature": a comprehensive (and respectful) taxonomy for pseudogenes and other "junk DNA". *Proc. Natl. Acad. Sci. U.S.A.* 89, 10706–10710.

Brouha, B., Schustak, J., Badge, R. M., Lutz-Prigge, S., Farley, A. H., Moran, J. V., et al. (2003). Hot L1s account for the bulk of retrotransposition in the human population. *Proc. Natl. Acad. Sci. U.S.A.* 100, 5280–5285.

Callinan, P. A., and Batzer, M. A. (2006). Retrotransposable elements and human disease. *Genome Dyn.* 1, 104–115.

Castro-Diaz, N., Ecco, G., Coluccio, A., Kapopoulou, A., Yazdanpanah, B., Friedli, M., et al. (2014). Evolutionally dynamic L1 regulation in embryonic stem cells. *Genes Dev.* 28, 1397–1409. doi: 10.1101/gad.241661.114

Catania, F. (2017). From intronization to intron loss: how the interplay between mRNA-associated processes can shape the architecture and the expression of eukaryotic genes. *Int. J. Biochem. Cell Biol.* 91, 136–144. doi: 10.1016/j.biocel.2017.06.017

Chow, H.-M., and Herrup, K. (2015). Genomic integrity and the ageing brain. *Nat. Rev. Neurosci.* 16, 672–684. doi: 10.1038/nrn4020

Cordaux, R., and Batzer, M. A. (2009). The impact of retrotransposons on human genome evolution. *Nat. Rev. Genet.* 10, 691–703. doi: 10.1038/nrg2640

Coufal, N. G., Garcia-Perez, J. L., Peng, G. E., Marchetto, M. C. N., Muotri, A. R., Mu, Y., et al. (2011). Ataxia telangiectasia mutated (ATM) modulates long interspersed element-1 (L1) retrotransposition in human neural stem cells. *Proc. Natl. Acad. Sci. U.S.A.* 108, 20382–20387. doi: 10.1073/pnas.1100273108

Davie, K., Janssens, J., Koldere, D., Waegeneer, M., de Pech, U., Kreft, L., et al. (2018). A single-cell transcriptome atlas of the aging drosophila brain. *Cell* 174, 982.e–998.e. doi: 10.1016/j.cell.2018.05.057

Djebali, S., Davis, C. A., Merkel, A., Dobin, A., Lassmann, T., Mortazavi, A., et al. (2012). Landscape of transcription in human cells. *Nature* 489, 101–108. doi: 10.1038/nature11233

Doucet, A. J., Wilusz, J. E., Miyoshi, T., Liu, Y., and Moran, J. V. (2015). A 3' Poly(A) tract is required for LINE-1 retrotransposition. *Mol. Cell.* 60, 728–741. doi: 10.1016/j.molcel.2015.10.012

Erwin, J. A., Paquola, A. C. M., Singer, T., Gallina, I., Novotny, M., Quayle, C., et al. (2016). L1-associated genomic regions are deleted in somatic cells of the healthy human brain. *Nat. Neurosci.* doi: 10.1038/nn.4388 [Epub ahead of print].

Esnault, C., Maestre, J., and Heidmann, T. (2000). Human LINE retrotransposons generate processed pseudogenes. *Nat. Genet.* 24, 363–367.

Evrony, G. D., Cai, X., Lee, E., Hills, L. B., Elhosary, P. C., Lehmann, H. S., et al. (2012). Single-neuron sequencing analysis of L1 retrotransposition and somatic mutation in the human brain. *Cell* 151, 483–496.

Evrony, G. D., Lee, E., Mehta, B. K., Benjamini, Y., Johnson, R. M., Cai, X., et al. (2015). Cell lineage analysis in human brain using endogenous retroelements. *Neuron* 85, 49–59. doi: 10.1016/j.neuron.2014.12.028

Feng, Q., Moran, J. V., Kazazian, H. H. J. R., and Boeke, J. D. (1996). Human L1 retrotransposon encodes a conserved endonuclease required for retrotransposition. *Cell* 87, 905–916.

Fischer, H.-G., Morawski, M., Bruckner, M. K., Mittag, A., Tarnok, A., and Arendt, T. (2012). Changes in neuronal DNA content variation in the human brain during aging. *Aging Cell* 11, 628–633. doi: 10.1111/j.1474-9726.2012.00826.x

Forsberg, L. A., Gisselsson, D., and Dumanski, J. P. (2017). Mosaicism in health and disease - clones picking up speed. *Nat. Rev. Genet.* 18, 128–142. doi: 10.1038/nrg.2016.145

Gagliardi, D., and Dziembowski, A. (2018). 5' and 3' modifications controlling RNA degradation: from safeguards to executioners. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* doi: 10.1098/rstb.2018.0160 [Epub ahead of print].

Garcia-Muse, T., and Aguilera, A. (2019). R loops: from physiological to pathological roles. *Cell* 179, 604–618. doi: 10.1016/j.cell.2019.08.055

Geller, L. N., and Potter, H. (1999). Chromosome missegregation and trisomy 21 mosaicism in Alzheimer's disease. *Neurobiol. Dis.* 6, 167–179.

Gomez-Herreros, F., Schuurs-Hoeijmakers, J. H. M., McCormack, M., Greally, M. T., Rulten, S., Romero-Granados, R., et al. (2014). TDP2 protects transcription from abortive topoisomerase activity and is required for normal neural function. *Nat. Genet.* 46, 516–521. doi: 10.1038/ng.2929

Green, P., Ewing, B., Miller, W., Thomas, P. J., and Green, E. D. (2003). Transcription-associated mutational asymmetry in mammalian evolution. *Nat. Genet.* 33, 514–517.

Greenberg, M. V. C., and Bourc'his, D. (2019). The diverse roles of DNA methylation in mammalian development and disease. Nature reviews. *Mol. Cell Biol.* 20, 590–607. doi: 10.1038/s41580-019-0159-6

Grzybowska, E. A. (2012). Human intronless genes: functional groups, associated diseases, evolution, and mRNA processing in absence of splicing. *Biochem. Biophys. Res. Commun.* 424, 1–6. doi: 10.1016/j.bbrc.2012.06.092

Guo, C., Jeong, H.-H., Hsieh, Y.-C., Klein, H.-U., Bennett, D. A., Jager, P. L., et al. (2018). Tau activates transposable elements in Alzheimer's disease. *Cell Rep.* 23, 2874–2880. doi: 10.1016/j.celrep.2018.05.004

Hammarlund, M., Hobert, O., Miller, D. M. III, and Sestan, N. (2018). The CeNGEN project: the complete gene expression map of an entire nervous system. *Neuron* 99, 430–433. doi: 10.1016/j.neuron.2018.07.042

Heyn, P., Kalinka, A. T., Tomancak, P., and Neugebauer, K. M. (2015). Introns and gene expression: cellular constraints, transcriptional regulation, and evolutionary consequences. *Bioessays* 37, 148–154. doi: 10.1002/bies.201400138

Hill, A. E., and Sorscher, E. J. (2006). The non-random distribution of intronless human genes across molecular function categories. *FEBS Lett.* 580, 4303–4305.

Holland, P. W. H., Marletaz, F., Maeso, I., Dunwell, T. L., and Paps, J. (2017). New genes from old: asymmetric divergence of gene duplicates and the evolution of development. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* doi: 10.1098/rstb.2015.0480 [Epub ahead of print].

Hube, F., and Francastel, C. (2015). Mammalian introns: when the junk generates molecular diversity. *Int. J. Mol. Sci.* 16, 4429–4452. doi: 10.3390/ijms16034429

Ingolia, N. T., Brar, G. A., Stern-Ginossar, N., Harris, M. S., Talhouarne, G. J. S., Jackson, S. E., et al. (2014). Ribosome profiling reveals pervasive translation outside of annotated protein-coding genes. *Cell Rep.* 8, 1365–1379. doi: 10.1016/j.celrep.2014.07.045

Iourov, I. Y., Vorsanova, S. G., Liehr, T., and Yurov, Y. B. (2009). Aneuploidy in the normal, Alzheimer's disease and ataxia-telangiectasia brain: differential expression and pathological meaning. *Neurobiol. Dis.* 34, 212–220. doi: 10.1016/j.nbd.2009.01.003

Ji, Z., Song, R., Regev, A., and Struhl, K. (2015). Many lncRNAs, 5'UTRs, and pseudogenes are translated and some are likely to express functional proteins. *eLife* 4:e08890. doi: 10.7554/eLife.08890

Jonsson, M. E., Ludvik Brattas, P., Gustafsson, C., Petri, R., Yudovich, D., Pircs, K., et al. (2019). Activation of neuronal genes via LINE-1 elements upon global DNA demethylation in human neural progenitors. *Nat. Commun.* 10:3182. doi: 10.1038/s41467-019-11150-8

Kaluski, S., Portillo, M., Besnard, A., Stein, D., Einav, M., Zhong, L., et al. (2017). Neuroprotective functions for the histone deacetylase SIRT6. *Cell Rep.* 18, 3052–3062. doi: 10.1016/j.celrep.2017.03.008

Kawamura, Y., Sanchez Calle, A., Yamamoto, Y., Sato, T.-A., and Ochiya, T. (2019). Extracellular vesicles mediate the horizontal transfer of an active LINE-1 retrotransposon. *J. Extracell. Vesicle* 8:1643214. doi: 10.1080/20013078.2019.1643214

Kazazian, H. H. JR., and Moran, J. V. (2017). Mobile DNA in Health and Disease. *N. Engl. J. Med.* 377, 361–370.

Kim, N., and Jinks-Robertson, S. (2012). Transcription as a source of genome instability. *Nat. Rev. Genet.* 13, 204–214. doi: 10.1038/nrg3152

Kokaeva, F. F., Den'mukhametova, S. V., Kanapin, A. A., Godukhin, O. V., Il'in, Y., and Ivanov, V. A. (2002). Antisence oligodeoxyribonucleotides for fragments of the reverse transcriptase gene of the LINE-1 element of rats disturb the formation of long-term memory. *Dokl. Biochem. Biophys.* 383, 93–95.

Kolisnyk, B., Al-Onaizi, M., Soreq, L., Barbash, S., Bekenstein, U., Haberman, N., et al. (2016). Cholinergic Surveillance over Hippocampal RNA Metabolism and Alzheimer's-Like Pathology. *Cereb. Cortex* doi: 10.1093/cercor/bhw177 [Epub ahead of print].

Kovalenko, T. F., and Patrushev, L. I. (2018). Pseudogenes as functionally significant elements of the genome. *Biochem. Biokhim.* 83, 1332–1349. doi: 10.1134/S0006297918110044

Kurnosov, A. A., Ustyugova, S. V., Nazarov, V. I., Minervina, A. A., Komkov, A. Y., Shugay, M., et al. (2015). The evidence for increased L1 activity in the site of human adult brain neurogenesis. *PLoS One* 10:e0117854. doi: 10.1371/journal.pone.0117854

Lake, B. B., Ai, R., Kaeser, G. E., Salathia, N. S., Yung, Y. C., Liu, R., et al. (2016). Neuronal subtypes and diversity revealed by single-nucleus RNA sequencing of the human brain. *Science* 352, 1586–1590. doi: 10.1126/science.aaf1204

Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., et al. (2001). Initial sequencing and analysis of the human genome. *Nature* 409, 860–921.

Lee, J. Y., Ji, Z., and Tian, B. (2008). Phylogenetic analysis of mRNA polyadenylation sites reveals a role of transposable elements in evolution of the 3'-end of genes. *Nucleic Acids Res.* 36, 5581–5590. doi: 10.1093/nar/gkn540

Lee, M.-H., and Chun, J. (2019). Mosaic APP gene recombination in Alzheimer's disease-what's next? *J. Exp. Neurosci.* 13:1179069519849669. doi: 10.1177/1179069519849669

Lee, M.-H., Siddoway, B., Kaeser, G. E., Segota, I., Rivera, R., Romanow, W. J., et al. (2018). Somatic APP gene recombination in Alzheimer's disease and normal neurons. *Nature* 563, 639–645. doi: 10.1038/s41586-018-0718-6

Lim, L., Mi, D., Llorca, A., and Marin, O. (2018). Development and functional diversification of cortical interneurons. *Neuron* 100, 294–313. doi: 10.1016/j.neuron.2018.10.009

Liu, W., Bai, X., Zhang, A., Huang, J., Xu, S., and Zhang, J. (2019). Role of exosomes in central nervous system diseases. *Front. Mol. Neurosci.* 12:240. doi: 10.3389/fnmol.2019.00240

Louhichi, A., Fourati, A., and Rebai, A. (2011). IGD: a resource for intronless genes in the human genome. *Gene* 488, 35–40. doi: 10.1016/j.gene.2011.08.013

MacGowan, D. J. L., Scelsa, S. N., Imperato, T. E., Liu, K.-N., Baron, P., and Polsky, B. (2007). A controlled study of reverse transcriptase in serum and CSF of HIV-negative patients with ALS. *Neurology* 68, 1944–1946.

Macia, A., Widmann, T. J., Heras, S. R., Ayllon, V., Sanchez, L., Benkaddour-Boumzaouad, M., et al. (2017). Engineered LINE-1 retrotransposition in nondividing human neurons. *Genome Res.* 27, 335–348. doi: 10.1101/gr.206805.116

Madabhushi, R., Gao, F., Pfenning, A. R., Pan, L., Yamakawa, S., Seo, J., et al. (2015). Activity-induced DNA breaks govern the expression of neuronal early-response genes. *Cell* 161, 1592–1605. doi: 10.1016/j.cell.2015.05.032

Majewski, J. (2003). Dependence of mutational asymmetry on gene-expression levels in the human genome. *Am. J. Hum. Genet.* 73, 688–692.

Marques, A. C., Dupanloup, I., Vinckenbosch, N., Reymond, A., and Kaessmann, H. (2005). Emergence of young human genes after a burst of retroposition in primates. *PLoS Biol.* 3:e357.

Mastroeni, D., Grover, A., Delvaux, E., Whiteside, C., Coleman, P. D., and Rogers, J. (2010). Epigenetic changes in Alzheimer's disease: decrements in DNA methylation. *Neurobiol. Aging* 31, 2025–2037. doi: 10.1016/j.neurobiolaging.2008.12.005

Mathias, S. L., Scott, A. F., Kazazian, H. H. J. R., Boeke, J. D., and Gabriel, A. (1991). Reverse transcriptase encoded by a human transposable element. *Science* 254, 1808–1810.

McConnell, M. J., Lindberg, M. R., Brennand, K. J., Piper, J. C., Voet, T., Cowing-Zitron, C., et al. (2013). Mosaic copy number variation in human neurons. *Science* 342, 632–637. doi: 10.1126/science.1243472

McConnell, M. J., Moran, J. V., Abyzov, A., Akbarian, S., Bae, T., Cortes-Ciriano, I., et al. (2017). Intersection of diverse neuronal genomes and neuropsychiatric disease: The Brain Somatic Mosaicism Network. *Science* doi: 10.1126/science.aal1641 [Epub ahead of print].

Meischl, C., Boer, M., Ahlin, A., and Roos, D. (2000). A new exon created by intronic insertion of a rearranged LINE-1 element as the cause of chronic granulomatous disease. *EJHG* 8, 697–703.

Miura, P., Shenker, S., Andreu-Agullo, C., Westholm, J. O., and Lai, E. C. (2013). Widespread and extensive lengthening of 3' UTRs in the mammalian brain. *Genome Res.* 23, 812–825. doi: 10.1101/gr.146886.112

Mladenov, E., Magin, S., Soni, A., and Iliakis, G. (2016). DNA double-strand-break repair in higher eukaryotes and its role in genomic instability and cancer: cell cycle and proliferation-dependent regulation. *Semin. Cancer Biol.* 3, 51–64. doi: 10.1016/j.semcancer.2016.03.003

Moran, J. V., Holmes, S. E., Naas, T. P., DeBerardinis, R. J., Boeke, J. D., and Kazazian, H. H. J. R. (1996). High frequency retrotransposition in cultured mammalian cells. *Cell* 87, 917–927.

Mosch, B., Mittag, A., Lenz, D., Arendt, T., and Tarnok, A. (2006). Laser scanning cytometry in human brain slices. *Cytometry A* 69, 135–138.

Mosch, B., Morawski, M., Mittag, A., Lenz, D., Tarnok, A., and Arendt, T. (2007). Aneuploidy and DNA replication in the normal human brain and Alzheimer's disease. *J. Neurosci.* 27, 6859–6867.

Muotri, A. R., Chu, V. T., Marchetto, M. C. N., Deng, W., Moran, J. V., and Gage, F. H. (2005). Somatic mosaicism in neuronal precursor cells mediated by L1 retrotransposition. *Nature* 435, 903–910.

Muotri, A. R., Zhao, C., Marchetto, M. C. N., and Gage, F. H. (2009). Environmental influence on L1 retrotransposons in the adult hippocampus. *Hippocampus* 19, 1002–1007. doi: 10.1002/hipo.20564

Myers, J. S., Vincent, B. J., Udall, H., Watkins, W. S., Morrish, T. A., Kilroy, G. E., et al. (2002). A comprehensive analysis of recently integrated human Ta L1 elements. *Am. J. Hum. Genet.* 71, 312–326.

Neil, A. J., Liang, M. U., Khristich, A. N., Shah, K. A., and Mirkin, S. M. (2018). RNA-DNA hybrids promote the expansion of Friedreich's ataxia (GAA)n repeats via break-induced replication. *Nucleic Acids Res.* 46, 3487–3497. doi: 10.1093/nar/gky099

Nilsen, T. W., and Graveley, B. R. (2010). Expansion of the eukaryotic proteome by alternative splicing. *Nature* 463, 457–463. doi: 10.1038/nature08909

Orlowski, C., Mah, L.-J., Vasireddy, R. S., El-Osta, A., and Karagiannis, T. C. (2011). Double-strand breaks and the concept of short- and long-term epigenetic memory. *Chromosoma* 120, 129–149. doi: 10.1007/s00412-010-0305-6

Paco, A., Adega, F., and Chaves, R. (2015). LINE-1 retrotransposons: from 'parasite' sequences to functional elements. *J. Appl. Genet.* 56, 133–145. doi: 10.1007/s13353-014-0241-x

Paquola, A. C. M., Erwin, J. A., and Gage, F. H. (2017). Insights into the role of somatic mosaicism in the brain. *Curr. Opin. Syst. Biol.* 1, 90–94. doi: 10.1016/j.coisb.2016.12.004

Park, S. Y., Seo, J., and Chun, Y. S. (2019). Targeted Downregulation of kdm4a Ameliorates Tau-engendered Defects in Drosophila melanogaster. *J. Korean Med. Sci.* 34, e225. doi: 10.3346/jkms.2019.34.e225

Polak, P., and Arndt, P. F. (2008). Transcription induces strand-specific mutations at the 5' end of human genes. *Genome Res.* 18, 1216–1223. doi: 10.1101/gr.076570.108

Poliseno, L. (2012). Pseudogenes: newly discovered players in human cancer. *Sci. Signal.* 5:re5.

Potter, H. (1991). Review and hypothesis: Alzheimer disease and Down syndrome–chromosome 21 nondisjunction may underlie both disorders. *Am. J. Hum. Genet.* 48, 1192–1200.

Potter, H., and Geller, L. N. (1996). Alzheimer's disease. Down's syndrome, and chromosome segregation. *Lancet* 348:66.

Protasova, M. S., Gusev, F. E., Grigorenko, A. P., Kuznetsova, I. L., Rogaev, E. I., and Andreeva, T. V. (2017). Quantitative analysis of L1-retrotransposons in Alzheimer's disease and aging. *Biochem. Biokhim.* 82, 962–971. doi: 10.1134/S0006297917080120

Raj, T., Li, Y. I., Wong, G., Humphrey, J., Wang, M., Ramdhani, S., et al. (2018). Integrative transcriptome analyses of the aging brain implicate altered splicing in Alzheimer's disease susceptibility. *Nat. Genet.* 50, 1584–1592. doi: 10.1038/s41588-018-0238-1

Rehen, S. K., McConnell, M. J., Kaushal, D., Kingsbury, M. A., Yang, A. H., and Chun, J. (2001). Chromosomal variation in neurons of the developing and adult mammalian nervous system. *Proc. Natl. Acad. Sci. U.S.A* 98, 13361–13366.

Rehen, S. K., Yung, Y. C., McCreight, M. P., Kaushal, D., Yang, A. H., Almeida, B. S., et al. (2005). Constitutional aneuploidy in the normal human brain. *J. Neurosci.* 25, 2176–2180.

Renthal, W., Boxer, L. D., Hrvatin, S., Li, E., Silberfeld, A., Nagy, M. A., et al. (2018). Characterization of human mosaic Rett syndrome brain tissue by single-nucleus RNA sequencing. *Nat. Neurosci.* 21, 1670–1679. doi: 10.1038/s41593-018-0270-6

Richard, P., and Manley, J. L. (2017). R loops and links to human disease. *J. Mol. Biol.* 429, 3168–3180. doi: 10.1016/j.jmb.2016.08.031

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

Rohrback, S., Siddoway, B., Liu, C. S., and Chun, J. (2018b). Genomic mosaicism in the developing and adult brain. *Dev. Neurobiol.* 78, 1026–1048. doi: 10.1002/dneu.22626

Rondon, A. G., and Aguilera, A. (2019). What causes an RNA-DNA hybrid to compromise genome integrity? *DNA Repair* 81:102660. doi: 10.1016/j.dnarep.2019.102660

Rouhana, L., Wang, L., Buter, N., Kwak, J. E., Schiltz, C. A., Gonzalez, T., et al. (2005). Vertebrate GLD2 poly(A) polymerases in the germline and the brain. *RNA* 11, 1117–1130.

Rowe, H. M., Jakobsson, J., Mesnard, D., Rougemont, J., Reynard, S., Aktas, T., et al. (2010). KAP1 controls endogenous retroviruses in embryonic stem cells. *Nature* 463, 237–240. doi: 10.1038/nature08674

Ryu, J. Y., Kim, H. U., and Lee, S. Y. (2015). Human genes with a greater number of transcript variants tend to show biological features of housekeeping and essential genes. *Mol. Biosyst.* 11, 2798–2807. doi: 10.1039/c5mb00322a

Saleh, A., Macia, A., and Muotri, A. R. (2019). Transposable elements. inflammation, and neurological disease. *Front. Neurol.* 10:894. doi: 10.3389/fneur.2019.00894

Salvador-Palomeque, C., Sanchez-Luque, F. J., Fortuna, P. R. J., Ewing, A. D., Wolvetang, E. J., Richardson, S. R., et al. (2019). Dynamic methylation of an L1 transduction family during reprogramming and neurodifferentiation. *Mol. Cell. Biol.* doi: 10.1128/MCB.00499-18 [Epub ahead of print].

Schwahn, U., Lenzner, S., Dong, J., Feil, S., Hinzmann, B., van Duijnhoven, G., et al. (1998). Positional cloning of the gene for X-linked retinitis pigmentosa 2. *Nat. Genet.* 19, 327–332.

Shabalina, S. A., Ogurtsov, A. Y., Spiridonov, A. N., Novichkov, P. S., Spiridonov, N. A., and Koonin, E. V. (2010). Distinct patterns of expression and evolution of intronless and intron-containing mammalian genes. *Mol. Biol. Evol.* 27, 1745–1749. doi: 10.1093/molbev/msq086

Shaul, O. (2017). How introns enhance gene expression. *Int. J. Of Biochem. Cell Biol.* 91, 145–155. doi: 10.1016/j.biocel.2017.06.016

Singer, T., McConnell, M. J., Marchetto, M. C. N., Coufal, N. G., and Gage, F. H. (2010). LINE-1 retrotransposons: mediators of somatic variation in neuronal genomes? *Trends Neurosci.* 33, 345–354. doi: 10.1016/j.tins.2010.04.001

Steele, A. J., Al-Chalabi, A., Ferrante, K., Cudkowicz, M. E., Brown, R. H. J. R., and Garson, J. A. (2005). Detection of serum reverse transcriptase activity in patients with ALS and unaffected blood relatives. *Neurology* 64, 454–458.

Steranka, J. P., Tang, Z., Grivainis, M., Huang, C. R. L., Payer, L. M., Rego, F. O. R., et al. (2019). Transposon insertion profiling by sequencing (TIPseq) for mapping LINE-1 insertions in the human genome. *Mob. DNA* 10:8.

Suberbielle, E., Sanchez, P. E., Kravitz, A. V., Wang, X., Ho, K., Eilertson, K., et al. (2013). Physiologic brain activity causes DNA double-strand breaks in neurons, with exacerbation by amyloid-beta. *Nat. Neurosci.* 16, 613–621. doi: 10.1038/nn.3356

Sun, Y., Li, Y., Luo, D., and Liao, D. J. (2012). Pseudogenes as weaknesses of ACTB (Actb) and GAPDH (Gapdh) used as reference genes in reverse transcription and polymerase chain reactions. *PLoS One* 7:e41659. doi: 10.1371/journal.pone.0041659

Swanger, S. A., He, Y. A., Richter, J. D., and Bassell, G. J. (2013). Dendritic GluN2A synthesis mediates activity-induced NMDA receptor insertion. *J. Neurosci.* 33, 8898–8908. doi: 10.1523/JNEUROSCI.0289-13.2013

Swergold, G. D. (1990). Identification, characterization, and cell specificity of a human LINE-1 promoter. *Mol. Cell. Biol.* 10, 6718–6729.

Taliaferro, J. M., Vidaki, M., Oliveira, R., Olson, S., Zhan, L., Saxena, T., et al. (2016). Distal alternative last exons localize mRNAs to neural projections. *Mol. Cell.* 61, 821–833. doi: 10.1016/j.molcel.2016.01.020

Tian, B., and Manley, J. L. (2017). Alternative polyadenylation of mRNA precursors. Nature reviews. *Mol. Cell Biol.* 18, 18–30. doi: 10.1038/nrm.2016.116

Tudek, A., Lloret-Llinares, M., and Jensen, T. H. (2018). The multitasking polyA tail: nuclear RNA maturation, degradation and export. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* doi: 10.1098/rstb.2018.0169 [Epub ahead of print].

Upton, K. R., Baillie, J. K., and Faulkner, G. J. (2011). Is somatic retrotransposition a parasitic or symbiotic phenomenon? *Mob. Genet. Elements* 1, 279–282.

Upton, K. R., Gerhardt, D. J., Jesuadian, J. S., Richardson, S. R., Sanchez-Luque, F. J., Bodea, G. O., et al. (2015). Ubiquitous L1 mosaicism in hippocampal neurons. *Cell* 161, 228–239. doi: 10.1016/j.cell.2015.03.026

van Meter, M., Kashyap, M., Rezazadeh, S., Geneva, A. J., Morello, T. D., Seluanov, A., et al. (2014). SIRT6 represses LINE1 retrotransposons by ribosylating KAP1 but this repression fails with stress and age. *Nat. Commun.* 5:5011. doi: 10.1038/ncomms6011

Verheijen, B. M., Vermulst, M., and van Leeuwen, F. W. (2018). Somatic mutations in neurons during aging and neurodegeneration. *Acta Neuropathol.* 135, 811–826. doi: 10.1007/s00401-018-1850-y

Wang, Y., Xie, Y., Li, L., He, Y., Zheng, D., Yu, P., et al. (2018). EZH2 RIP-seq identifies tissue-specific long non-coding RNAs. *Curr.Gene Ther.* doi: 10.2174/1566523218666181008125010 [Epub ahead of print].

Warkocki, Z., Liudkovska, V., Gewartowska, O., Mroczek, S., and Dziembowski, A. (2018). Terminal nucleotidyl transferases (TENTs) in mammalian RNA metabolism. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* doi: 10.1098/rstb.2018.0162 [Epub ahead of print].

Wei, W., Gilbert, N., Ooi, S. L., Lawler, J. F., Ostertag, E. M., Kazazian, H. H., et al. (2001). Human L1 retrotransposition: cis preference versus trans complementation. *Mol. Cell. Biol.* 21, 1429–1439.

Westra, J. W., Rivera, R. R., Bushman, D. M., Yung, Y. C., Peterson, S. E., Barral, S., et al. (2010). Neuronal DNA content variation (DCV) with regional and individual differences in the human brain. *J. Comp. Neurol.* 518, 3981–4000. doi: 10.1002/cne.22436

Xiang, Y., Yan, K., Zheng, Q., Ke, H., Cheng, J., Xiong, W., et al. (2019). Histone Demethylase KDM4B promotes DNA damage by activating long interspersed nuclear element-1. *Cancer Res.* 79, 86–98. doi: 10.1158/0008-5472.CAN-18-1310

Xiong, J., Jiang, X., Ditsiou, A., Gao, Y., Sun, J., Lowenstein, E. D., et al. (2018). Predominant patterns of splicing evolution on human, chimpanzee and

macaque evolutionary lineages. *Hum. Mol. Genet.* 27, 1474–1485. doi: 10.1093/hmg/ddy058

Yu, K., Chedin, F., Hsieh, C.-L., Wilson, T. E., and Lieber, M. R. (2003). R-loops at immunoglobulin class switch regions in the chromosomes of stimulated B cells. *Nat. Immunol.* 4, 442–451.

Zhao, C., and Pyle, A. M. (2017). The group II intron maturase: a reverse transcriptase and splicing factor go hand in hand. *Curr. Opin. Struct. Biol* 47, 30–39. doi: 10.1016/j.sbi.2017.05.002

Zhong, Y., Du, Y., Yang, X., Mo, Y., Fan, C., Xiong, F., et al. (2018). Circular RNAs function as ceRNAs to regulate and control human cancer progression. *Mol. Cancer* 17:79. doi: 10.1186/s12943-018-0827-8

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Genetics

PERSPECTIVE
published: 07 May 2020
doi: 10.3389/fgene.2020.00390

Mosaic Somatic Gene Recombination as a Potentially Unifying Hypothesis for Alzheimer's Disease

*Gwendolyn E. Kaeser and Jerold Chun**

Sanford Burnham Prebys Medical Discovery Institute, La Jolla, CA, United States

The recent identification of somatic gene recombination (SGR) in human neurons affecting the well-known Alzheimer's disease (AD) pathogenic gene, amyloid precursor protein (APP), has implications for the normal and the diseased human brain. The amyloid hypothesis has been the prevailing theory for sporadic AD (SAD) pathogenesis since the discovery of *APP* gene involvement in familial AD and Down syndrome. Yet, despite enormous scientific and clinical effort, no disease-modifying therapy has emerged. SGR offers a novel mechanism to explain AD pathogenesis and the failures of amyloid-related clinical trials, while maintaining consistency with most aspects of the amyloid hypothesis and additionally supporting possible roles for tau, oxidative stress, inflammation, infection, and prions. SGR retro-inserts novel "genomic complementary DNAs" (gencDNAs) into neuronal genomes and becomes dysregulated in SAD, producing numerous mosaic *APP* variants, including DNA mutations observed in familial AD. Notably, SGR requires gene transcription, DNA strand-breaks, and reverse transcriptase (RT) activity, all of which may be promoted by well-known AD risk factors and provide a framework for the pursuit of new SGR-based therapeutics. In this perspective, we review evidence for *APP* SGR in AD pathogenesis and discuss its possible relevance to other AD-related dementias. Further, SGR's requirement for RT activity and the relative absence of AD in aged HIV -infected patients exposed to RT inhibitors suggest that these Food and Drug Administration (FDA)-approved drugs may represent a near-term disease-modifying therapy for AD.

Keywords: Alzheimer's disease, mosaicism, somatic gene recombination, amyloid cascade hypothesis, gencDNA, amyloid precursor protein, APP

OPEN ACCESS

Edited by:
Svetlana G. Vorsanova, Veltischev Research and Clinical Institute for Pediatrics of the Pirogov Russian National Research Medical University, Russia

Reviewed by:
Thomas Liehr, Friedrich Schiller University Jena, Germany
Fabio Coppedè, University of Pisa, Italy
Antoneta Granic, Newcastle University, United Kingdom

***Correspondence:**
Jerold Chun
jchun@sbpdiscovery.org

Specialty section:
This article was submitted to Human Genomics, a section of the journal Frontiers in Genetics

Received: *12 December 2019*
Accepted: *27 March 2020*
Published: *07 May 2020*

Citation:
Kaeser GE and Chun J (2020) Mosaic Somatic Gene Recombination as a Potentially Unifying Hypothesis for Alzheimer's Disease. Front. Genet. 11:390. doi: 10.3389/fgene.2020.00390

GENOMIC MOSAICISM AT THE *APP* LOCUS

We first speculated that SGR might exist in the brain based upon the expression of immunological recombination genes, as described over a quarter century ago for recombination activating gene-1 (Chun et al., 1991) and later, non-homologous end-joining genes (Gao et al., 1998). Subsequent studies to identify somatically generated genomic mosaicism in the human brain identified chromosomal aneuploidies that represent large CNVs

Abbreviations: AD, Alzheimer's disease; APP, amyloid precursor protein; CNV, copy number variation; DISH, DNA *in situ* hybridization; gencDNA, genomic complementary DNA; PNA-FISH, peptide nucleic acid fluorescent *in situ* hybridization; RT, reverse transcriptase; SAD, sporadic AD; SGR, somatic gene recombination.

(Rehen et al., 2001). The application of newer technologies including fluorescence-activated nuclear sorting (Rehen et al., 2005; Westra et al., 2010) and single-cell sequencing expanded the discovery of somatically arising genomic mosaicism forms, revealing an immense diversity of DNA sequence differences present among single cells (reviewed in Rohrback et al., 2018). This includes Jackson Pollock-like displays reflective of enormous single-cell transcriptome diversity in the brain (Lake et al., 2016, 2018) that is consistent with genomic mosaicism. Neuronal genomic mosaicism takes many forms including aneuploidies, CNVs, single nucleotide variations (SNVs), and long interspersed nuclear element 1 (LINE1). Some of these have been associated with neurodegenerative (including AD) and neuropsychiatric disorders, which have been reviewed extensively and will not be the subject of this perspective (Arendt et al., 2009; Leija-Salazar et al., 2018; Rohrback et al., 2018; Shepherd et al., 2018; Iourov et al., 2019; Potter et al., 2019).

Although the existence of genomic mosaicism is now established, its functions are less clear. Roles in transcriptomic regulation (Kaushal et al., 2003), cell survival (Peterson et al., 2012), and neural circuits (Kingsbury et al., 2005) have been reported, and others have speculated on the importance of genomic mosaicism in the creation of neuronal diversity (Rehen et al., 2001, 2005; Muotri and Gage, 2006; Gericke, 2008), yet these general phenomena did not reveal effects on specific genes or DNA alterations that might be analogous to V(D)J recombination in the immune system (Papavasiliou and Schatz, 2002). However, a candidate gene emerged when we observed increases in a major sub-type of mosaicism called "DNA content variation" (Westra et al., 2010) in SAD neurons of the prefrontal cerebral cortex, where SAD neurons contained ∼500 megabase pairs more DNA than the non-diseased controls (Bushman et al., 2015). We reasoned that the increase could affect *APP*, a key gene in AD pathogenesis that is causal in familial AD and Down syndrome through mutations and, in particular, CNVs: mosaically increased *APP* CNVs in SAD brains may drive pathology. This possibility was confirmed using multiple approaches including PNA-FISH, small-population qPCR, and single-neuron qPCR, which demonstrated that somatic and mosaic changes to the *APP* locus were enriched in SAD neurons over non-diseased controls and were not associated with trisomy of chromosome 21 (Bushman et al., 2015). Interestingly, PNA-FISH targeting individual *APP* exons and exon–exon copy number discordance by single-cell qPCR suggested that the physical arrangement of *APP* CNVs could be non-uniform (Bushman et al., 2015).

Additional studies confirmed this possibility and revealed SGR at the *APP* locus (**Figure 1A**), occurring as variant *APP* coding sequences that lacked introns and were akin to complementary DNA (cDNA) sequences except that they were present in genomic DNA and were therefore termed "gencDNAs" (**Figure 1B**) (Lee et al., 2018). These novel gencDNAs were further characterized by intra-exonic junctions with shared microhomology regions between the two joined exonic regions. Identical forms were also documented in mRNAs. The formation of *APP* gencDNAs *in vitro* required *APP* transcription, DNA strand breakage, and RT activity. Neuronal SGR represents a novel mechanism to produce genomic mosaicism that has functional implications, particularly for AD pathogenesis and therapeutics, while suggesting a more general paradigm underlying sporadic brain diseases through dysregulated SGR of both known and unknown pathogenic genes.

SGR AFFECTING *APP* IS DYSREGULATED IN SAD BRAINS

At least 12 distinct approaches, including non-targeted and unbiased methods, were used to identify and validate somatic mosaic events at the *APP* locus (Bushman et al., 2015; Lee et al., 2018; Lee et al., 2019). SGR was identified in both normal and diseased brains but appears to be dysregulated with disease, resulting in dramatic increases to both the number and the form of *APP* gencDNAs in SAD neurons. Novel approaches were utilized, including DISH and high-fidelity, long-read sequencing to establish disease alterations.

DISH was developed by modifying BaseScope (Advanced Cell Diagnostics, Fremont, CA, United States) technology that can detect SNVs (Baker et al., 2017) and exon:exon junctions in RNA but was adapted for use on genomic DNA. Probes were designed to target multiple gencDNA sequences, including the exon16:exon17 junction and the intra-exonic junction formed by the microhomology fusion of exon 3 to exon 16 (**Figure 1B**). Several parallel approaches were employed, including sense and antisense probes that demonstrated DNA specificity (*vs.* RNA that is not recognized by sense-strand probes), targeted restriction enzyme digestion that effectively destroyed the DNA target locus and dependent signal, DISH double-labeling that indicated that gencDNA loci are distinct from the endogenous alleles, and the use of synthetic targets in cell culture that confirmed probe specificity (Lee et al., 2018). SAD brains exhibited an average of 1.2–1.8 gencDNAs per nucleus, with 60–70% of prefrontal SAD cortical neuronal nuclei having at least one signal. In contrast, the control brains averaged 0.4 gencDNAs per nucleus, and only 25% of nuclei had at least one signal. The three- to fourfold increase in SAD gencDNA number was consistent throughout all biological and technical replicates (six SAD and six non-diseased brains; three experiments per brain). Notably, detection by this technique is limited to the targeted exon:exon sequence or intra-exonic junction and, therefore, does not capture the full diversity of possible gencDNAs, including full-length or more complex structural variants.

The second novel approach to assess gencDNAs and identify disease-related differences was high-fidelity, long-read sequencing with Pacific Biosciences' single-molecule real-time circular consensus sequencing (SMRT CCS or "PacBio sequencing") (Eid et al., 2009; Hebert et al., 2018), which is capable of identifying SNVs with 99.999999% confidence. *APP*-targeted PCR products were amplified from neurons (five SAD and five non-diseased brains) and sequenced. These experiments revealed enormous gencDNA diversity involving thousands of unique species. Importantly, gencDNA sequences changed significantly with disease, despite identical PCR targeting that involved amplification with exon 1 and 18 primers (myriad other

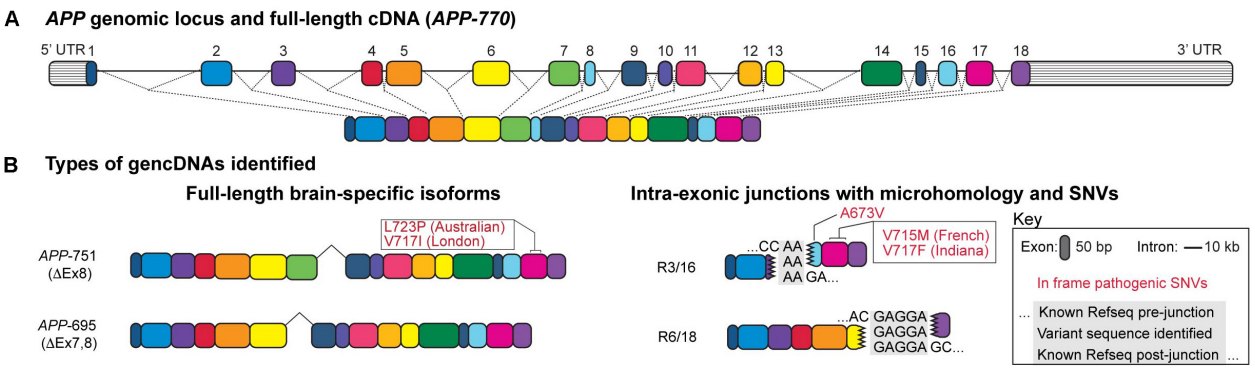

FIGURE 1 | Structure of a gencDNA. **(A)** The *APP* genomic locus and the exons within a full-length cDNA. **(B)** Two types of gencDNAs were identified in both RNA and DNA: full-length brain-specific isoforms (APP-751 and APP-695) and truncated sequences with intra-exonic junctions and microhomology domains (R3/16 and R6/18). Known pathogenic SNVs were also identified in some variants (in-frame examples include Australian, London, French, Indiana, and A673V; shown in red). Figure modified from Lee et al. (2018).

species may exist). The SAD brains had 10 times more unique reads per neuronal nucleus, and we identified 45 unique intra-exonic junctions in SAD brains, contrasting with just 20 unique intra-exonic junctions found in non-diseased brains despite using ∼70% more neuronal nuclei. Most remarkably, PacBio sequencing identified 11 mosaic SNVs in or around the Aβ encoding region that are considered to be disease-causing in familial AD, present only in SAD neurons. The results from DISH and PacBio sequencing together confirm that normal human neurons display a baseline level of *APP* gencDNAs that is increased and fundamentally altered in number and form with AD, including the formation of pathogenic SNVs. Independent support for gencDNAs was recently published by an unrelated laboratory (Lee et al., 2019; Park et al., 2019). Preliminary analyses of these data identified diverse integration sites for gencDNAs on multiple chromosomes (Lee et al., 2019) and are consistent with DISH signals that were distinct from wild-type (chromosome 21) alleles (Lee et al., 2018).

MOLECULAR DIVERSITY PRODUCED BY SGR MAY LINK MULTIPLE AD HYPOTHESES

Somatic gene recombination likely has normal functions; however, it appears to be dysregulated in AD, most likely during the proposed cellular phase of AD (critically reviewed in De Strooper and Karran, 2016). SGR could create variant *APP* sequences that become translated into heterogeneous populations of APP variant and Aβ-like proteins—in addition to serving as more classical secretase substrates to generate Aβ—that could result in myriad downstream biochemical processes, as was reported for AD. SGR of *APP* could have effects on primary, secondary, tertiary, or quaternary protein structure and therefore could have a vast array of functional effects, including those related to prions. The heterogeneity of *APP* forms produced by SGR invokes modification of the amyloid hypothesis to integrate this new feature while still maintaining decades of supportive observations. SGR also accounts for experimental discrepancies and clinical trial failures. In doing so, it may unify other hypotheses of SAD etiology and pathogenesis *via* a modified amyloid hypothesis (**Figure 2**). Other pathogenic actions of SGR, such as those produced by the integration of mobile elements, may also be relevant. The initial views on the implications of SGR in AD (Castro et al., 2019; Lee and Chun, 2019) are expanded upon next.

Amyloid Cascade Hypothesis Modified by Somatic Gene Recombination

The amyloid cascade hypothesis (or the amyloid hypothesis) has been the predominant AD theory for decades (Hardy and Higgins, 1992; Hardy and Selkoe, 2002; Selkoe and Hardy, 2016), having emerged through the identification of amyloid-β (Aβ) as the plaque-forming peptide from AD and Down syndrome brains ∼35 years ago (Glenner and Wong, 1984a,b), which allowed the subsequent identification of *APP* as the gene locus responsible for Aβ (Goldgaber et al., 1987; Tanzi et al., 1987b). The strongest evidence for the involvement of *APP* and its cleavage product Aβ in AD comes from familial AD and Down syndrome studies. Familial AD exhibits typical Mendelian inheritance of mutations or CNVs in *APP* or mutations in the secretase genes, PSEN1 (Sherrington et al., 1995) and PSEN2 (Levy-Lahad et al., 1995; Rogaev et al., 1995) that alter Aβ processing and lead to early-onset AD. *APP* shows a clear gene dosage effect, where three copies of *APP* in Down syndrome (Wiseman et al., 2015) or rare familial AD cases (Rovelet-Lecrux et al., 2006; Sleegers et al., 2006; Hooli et al., 2012) are sufficient to produce AD neuropathology and/or symptomology. In the amyloid hypothesis, the accumulation of Aβ in the brain and its aggregation into plaques result in downstream processes that lead to hyperphosphorylation of tau, resultant neurofibrillary tangles (NFTs), synaptic dysfunction, cell death, and ultimately AD.

One major criticism of the amyloid hypothesis is the timing of plaque deposition. Aβ plaques do not necessarily correlate well with cognitive impairment, and many individuals have

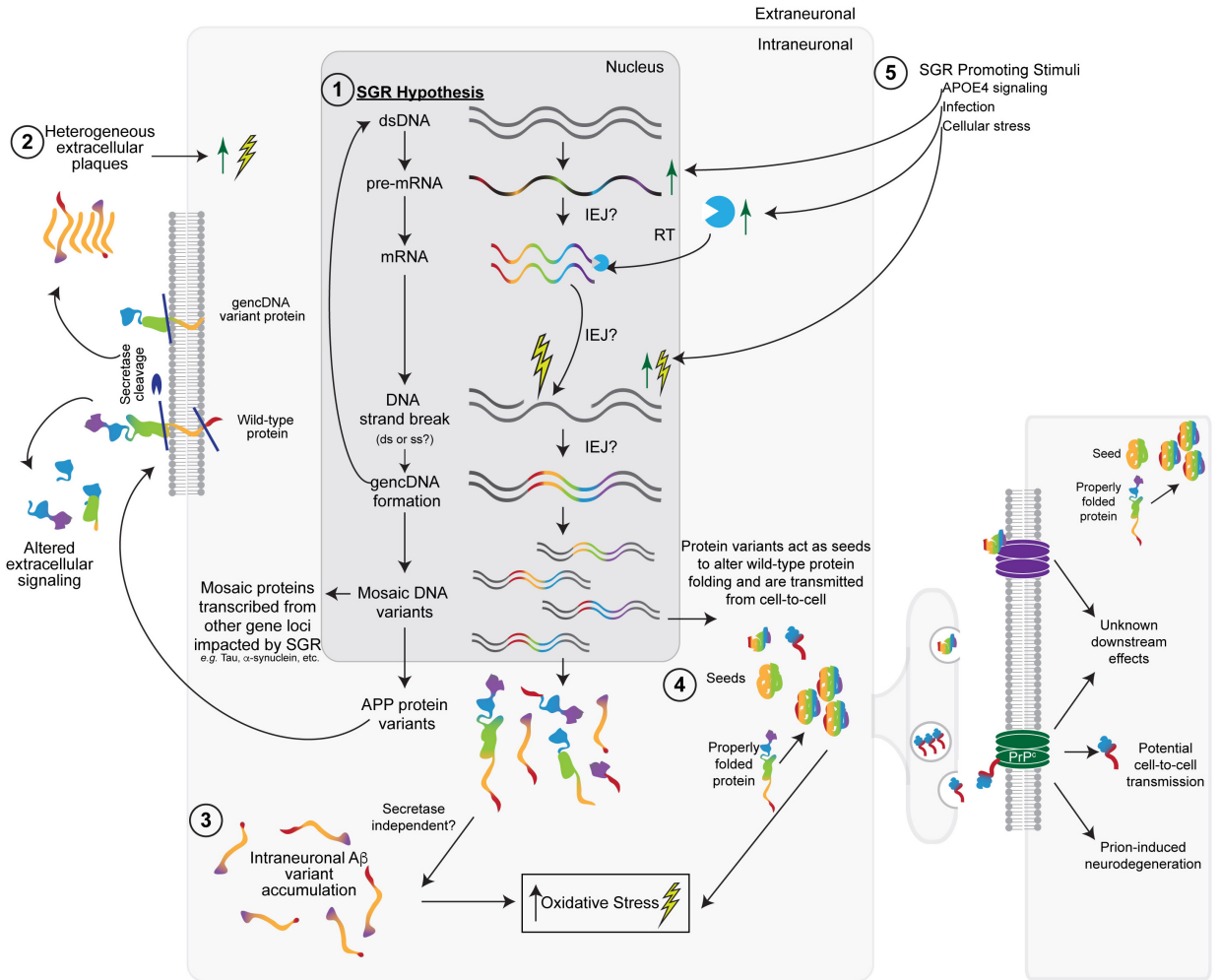

FIGURE 2 | The somatic gene recombination (SGR) hypothesis for SAD. **(1)** Dysregulated SGR of *APP* through the insertion of reverse-transcribed mRNA leads to mosaic genomic *APP* variants that result in variant proteins with a number of downstream effects. Other gene loci may also be impacted by SGR. **(2)** In keeping with the Aβ-hypothesis, some variants would be transported to the cellular membrane, where wild-type, SNV-containing, and gencDNA variant proteins may or may not be cleaved by the traditional secretase pathways to produce heterogeneous extracellular plaques and altered extracellular signaling pathways. **(3)** Intracellular variants may also accumulate without the need for secretase cleavage. The accumulation of intraneuronal Aβ variants likely increases cellular (oxidative) stress, leading to an increase in DNA strand breaks, the insertion of gencDNAs, and the production of variants, creating a feed-forward loop that promotes a disease. **(4)** Variant proteins may also act as "seeds," which alter the conformation of wild-type APP or other gencDNA variant proteins to create more aggregates. These may be propagated from cell-to-cell and cause prion-like transfer and neurodegeneration through the possible involvement of prion protein receptors (PrP). **(5)** Various stimuli (e.g., APOE4, infection, and cellular stress) could promote SGR *via* actions at multiple steps, including increased *APP* transcription.

abundant Aβ deposits at death and yet were cognitively normal antemortem (Crystal et al., 1988; Katzman et al., 1988; Troncoso et al., 1996). These plaques tend to be more diffuse with lower levels of Aβ oligomers (Esparza et al., 2013), which suggests that pathogenic plaques sequester toxic oligomers (Selkoe and Hardy, 2016). Additionally, there is evidence that the duration of plaque deposition is more predictive of SAD (Insel et al., 2017) rather than the presence of plaques *per se*. Another major criticism of the amyloid hypothesis has been its failure to yield any disease-modifying therapies despite the many clinical trials targeting Aβ components (Carlsson, 2008; Cummings et al., 2014; Anderson et al., 2017; Mehta et al., 2017; Burki, 2018; Egan et al., 2018). These discrepancies and others have led to calls to abandon or fundamentally rethink the amyloid hypothesis (Herrup, 2015; Morris et al., 2018).

However, these major discrepancies may be explained by SGR. SGR may generate diverse APP protein products, including Aβ and Aβ-like molecules, on a continuum of toxicity within plaques and as soluble proteins. Therefore, SGR of *APP* incorporates supportive data for the amyloid hypothesis in SAD by vastly expanding the gene forms, including those containing pathogenic SNVs, and resulting protein products associated with amyloid (**Figure 2**). These products would access downstream pathogenic cellular mechanisms observed in familial AD but doing so mosaically and somatically in SAD. SGR affecting *APP* also provides an opportunity to reconsider clinical trial failures. All

therapeutic antibodies against Aβ are monoclonal, which target distinct epitopes of a conserved amino acid sequence of Aβ (Arndt et al., 2018), and many antibodies are effective at clearing the plaques formed by this amino acid sequence. However, SGR creates myriad different *APP* variant genes, transcripts, and predicted amyloid proteins that may not be recognized by mono-specific Aβ-antibodies. These variant species would therefore remain in the brain in various potential forms (e.g., other plaques, fibrils, prions, and soluble products). Notably, some forms arising from gencDNAs will share conserved epitopes that may be accessed by mono-molecular agents to affect a subset of gencDNAs providing partial efficacy. Similarly, SGR may create products that do not require secretase cleavage, a view supported by the small size of some variant genes and RNA transcripts (**Figure 2**). Taken together, SGR maintains the central genetic importance of *APP* in familial forms of AD and Down syndrome but significantly extends it to account for SAD without requiring germline changes in *APP*.

Any AD hypothesis must account for statistical relationships to AD risk genes (Kunkle et al., 2019). In this regard, SGR could be augmented by risk genes like APOE4, the major risk allele for AD (Saunders et al., 1993; Strittmatter et al., 1993) that has been shown to increase *APP* transcription and Aβ deposition (Huang et al., 2017; Liu et al., 2017). This function of APOE4 is highly consistent with the SGR hypothesis, where *APP* transcription was shown to be required for gencDNA formation in culture and in J20 mouse neurons. Increases in *APP* transcription could therefore increase the probability of *APP* SGR occurring, and gencDNA production could lead to SGR-dependent "seeds" that promote toxic plaque deposition. Other risk genes could similarly interface with SGR to produce somatic, disease-promoting genomic changes. Notably, PSEN1 variants were not identified using the same detection pipeline that identified *APP* gencDNAs. One possibility is that the gain-of-function end-points that appear to occur for APP are not relevant to the promotion of AD by PSEN1 mutations, a view consistent with the scientific literature that identifies loss-of-function rather than gain-of-function mechanisms (Kelleher and Shen, 2017; Sun et al., 2017).

Tau Hypothesis Compatibility With SGR

Tau is a microtubule-associated protein that becomes hyperphosphorylated in disease and can aggregate to form NFTs, the second major pathological hallmark of AD (Braak and Braak, 1991; Braak et al., 2006). Tau pathology is closely correlated with neurodegeneration and clinical symptoms (Ossenkoppele et al., 2016; Scholl et al., 2016; Okamura and Yanai, 2017) and may be a key initiator of stressors leading ultimately to cell death in both the Aβ and tau hypotheses (Selkoe and Hardy, 2016; Kametani and Hasegawa, 2018).

Somatic gene recombination is consistent with the tau hypothesis in two distinct and non-mutually exclusive ways. First, SGR generates APP protein variants that could alter tau phosphorylation and processing: a function proposed for Aβ (Rapoport et al., 2002; Dolan and Johnson, 2010; Jin et al., 2011; Moore et al., 2015). Second, SGR might act on the gene for tau—*MAPT*—in the same manner as *APP*, thereby creating myriad and mosaic *MAPT* variants (**Figure 2**). *MAPT* mutations are known to cause autosomal dominant forms of frontotemporal lobar degeneration (FTLD) and Parkinson's disease (PD), with over 40 pathogenic mutations identified to date (Rademakers et al., 2004; Ghetti et al., 2015), and SGR could generate related SNVs. There is some debate on whether germline *MAPT* mutations also represent an increased risk for developing AD; however, a recent meta-analysis of a subset of SNVs reported a significantly increased risk for AD that was furthered by APOE4 carrier status (Zhou and Wang, 2017). If *MAPT* were mosaically altered in SAD brains, but perhaps in different cells or brain regions, it might contribute to AD progression, and importantly, explain the high co-morbidity between AD and PD/other proteinopathies (Kovacs et al., 2013; Brenowitz et al., 2017; Kapasi et al., 2017); further study is warranted.

Prion Hypothesis Relevance to SGR

Prions are misfolded proteins that are able to transmit disease in a fashion similar to an infection, *via* transfer of proteins from cell to cell (Watts and Prusiner, 2018). Evidence has been steadily mounting for the involvement of prions in SAD and other neurodegenerative diseases (e.g., PD and FTLD) in which misfolded proteins are prone to accumulation. Both Aβ (Jaunmuktane et al., 2015; Olsson et al., 2018) and tau (Holmes and Diamond, 2014; Alonso et al., 2016; Kaufman et al., 2016) have been implicated as prion proteins in AD. The enormous potential protein heterogeneity encoded by SGR gene variants is well suited to create DNA sequences encoding mutant prion-like proteins (**Figure 2**). Such proteins may have seeding effects, leading to the misfolding of normal APP and Aβ—or other SGR-derived proteins—which could then act as prion-like transmissible agents. Were this to occur, SGR variants may impact neighboring cells, perhaps *via* the prion protein receptor itself (Lauren et al., 2009; Gimbel et al., 2010), thereby amplifying the spread of pathogenic SGR products. This mechanism might enable propagation throughout the AD brain to promote the documented neuroanatomical progression of plaques and tangles (Arnold et al., 2013). SGR might also enable the identification of key pathological amino acid sequences in prion-like proteins.

Inflammation and Cellular Stress Hypotheses and SGR

There are multiple hypotheses for SAD that incorporate some component of the inflammation pathway, oxidative stress, biometal accumulation, and/or mitochondrial dysfunction. These mechanisms likely combine to accelerate neurodegeneration. However, there is debate about whether these processes cause neurodegeneration or are the result of it (Andersen, 2004). Indeed during an inflammatory response, glia produce high levels of free radicals that promote cellular damage and augment neuroinflammation (Solleiro-Villavicencio and Rivas-Arancibia, 2018); thus, the mechanisms underlying such hypotheses are likely to involve a multifaceted, cyclical mechanism of neurodegeneration.

As a class, transposable elements have complex roles in cellular stress (Horvath et al., 2017) that may be emulated or impacted by gencDNAs. Cellular stress causes nucleic acid oxidation which often results in strand breaks. DNA strand breaks were shown to be required for SGR retro-insertion and gencDNA formation in cell culture (Lee et al., 2018). It is therefore possible that a feed-forward mechanism exists where cellular stress causes the strand breaks that enable SGR of *APP* gencDNAs, whose products would, in turn, increase oxidative stress (**Figure 2**). Additionally, both DNA and RNA oxidation occur in AD brains (Nunomura et al., 1999, 2012); the resultant DNA strand breaks may promote dysregulated gencDNA retro-insertion and could contribute to the formation of intra-exonic junctions.

Trisomy 21 Hypothesis and SGR

Trisomy of chromosome 21 has long been associated with AD through Down syndrome and the first identification of APP, leading to early hypotheses that SAD might be caused by trisomy 21. This hypothesis was rigorously tested in 1987 and no duplication of chromosome 21 (St George-Hyslop et al., 1987) or the *APP* gene (Tanzi et al., 1987a) was found in bulk samples of SAD brains. Reports on linkage between global mosaic trisomy 21 and SAD have been reviewed elsewhere (Potter, 1991; Potter et al., 2016, 2019). It is notable that mosaic aneuploidies involving all chromosomes are found throughout the normal vertebrate brain, including humans, independent of AD (reviewed in Arendt et al., 2009; Leija-Salazar et al., 2018; Rohrback et al., 2018; Shepherd et al., 2018; Iourov et al., 2019; Potter et al., 2019). While some studies have reported an increase in brain aneuploidies associated with AD (Iourov et al., 2009; Arendt et al., 2010; Yurov et al., 2014), others have not identified disease associations (Thomas and Fenech, 2008; Westra et al., 2009; Bushman et al., 2015). Critically, sampling issues affect all studies of aneuploidy because of the minute fraction of interrogated cells utilized compared to the total number of cells within the brain. Interestingly, increased gene transcription could theoretically increase the probability of SGR for a given gene, in support of a link between chromosomal gains that promote transcription and SGR.

Infection Hypothesis and SGR

The infection or pathogen hypothesis proposes that viral, fungal, and/or bacterial infections may trigger AD. These hypotheses are based on reports of the presence of viruses, fungi, or bacteria (or their remnant signatures) within the SAD brain (Hill et al., 2014; Itzhaki et al., 2016; Harding et al., 2017). Some viruses, including human immunodeficiency virus (HIV) and hepatitis B virus (HBV) (Mastroeni et al., 2018), as well as certain bacteria (Lampson et al., 2005), possess demonstrated RT activity—a requirement for SGR. In addition, the vast diversity of *APP* gencDNAs could conceivably produce proteins that could bind to and possibly neutralize infectious agents by analogy to immunoglobulins in the immune system (Eimer et al., 2018). However, the causality, specificity, and presence of these infectious agents require further study, as underscored by reports of bacterial contamination as artifacts in human microbiome studies (reviewed in Eisenhofer et al., 2019).

The SGR Hypothesis in AD

The preceding discussion outlines concepts and hypotheses that could be accessed by SGR. Normally, *APP* SGR acts first upon mRNAs transcribed from the wild-type locus producing varied *APP* gencDNAs *via* an RNA intermediate and RT activity, requiring DNA strand breaks to enable retro-insertion. These gencDNA sequences then retro-insert into genomic DNA, generating cDNA-like sequences that lack introns. They may be full-length DNA copies of known *APP* splice variants (*APP* 571 and 695) or appear as truncated forms containing intra-exonic junctions. The insertion sites appear most commonly outside of wild-type loci, based upon DISH and initial insertion site analyses, with relatively few cells containing one or more copies. Normal SGR may represent a form of cellular memory, where activity-dependent transcription and DNA breaks of multiple etiologies enable the incorporation of gencDNAs, particularly as preferred and already-spliced forms that could later be re-expressed using similar or different promoters.

In disease, dysregulation of SGR occurs. It appears to involve coordinate actions of at least three SGR components: gene transcription, RT activity, and DNA strand breaks. Dysregulation then produces myriad numbers and forms of gencDNAs that could be neurotoxic *via* retro-insertion (as documented for other mobile elements; Horvath et al., 2017), other non-coding disruptions of RNA and DNA, and/or pathogenic APP proteins with altered primary, secondary, tertiary, or quaternary structure that would impact the functionality of APP, Aβ, and prion-like proteins. Known risk genes could be involved, for example, with APOE4 increasing *APP* transcription (Huang et al., 2017) that would, in turn, increase SGR. In this view, some classes of risk factor genes would promote SGR actions on AD "driver" or causal genes. Similarly, increased inflammation or oxidative stress would increase DNA strand breaks, resulting in more gencDNA retro-insertions into new, potentially deleterious genomic locations. Additional risk factors would increase pathogenic gencDNA variants in a feed-forward loop to increase gencDNA production that includes pathogenic SNVs known from familial cases, passing through a disease threshold. Other pathogenic SNVs not compatible with life (and familial AD manifestation)—and other genes—could also be produced in SAD, while familial AD and Down syndrome pathology may also involve SGR, which could explain the decades of life still required to produce AD in these genetic disorders. SGR could generate prion-like sequences producing toxic protein accumulations in neuroanatomically defined patterns. Critically, SGR utilizes RT that appears to create SNVs through imprecise template copying and also identifies an accessible therapeutic strategy using Food and Drug Administration (FDA)-approved RT inhibitors (Lee and Chun, 2019).

SGR AND OTHER BRAIN DISEASES

The existence of SGR in the normal and the AD brain could potentially unify mechanisms for neurological and possibly neuropsychiatric sporadic brain diseases, where somatic, mosaic changes in DNA sequences generate pathogenic loci. Most other

neurodegenerative and neuropsychiatric disorders, including AD-related dementias like FTLD and PD dementia/Lewy body dementia, present most commonly as a sporadic disease. SGR in theory could act on any number of gene loci, including those identified in familial disease to produce mosaic genomic variations that drive sporadic disease. Notably, new mutations in known pathogenic genes as well as in unrecognized genes could be somatically and mosaically altered in the brain, which again might be incompatible with life if present in the germline. SGR dysregulation could also explain the multiple decades it takes for most neurodegenerative diseases to progress as well as patient-to-patient variability in disease progression, wherein the generation and the accumulation of pathogenic gene variants occur mosaically over time. Similarly, SGR can explain the comorbidity of mixed dementias, where ∼50–75% of patients with dementia have neuropathology from at least two of the AD-related dementias (Kovacs et al., 2013; James et al., 2016; Brenowitz et al., 2017; Kapasi et al., 2017). In this scenario, SGR could act on different genes within the same brain, affecting various cell types and neuroanatomical regions.

SGR AND REVERSE TRANSCRIPTASE ACTIVITY

The origin of brain RT activity is not yet clear. RTs were first discovered in retroviruses (Baltimore, 1970; Temin and Mizutani, 1970). However, a more likely source of endogenous RT activity in humans resides in the genome, within what was once called "junk DNA" (Ohno, 1972) and includes LINE1 sequences, human endogenous retrovirus (HERVs), and sequences encoding telomerases. LINE1 sequences account for ∼17% of the human genome and include over 500,000 copies (Lander et al., 2001). Two open reading frames within LINE1 are ORF1, thought to encode a high-affinity RNA binding protein, and ORF2 that encodes an RT and an endonuclease. A vast majority of these sequences are thought to be inactive; however, some have been shown to enable retro-transposition in cancer and have been implicated as a driver of evolution (Lee et al., 2012). LINE1 has also been hypothesized to contribute to neuronal diversity by disrupting existing genes or DNA elements upon re-insertion of LINE1 sequences during neurogenesis (Muotri et al., 2005), a concept that remains under active investigation (Evrony et al., 2012 *vs*. Upton et al., 2015). Another 8% of the genome is made up of HERVs that contain a possible RT within its *pol* gene, albeit with limited expected activity in the human genome (Nelson et al., 2003, 2004; Thomas et al., 2018). Human telomerase (encoded by the TERT gene) also has RT activity (Leao et al., 2018) and may further provide an RT source, as could other unknown enzymes.

The Clinical Potential of SGR Inhibition in AD and Other Brain Diseases

The demonstrated involvement of RT activity in SGR implicates its inhibition as a possible AD preventative and/or therapeutic intervention. Critically, multiple FDA-approved RT inhibitors have been developed for the treatment of HIV (and later, HBV), with over three decades of continuous use as part of combination anti-retroviral therapies, which may provide real-world evidence of their efficacy in the prevention and/or treatment of AD. The number of treated HIV-infected patients who are also at risk of SAD (being ∼65 years or older) currently number up to ∼80,000–100,000 patients in the United States (CDC, 2018), which would yield an expected AD prevalence in the thousands (10% of all persons age 65 or older have AD; Alzheimer's Association, 2019). A surveillance of these patients has occurred for over a decade in anticipation of an increase of AD in HIV-infected patients (Alisky, 2007), yet only one documented AD/HIV-infected case has appeared in the peer-reviewed literature (Turner et al., 2016). The limitations of *post hoc* epidemiology are acknowledged, and prospective AD clinical trials are needed. However, even if confounders resulted in these numbers being off by a factor of 100, there would still be a significant difference between the number of reported HIV-infected SAD cases *vs*. the expected prevalence of SAD in this age group. Since approved RT inhibitors have over 30 years of real-world human safety data, legal, off-label prescription by a licensed physician represents a promising option for AD patients where no effective and safe therapy currently exists.

CONCLUSION

Genomic mosaicism in the human brain is a biological fact that manifests through multiple forms of DNA sequence changes within single cells, from aneuploidies through SNVs. The recent discovery of SGR acting on *APP* provides functionality for genomic mosaicism through actions on a single gene, with both normal and disease implications. Normally, SGR may act as a form of cellular memory (Crick, 1984; Davis and Squire, 1984), where transcriptional activity and resulting DNA breaks may enable the retro-insertion of gencDNAs ready for re-expression as pre-spliced and varied mRNAs and diverse protein products: a form of long-lasting memory. SGR may also resemble forms of genomic "streamlining" that have been documented through phylogeny and contribute to species evolution (Roy and Gilbert, 2006). The dysregulation of SGR produces disease through increased numbers and forms of toxic gencDNAs, as illustrated by somatic, genomic changes documented in SAD brains. Importantly, the SGR hypothesis in AD does not reject the amyloid hypothesis outright but rather incorporates major features to modify the hypothesis while also accommodating other distinct hypotheses and explaining discrepancies in the scientific and clinical trial literature through the generation of *APP* variants and downstream molecular diversity. SGR presents a new source of potential therapeutics, some with near-term implications for the treatment and/or prevention of AD by use of FDA-approved medicines targeting endogenous brain RTs, an approach supported by human epidemiological data on older HIV-infected patients. SGR and its roles in AD represent a new step toward understanding the functions of genomic mosaicism in the normal, aging, and diseased brain.

AUTHOR CONTRIBUTIONS

GK and JC wrote this manuscript.

FUNDING

The research reported in this publication was supported by the NIA of the National Institutes of Health under award number R56AG067489 (JC). This work was supported by non-Federal funds from The Shaffer Family Foundation and The Bruce Ford and Anne Smith Bundy Foundation (JC). 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 Dr. Laura Wolszon and Danielle Jones for their assistance in editing this manuscript.

REFERENCES

Alisky, J. M. (2007). The coming problem of HIV-associated Alzheimer's disease. *Med. Hypotheses* 69, 1140–1143. doi: 10.1016/j.mehy.2007.02.030

Alonso, A. D., Beharry, C., Corbo, C. P., and Cohen, L. S. (2016). Molecular mechanism of prion-like tau-induced neurodegeneration. *Alzheimers Dement.* 12, 1090–1097. doi: 10.1016/j.jalz.2015.12.014

Alzheimer's Association (2019). Alzheimer's disease facts and figures. *Alzheimers Dement.* 15, 321–387. doi: 10.1016/j.jalz.2019.01.010

Andersen, J. K. (2004). Oxidative stress in neurodegeneration: cause or consequence? *Nat. Med.* 10(Suppl.), S18–S25. doi: 10.1038/nrn1434

Anderson, R. M., Hadjichrysanthou, C., Evans, S., and Wong, M. M. (2017). Why do so many clinical trials of therapies for Alzheimer's disease fail? *Lancet* 390, 2327–2329. doi: 10.1016/S0140-6736(17)32399-91

Arendt, T., Bruckner, M. K., Mosch, B., and Losche, A. (2010). Selective cell death of hyperploid neurons in Alzheimer's disease. *Am. J. Pathol.* 177, 15–20. doi: 10.2353/ajpath.2010.090955

Arendt, T., Mosch, B., and Morawski, M. (2009). Neuronal aneuploidy in health and disease: a cytomic approach to understand the molecular individuality of neurons. *Int. J. Mol. Sci.* 10, 1609–1627. doi: 10.3390/ijms10041609

Arndt, J. W., Qian, F., Smith, B. A., Quan, C., Kilambi, K. P., Bush, M. W., et al. (2018). Structural and kinetic basis for the selectivity of aducanumab for aggregated forms of amyloid-beta. *Sci. Rep.* 8:6412. doi: 10.1038/s41598-018-24501-0

Arnold, S. E., Toledo, J. B., Appleby, D. H., Xie, S. X., Wang, L. S., Baek, Y., et al. (2013). Comparative survey of the topographical distribution of signature molecular lesions in major neurodegenerative diseases. *J. Comp. Neurol.* 521, 4339–4355. doi: 10.1002/cne.23430

Baker, A. M., Huang, W., Wang, X. M., Jansen, M., Ma, X. J., Kim, J., et al. (2017). Robust RNA-based in situ mutation detection delineates colorectal cancer subclonal evolution. *Nat. Commun.* 8:1998. doi: 10.1038/s41467-017-02295-5

Baltimore, D. (1970). RNA-dependent DNA polymerase in virions of RNA tumour viruses. *Nature* 226, 1209–1211. doi: 10.1038/2261209a0

Braak, H., Alafuzoff, I., Arzberger, T., Kretzschmar, H., and Del Tredici, K. (2006). Staging of Alzheimer disease-associated neurofibrillary pathology using paraffin sections and immunocytochemistry. *Acta Neuropathol.* 112, 389–404. doi: 10.1007/s00401-006-0127-z

Braak, H., and Braak, E. (1991). Neuropathological stageing of Alzheimer-related changes. *Acta Neuropathol.* 82, 239–259. doi: 10.1007/bf00308809

Brenowitz, W. D., Keene, C. D., Hawes, S. E., Hubbard, R. A., Longstreth, W. T. Jr., Woltjer, R. L., et al. (2017). Alzheimer's disease neuropathologic change, Lewy body disease, and vascular brain injury in clinic- and community-based samples. *Neurobiol. Aging* 53, 83–92. doi: 10.1016/j.neurobiolaging.2017.01.017

Burki, T. (2018). Alzheimer's disease research: the future of bace inhibitors. *Lancet* 391:2486. doi: 10.1016/S0140-6736(18)31425-31429

Bushman, D. M., Kaeser, G. E., Siddoway, B., Westra, J. W., Rivera, R. R., Rehen, S. K., et al. (2015). Genomic mosaicism with increased amyloid precursor protein (APP) gene copy number in single neurons from sporadic Alzheimer's disease brains. *eLife* 4:e05116. doi: 10.7554/eLife.05116

Carlsson, C. M. (2008). Lessons learned from failed and discontinued clinical trials for the treatment of Alzheimer's disease: future directions. *J. Alzheimers Dis.* 15, 327–338. doi: 10.3233/jad-2008-15214

Castro, M. A., Hadziselimovic, A., and Sanders, C. R. (2019). The vexing complexity of the amyloidogenic pathway. *Protein Sci.* 28, 1177–1193. doi: 10.1002/pro.3606

CDC (2018). *Center for Disease Control and Prevention. HIV Surveillance Report, 2017*. Available: http://www.cdc.gov/hiv/library/reports/hiv-surveillance.html (accessed October 1, 2019).

Chun, J. J., Schatz, D. G., Oettinger, M. A., Jaenisch, R., and Baltimore, D. (1991). The recombination activating gene-1 (RAG-1) transcript is present in the murine central nervous system. *Cell* 64, 189–200. doi: 10.1016/0092-8674(91)90220-s

Crick, F. (1984). Memory and molecular turnover. *Nature* 312:101. doi: 10.1038/312101a0

Crystal, H., Dickson, D., Fuld, P., Masur, D., Scott, R., Mehler, M., et al. (1988). Clinico-pathologic studies in dementia: nondemented subjects with pathologically confirmed Alzheimer's disease. *Neurology* 38, 1682–1687. doi: 10.1212/wnl.38.11.1682

Cummings, J. L., Morstorf, T., and Zhong, K. (2014). Alzheimer's disease drug-development pipeline: few candidates, frequent failures. *Alzheimers Res. Ther.* 6:37. doi: 10.1186/alzrt269

Davis, H. P., and Squire, L. R. (1984). Protein synthesis and memory: a review. *Psychol. Bull.* 96, 518–559.

De Strooper, B., and Karran, E. (2016). The cellular phase of Alzheimer's disease. *Cell* 164, 603–615. doi: 10.1016/j.cell.2015.12.056

Dolan, P. J., and Johnson, G. V. (2010). The role of tau kinases in Alzheimer's disease. *Curr. Opin. Drug Discov. Dev.* 13, 595–603. doi: 10.6026/97320630091023

Egan, M. F., Kost, J., Tariot, P. N., Aisen, P. S., Cummings, J. L., Vellas, B., et al. (2018). Randomized trial of verubecestat for mild-to-moderate Alzheimer's disease. *N. Engl. J. Med.* 378, 1691–1703. doi: 10.1056/NEJMoa1706441

Eid, J., Fehr, A., Gray, J., Luong, K., Lyle, J., Otto, G., et al. (2009). Real-time DNA sequencing from single polymerase molecules. *Science* 323, 133–138. doi: 10.1126/science.1162986

Eimer, W. A., Vijaya Kumar, D. K., Navalpur Shanmugam, N. K., Rodriguez, A. S., Mitchell, T., Washicosky, K. J., et al. (2018). Alzheimer's disease-associated beta-amyloid is rapidly seeded by herpesviridae to protect against brain infection. *Neuron* 99, 56–63.e3. doi: 10.1016/j.neuron.2018.06.030

Eisenhofer, R., Minich, J. J., Marotz, C., Cooper, A., Knight, R., and Weyrich, L. S. (2019). Contamination in low microbial biomass microbiome studies: issues and recommendations. *Trends Microbiol.* 27, 105–117. doi: 10.1016/j.tim.2018.11.003

Esparza, T. J., Zhao, H., Cirrito, J. R., Cairns, N. J., Bateman, R. J., Holtzman, D. M., et al. (2013). Amyloid-beta oligomerization in Alzheimer dementia versus high-pathology controls. *Ann. Neurol.* 73, 104–119. doi: 10.1002/ana.23748

Evrony, G. D., Cai, X., Lee, E., Hills, L. B., Elhosary, P. C., Lehmann, H. S., et al. (2012). Single-neuron sequencing analysis of L1 retrotransposition and somatic mutation in the human brain. *Cell* 151, 483–496. doi: 10.1016/j.cell.2012.09.035

Gao, Y., Sun, Y., Frank, K. M., Dikkes, P., Fujiwara, Y., Seidl, K. J., et al. (1998). A critical role for DNA end-joining proteins in both lymphogenesis and neurogenesis. *Cell* 95, 891–902. doi: 10.1016/s0092-8674(00)81714-81716

Gericke, G. S. (2008). An integrative view of dynamic genomic elements influencing human brain evolution and individual neurodevelopment. *Med. Hypotheses* 71, 360–373. doi: 10.1016/j.mehy.2008.03.048

Ghetti, B., Oblak, A. L., Boeve, B. F., Johnson, K. A., Dickerson, B. C., and Goedert, M. (2015). Invited review: frontotemporal dementia caused by

microtubule-associated protein tau gene (MAPT) mutations: a chameleon for neuropathology and neuroimaging. *Neuropathol. Appl. Neurobiol.* 41, 24–46. doi: 10.1111/nan.12213

Gimbel, D. A., Nygaard, H. B., Coffey, E. E., Gunther, E. C., Lauren, J., Gimbel, Z. A., et al. (2010). Memory impairment in transgenic Alzheimer mice requires cellular prion protein. *J. Neurosci.* 30, 6367–6374. doi: 10.1523/JNEUROSCI.0395-10.2010

Glenner, G. G., and Wong, C. W. (1984a). Alzheimer's disease and Down's syndrome: sharing of a unique cerebrovascular amyloid fibril protein. *Biochem. Biophys. Res. Commun.* 122, 1131–1135. doi: 10.1016/0006-291x(84)91209

Glenner, G. G., and Wong, C. W. (1984b). Alzheimer's disease: initial report of the purification and characterization of a novel cerebrovascular amyloid protein. *Biochem. Biophys. Res. Commun.* 120, 885–890. doi: 10.1016/s0006-291x(84)80190-80194

Goldgaber, D., Lerman, M. I., McBride, O. W., Saffiotti, U., and Gajdusek, D. C. (1987). Characterization and chromosomal localization of a cDNA encoding brain amyloid of Alzheimer's disease. *Science* 235, 877–880. doi: 10.1126/science.3810169

Harding, A., Gonder, U., Robinson, S. J., Crean, S., and Singhrao, S. K. (2017). Exploring the association between Alzheimer's disease, oral health, microbial endocrinology and nutrition. *Front. Aging Neurosci.* 9:398. doi: 10.3389/fnagi.2017.00398

Hardy, J., and Selkoe, D. J. (2002). The amyloid hypothesis of Alzheimer's disease: progress and problems on the road to therapeutics. *Science* 297, 353–356. doi: 10.1126/science.1072994

Hardy, J. A., and Higgins, G. A. (1992). Alzheimer's disease: the amyloid cascade hypothesis. *Science* 256, 184–185. doi: 10.1126/science.1566067

Hebert, P. D. N., Braukmann, T. W. A., Prosser, S. W. J., Ratnasingham, S., deWaard, J. R., Ivanova, N. V., et al. (2018). A sequel to sanger: amplicon sequencing that scales. *BMC Genomics* 19:219. doi: 10.1186/s12864-018-4611-4613

Herrup, K. (2015). The case for rejecting the amyloid cascade hypothesis. *Nat. Neurosci.* 18, 794–799. doi: 10.1038/nn.4017

Hill, J. M., Clement, C., Pogue, A. I., Bhattacharjee, S., Zhao, Y., and Lukiw, W. J. (2014). Pathogenic microbes, the microbiome, and Alzheimer's disease (AD). *Front. Aging Neurosci.* 6:127. doi: 10.3389/fnagi.2014.00127

Holmes, B. B., and Diamond, M. I. (2014). Prion-like properties of Tau protein: the importance of extracellular Tau as a therapeutic target. *J. Biol. Chem.* 289, 19855–19861. doi: 10.1074/jbc.R114.549295

Hooli, B. V., Mohapatra, G., Mattheisen, M., Parrado, A. R., Roehr, J. T., Shen, Y., et al. (2012). Role of common and rare APP DNA sequence variants in Alzheimer disease. *Neurology* 78, 1250–1257. doi: 10.1212/WNL.0b013e3182515972

Horvath, V., Merenciano, M., and Gonzalez, J. (2017). Revisiting the relationship between transposable elements and the eukaryotic stress response. *Trends Genet.* 33, 832–841. doi: 10.1016/j.tig.2017.08.007

Huang, Y. A., Zhou, B., Wernig, M., and Sudhof, T. C. (2017). ApoE2, ApoE3, and ApoE4 differentially stimulate APP transcription and abeta secretion. *Cell* 168, 427–441.e21. doi: 10.1016/j.cell.2016.12.044

Insel, P. S., Ossenkoppele, R., Gessert, D., Jagust, W., Landau, S., Hansson, O., et al. (2017). Time to amyloid positivity and preclinical changes in brain metabolism, atrophy, and cognition: evidence for emerging amyloid pathology in Alzheimer's disease. *Front. Neurosci.* 11:281. doi: 10.3389/fnins.2017.00281

Iourov, I. Y., Vorsanova, S. G., Liehr, T., and Yurov, Y. B. (2009). Aneuploidy in the normal, Alzheimer's disease and ataxia-telangiectasia brain: differential expression and pathological meaning. *Neurobiol. Dis.* 34, 212–220. doi: 10.1016/j.nbd.2009.01.003

Iourov, I. Y., Vorsanova, S. G., Yurov, Y. B., and Kutsev, S. I. (2019). Ontogenetic and pathogenetic views on somatic chromosomal mosaicism. *Genes* 10:379. doi: 10.3390/genes10050379

Itzhaki, R. F., Lathe, R., Balin, B. J., Ball, M. J., Bearer, E. L., Braak, H., et al. (2016). Microbes and Alzheimer's disease. *J. Alzheimers Dis.* 51, 979–984. doi: 10.3233/JAD-160152

James, B. D., Wilson, R. S., Boyle, P. A., Trojanowski, J. Q., Bennett, D. A., and Schneider, J. A. (2016). TDP-43 stage, mixed pathologies, and clinical Alzheimer's-type dementia. *Brain* 139, 2983–2993. doi: 10.1093/brain/aww224

Jaunmuktane, Z., Mead, S., Ellis, M., Wadsworth, J. D., Nicoll, A. J., Kenny, J., et al. (2015). Evidence for human transmission of amyloid-beta pathology and cerebral amyloid angiopathy. *Nature* 525, 247–250. doi: 10.1038/nature15369

Jin, M., Shepardson, N., Yang, T., Chen, G., Walsh, D., and Selkoe, D. J. (2011). Soluble amyloid beta-protein dimers isolated from Alzheimer cortex directly induce Tau hyperphosphorylation and neuritic degeneration. *Proc. Natl. Acad. Sci. U.S.A.* 108, 5819–5824. doi: 10.1073/pnas.1017033108

Kametani, F., and Hasegawa, M. (2018). Reconsideration of amyloid hypothesis and Tau hypothesis in Alzheimer's disease. *Front. Neurosci.* 12:25. doi: 10.3389/fnins.2018.00025

Kapasi, A., DeCarli, C., and Schneider, J. A. (2017). Impact of multiple pathologies on the threshold for clinically overt dementia. *Acta Neuropathol.* 134, 171–186. doi: 10.1007/s00401-017-1717-1717

Katzman, R., Terry, R., DeTeresa, R., Brown, T., Davies, P., Fuld, P., et al. (1988). Clinical, pathological, and neurochemical changes in dementia: a subgroup with preserved mental status and numerous neocortical plaques. *Ann. Neurol.* 23, 138–144. doi: 10.1002/ana.410230206

Kaufman, S. K., Sanders, D. W., Thomas, T. L., Ruchinskas, A. J., Vaquer-Alicea, J., Sharma, A. M., et al. (2016). Tau prion strains dictate patterns of cell pathology, progression rate, and regional vulnerability in vivo. *Neuron* 92, 796–812. doi: 10.1016/j.neuron.2016.09.055

Kaushal, D., Contos, J. J., Treuner, K., Yang, A. H., Kingsbury, M. A., Rehen, S. K., et al. (2003). Alteration of gene expression by chromosome loss in the postnatal mouse brain. *J. Neurosci.* 23, 5599–5606. doi: 10.1523/JNEUROSCI.23-13-05599.2003

Kelleher, R. J. III, and Shen, J. (2017). Presenilin-1 mutations and Alzheimer's disease. *Proc. Natl. Acad. Sci. U.S.A.* 114, 629–631. doi: 10.1073/pnas.1619574114

Kingsbury, M. A., Friedman, B., McConnell, M. J., Rehen, S. K., Yang, A. H., Kaushal, D., et al. (2005). Aneuploid neurons are functionally active and integrated into brain circuitry. *Proc. Natl. Acad. Sci. U.S.A.* 102, 6143–6147. doi: 10.1073/pnas.0408171102

Kovacs, G. G., Milenkovic, I., Wohrer, A., Hoftberger, R., Gelpi, E., Haberler, C., et al. (2013). Non-Alzheimer neurodegenerative pathologies and their combinations are more frequent than commonly believed in the elderly brain: a community-based autopsy series. *Acta Neuropathol.* 126, 365–384. doi: 10.1007/s00401-013-1157-y

Kunkle, B. W., Grenier-Boley, B., Sims, R., Bis, J. C., Damotte, V., Naj, A. C., et al. (2019). Genetic meta-analysis of diagnosed Alzheimer's disease identifies new risk loci and implicates Abeta, tau, immunity and lipid processing. *Nat. Genet.* 51, 414–430. doi: 10.1038/s41588-019-0358-352

Lake, B. B., Ai, R., Kaeser, G. E., Salathia, N. S., Yung, Y. C., Liu, R., et al. (2016). Neuronal subtypes and diversity revealed by single-nucleus RNA sequencing of the human brain. *Science* 352, 1586–1590. doi: 10.1126/science.aaf1204

Lake, B. B., Chen, S., Sos, B. C., Fan, J., Kaeser, G. E., Yung, Y. C., et al. (2018). Integrative single-cell analysis of transcriptional and epigenetic states in the human adult brain. *Nat. Biotechnol.* 36, 70–80. doi: 10.1038/nbt.4038

Lampson, B. C., Inouye, M., and Inouye, S. (2005). Retrons, msDNA, and the bacterial genome. *Cytogenet. Genome Res.* 110, 491–499. doi: 10.1159/000084982

Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., et al. (2001). Initial sequencing and analysis of the human genome. *Nature* 409, 860–921. doi: 10.1038/35057062

Lauren, J., Gimbel, D. A., Nygaard, H. B., Gilbert, J. W., and Strittmatter, S. M. (2009). Cellular prion protein mediates impairment of synaptic plasticity by amyloid-beta oligomers. *Nature* 457, 1128–1132. doi: 10.1038/nature07761

Leao, R., Apolonio, J. D., Lee, D., Figueiredo, A., Tabori, U., and Castelo-Branco, P. (2018). Mechanisms of human telomerase reverse transcriptase (hTERT) regulation: clinical impacts in cancer. *J. Biomed. Sci.* 25:22. doi: 10.1186/s12929-018-0422-428

Lee, E., Iskow, R., Yang, L., Gokcumen, O., Haseley, P., Luquette, L. J., et al. (2012). Landscape of somatic retrotransposition in human cancers. *Science* 337, 967–971. doi: 10.1126/science.1222077

Lee, M. H., and Chun, J. (2019). Mosaic APP gene recombination in Alzheimer's disease-what's next? *J. Exp. Neurosci.* 13:1179069519849669. doi: 10.1177/1179069519849669

Lee, M.-H., Liu, C. S., Zhu, Y., Kaeser, G. E., Rivera, R., Romanow, W. J., et al. (2019). Reply: evidence that APP gene copy number changes reflect recombinant vector contamination. *bioRxiv* [preprint]. doi: 10.1101/730291

Lee, M. H., Siddoway, B., Kaeser, G. E., Segota, I., Rivera, R., Romanow, W. J., et al. (2018). Somatic APP gene recombination in Alzheimer's disease and normal neurons. *Nature* 563, 639–645. doi: 10.1038/s41586-018-0718-716

Leija-Salazar, M., Piette, C., and Proukakis, C. (2018). Review: somatic mutations in neurodegeneration. *Neuropathol. Appl. Neurobiol.* 44, 267–285. doi: 10.1111/nan.12465

Levy-Lahad, E., Wasco, W., Poorkaj, P., Romano, D. M., Oshima, J., Pettingell, W. H., et al. (1995). Candidate gene for the chromosome 1 familial Alzheimer's disease locus. *Science* 269, 973–977. doi: 10.1126/science.7638622

Liu, C. C., Zhao, N., Fu, Y., Wang, N., Linares, C., Tsai, C. W., et al. (2017). ApoE4 accelerates early seeding of amyloid pathology. *Neuron* 96, 1024–1032.e3. doi: 10.1016/j.neuron.2017.11.013

Mastroeni, D., Nolz, J., Sekar, S., Delvaux, E., Serrano, G., Cuyugan, L., et al. (2018). Laser-captured microglia in the Alzheimer's and Parkinson's brain reveal unique regional expression profiles and suggest a potential role for hepatitis B in the Alzheimer's brain. *Neurobiol. Aging* 63, 12–21. doi: 10.1016/j.neurobiolaging.2017.10.019

Mehta, D., Jackson, R., Paul, G., Shi, J., and Sabbagh, M. (2017). Why do trials for Alzheimer's disease drugs keep failing? A discontinued drug perspective for 2010-2015. *Expert Opin. Investig. Drugs* 26, 735–739. doi: 10.1080/13543784.2017.1323868

Moore, S., Evans, L. D., Andersson, T., Portelius, E., Smith, J., Dias, T. B., et al. (2015). APP metabolism regulates tau proteostasis in human cerebral cortex neurons. *Cell Rep.* 11, 689–696. doi: 10.1016/j.celrep.2015.03.068

Morris, G. P., Clark, I. A., and Vissel, B. (2018). Questions concerning the role of amyloid-beta in the definition, aetiology and diagnosis of Alzheimer's disease. *Acta Neuropathol.* 136, 663–689. doi: 10.1007/s00401-018-1918-1918

Muotri, A. R., Chu, V. T., Marchetto, M. C., Deng, W., Moran, J. V., and Gage, F. H. (2005). Somatic mosaicism in neuronal precursor cells mediated by L1 retrotransposition. *Nature* 435, 903–910. doi: 10.1038/nature03663

Muotri, A. R., and Gage, F. H. (2006). Generation of neuronal variability and complexity. *Nature* 441, 1087–1093. doi: 10.1038/nature04959

Nelson, P. N., Carnegie, P. R., Martin, J., Davari Ejtehadi, H., Hooley, P., Roden, D., et al. (2003). Demystified. Human endogenous retroviruses. *Mol. Pathol.* 56, 11–18. doi: 10.1136/mp.56.1.11

Nelson, P. N., Hooley, P., Roden, D., Davari Ejtehadi, H., Rylance, P., Warren, P., et al. (2004). Human endogenous retroviruses: transposable elements with potential? *Clin. Exp. Immunol.* 138, 1–9. doi: 10.1111/j.1365-2249.2004.02592.x

Nunomura, A., Moreira, P. I., Castellani, R. J., Lee, H. G., Zhu, X., Smith, M. A., et al. (2012). Oxidative damage to RNA in aging and neurodegenerative disorders. *Neurotox. Res.* 22, 231–248. doi: 10.1007/s12640-012-9331-x

Nunomura, A., Perry, G., Pappolla, M. A., Wade, R., Hirai, K., Chiba, S., et al. (1999). RNA oxidation is a prominent feature of vulnerable neurons in Alzheimer's disease. *J. Neurosci.* 19, 1959–1964. doi: 10.1523/JNEUROSCI.19-06-01959.1999

Ohno, S. (1972). So much "junk" DNA in our genome. *Brookhaven Symp. Biol.* 23, 366–370.

Okamura, N., and Yanai, K. (2017). Brain imaging: applications of tau PET imaging. *Nat. Rev. Neurol.* 13, 197–198. doi: 10.1038/nrneurol.2017.38

Olsson, T. T., Klementieva, O., and Gouras, G. K. (2018). Prion-like seeding and nucleation of intracellular amyloid-beta. *Neurobiol. Dis.* 113, 1–10. doi: 10.1016/j.nbd.2018.01.015

Ossenkoppele, R., Schonhaut, D. R., Scholl, M., Lockhart, S. N., Ayakta, N., Baker, S. L., et al. (2016). Tau PET patterns mirror clinical and neuroanatomical variability in Alzheimer's disease. *Brain* 139(Pt 5), 1551–1567. doi: 10.1093/brain/aww027

Papavasiliou, F. N., and Schatz, D. G. (2002). Somatic hypermutation of immunoglobulin genes: merging mechanisms for genetic diversity. *Cell* 109(Suppl.), S35–S44. doi: 10.1016/s0092-8674(02)00706-707

Park, J. S., Lee, J., Jung, E. S., Kim, M. H., Kim, I. B., Son, H., et al. (2019). Brain somatic mutations observed in Alzheimer's disease associated with aging and dysregulation of tau phosphorylation. *Nat. Commun.* 10:3090. doi: 10.1038/s41467-019-11000-11007

Peterson, S. E., Yang, A. H., Bushman, D. M., Westra, J. W., Yung, Y. C., Barral, S., et al. (2012). Aneuploid cells are differentially susceptible to caspase-mediated death during embryonic cerebral cortical development. *J. Neurosci.* 32, 16213–16222. doi: 10.1523/JNEUROSCI.3706-12.2012

Potter, H. (1991). Review and hypothesis: Alzheimer disease and Down syndrome–chromosome 21 nondisjunction may underlie both disorders. *Am. J. Hum. Genet.* 48, 1192–1200.

Potter, H., Chial, H. J., Caneus, J., Elos, M., Elder, N., Borysov, S., et al. (2019). Chromosome instability and mosaic aneuploidy in neurodegenerative and neurodevelopmental disorders. *Front. Genet.* 10:1092. doi: 10.3389/fgene.2019.01092

Potter, H., Granic, A., and Caneus, J. (2016). Role of Trisomy 21 mosaicism in sporadic and familial Alzheimer's disease. *Curr. Alzheimer Res.* 13, 7–17. doi: 10.2174/1567205013011512071006 16

Rademakers, R., Cruts, M., and van Broeckhoven, C. (2004). The role of tau (MAPT) in frontotemporal dementia and related tauopathies. *Hum. Mutat.* 24, 277–295. doi: 10.1002/humu.20086

Rapoport, M., Dawson, H. N., Binder, L. I., Vitek, M. P., and Ferreira, A. (2002). Tau is essential to beta -amyloid-induced neurotoxicity. *Proc. Natl. Acad. Sci. U.S.A.* 99, 6364–6369. doi: 10.1073/pnas.092136199

Rehen, S. K., McConnell, M. J., Kaushal, D., Kingsbury, M. A., Yang, A. H., and Chun, J. (2001). Chromosomal variation in neurons of the developing and adult mammalian nervous system. *Proc. Natl. Acad. Sci. U.S.A.* 98, 13361–13366. doi: 10.1073/pnas.231487398

Rehen, S. K., Yung, Y. C., McCreight, M. P., Kaushal, D., Yang, A. H., Almeida, B. S., et al. (2005). Constitutional aneuploidy in the normal human brain. *J. Neurosci.* 25, 2176–2180. doi: 10.1523/JNEUROSCI.4560-04.2005

Rogaev, E. I., Sherrington, R., Rogaeva, E. A., Levesque, G., Ikeda, M., Liang, Y., et al. (1995). Familial Alzheimer's disease in kindreds with missense mutations in a gene on chromosome 1 related to the Alzheimer's disease type 3 gene. *Nature* 376, 775–778. doi: 10.1038/376775a0

Rohrback, S., Siddoway, B., Liu, C. S., and Chun, J. (2018). Genomic mosaicism in the developing and adult brain. *Dev. Neurobiol.* 78, 1026–1048. doi: 10.1002/dneu.22626

Rovelet-Lecrux, A., Hannequin, D., Raux, G., Le Meur, N., Laquerriere, A., Vital, A., et al. (2006). APP locus duplication causes autosomal dominant early-onset Alzheimer disease with cerebral amyloid angiopathy. *Nat. Genet.* 38, 24–26. doi: 10.1038/ng1718

Roy, S. W., and Gilbert, W. (2006). The evolution of spliceosomal introns: patterns, puzzles and progress. *Nat. Rev. Genet.* 7, 211–221. doi: 10.1038/nrg1807

Saunders, A. M., Strittmatter, W. J., Schmechel, D., George-Hyslop, P. H., Pericak-Vance, M. A., Joo, S. H., et al. (1993). Association of apolipoprotein E allele epsilon 4 with late-onset familial and sporadic Alzheimer's disease. *Neurology* 43, 1467–1472. doi: 10.1212/wnl.43.8.1467

Scholl, M., Lockhart, S. N., Schonhaut, D. R., O'Neil, J. P., Janabi, M., Ossenkoppele, R., et al. (2016). PET imaging of tau deposition in the aging human brain. *Neuron* 89, 971–982. doi: 10.1016/j.neuron.2016.01.028

Selkoe, D. J., and Hardy, J. (2016). The amyloid hypothesis of Alzheimer's disease at 25 years. *EMBO Mol. Med.* 8, 595–608. doi: 10.15252/emmm.201606210

Shepherd, C. E., Yang, Y., and Halliday, G. M. (2018). Region- and cell-specific aneuploidy in brain aging and neurodegeneration. *Neuroscience* 374, 326–334. doi: 10.1016/j.neuroscience.2018.01.050

Sherrington, R., Rogaev, E. I., Liang, Y., Rogaeva, E. A., Levesque, G., Ikeda, M., et al. (1995). Cloning of a gene bearing missense mutations in early-onset familial Alzheimer's disease. *Nature* 375, 754–760. doi: 10.1038/375754a0

Sleegers, K., Brouwers, N., Gijselinck, I., Theuns, J., Goossens, D., Wauters, J., et al. (2006). APP duplication is sufficient to cause early onset Alzheimer's dementia with cerebral amyloid angiopathy. *Brain* 129(Pt 11), 2977–2983. doi: 10.1093/brain/awl203

Solleiro-Villavicencio, H., and Rivas-Arancibia, S. (2018). Effect of chronic oxidative stress on neuroinflammatory response mediated by CD4(+)T cells in neurodegenerative diseases. *Front. Cell Neurosci.* 12:114. doi: 10.3389/fncel.2018.00114

St George-Hyslop, P. H., Tanzi, R. E., Polinsky, R. J., Neve, R. L., Pollen, D., Drachman, D., et al. (1987). Absence of duplication of chromosome 21 genes in familial and sporadic Alzheimer's disease. *Science* 238, 664–666. doi: 10.1126/science.2890206

Strittmatter, W. J., Saunders, A. M., Schmechel, D., Pericak-Vance, M., Enghild, J., Salvesen, G. S., et al. (1993). Apolipoprotein E: high-avidity binding to beta-amyloid and increased frequency of type 4 allele in late-onset familial Alzheimer

disease. *Proc. Natl. Acad. Sci. U.S.A.* 90, 1977–1981. doi: 10.1073/pnas.90.5.1977

Sun, L., Zhou, R., Yang, G., and Shi, Y. (2017). Analysis of 138 pathogenic mutations in presenilin-1 on the in vitro production of Abeta42 and Abeta40 peptides by gamma-secretase. *Proc. Natl. Acad. Sci. U.S.A.* 114, E476–E485. doi: 10.1073/pnas.1618657114

Tanzi, R. E., Bird, E. D., Latt, S. A., and Neve, R. L. (1987a). The amyloid beta protein gene is not duplicated in brains from patients with Alzheimer's disease. *Science* 238, 666–669. doi: 10.1126/science.2890207

Tanzi, R. E., Gusella, J. F., Watkins, P. C., Bruns, G. A., St George-Hyslop, P., Van Keuren, M. L., et al. (1987b). Amyloid beta protein gene: cDNA, mRNA distribution, and genetic linkage near the Alzheimer locus. *Science* 235, 880–884. doi: 10.1126/science.2949367

Temin, H. M., and Mizutani, S. (1970). RNA-dependent DNA polymerase in virions of Rous sarcoma virus. *Nature* 226, 1211–1213. doi: 10.1038/2261211a0

Thomas, J., Perron, H., and Feschotte, C. (2018). Variation in proviral content among human genomes mediated by LTR recombination. *Mob DNA* 9:36. doi: 10.1186/s13100-018-0142-143

Thomas, P., and Fenech, M. (2008). Chromosome 17 and 21 aneuploidy in buccal cells is increased with ageing and in Alzheimer's disease. *Mutagenesis* 23, 57–65. doi: 10.1093/mutage/gem044

Troncoso, J. C., Martin, L. J., Dal Forno, G., and Kawas, C. H. (1996). Neuropathology in controls and demented subjects from the Baltimore longitudinal study of aging. *Neurobiol. Aging* 17, 365–371. doi: 10.1016/0197-4580(96)00028-20

Turner, R. S., Chadwick, M., Horton, W. A., Simon, G. L., Jiang, X., and Esposito, G. (2016). An individual with human immunodeficiency virus, dementia, and central nervous system amyloid deposition. *Alzheimers Dement.* 4, 1–5. doi: 10.1016/j.dadm.2016.03.009

Upton, K. R., Gerhardt, D. J., Jesuadian, J. S., Richardson, S. R., Sanchez-Luque, F. J., Bodea, G. O., et al. (2015). Ubiquitous L1 mosaicism in hippocampal neurons. *Cell* 161, 228–239. doi: 10.1016/j.cell.2015.03.026

Watts, J. C., and Prusiner, S. B. (2018). beta-Amyloid prions and the pathobiology of Alzheimer's disease. *Cold Spring Harb. Perspect. Med.* 8:a023507. doi: 10.1101/cshperspect.a023507

Westra, J. W., Barral, S., and Chun, J. (2009). A reevaluation of tetraploidy in the Alzheimer's disease brain. *Neurodegener. Dis.* 6, 221–229. doi: 10.1159/000236901

Westra, J. W., Rivera, R. R., Bushman, D. M., Yung, Y. C., Peterson, S. E., Barral, S., et al. (2010). Neuronal DNA content variation (DCV) with regional and individual differences in the human brain. *J. Comp. Neurol.* 518, 3981–4000. doi: 10.1002/cne.22436

Wiseman, F. K., Al-Janabi, T., Hardy, J., Karmiloff-Smith, A., Nizetic, D., Tybulewicz, V. L., et al. (2015). A genetic cause of Alzheimer disease: mechanistic insights from Down syndrome. *Nat. Rev. Neurosci.* 16, 564–574. doi: 10.1038/nrn3983

Yurov, Y. B., Vorsanova, S. G., Liehr, T., Kolotii, A. D., and Iourov, I. Y. (2014). X chromosome aneuploidy in the Alzheimer's disease brain. *Mol. Cytogenet.* 7:20. doi: 10.1186/1755-8166-7-20

Zhou, F., and Wang, D. (2017). The associations between the MAPT polymorphisms and Alzheimer's disease risk: a meta-analysis. *Oncotarget* 8, 43506–43520. doi: 10.18632/oncotarget.16490

Conflict of Interest: JC is the co-founder of Mosaic Pharmaceuticals.

GK declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Genetics

ORIGINAL RESEARCH
published: 16 September 2020
doi: 10.3389/fgene.2020.519206

Establishment of Quantitative PCR Assays for Active Long Interspersed Nuclear Element-1 Subfamilies in Mice and Applications to the Analysis of Aging-Associated Retrotransposition

*Ryota Kuroki[1†], Yui Murata[1†], Satoshi Fuke[2‡], Yutaka Nakachi[1], Jun Nakashima[1], Gregory C. Kujoth[3], Tomas A. Prolla[3], Miki Bundo[1,4], Tadafumi Kato[2] and Kazuya Iwamoto[1]**

[1]Department of Molecular Brain Science, Graduate School of Medical Sciences, Kumamoto University, Kumamoto, Japan, [2]Lab for Molecular Dynamics of Mental Disorders, RIKEN Center for Brain Science, Wako, Japan, [3]Department of Genetics and Medical Genetics, University of Wisconsin, Madison, WI, United States, [4]PRESTO, Japan Science and Technology Agency, Saitama, Japan

The retrotransposon long interspersed nuclear element-1 (LINE-1) can autonomously increase its copy number within a host genome through the retrotransposition process. LINE-1 is active in the germline and in neural progenitor cells, and its somatic retrotransposition activity has a broad impact on neural development and susceptibility to neuropsychiatric disorders. The method to quantify the genomic copy number of LINE-1 would be important in unraveling the role of retrotransposition, especially in the brain. However, because of the species-specific evolution of LINE-1 sequences, methods for quantifying the copy number should be independently developed. Here, we developed a quantitative PCR (qPCR) assay to measure the copy number of active LINE-1 subfamilies in mice. Using the assay, we investigated aging-associated alterations of LINE-1 copy number in several brain regions in wild-type mice and *Polg*$^{+/D257A}$ mice as a model for accelerated aging. We found that aged *Polg*$^{+/D257A}$ mice showed higher levels of the type GfII LINE-1 in the basal ganglia than the wild-type mice did, highlighting the importance of assays that focus on an individual active LINE-1 subfamily.

Keywords: retrotransposition, mitochondrial DNA, non-LTR, basal ganglia, somatic mosaicism, *POLG*, aging

OPEN ACCESS

Edited by:
Henry H. Q. Heng, Wayne State University, United States

Reviewed by:
Sandra Rose Richardson, University of Queensland, Australia
Giovanni Pascarella, RIKEN Center for Integrative Medical Sciences (IMS), Japan

***Correspondence:**
Kazuya Iwamoto
iwamotok@kumamoto-u.ac.jp

†These authors have contributed equally to this work

‡Present address:
Satoshi Fuke
Research Unit/Neuroscience, Sohyaku. Innovative Research Division, Mitsubishi Tanabe Pharma Corporation, Yokohama, Japan

Specialty section:
This article was submitted to Human Genomics, a section of the journal Frontiers in Genetics

Received: *11 December 2019*
Accepted: *20 August 2020*
Published: *16 September 2020*

Citation:
Kuroki R, Murata Y, Fuke S, Nakachi Y, Nakashima J, Kujoth GC, Prolla TA, Bundo M, Kato T and Iwamoto K (2020) Establishment of Quantitative PCR Assays for Active Long Interspersed Nuclear Element-1 Subfamilies in Mice and Applications to the Analysis of Aging-Associated Retrotransposition. Front. Genet. 11:519206. doi: 10.3389/fgene.2020.519206

INTRODUCTION

Long interspersed nuclear element-1 (LINE-1) is a retrotransposon with a length of approximately 6 kb. It occupies approximately 17 and 19% of the human genome and mouse genome, respectively (Lander et al., 2001; Mouse Genome Sequencing Consortium et al., 2002). Full-length LINE-1 is composed of a 5' untranslated region (UTR), open reading frame (ORF) 1, ORF2, a 3'UTR, and a poly-A tail. ORF1 encodes an RNA-binding protein (Holmes et al., 1992;

Hohjoh and Singer, 1997), and ORF2 encodes the protein with reverse transcriptase and endonuclease activity (Mathias et al., 1991; Feng et al., 1996). LINE-1 can increase its copy number within the host genome autonomously by a process called retrotransposition, which involves transcription of LINE-1, translation of ORFs, and translocating LINE-1 transcripts to the nucleus for their reverse transcription (Hohjoh and Singer, 1997; Cost et al., 2002). Retrotransposon activity is known to occur in germline cells and during early embryogenesis. When the newly transcribed copy of LINE-1 is inserted into genomic regions, it often affects genome stability and gene expression, resulting in a number of Mendelian diseases (Goodier and Kazazian, 2008; Cordaux and Batzer, 2009; Hancks and Kazazian, 2016). In addition to germline cells, recent findings suggest that LINE-1 is also active in neural precursor cells during early neurodevelopment and adult neurogenesis in the hippocampus, resulting in somatic mosaicism in brain cells (Erwin et al., 2014; Evrony et al., 2016; Faulkner and Billon, 2018; Saleh et al., 2019). Somatic LINE-1 retrotransposition in neurons is considered to be involved in the pathophysiology of neuropsychiatric disorders (Muotri et al., 2010; Coufal et al., 2011; Bundo et al., 2014; Iwamoto, 2019; Saleh et al., 2019).

In the other cell types and in most of the developmental stages, LINE-1 activity is strictly suppressed by multiple mechanisms, including genetic, epigenetic, posttranscriptional, and posttranslational regulation, depending on the type and evolutionary origin of LINE-1 (Goodier and Kazazian, 2008; Goodier, 2016). However, in addition to cancer (Rodic, 2018), accumulating evidence suggests that aging may be associated with increased LINE-1 activity (Saleh et al., 2019). The expression level and copy number of LINE-1 are increased with aging in liver and muscle tissue in mice (De Cecco et al., 2013; Min et al., 2019) and in senescent cells (De Cecco et al., 2019). An increase in LINE-1 copy number was also reported in the brains of adult rats compared to those of younger rats (Giorgi et al., 2018), and a similar increase was observed in mice with a deficiency in SIRT6, which is a regulator of longevity (Liao and Kennedy, 2016; Simon et al., 2019). Whether LINE-1 in nondividing mature neurons exhibits retrotransposition remains unclear, a study showed that engineered LINE-1 can retrotranspose in human neurons (Macia et al., 2017).

Estimation of the active LINE-1 copy number in human and model animals will provide important information for understanding the role of retrotransposition. For this purpose, a quantitative PCR (qPCR)-based estimation technique has been used, because it allows high-throughput measuring in a cost-effective manner. However, the structure and evolutionary characteristics of LINE-1 differ between humans and model animals, such as mice. In humans, only the most evolutionarily young LINE-1 subfamily, Hs, retains retrotransposition activity (Skowronski, et al., 1988), while at least three subfamilies (A, Gf, and Tf) retain activity in mice (Sookdeo et al., 2013). These three subfamilies are further classified into three A types (AI, AII, and AIII), two Gf types (GfI and GfII), and three Tf types (TfI, TfII, and TfIII). In addition, in mice, LINE-1 contains repeat sequences called monomers in the 5'UTR, which are not present in human LINE-1. Given that different active subfamilies in mice have different transcriptional activity and epigenetic profiles (DeBerardinis and Kazazian, 1999; Bulut-Karslioglu et al., 2014; Murata et al., 2017), detailed analysis of specific subfamilies is critically important.

Here, we developed a qPCR-based assay to quantify the copy number of active LINE-1 subfamilies in mice. Using this assay, we investigated aging-associated LINE-1 copy number change in *Polg*$^{+/D257A}$ mice, which we considered an animal model of chronic progressive external ophthalmoplegia (CPEO) that exhibits a premature aging characterized by accumulation of deleted mtDNA and motor dysfunction (Fuke et al., 2014). *POLG* is a nuclear-encoded mitochondrial DNA (mtDNA) polymerase, and its mutations are known to cause CPEO and associate with psychiatric disorders (Kasahara et al., 2017; Kato, 2019). Mice carrying a D257A knock-in mutation in *Polg* (*Polg*$^{D257A/D257A}$) lost proofreading activity of mtDNA and showed drastic accelerated aging phenotypes, including weight loss, reduced subcutaneous fat, hair loss, curvature of the spine, osteoporosis, and a reduced life span (Trifunovic et al., 2004; Kujoth et al., 2005). Although the mice carrying the heterozygous *Polg*^{D257A} (*Polg*$^{+/D257A}$) were reportedly normal (Trifunovic et al., 2004; Kujoth et al., 2005), we previously found that they showed age-dependent increased accumulation of mtDNA deletions and behavioral alterations, including motor dysfunction (Fuke et al., 2014). In this study, we found a subfamily-specific increase in the LINE-1 copy number in the basal ganglia of aged *Polg*$^{+/D257A}$ mice, showing the importance of a specific assay focusing on an individual member of the LINE-1 subfamilies.

MATERIALS AND METHODS

Primer Design

Consensus sequences of LINE-1 subfamilies in mice, including active subfamilies (Tf, A, and Gf) were retrieved from Repbase (Bao et al., 2015; Kojima, 2018). We designed both forward and reverse PCR primers with unique sequences for each active subfamily at their 3' ends. Primer sequences were searched for homologous consensus sequences using GENETYX ver.13 (GENETYX, Tokyo, Japan) to rule out the possibility of incorrect annealing. For the purpose of quality control, initial PCR was performed using rTaq DNA Polymerase (TOYOBO, Osaka, Japan) with a total of 5 ng of mouse genomic DNA as a template. PCR conditions were as follows: 1 min at 94°C followed by 40 cycles of 15 s at 95°C and 45 s at 65°C. Electrophoresis was performed on 2% agarose gel and visualized using GelGreen (COSMO BIO, Tokyo, Japan). Direct Sanger sequencing of PCR products was performed on all the candidate products after ExoSAP-IT Express PCR Cleanup Reagents (Thermo Fisher SCIENTIFIC, Waltham, Massachusetts, United States) were used to purify the DNA (Eurofins Genomics Inc., Tokyo, Japan).

Quantitative PCR

qPCR was performed using THUNDERBIRD SYBR qPCR mix (TOYOBO) and a total of 500 pg of genomic DNA as a template; reactions were carried out on a Quantstudio® 5

Real-Time PCR System (Thermo Fisher SCIENTIFIC). All primer pairs were used at a 5 μM concentration. qPCR conditions were the same as those listed above. The melting curve analysis conditions were as follows: 15 s at 95°C, 15 s at 60°C, and 15 s at 95°C. LINE-1 copy number was adjusted using internal control, 5srRNA, used previously (Muotri et al., 2010; Bundo et al., 2014). Quantification was performed in triplicate per sample. Raw Ct data are available upon request.

TA Cloning and Sequencing of Single Colonies

PCR products amplified with the GfII_ORF1 and GfI_5'UTR-ORF1 primer pairs (**Table 1**) were TA-cloned into a pCR4-TOPO vector using a TOPO TA cloning kit (Thermo Fisher SCIENTIFIC). We then transformed a DH5α strain with the vector samples and sequenced plasmids derived from single colonies.

Dendritic Tree

Dendritic trees were drawn using the mouse LINE-1 consensus sequences, using NJplot (Perrière and Gouy, 1996) based on maximum likelihood phylogenetic tree by PsyML (Dereeper et al., 2008). The tree of PCR amplicons (N = 65 for GFII_ORF1 and N = 49 for GfI_5'UTR-ORF1) was derived from TA-cloning, using GENETYX ver.13 based on neighbor joining method and a Kimura 2-parameter model. Three consensus sequences (L1MM_F, L1VL1_5, and L1VL2_5) were removed from these analyses because they did not have a corresponding sequence to the target region.

Animal Model

Polg^{D257A} mice were described previously (Kujoth et al., 2005; Fuke et al., 2014). In brief, *Polg*$^{+/D257A}$ mice (Kujoth et al., 2005) were backcrossed with C57BL/6JJcl mice for more than six generations (Fuke et al., 2014). Five brain regions, frontal lobe, hippocampus, posterior cortex, basal ganglia, and cerebellum were dissected, and genomic DNAs were extracted from them as described (Fuke et al., 2014).

TABLE 1 | List of primer pairs used in this study.

Subfamily	Primer name		Sequence (5' -> 3')
Universal	m5UTR	F	TAAGAGAGCTTGCCAGCAGAGA
		R	GCAGACCTGGGAGACAGATTCT
	mORF1	F	TGGAAGAGAGAATCTCAGGTGC
		R	TTGTGCCGATGTTCTCTATGG
	mORF2	F	CTGGCGAGGATGTGGAGAA
		R	CCTGCAATCCCACCAACAAT
AI, AII	A_ORF2_1	F	CACTTTAGTAAAGCTCAAAGCAT
		R	ATGTTCTGTAGATATCTGTCAGG
AI, AII, AIII	A_ORF1	F	GACCAAACCTACGGATAATAGGAATT
		R	GATCATGGGCATCTCTTTTTTTAT
	A_ORF2_2	F	TTGGCGTGACTCTAACTAAGGAG
		R	CCTAGGTTTTTTGTTATTCCAGACA
GfI	GfI_5'UTR-ORF1	F	AGAGAGCTTGTCTCCCACGC
		R	CATGAGATATGCTTTTAAATCCAGGTCTAC
GfII	GfII_ORF1	F	AACCCAAAGTGAGGCAACAG
		R	CATCCACTCCTA TTATCCGTAGGTTC
TfII	TfII_3'UTR	F	GGGATCCACCCCATAATCAGCTTCCAAAT
		R	TCCCCTGTACCGGGGCACAC
Internal control	m5srRNA	F	ACGGCCATACCACCCTGAA
		R	GGTCTCCCATCCAAGTACTAACCA

Universal and internal control pairs were previously reported (Muotri et al., 2010; Bundo et al., 2014).

Statistical Analysis

Welch's t-test was employed for comparison between two groups. $p < 0.05$ was considered significant. We considered a robust change to occur only if changes in both tissues were supported by statistical significance (Welch's t-test, $p < 0.05$ in both tissues).

RESULTS

Establishment of qPCR Assays for Individual Active LINE-1 Subfamilies in Mice

We retrieved a total of 34 mouse LINE-1 consensus sequences from Repbase, which is a database of repetitive DNA elements (Bao et al., 2015; Kojima, 2018). Based on the consensus sequences, primer pairs that can specifically amplify the active LINE-1 subfamily (A, Gf, and Tf) were designed. Because each active subfamily was further divided into subtypes, i.e., A for AI, AII, and AIII, Gf for GfI and GfII, and Tf for TfI, TfII, and TfIII (**Supplementary Figure S1**), we first tried to design primer pairs for each type and then designed primers to include several types within the same subfamily. After excluding the primer pairs that may incorrectly anneal to other LINE-1 locations, we designed a total of 28 primer pairs, including four pairs for AI, AII, and AIII, one pair for AI and AII, four pairs for GfI, eight pairs for GfII, two pairs for TfI, one pair for TfII, three pairs for TfIII, and five pairs for TfI and TfII. We then determined whether the designed primer pair produced a single amplicon by agarose gel electrophoresis (**Figure 1A**), followed by direct Sanger sequencing. Representative data of Gf_II ORF1 were shown in **Supplementary Figure S2**. Primer pairs were then tested using a melting curve analysis in a qPCR context (**Figure 1A**). Amplicons from the primer pairs, GfII_ORF1, and GfI_5'UTR-ORF1 (**Figures 1B,C** and **Table 1**) were further analyzed by TA-cloning. In GfII_ORF1, sequences obtained from 65 individual bacterial colonies revealed that 81.5% (53/65) of amplicons showed high similarity with the GfII consensus sequence. Other amplicons (12/65) were also considered to be GfII variants because all of them contained unique GfII-specific sequences (**Supplementary Figure S2**). Similarly, in GfI_5'UTR-ORF1, sequences obtained from 49 individual bacterial colonies revealed that 79.6% (39/49) of amplicons showed high similarity with the GfI consensus sequence. All other amplicons (10/49) were also considered to be GfI variants because they contained unique GfI-specific sequences (data not shown). Finally, we obtained a total of six primer pairs that were highly specific for the target active LINE-1 subfamilies. They included two pairs for all three active type A (I, II, and III), one pair for AI and AII, and one pair each for GfI, GfII, and TfII (**Figure 1C** and **Table 1**).

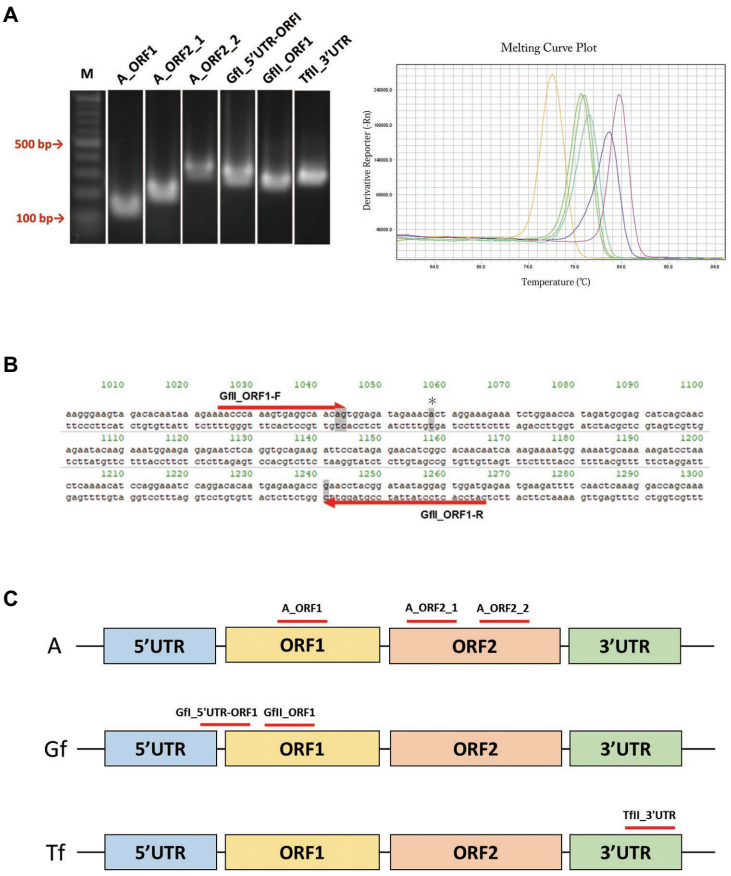

FIGURE 1 | PCR primer pairs specific for individual active long interspersed nuclear element-1 (LINE-1) subfamilies. **(A)** Agarose gel electrophoresis of PCR products, *left*. Melting curve analysis of PCR products in the quantitative PCR (qPCR) context, *right*. M, molecular size markers. **(B)** Sequences of the GfII_ORF1 primer pair. Primer sequences aligned with the consensus sequence of GfII are shown. The unique sequence in GfII is highlighted in gray. *Indicates the unique sequence used for sequencing analysis of PCR products. **(C)** Location of the validated primer pairs. Monomer sequencers are omitted from the illustrations.

Subfamily-Specific LINE-1 Copy Number Analysis of Various Brain Regions of *Polg*$^{+/D257A}$ Mice

We measured LINE-1 copy number in various brain regions (basal ganglia, cerebellum, frontal lobes, hippocampus, and posterior cortex) from the aged *Polg*$^{+/D257A}$ mice; these mice were shown to exhibit accumulation of mtDNA deletions during aging (Fuke et al., 2014). We examined the LINE-1 copy number in aged mice (84 weeks old) with conventional primer pairs that do not target specific LINE-1 subfamilies (Muotri et al., 2010; Bundo et al., 2014) and those we developed in this study (**Table 1**). Due to the multiple statistical testing methods and the limited number of samples, we used two different tissues, heart and skeletal muscles, as references. We considered a robust change to occur only if changes in both tissues were supported by statistical significance (Welch's *t*-test, $p < 0.05$ in both tissues). We found that the conventional primer pairs did not detect copy number changes in the tested brain regions from *Polg*$^{+/D257A}$ mice (**Figure 2A**). However, among the developed primer pairs, we found consistently higher GfII LINE-1 copy numbers in the basal ganglia in *Polg*$^{+/D257A}$ mice than in wild-type mice (**Figures 2A,B**). All the comparisons were listed in **Supplementary Figure S3**.

DISCUSSION

Here, we developed subfamily-specific LINE-1 copy number assays in mice and investigated age-related changes in LINE-1 copy number in the brains of *Polg*$^{+/D257A}$ mice. We found that aged *Polg*$^{+/D257A}$ mice showed an increase of GfII in the basal ganglia over what was seen in wild-type mice, highlighting the importance of specific assays focusing on individual active LINE-1 subfamilies.

We successfully generated a total of six primer pairs that were highly specific to target subfamilies. Copy number and expression analyses specifically targeting active LINE-1 subfamilies in mice were previously reported (Jachowicz et al., 2017; Bedrosian et al., 2018). However, the primers in those studies were designed to amplify conserved regions among three active subfamilies

A

		/ Heart					/ Skeletal muscle				
		Fl	Bg	Hp	Pc	Ce	Fl	Bg	Hp	Pc	Ce
Universal	m5UTR	0.99	1.08	1.00	1.06	0.96	1.02	1.11	1.03	1.10	1.03
	mORF1	1.01	0.96	1.01	1.03	0.96	1.02	0.97	1.02	1.04	0.97
	mORF2	1.03	1.03	1.01	1.06	0.94	1.05	1.04	1.01	1.07	0.94
Subfamily specific	A_ORF1	0.97	1.06	0.97	1.00	1.06	0.86	0.94	0.87	0.89	0.94
	A_ORF2_1	1.08	1.12	1.08	1.01	1.09	0.94	0.97	0.94	0.89	0.95
	A_ORF2_2	1.09	1.14	1.07	1.04	0.96	0.99	1.03	0.97	0.94	0.88
	Gfl_5'UTR-ORF1	1.00	1.04	1.05	1.01	1.04	0.93	0.97	0.99	0.94	0.97
	GfII_ORF1	1.06	1.22*	1.09	1.08	1.13	1.01	1.17*	1.04	1.03	1.07
	TfII_3'UTR	1.09	1.03	1.01	1.00	1.04	1.01	0.95	0.93	0.93	0.96

Fold change: > 1.10 | 1.10-1.00 | 1.00-0.90 | 0.90 >

B

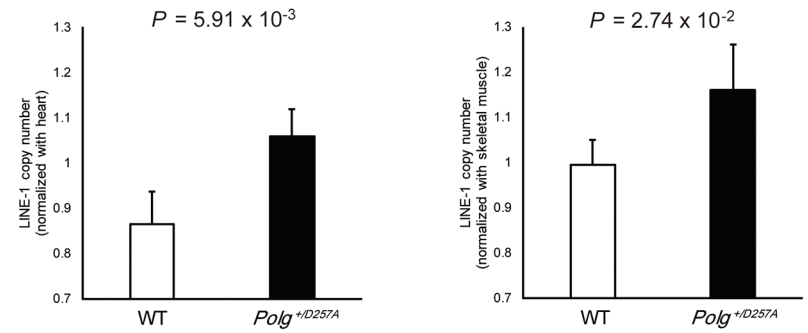

FIGURE 2 | LINE-1 copy number detected in 84-week-old *Polg*$^{+/D257A}$ mouse brain. **(A)** Changes in LINE-1 content in *Polg*$^{+/D257A}$ mice (n = 4) and wild-type mice (n = 4). The fold change relative to wild-type mice is shown in each reference tissue. Color indicates the extent of fold change with a nominal significant difference (Welch's t-test, $p < 0.05$). *Indicates a robust change defined as significant in both references (Welch's t-test, $p < 0.05$ in both tissues). Fl, frontal lobe; Hp, hippocampus; Pc, posterior cortex; Bg, basal ganglia; Ce, cerebellum. **(B)** LINE-1 copy number in the basal ganglia measured using the primer pair GfII_ORF1. The copy number in the brain was normalized to the number in the heart (left panel) or skeletal muscle (right panel). Data are represented as the mean ± standard deviation. WT, wild-type mice. All the comparisons were listed in **Supplementary Figure S3**.

(Bedrosian et al., 2018) or monomer regions (Jachowicz et al., 2017), which were located in the upstream region of the 5'-UTR of LINE-1; thus, those regions were not suitable for measuring somatic retrotransposition because the reverse transcription process usually stops prematurely.

In quantifying repetitive sequences such as LINE-1 by qPCR, the sequence specificity has been the critical confounding factor (Evrony et al., 2016). Based on previous reports, the total target LINE-1 copy numbers are estimated to be 3,466 for type A, 615 for GfI, 368 for GfII, and 1,282 for TfII in the full-length context (Sookdeo et al., 2013). Moreover, subfamily specificity of LINE-1 is important from a functional point of view. Transcriptional level of LINE-1 is proportional to the number of monomers in the 5'UTR (DeBerardinis and Kazazian, 1999), which are different in each subfamily. Epigenetic status, including DNA methylation and histone markers, is distinct in each LINE-1 subfamily (Bulut-Karslioglu et al., 2014; Murata et al., 2017). The subfamily specificity found in this study further supports the distinct regulation of LINE-1 retrotransposition activity in mouse brain cells.

It is noteworthy that in the protocol described here, we used 500 pg of genomic DNA as a template for qPCR. However, we confirmed that stable quantification data can be obtained from 100 pg of genomic DNA. Thus, the analysis is possible in more specific anatomical brain regions or in smaller cell populations. We also confirmed that the described primer pairs can be used for measuring subfamily-specific expression levels (Murata et al., unpublished data).

Several limitations should be kept in mind in applying the primers, however. First, because we put the highest priority in selecting primer sequences with high specificity for a target subfamily, some types within a subfamily were not assessed, or they were measured together. In the A subfamily, we obtained primer pairs common to AI, AII, and AIII, and a primer pair common to AI and AII. In the Gf subfamily, the primer pairs specific for GfI or GfII were independently established. In the Tf subfamily, the established primer pair measured TfII but not TfI or TfIII. Therefore, the results should be interpreted depending on the covered types. Second, the locations of the amplified regions differed among the primer pairs. Considering that the reverse transcription process is immaturely ended in general, the primer pairs targeted for the 3' end of LINE-1 would have more sensitivity for detecting retrotransposition events, whereas those targeting the upstream region of LINE-1 could examine more functional retrotransposition. Therefore, the sensitivity of the measured data will be different based on the target location of the primers. Third, although SYBR-based qPCR has been

widely used for LINE-1 copy number assay (Muotri et al., 2010; Bundo et al., 2014), other quantification approaches such as the Taqman-probes, the peptide nucleic acid-probes, and the droplet digital PCR technique (Newkirk et al., 2020) will improve the sensitivity and the specificity of LINE-1 copy number assay.

Increased activity of LINE-1 in aging and senescent cells has been reported (De Cecco et al., 2013, 2019; Liao and Kennedy, 2016; Giorgi et al., 2018; Min et al., 2019; Saleh et al., 2019; Simon et al., 2019). Our data showing an increased GfII copy number in the basal ganglia of aged *Polg*$^{+/D257A}$ mice seemed to be in accordance with these previous reports. *Polg*$^{D257A/D257A}$ mice showed a severe phenotype of premature aging, resulting in premature death starting at 40 weeks (Trifunovic et al., 2004; Kujoth et al., 2005), so we analyzed the heterozygous mutant in this study. Although the phenotypes of *Polg*$^{+/D257A}$ mice were reportedly normal, we previously observed the presence of mild motor dysfunction at 34 weeks and the accumulation of deleted mtDNAs from 48 weeks in the basal ganglia without reducing the life span.

Among the various brain regions we analyzed, we detected robust LINE-1 copy number change in the basal ganglia. Basal ganglia have a relatively higher number of mtDNAs compared to other brain regions (Fuke et al., 2014); thus, it may be a susceptible brain region to aging-related LINE-1 copy number change. Each active LINE-1 subfamily harbors unique structures of monomers, tandem repeats in the promoter regions and different epigenetic status in brain (Murata et al., 2017). These suggested that they have different expression pattern and distinctive roles. Therefore, increased GfII in basal ganglia during aging suggests that there might be GfII-specific regulators in basal ganglia, whose expressions were altered during aging.

We detected 1.1-fold change in GfII by qPCR. The standard curve analysis indicated that Ct values showed a linear relationship around this magnitude of change (data not shown). By roughly estimation, this change corresponds to increase of about 37 copies of GfII per cell. The copy number change of this magnitude has been often reported by qPCR analyses of LINE-1 (Coufal et al., 2009, for example). However, genome analyses of single neurons reported much smaller extent of changes that cannot be theoretically detected by qPCR (Evrony et al., 2012, 2016; Sanchez-Luque et al., 2019). Other approaches such as deep sequencing analysis will help to interpret the possible discrepancy.

Accumulation of deleted mtDNA has been observed in heart and skeletal muscles (Fuke et al., 2014). Because we used these tissues as references in this study, our copy number estimation in brain may be confounded, if these tissues showed altered LINE-1 activities. However, we did not detect LINE-1 copy number change in heart normalized by skeletal muscle (and vice versa) between *Polg*$^{+/D257A}$ and wild-type mice (**Supplementary Figure S4**).

In senescent cells, an increased LINE-1 copy number is concomitant with increased expression of LINE-1, which is driven by increased expression of the activator FOXA1, decreased expression of the repressor RB1, and LINE-1 demethylation (De Cecco et al., 2019; Min et al., 2019). In addition, LINE-1 copy number in the cytosol is further increased by decreased TREX1 3' exonuclease (Thomas et al., 2017; De Cecco et al., 2019). A similar scenario in the brain of this aging mice model might be applicable. Future experiments will include examination of expression levels of the relevant genes, epigenetic status of LINE-1, and LINE-1 copy number in cytosolic DNA for better understanding of the role of retrotransposition and aging.

DATA AVAILABILITY STATEMENT

Data used in preparing this article can be available upon request.

ETHICS STATEMENT

The animal study was reviewed and approved by the Animal Experiment Committee of RIKEN (Wako, Saitama, Japan) and Kumamoto University (Kumamoto City, Kumamoto, Japan).

AUTHOR CONTRIBUTIONS

RK and YM equally contributed to the work. RK, YM, YN, and JN performed the experiments and data analyses. SF, GK, and TP provided the materials. TK, MB, and KI supervised the study. RK, YM, and KI wrote the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was partly supported by JSPS KAKENHI (16H06395, 16H06399, 16K21720, 17H01573, 18H02753, 18H05428, 18H05430, 18H05435, and 18K07567) and AMED (JP19dm0107123 and JP19dm0207074).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/article/10.3389/fgene.2020.519206/full#supplementary-material

REFERENCES

Bao, W., Kojima, K. K., and Kohany, O. (2015). Repbase Update, a database of repetitive elements in eukaryotic genomes. *Mob. DNA* 6:11. doi: 10.1186/s13100-015-0041-9

Bedrosian, T. A., Quayle, C., Novaresi, N., and Gage, F. H. (2018). Early life experience drives structural variation of neural genomes in mice. *Science* 359, 1395–1399. doi: 10.1126/science.aah3378

Bulut-Karslioglu, A., De La Rosa-Velázquez, I. A., Ramirez, F., Barenboim, M., Onishi-Seebacher, M., Arand, J., et al. (2014). Suv39h-dependent H3K9me3 marks intact retrotransposons and silences LINE elements in mouse embryonic stem cells. *Mol. Cell* 55, 277–290. doi: 10.1016/j.molcel.2014.05.029

Bundo, M., Toyoshima, M., Okada, Y., Akamatsu, W., Ueda, J., Nemoto-Miyauchi, T., et al. (2014). Increased l1 retrotransposition in the neuronal genome in schizophrenia. *Neuron* 81, 306–313. doi: 10.1016/j.neuron.2013.10.053

Cordaux, R., and Batzer, M. A. (2009). The impact of retrotransposons on human genome evolution. *Nat. Rev. Genet.* 10, 691–703. doi: 10.1038/nrg2640

Cost, G. J., Feng, Q., Jacquier, A., and Boeke, J. D. (2002). Human L1 element target-primed reverse transcription in vitro. *EMBO J.* 21, 5899–5910. doi: 10.1093/emboj/cdf592

Coufal, N. G., Garcia-Perez, J. L., Peng, G. E., Marchetto, M. C., Muotri, A. R., Mu, Y., et al. (2011). Ataxia telangiectasia mutated (ATM) modulates long interspersed element-1 (L1) retrotransposition in human neural stem cells. *Proc. Natl. Acad. Sci. U. S. A.* 108, 20382–20387. doi: 10.1073/pnas.1100273108

Coufal, N. G., Garcia-Perez, J. L., Peng, G. E., Yeo, G. W., Mu, Y., Lovci, M. T., et al. (2009). L1 retrotransposition in human neural progenitor cells. *Nature* 460, 1127–1131. doi: 10.1038/nature08248

DeBerardinis, R. J., and Kazazian, H. H. Jr. (1999). Analysis of the promoter from an expanding mouse retrotransposon subfamily. *Genomics* 56, 317–323. doi: 10.1006/geno.1998.5729

De Cecco, M., Criscione, S. W., Peterson, A. L., Neretti, N., Sedivy, J. M., and Kreiling, J. A. (2013). Transposable elements become active and mobile in the genomes of aging mammalian somatic tissues. *Aging* 5, 867–883. doi: 10.18632/aging.100621

De Cecco, M., Ito, T., Petrashen, A. P., Elias, A. E., Skvir, N. J., Criscione, S. W., et al. (2019). L1 drives IFN in senescent cells and promotes age-associated inflammation. *Nature* 566, 73–78. doi: 10.1038/s41586-018-0784-9

Dereeper, A., Guignon, V., Blanc, G., Audic, S., Buffet, S., Chevenet, F., et al. (2008). Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Res.* 36, W465–W469. doi: 10.1093/nar/gkn180

Erwin, J. A., Marchetto, M. C., and Gage, F. H. (2014). Mobile DNA elements in the generation of diversity and complexity in the brain. *Nat. Rev. Neurosci.* 15, 497–506. doi: 10.1038/nrn3730

Evrony, G. D., Cai, X., Lee, E., Hills, L. B., Elhosary, P. C., Lehmann, H. S., et al. (2012). Single-neuron sequencing analysis of L1 retrotransposition and somatic mutation in the human brain. *Cell* 151, 483–496. doi: 10.1016/j.cell.2012.09.035

Evrony, G. D., Lee, E., Park, P. J., and Walsh, C. A. (2016). Resolving rates of mutation in the brain using single-neuron genomics. *elife* 5:e12966. doi: 10.7554/eLife.12966

Faulkner, G. J., and Billon, V. (2018). L1 retrotransposition in the soma: a field jumping ahead. *Mob. DNA* 9:22. doi: 10.1186/s13100-018-0128-1

Feng, Q., Moran, J. V., Kazazian, H. H. Jr., and Boeke, J. D. (1996). Human L1 retrotransposon encodes a conserved endonuclease required for retrotransposition. *Cell* 87, 905–916. doi: 10.1016/S0092-8674(00)81997-2

Fuke, S., Kametani, M., Yamada, K., Kasahara, T., Kubota-Sakashita, M., Kujoth, G. C., et al. (2014). Heterozygous Polg mutation causes motor dysfunction due to mtDNA deletions. *Ann. Clin. Transl. Neurol.* 1, 909–920. doi: 10.1002/acn3.133

Giorgi, G., Virgili, M., Monti, B., and Del Re, B. (2018). Long interspersed nuclear elements (LINEs) in brain and non-brain tissues of the rat. *Cell Tissue Res.* 374, 17–24. doi: 10.1007/s00441-018-2843-9

Goodier, J. L. (2016). Restricting retrotransposons: a review. *Mob. DNA* 7:16. doi: 10.1186/s13100-016-0070-z

Goodier, J. L., and Kazazian, H. H. Jr. (2008). Retrotransposons revisited: the restraint and rehabilitation of parasites. *Cell* 135, 23–35. doi: 10.1016/j.cell.2008.09.022

Hancks, D. C., and Kazazian, H. H. Jr. (2016). Roles for retrotransposon insertions in human disease. *Mob. DNA* 7:9. doi: 10.1186/s13100-016-0065-9

Hohjoh, H., and Singer, M. F. (1997). Sequence-specific single-strand RNA binding protein encoded by the human LINE-1 retrotransposon. *EMBO J.* 16, 6034–6043. doi: 10.1093/emboj/16.19.6034

Holmes, S. E., Singer, M. F., and Swergold, G. D. (1992). Studies on p40, the leucine zipper motif-containing protein encoded by the first open reading frame of an active human LINE-1 transposable element. *J. Biol. Chem.* 267, 19765–19768.

Iwamoto, K. (2019). Understanding the epigenetic architecture of psychiatric disorders: modifications and beyond. *Psychiatry Clin. Neurosci.* 72:194. doi: 10.1111/pcn.12646

Jachowicz, J. W., Bing, X., Pontabry, J., Bošković, A., Rando, O. J., and Torres-Padilla, M. E. (2017). LINE-1 activation after fertilization regulates global chromatin accessibility in the early mouse embryo. *Nat. Genet.* 49, 1502–1510. doi: 10.1038/ng.3945

Kasahara, T., Ishiwata, M., Kakiuchi, C., Fuke, S., Iwata, N., Ozaki, N., et al. (2017). Enrichment of deleterious variants of mitochondrial DNA polymerase gene (POLG1) in bipolar disorder. *Psychiatry Clin. Neurosci.* 71, 518–529. doi: 10.1111/pcn.12496

Kato, T. (2019). Current understanding of bipolar disorder: toward integration of biological basis and treatment strategies. *Psychiatry Clin. Neurosci.* 73, 526–540. doi: 10.1111/pcn.12852

Kojima, K. K. (2018). Human transposable elements in Repbase: genomic footprints from fish to humans. *Mob. DNA* 9:2. doi: 10.1186/s13100-017-0107-y

Kujoth, G. C., Hiona, A., Pugh, T. D., Someya, S., Panzer, K., Wohlgemuth, S. E., et al. (2005). Mitochondrial DNA mutations, oxidative stress, and apoptosis in mammalian aging. *Science* 309, 481–484. doi: 10.1126/science.1112125

Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., et al. (2001). Initial sequencing and analysis of the human genome. *Nature* 409, 860–921. doi: 10.1038/35057062

Liao, C. Y., and Kennedy, B. K. (2016). SIRT6, oxidative stress, and aging. *Cell Res.* 26, 143–144. doi: 10.1038/cr.2016.8

Macia, A., Widmann, T. J., Heras, S. R., Ayllon, V., Sanchez, L., Benkaddour-Boumzaouad, M., et al. (2017). Engineered LINE-1 retrotransposition in nondividing human neurons. *Genome Res.* 27, 335–348. doi: 10.1101/gr.206805.116

Mathias, S. L., Scott, A. F., Kazazian, H. H. Jr., Boeke, J. D., and Gabriel, A. (1991). Reverse transcriptase encoded by a human transposable element. *Science* 254, 1808–1810. doi: 10.1126/science.1722352

Min, B., Jeon, K., Park, J. S., and Kang, Y. K. (2019). Demethylation and derepression of genomic retroelements in the skeletal muscles of aged mice. *Aging Cell* 18:e13042. doi: 10.1111/acel.13042

Mouse Genome Sequencing Consortium, Waterston, R. H., Lindblad-Toh, K., Birney, E., Rogers, J., Abril, J. F., et al. (2002). Initial sequencing and comparative analysis of the mouse genome. *Nature* 420, 520–562. doi: 10.1038/nature01262

Muotri, A. R., Marchetto, M. C., Coufal, N. G., Oefner, R., Yeo, G., Nakashima, K., et al. (2010). L1 retrotransposition in neurons is modulated by MeCP2. *Nature* 468, 443–446. doi: 10.1038/nature09544

Murata, Y., Bundo, M., Ueda, J., Kubota-Sakashita, M., Kasai, K., Kato, T., et al. (2017). DNA methylation and hydroxymethylation analyses of the active LINE-1 subfamilies in mice. *Sci. Rep.* 7:13624. doi: 10.1038/s41598-017-14165-7

Newkirk, S. J., Kong, L., Jones, M. M., Habben, C. E., Bishop, V., Ye, P., et al. (2020). Subfamily-specific quantification of endogenous mouse L1 retrotransposons by droplet digital PCR. *Anal. Biochem.* 601:113779. doi: 10.1016/j.ab.2020.113779

Perrière, G., and Gouy, M. (1996). WWW-query: an on-line retrieval system for biological sequence banks. *Biochimie* 78, 364–369. doi: 10.1016/0300-9084(96)84768-7

Rodic, N. (2018). LINE-1 activity and regulation in cancer. *Front. Biosci.* 23:4666. doi: 10.2741/4666

Saleh, A., Macia, A., and Muotri, A. R. (2019). Transposable elements, inflammation, and neurological disease. *Front. Neurol.* 10:894. doi: 10.3389/fneur.2019.00894

Sanchez-Luque, F. J., Kempen, M. H. C., Gerdes, P., Vargas-Landin, D. B., Richardson, S. R., Troskie, R. L., et al. (2019). LINE-1 evasion of epigenetic repression in humans. *Mol. Cell* 75, 590.e12–604.e12. doi: 10.1016/j.molcel.2019.05.024

Simon, M., Van Meter, M., Ablaeva, J., Ke, Z., Gonzalez, R. S., Taguchi, T., et al. (2019). LINE1 derepression in aged wild-type and SIRT6-deficient mice drives inflammation. *Cell Metab.* 29, 871.e5–885.e5. doi: 10.1016/j.cmet.2019.02.014

Skowronski, J., Fanning, T. G., and Singer, M. F. (1988). Unit-length line-1 transcripts in human teratocarcinoma cells. *Mol. Cell. Biol.* 8, 1385–1397. doi: 10.1128/MCB.8.4.1385

Sookdeo, A., Hepp, C. M., McClure, M. A., and Boissinot, S. (2013). Revisiting the evolution of mouse LINE-1 in the genomic era. *Mob. DNA* 4:3. doi: 10.1186/1759-8753-4-3

Thomas, C. A., Tejwani, L., Trujillo, C. A., Negraes, P. D., Herai, R. H., Mesci, P., et al. (2017). Modeling of TREX-dependent autoimmune disease using human stem cells highlights L1 accumulation as a source of neuroinflammation. *Cell Stem Cell* 21, 319–331. doi: 10.1016/j.stem.2017.07.009

Trifunovic, A., Wredenberg, A., Falkenberg, M., Spelbrink, J. N., Rovio, A. T., Bruder, C. E., et al. (2004). Premature ageing in mice expressing defective mitochondrial DNA polymerase. *Nature* 429, 417–423. doi: 10.1038/nature02517

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Genetics

METHODS
published: 16 November 2020
doi: 10.3389/fgene.2020.505441

SCCNV: A Software Tool for Identifying Copy Number Variation From Single-Cell Whole-Genome Sequencing

Xiao Dong[1], Lei Zhang[1], Xiaoxiao Hao[1], Tao Wang[2] and Jan Vijg[1,3]**

[1] Department of Genetics, Albert Einstein College of Medicine, Bronx, NY, United States, [2] Department of Epidemiology and Population Health, Albert Einstein College of Medicine, Bronx, NY, United States, [3] Center for Single-Cell Omics in Aging and Disease, School of Public Health, Shanghai Jiao Tong University School of Medicine, Shanghai, China

OPEN ACCESS

Edited by:
Ivan Y. Iourov, Mental Health Research Center of Russian Academy of Medical Sciences, Russia

Reviewed by:
Oxana Kurinnaia, Veltischev Research and Clinical Institute for Pediatrics of the Pirogov Russian National Research Medical University, Russia
Luis Alberto Méndez-Rosado, Centro Nacional de Genética Médica, Cuba

***Correspondence:**
Xiao Dong
biosinodx@gmail.com;
xiao.dong@einsteinmed.org
Jan Vijg
jan.vijg@einsteinmed.org

Specialty section:
This article was submitted to Human Genomics, a section of the journal Frontiers in Genetics

Received: *25 October 2019*
Accepted: *28 October 2020*
Published: *16 November 2020*

Citation:
Dong X, Zhang L, Hao X, Wang T and Vijg J (2020) SCCNV: A Software Tool for Identifying Copy Number Variation From Single-Cell Whole-Genome Sequencing. Front. Genet. 11:505441. doi: 10.3389/fgene.2020.505441

Identification of *de novo* copy number variations (CNVs) across the genome in single cells requires single-cell whole-genome amplification (WGA) and sequencing. Although many experimental protocols of amplification methods have been developed, all suffer from uneven distribution of read depth across the genome after sequencing of DNA amplicons, which constrains the usage of conventional CNV calling methodologies. Here, we present SCCNV, a software tool for detecting CNVs from whole genome-amplified single cells. SCCNV is a read-depth based approach with adjustment for the WGA bias. We demonstrate its performance by analyzing data obtained with most of the single-cell amplification methods that have been employed for CNV analysis, including DOP-PCR, MDA, MALBAC, and LIANTI. SCCNV is freely available at https://github.com/biosinodx/SCCNV.

Keywords: single-cell whole-genome sequencing, single-cell whole-genome amplification, amplification bias, copy number variation, software development

INTRODUCTION

Each single cell in a tissue or cell population has its own unique genome due to accumulating *de novo* mutations, such as single-nucleotide variations (SNVs), structural variations (SVs), copy number variations (CNVs) and aneuploidies. The frequency and spectrum of the mutations reflect the loss of genome integrity of a cell population, critically important to cancer and aging (Vijg and Dong, 2020). To detect the mutations unique to a single cell, single-cell whole-genome sequencing (SCWGS) is necessary. SCWGS requires whole-genome amplification (WGA), which is often biased, leading to uneven distribution of DNA content across the genome or differences between alleles. This essentially constrain the usage of variant callers designed for non-amplified bulk DNA. We recently developed a new software tool, SCcaller, that uses heterozygous SNPs to correct for the allelic bias hampering SNV calling (Dong et al., 2017).

CNV calling is typically based on variation of sequencing depth across the genome. However, for a single cell amplicon, variation of sequencing depth increases dramatically due to the locus-specific amplification bias (Navin et al., 2011; Zong et al., 2012; Chen et al., 2017). To solve this issue computationally, we developed SCCNV, a software tool to identify CNVs from SCWGS. SCCNV is also based on a read-depth approach: it controls not only bias during sequencing and alignment, e.g., bias associated with mappability and GC content,

but also the locus-specific amplification bias. We demonstrate the performance of SCCNV using SCWGS data of multiple experimental protocols, i.e., DOP-PCR (degenerative-oligonucleotide PCR), MDA (multiple displacement amplification), MALBAC (multiple annealing and looping–based amplification cycles), and LIANTI (linear amplification via transposon insertion) (Navin et al., 2011; Gundry et al., 2012; Zong et al., 2012; Chen et al., 2017; Dong et al., 2017).

MATERIALS AND METHODS

SCCNV

Our software tool for analyzing single-cell copy number variation (SCCNV) was written in Python. Its source code is freely available with a usage description and an example at Github[1] under the GNU Affero General Public License v3.0. It uses SCWGS data after alignment as input (i.e., a bam file per single cell). Of note, SCCNV cannot take sequencing data of a pool of single cells (a bam file composed of thousands of single cells data), e.g., the 10× Genomics single-cell copy number data, as input.

First, SCCNV divides the genome into bins of equal size (500 kb as default), and counts the numbers of reads per bin of a cell. This step is relatively time-consuming, and we suggest users to use samtools on a high-performance computer cluster in parallel for all samples to be time-efficient (see instructions on Github). The remaining major steps of SCCNV do not require much computational resources – most modern desktop computers should work well.

SCCNV then normalizes mappability, which indicates the efficiency of the alignment to a genomic region. For a bin b of a cell, SCCNV adjusts the raw number of reads, denoted by NR_{raw}, by dividing over the mappability M,

$$NR_{\mathrm{map},b} = {NR_{\mathrm{raw},b}}/{M_b} \quad (1)$$

where mappability M is a value ranging from 0 to 1. SCCNV uses Encode Align100mer mappability score, downloaded from the UCSC genome browser, and calculates the mappability of each bin by using their weighted average.

Then, SCCNV normalizes for GC content. For a cell, SCCNV calculates the percentile of GC content of each bin. For a bin b of the cell, its number of aligned reads after normalizing GC content, $NR_{GC,b}$, is,

$$NR_{GC,b} = NR_{\mathrm{map},b} \times {NR_{\mathrm{map,genome}}}/{NR_{\mathrm{map},b,\mathrm{percentile}}} \quad (2)$$

where $NR_{map,genome}$ is the average NR_{map} per bin of all bins from the cell; $NR_{map,b,percentile}$ is the average NR_{map} per bin of bins in the same GC percentile as bin b.

After normalization for mappability and GC content, a pattern of sequencing read depth emerges that is consistent across different cells amplified using the same experimental protocol, i.e., the locus-specific amplification bias. Therefore, the bias is normalized across all cells in a particular batch and experiment. First, to make the $NR_{GC,b}$ comparable across cells, SCCNV converts it to a raw copy number estimate, denoted by $CN_{raw,b}$ for bin b of cell c, as follows,

$$CN_{\mathrm{raw},b,c} = {NR_{GC,b,c}}/{NR_{GC,\mathrm{genome},c}} \times \mathrm{ploidy} \quad (3)$$

where $NR_{GC,genome,c}$ is the median $NR_{GC,c}$ per bin in the genome of cell c; ploidy is 2 by default. Second, the adjusted copy number is estimated as,

$$CN_{\mathrm{adjusted},b,c} = {CN_{\mathrm{raw},b,c}}/{CN_{\mathrm{raw},b,-c}} \times \mathrm{ploidy} \quad (4)$$

where $CN_{raw,b,-c}$ denotes the average CN_{raw} for bin b across all cells except cell c. Of note, with this step SCCNV aims to discover the difference between the cell c and the other cells. When analyzing CNVs of multiple tumor cells, it is not appropriate to use all tumor cells as input of SCCNV; instead, one should use one tumor cell with two or more normal diploid cells as the input.

Then SCCNV uses a sliding window approach to further minimize amplification noise. By default, a window includes 11,500-kb bins, i.e., 5.5 Mb of DNA sequence in total, with a 500-kb step size between two neighboring windows,

$$CN_{smoothed,b,c} = \frac{1}{11}\sum_{i=b-5}^{b+5} CN_{\mathrm{adjusted},i,c} \quad (5)$$

SCCNV then models the distribution of $CN_{smoothed,b,c}$ of all bins in autosomes of a cell c as a normal distribution $N(\mu, {\sigma_c}^2)$. The $\mu = 2$, and σ is estimated as,

$$\sigma_c = |CN_{\mathrm{smoothed},30.9\%,c} - \mu| + |CN_{\mathrm{smoothed},69.1\%,c} - \mu| \quad (6)$$

where $CN_{smoothed,30.9\%,c}$ and $CN_{smoothed,69.1\%,c}$ are the 30.9 and 69.1% percentiles of the $CN_{smoothed,b,c}$ of all bins in the autosomes, corresponding to the $\mu - 0.5\sigma$ and $\mu + 0.5\sigma$ percentiles, respectively. Here, we did not use the observed s.d. of $CN_{smoothed,b,c}$ of all the bins because the normal distribution was to estimate amplification noise, not real variation in copy number across the genome. When a cell has several large CNVs, the s.d. will be high, even if its amplification noise remains low.

Assuming equally likely priors, for a bin b and a given possible copy number $k \in \{0, 1, 2, 3, 4\}$, its posterior probability is,

$$P(H_k|x) = {f_k(x)}\Big/{\sum_{i=0}^{4} f_k(x)} \quad (7)$$

where x is the $CN_{smoothed,b,c}$, and $f_i(x)$ is the probability density function of a normal distribution,

$$f_i(x) = \frac{1}{\sigma_c\sqrt{2\pi}}\exp\left(-\frac{(x-k)^2}{2\sigma_c^2}\right) \quad (8)$$

where the variance ${\sigma_c}^2$ is calculated according to Eq. (5). We only used $k \in \{0, 1, 2, 3, 4\}$ because the final copy number call was after multiple testing correction, i.e., Eq. (9) below, and we wished to minimize the number of hypotheses tested, i.e., five for a copy number of 0–4. However, this will result in an underestimation if the real copy number exceeds four. To resolve this issue, for bins with copy numbers $\geq$4 and $\leq$100, SCCNV reports the closest integer to the $CN_{smoothed,b,c}$.

[1] https://github.com/biosinodx/SCCNV

SCCNV allows <1 false positive per cell. Therefore, it determines bin b as a copy number variant when,

$$P(H_k|x) \geq 1 - \frac{1}{\text{GenomeSize}(3.2\,GB)/\text{WindowSize}(5\,\text{Mb})} = 0.998 \quad (9)$$

Sensitivity and False Positive Rate

To determine copy number, SCCNV is based on a statistical test described in equations (8) and (9) for a normal distribution and multiple testing correction separately. With a given value of coefficient of variation (CV) of $CN_{smoothed,b,c}$, sensitivity and FPR can be estimated as follows. Sensitivity equals the difference between two cumulative distribution functions (CDFs) of Eq. (8) at the upper and lower boundaries, which SCCNV provides the correct CNV call after the correction in Eq. (9). The percentage of FP out of all bins is equal to the sum of (a) CDF at the lower boundary of SCCNV providing an incorrect CN gain call; and (b) 1 – CDF at the upper boundary of SCCNV providing an incorrect CN loss call. Then FPR was estimated as the ratio of % of FP to the sum of % of FP and % of TN.

For example, under the assumption that the true copy number is 2, if SCCNV calls CN = 2 when $CN_{smoothed,b,c}$ is between 1.8 and 2.2, sensitivity = CDF(x = 2.2, μ = 2) – CDF(x = 1.8, μ = 2), in which CDF is the cumulative distribution function of Eq. (8). If SCCNV calls (a) CN = 1 when $CN_{smoothed,b,c}$ is between 0.8 and 1.2, and (b) CN = 3 when $CN_{smoothed,b,c}$ is between 2.8 and 3.2, then%FP = CDF(x = 1,2, μ = 2) + 1 – CDF(2.8, μ = 2).

Testing Datasets and Preprocessing of Data

Four SCWGS datasets were obtained for demonstrating and validating the performance of SCCNV (Zong et al., 2012; Lodato et al., 2015; Chen et al., 2017; Dong et al., 2017). The datasets included 8.2 TB SCWGS of 63 single human fibroblasts, neurons and cells of a tumor cell line amplified using eight different protocols, i.e., DOP-PCR (Sigma), Rubicon, MALBAC, LIANTI, and MDA (including Qiagen, GE, Lodato et al's MDA and

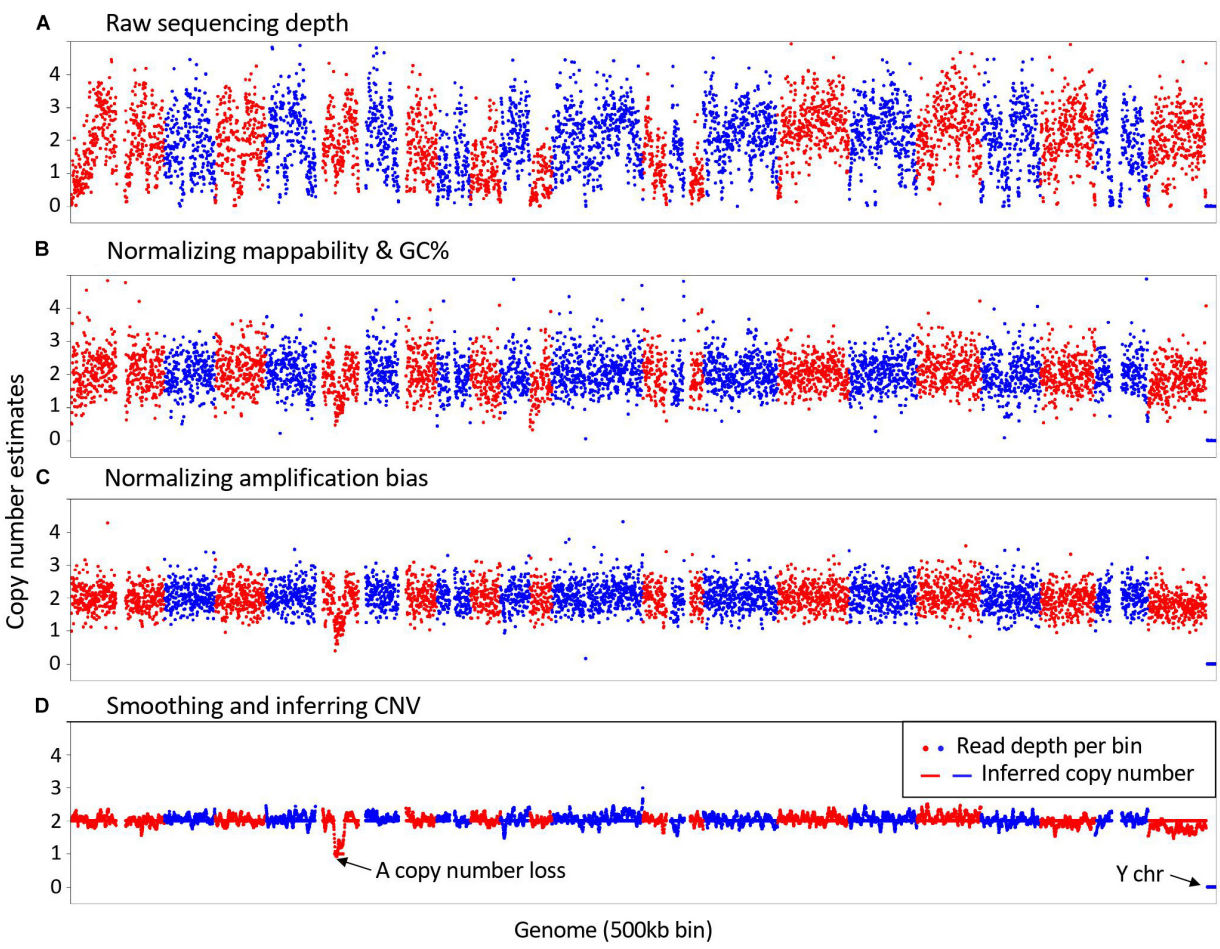

FIGURE 1 | Major steps in SCCNV. A example of copy number estimates of **(A)** Raw sequencing depth; **(B)** after normalizing mappability and GC content; **(C)** after normalizing amplification bias; and **(D)** final results is presented. The example is a normal neuronal nucleus amplified with MDA (SRA id: SRR2141574). Each dot presents a 500 kb bin in the genome. Red and blue colors indicate bins of different chromosomes in Red and blue colors present bins of different chromosomes in their lexicographic order.

SCMDA). **Supplementary Table 1** lists all the single-cell data used in this study.

Sequence alignment was performed using BWA and GATK as follows (Li and Durbin, 2009; McKenna et al., 2010). Raw sequencing data of each sample (single cell and bulk DNA) were obtained from the SRA database and subjected to quality control using FastQC (version 0.11.4;[2]) and trimming using Trim Galore (version 0.4.1;[3]) with default parameters. Then they were aligned to the human reference genome (version hg19) using BWA MEM (version 0.7.12; option: -t number of CPUs -M reference genome fasta file) (Li and Durbin, 2009). PCR duplications were removed using picard tools (version 1.119;[4]). The alignments were subjected to indel realignment and basepair recalibration using GATK (version 3.5; using options, RealignerTargetCreator, IndelRealigner, BaseRecalibrator, and PrintReads) (McKenna et al., 2010). The step above was used for generating an analysis-ready bam file for other types of variants, e.g., single nucleotide variants, small insertions and deletions, and this step is optional for large CNVs or aneuploidies using SCCNV. Reads with mapQ $<$ 30 were discarded. The number of reads per bin of each sample was calculated using samtools (version 1.3; option: bedcov) (Li et al., 2009). SCCNV (version 1.0) was used to estimate CNV of each cell.

RESULTS

Major Steps in SCCNV

As illustrated in **Figure 1**, SCCNV is composed of four major steps. It first calculates sequencing depth of the genome in bins of equal size (500-kb by default). Second, it normalizes the depth based on two features of the reference genome, including mappability and GC content. These two features are usually also considered by conventional CNV callers for bulk DNA sequencing. Next, it does further normalization across single cells of a same experimental batch. This step minimizes locus-specific bias due to WGA. Finally, it smooths the data (5 Mb by default) and infers copy number of each bin. Intermediate results between any two connecting steps can be generated by SCCNV for users to monitor its performance. We provide an example about the intermediate results of a normal neuronal nucleus amplified with MDA (SRA id: SRR2141574) in **Supplementary Figures 1–4**.

Performance on Real Datasets

To evaluate the performance of SCCNV, we obtained four SCWGS datasets from the SRA database, which includes 8.2 TB high-depth WGS data of 63 single human fibroblasts, neurons and cells of a tumor cell line amplified using eight different protocols, i.e., DOP-PCR (Sigma), Rubicon, MALBAC, LIANTI, and MDA (including four MDA protocols, Qiagen, GE, Lodato et al's MDA and SCMDA) (**Supplementary Table 1**; Zong et al., 2012; Lodato et al., 2015; Chen et al., 2017; Dong et al., 2017). The data were processed as described in the Materials and Methods.

We used the CV of sequencing depth across all genomic bins on autosomes as an indicator of performance, because it directly determines sensitivity and FPR of copy number calling step in SCCNV. We show sensitivity and FPR of the copy number calling in **Figures 2A,B**, respectively. As the CV decreased from 0.135 to 0.041, the sensitivity increased from 0 to 100% and the FPR decreased from 3.8×10^{-3} to 3.9×10^{-11}.

For the real datasets, we calculated the CV of raw data and normalized data after each step to demonstrate the performance of normalization in SCCNV (**Figure 3**). Almost all raw data (CV: 0.475 $\pm$ 0.135, avg. $\pm$ s.d.) are beyond the detection threshold, i.e., CV = 0.135. Each step of normalization decreased the CV by a significant fraction: on average, mappability normalization by 5%, GC content normalization by 33% percent, across-cell normalization by 22%, and smoothing by 55%. This shows the contributions of each normalization step to performance increase in the final variant calling. After all the normalization steps, the CVs are 0.107 $\pm$ 0.076 (avg. $\pm$ s.d.), corresponding to a

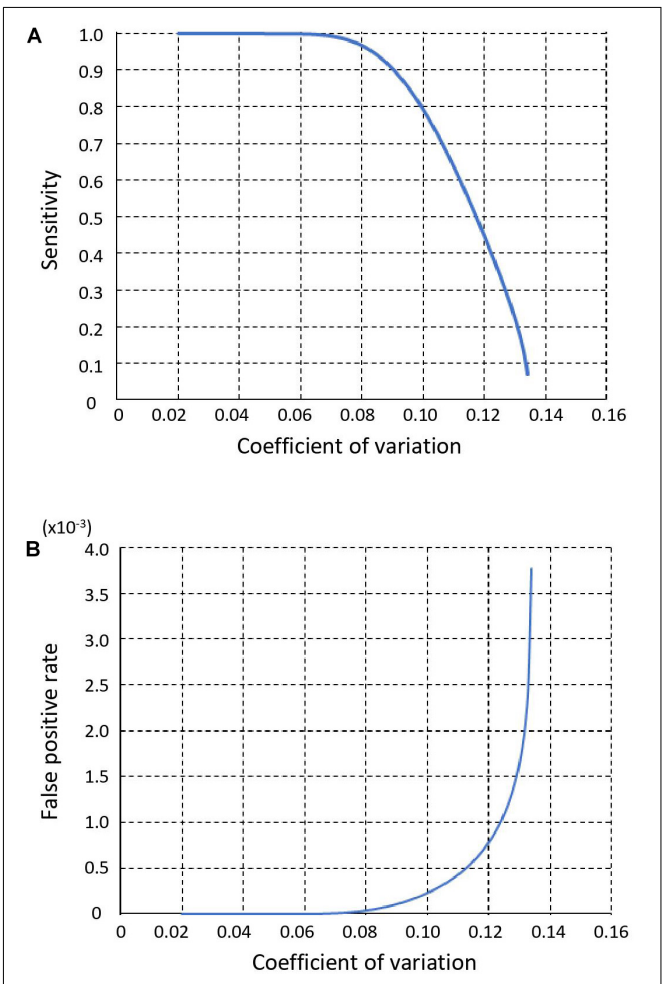

FIGURE 2 | Sensitivity **(A)** and False positive rate **(B)** of SCCNV. *Y*-axis value represents the coefficient of variation of the sequencing read depth (or normalized read depth) of all 500 kb bins across a single-cell genome (see section "Materials and Methods").

[2]https://www.bioinformatics.babraham.ac.uk/projects/fastqc/

[3]https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/

[4]http://broadinstitute.github.io/picard/index.html

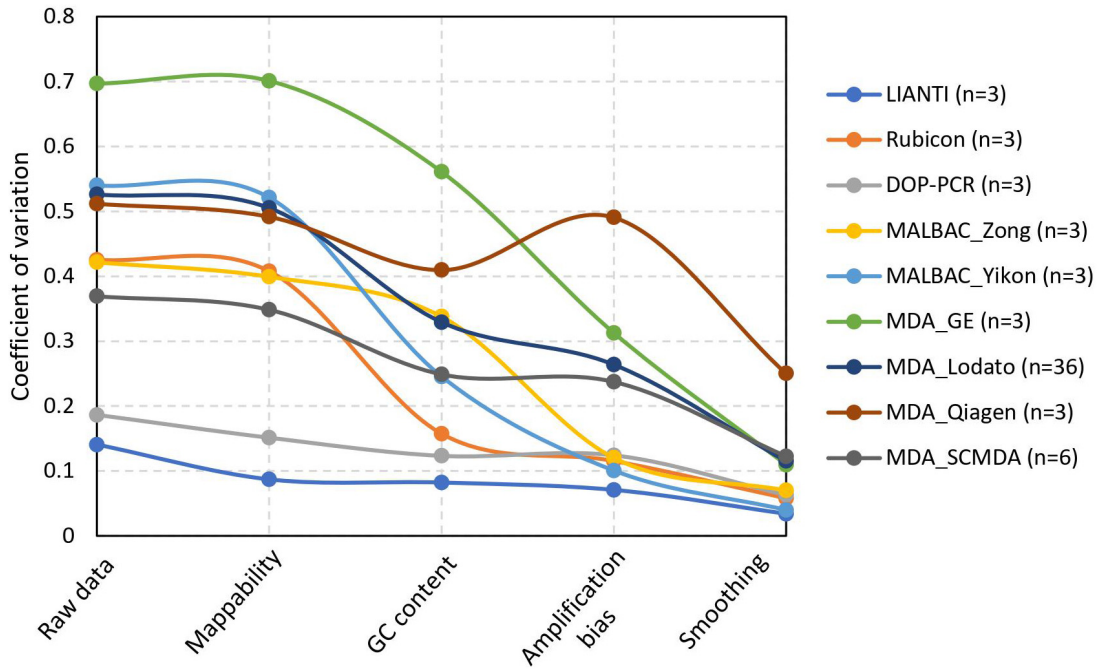

FIGURE 3 | Decreased coefficient of variation by normalization steps in SCCNV. Each line presents the average CV of normalized sequencing depths of multiple single cells amplified using the same experimental procedures. See **Supplementary Table 1** for citations and SRA ids of each experimental batch.

sensitivity of 68.6% and an FPR of 3.6×10^{-4} on average. Of note, different amplification protocols have significantly different performance when using SCCNV, likely due to differences in DNA amplification linearity among the protocols. As expected, LIANTI outperformed all the others (Chen et al., 2017). Protocols that included PCR steps, i.e., DOP-PCR, MALBAC and Rubicon, ranked in the middle. Although known as suffering from the least artifactual SNVs (Dong et al., 2017; Zhang et al., 2019), MDA-based protocols were ended last (**Figure 3**).

DISCUSSION

Identification of copy number variation and aneuploidy has been one of the major areas of genomics methods development. Several statistical models have been developed for analyzing initially microarray data and later sequencing data of bulk DNA, for example, Circular Binary Segmentation (CBS), Mean Shift-Based (MSB) model, Shifting Level Model (SLM), Expectation Maximization (EM) model, and Hidden Markov Model (HMM) as discussed in Zhao et al. (2013). Based on these models, multiple computational software tools have been developed, e.g., CBS, Copynumber, CNVnator, and HMMcopy (Olshen et al., 2004; Abyzov et al., 2011; Ha et al., 2012; Nilsen et al., 2012). To call CNVs, most of the methods rely on assessing either sequencing read depth or alternate allele fraction at heterozygous SNPs across the genome of one sample, i.e., across-genome normalization. Some of the methods have been applied directly for analyzing single-cell sequencing data with specific filtering for cells with too much bias after WGA.

A few new tools for single cell data were also developed recently under the same rationale (assessing one sample at a time, or across-genome normalization), such as AneuFinder, baseqCNV, Ginkgo and SCOPE (Garvin et al., 2015; Bakker et al., 2016; Fu et al., 2019; Wang et al., 2019). SCCNV was developed based on our observation that the locus-specific amplification bias is often the same in different cells within one experimental batch and amplified using the same protocol (e.g., **Supplementary Figure 3**); and we showed that normalization across multiple samples (cells) significantly contributed to the increase in variant calling performance for single cells amplified using most WGA protocols (**Figure 3**). Following the same across-sample normalization rationale, another software tool, SCNV, was developed (Wang et al., 2018). It differs from SCCNV that SCCNV performs normalization based on empirical data directly (Eq. 4) without any assumption on its distribution. Of note, with across-sample normalization, SCCNV essentially aims to identify differences among different cells in one input batch and, therefore, it is important to input cells of interest (e.g., tumor cells) together with cells with a standard diploid genome.

CONCLUSION

We developed SCCNV to identify copy number variations from whole-genome amplified single cells. We demonstrated its step-wise performance using most of the recent SCWGS datasets generated with 8 different amplification protocols.

DATA AVAILABILITY STATEMENT

Raw sequencing data of each sample were obtained from SRA database (SRA SRP067062, SRA SRA060929, SRA SRP102259, SRA SRP041470, and SRA SRP061939).

AUTHOR CONTRIBUTIONS

XD, LZ, and JV conceived the study. XD and TW developed the method. XD and XH analyzed the data. XD, LZ, and JV wrote the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work has been supported by the NIH grants P01 AG017242, P01 AG047200, P30 AG038072, K99 AG056656, U01 HL145560, and U01 ES029519, and the Paul F. Glenn Center for the Biology of Human Aging.

ACKNOWLEDGMENTS

This manuscript has been released as a Pre-Print at https:// www.biorxiv.org/ content/ 10.1101/ 535807v1 (Dong et al., 2019).

CODE AVAILABILITY STATEMENT

The source code of SCCNV can be found in the https://github.com/biosinodx/SCCNV.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2020.505441/full#supplementary-material

REFERENCES

Abyzov, A., Urban, A. E., Snyder, M., and Gerstein, M. (2011). CNVnator: an approach to discover, genotype, and characterize typical and atypical CNVs from family and population genome sequencing. *Genome Res.* 21, 974–984. doi: 10.1101/gr.114876.110

Bakker, B., Taudt, A., Belderbos, M. E., Porubsky, D., Spierings, D. C., De Jong, T. V., et al. (2016). Single-cell sequencing reveals karyotype heterogeneity in murine and human malignancies. *Genome Biol.* 17:115.

Chen, C., Xing, D., Tan, L., Li, H., Zhou, G., Huang, L., et al. (2017). Single-cell whole-genome analyses by linear amplification via transposon insertion (LIANTI). *Science* 356, 189–194. doi: 10.1126/science.aak9787

Dong, X., Zhang, L., Hao, X., Wang, T., and Vijg, J. (2019). SCCNV: a software tool for identifying copy number variation from single-cell whole-genome sequencing. *bioRxiv* [Preprint], doi: 10.1101/535807

Dong, X., Zhang, L., Milholland, B., Lee, M., Maslov, A. Y., Wang, T., et al. (2017). Accurate identification of single-nucleotide variants in whole-genome-amplified single cells. *Nat. Methods* 14, 491–493. doi: 10.1038/nmeth.4227

Fu, Y., Zhang, F., Zhang, X., Yin, J., Du, M., Jiang, M., et al. (2019). High-throughput single-cell whole-genome amplification through centrifugal emulsification and eMDA. *Commun. Biol.* 2:147.

Garvin, T., Aboukhalil, R., Kendall, J., Baslan, T., Atwal, G. S., Hicks, J., et al. (2015). Interactive analysis and assessment of single-cell copy-number variations. *Nat. Methods* 12, 1058–1060. doi: 10.1038/nmeth.3578

Gundry, M., Li, W., Maqbool, S. B., and Vijg, J. (2012). Direct, genome-wide assessment of DNA mutations in single cells. *Nucleic Acids Res.* 40, 2032–2040. doi: 10.1093/nar/gkr949

Ha, G., Roth, A., Lai, D., Bashashati, A., Ding, J., Goya, R., et al. (2012). Integrative analysis of genome-wide loss of heterozygosity and monoallelic expression at nucleotide resolution reveals disrupted pathways in triple-negative breast cancer. *Genome Res.* 22, 1995–2007. doi: 10.1101/gr.137570.112

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with burrows-wheeler transform. *Bioinformatics* 25, 1754–1760. doi: 10.1093/bioinformatics/btp324

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The sequence alignment/Map format and SAMtools. *Bioinformatics* 25, 2078–2079. doi: 10.1093/bioinformatics/btp352

Lodato, M. A., Woodworth, M. B., Lee, S., Evrony, G. D., Mehta, B. K., Lee, E., et al. (2015). Somatic mutation in single human neurons tracks developmental and transcriptional history. *Science* 350, 94–98. doi: 10.1126/science.aab1785

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., et al. (2010). The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 20, 1297–1303. doi: 10.1101/gr.107524.110

Navin, N., Kendall, J., Troge, J., Andrews, P., Rodgers, L., Mcindoo, J., et al. (2011). Tumour evolution inferred by single-cell sequencing. *Nature* 472, 90–94. doi: 10.1038/nature09807

Nilsen, G., Liestol, K., Van Loo, P., Moen Vollan, H. K., Eide, M. B., Rueda, O. M., et al. (2012). Copynumber: efficient algorithms for single- and multi-track copy number segmentation. *BMC Genom.* 13:591. doi: 10.1186/1471-2164-13-591

Olshen, A. B., Venkatraman, E. S., Lucito, R., and Wigler, M. (2004). Circular binary segmentation for the analysis of array-based DNA copy number data. *Biostatistics* 5, 557–572. doi: 10.1093/biostatistics/kxh008

Vijg, J., and Dong, X. (2020). Pathogenic mechanisms of somatic mutation and genome mosaicism in aging. *Cell* 182, 12–23. doi: 10.1016/j.cell.2020.06.024

Wang, R., Lin, D.-Y., and Jiang, Y. (2019). SCOPE: a normalization and copy number estimation method for single-cell DNA sequencing. *bioRxiv* [Preprint], doi: 10.1101/594267

Wang, X., Chen, H., and Zhang, N. R. (2018). DNA copy number profiling using single-cell sequencing. *Brief Bioinform.* 19, 731–736. doi: 10.1093/bib/bbx004

Zhang, L., Dong, X., Lee, M., Maslov, A. Y., Wang, T., and Vijg, J. (2019). Single-cell whole-genome sequencing reveals the functional landscape of somatic mutations in B lymphocytes across the human lifespan. *Proc. Natl. Acad. Sci. U.S.A.* 116, 9014–9019. doi: 10.1073/pnas.1902510116

Zhao, M., Wang, Q., Wang, Q., Jia, P., and Zhao, Z. (2013). Computational tools for copy number variation (CNV) detection using next-generation sequencing data: features and perspectives. *BMC Bioinform.* 14(Suppl. 11):S1. doi: 10.1186/1471-2105-14-S11-S1

Zong, C., Lu, S., Chapman, A. R., and Xie, X. S. (2012). Genome-wide detection of single-nucleotide and copy-number variations of a single human cell. *Science* 338, 1622–1626. doi: 10.1126/science.1229164

Conflict of Interest: XD, LZ, and JV are co-founders of SingulOmics Corp.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership