Mosaic *APP* Gene Recombination in Alzheimer's Disease—What's Next?

Ming-Hsiang Lee and Jerold Chun [ID]

Sanford Burnham Prebys Medical Discovery Institute, La Jolla, CA, USA.

ABSTRACT: A first example of somatic gene recombination (SGR) within the human brain was recently reported, involving the well-known Alzheimer's disease (AD)-related gene amyloid precursor protein (*APP*). SGR was characterized by the creation of *APP* genomic complementary DNA (gencDNA) sequences that were identified in prefrontal cortical neurons from both normal and sporadic Alzheimer's disease (SAD) brains. Notably, SGR in SAD appeared to become dysregulated, producing many more numbers and forms of *APP* gencDNAs, including 11 single-nucleotide variations (SNVs) that are considered pathogenic *APP* mutations when they occur in families, yet are present mosaically among SAD neurons. *APP* gene transcription, reverse transcriptase (RT) activity, and DNA strand-breaks were shown to be three key factors required for *APP* gencDNA production. Many mechanistic details remain to be determined, particularly how *APP* gencDNAs are involved in AD initiation and progression. The possibility of reducing disease-related SGR through the use of RT inhibitors that are already FDA-approved for HIV and Hepatitis B treatment represents both a testable hypothesis for AD clinical trials and a genuine therapeutic option, where none currently exists, for AD patients.

KEYWORDS: Genomic mosaicism, somatic gene recombination, genomic complementary DNA, Alzheimer's disease, reverse transcriptase, reverse transcriptase inhibitor, DNA breaks, DNA rearrangement, Alzheimer's disease therapeutics

Every cell in our body has historically been considered to have identical DNA. This same genetic information in different cell types is transcribed into RNA and then translated into proteins, producing diverse cell phenotypes and functions. However, the discovery of somatic genomic mosaicism (SGM) in the brain[1] indicates that within a given individual, not every brain cell has the same DNA sequence. These *genomic* changes can be distinguished from "genetic" ones that enter the germline and can thus be passed on to future generations; in contrast, SGM does not alter the germline. SGM encompasses all somatic changes altering DNA sequences, which are distinct from epigenetic changes that do not. The complete forms and functions of brain SGM are incompletely understood, but have been shown to impact gene expression, cell survival, cell lineage, and functional circuits within the brain, all supporting functional consequences of SGM.[2]

Outside of the brain, the best-known example of SGM has critical functions in the immune system through a fundamental process of somatic gene recombination (SGR) called "V(D)J recombination." This is responsible for generating the astronomical repertoire of immunoglobulin and T-cell receptors during the development of B and T cells of the adaptive immune system, which protects us from different kinds of pathogens. Could a similar process occur in the brain? This attractive idea received speculative discussion beginning in the 1960s, but evidence for SGR in the brain eluded scientists despite decades of active searching (reviewed in Rohrback et al[2]). This situation has recently changed with the discovery of SGR affecting the Alzheimer's disease (AD)-related gene, amyloid precursor protein (*APP*).[3]

AD is the most prevalent type of dementia. Inherited *APP* gene mutations or increased gene copy number has been shown to contribute to rare familial Alzheimer's disease (FAD)[4,5] and AD pathology in Down syndrome.[6] By contrast, the etiology of sporadic Alzheimer's disease (SAD) is not clear. Interestingly, DNA content and *APP* gene copy number, revealed by flow cytometry and single-neuron qPCR analyses, respectively, were both increased in neurons from postmortem SAD, compared with age-matched non-diseased (ND) prefrontal cortices.[7] *APP* gene in situ hybridization experiments revealed diverse morphology and intensity of signals, hinting at the possibility of non-uniform *APP* genomic amplification, which might be produced by SGR. This idea was borne out by close examination of the *APP* gene in small neuronal populations and single neurons; however, neuronal SGR was very different compared to what occurs in the immune system. In the brain, SGR was found to occur mainly in post-mitotic neurons, contrasting with V(D)J SGR which occurs in proliferating lymphocytes. Neuronal SGR produced genomic complementary DNAs (gencDNAs) that were copied from spliced RNA, resulting in thousands of *APP* gencDNAs characterized by recombined intra-exonic junctions (IEJs), single-nucleotide variations (SNVs), and insertions and deletions (Indels), all of which were enriched in SAD cortical neurons. Importantly, 11 somatic SNVs were identical to known FAD pathogenic mutations in SAD but not ND, strongly implicating a pathogenic role of *APP* gencDNAs in SAD.

We modeled *APP* gencDNA formation in culture and in J20 (*APP* transgenic) AD mice and concluded that *APP* gencDNA formation involves three factors: *APP* gene transcription, reverse transcriptase (RT) activity, and DNA strand-breaks. The proposed model for *APP* gencDNA production is this: *APP* is first transcribed—preferably at a high level—and spliced. It is then reverse transcribed into cDNA via RT activity, followed by "retro-insertion" back into the genome at the sites of DNA breakage. At some stage that is not yet known, IEJs are introduced into the gencDNAs along with SNVs likely produced by RT activity. *APP* gencDNA variants in the genome can then be re-expressed and retro-inserted again and again to generate multiple copies and myriad forms. The implications of neuronal SGR are potentially vast, and several are discussed below.

RT Activity Exists in Human Brains, Contributing to SGM and SGR

The identity of endogenous RTs in human brains is still unknown. At least three endogenous sources that might provide RT activity are present in the germline, including long interspersed nuclear element 1 (LINE1), human endogenous retrovirus (HERV), and telomerase (TERT). Since TERT is specialized for telomere elongation with its own RNA component, it is not discussed here. LINE1 is an autonomous mobile element composed of a 5′ untranslated region (UTR), open reading frame 1 (ORF1), ORF2, and 3′ UTR. ORF2 encodes a protein containing a putative RT.[8] HERV has a basic gene structure of a retrovirus, including a possibly functional "pol" RT gene. Both LINE1 and HERV are widely distributed in the human genome with many copies (3220 for HERVs[9] and over 500 000 for LINE1)[10]; most are considered inactive because of mutations; yet some are thought to be active. Mechanisms that can somatically introduce SNVs[11] into expressed genes may provide a pathway for reactivation.

The existence of SGR and roles for RTs may also be relevant to DNA content variation (DCV).[7,12,13] In our previous DCV analysis, we estimated about 250 Mbp gains of DNA in neurons from postmortem prefrontal cortices, compared with lymphocyte and cerebellar controls. Furthermore, the DNA content in SAD neurons was even higher than that of ND neurons. Although more studies are required to understand the nature of the gained DNA sequences, the reported 3- to 5-fold increase of *APP* gencDNAs (and possibly other genes) in SAD[3] may contribute to the higher DNA content. Perhaps RT activity is involved in producing this significant, yet incompletely understood, sub-genomic DNA increase. This scenario may also have relevance to controversies over adult neurogenesis[14-16] where nucleotide incorporation and neurogenesis markers might also reflect SGM/SGR.

DNA Breaks Provide Retro-insertion Sites

DNA breaks can be induced under different physiological conditions. Oxidative stress caused by reactive oxygen species (ROS) is considered to be the major cause of damage to DNA

in brains because of its high metabolic rate and its association with AD.[17] ROS introduces oxidized bases leading to DNA single-strand breaks (SSB) and also DNA double-strand breaks (DSB) at a lower frequency.[18] Specifically, DSB happens where ROS-induced lesions are close to each other or encounter active DNA replication or RNA transcription machinery.[19] DSB can also be formed at the promoter regions of early-response genes upon neural activity stimulation.[20] Understanding the types of DNA breaks involved in *APP* gencDNA formation is not only important for elucidating the mechanism itself, but also to gain insights into genomic integration sites, eg, promoters of early-response genes. Integration sites of *APP* gencDNAs likely play an important role in determining their expression, whereas retro-inserted *APP* gencDNAs may be controlled by different promoters. Using DNA in situ hybridization (DISH) analysis, we found that *APP* gencDNAs are integrated away from the two allelic *APP* loci in single neurons, but with distinct patterns in each neuron, supporting the existence of diverse integration sites. New approaches are being developed to identify these sites.

How do IEJs Form?

We identified IEJs of *APP* gencDNAs and mRNAs from SAD and ND prefrontal cortical neurons. However, the step at which IEJs form during *APP* gencDNA formation is still not clear. IEJs may be generated in RNA, DNA, or in both. The overlapping sequence homology regions, ranging from 2 to 20 nucleotides of recombined exons, implicate microhomology-mediated end-joining that occurs in DNA, providing at least one known mechanism that might be in play. More studies are needed to elucidate IEJ production mechanisms.

The Role of *APP* gencDNA in AD

Increases in the total number and diversity of sequences, including 11 SNVs identical to pathogenic mutations, were observed in *APP* gencDNAs from SAD vs ND cortical neurons. The mechanisms regarding how *APP* gencDNAs can cause SAD require further study. *APP* gencDNA in situ hybridization in a J20 mouse model showed that gencDNAs could be increased in post-mitotic neurons which accumulate with age. However, in humans, this accumulation is likely limited because it may be detrimental to neurons if the copy number passes a certain threshold. In addition, neural function may be affected both by SGR of *APP*, or possibly other genes, and through effects of retro-insertion, which itself might disrupt specific genes by acting as an insertional mutagen. Although some of the newly identified variants were shown to be toxic in culture, a series of experiments must be carried out to assess AD pathogenicity of endogenous *APP* gencDNAs and/or their gene products, to better understand *APP* gencDNA function. For example, *APP* gencDNA variants can be tested in vivo to see if they facilitate or exacerbate cognitive impairment, neural cell loss, Aβ plaque formation, gliosis, etc, in established AD mouse models.

Using Existing RT Inhibitors and Developing Novel Ones for New AD Treatments

Developing new AD medicines is a pressing need, accentuated by the current lack of effective treatments and the continuous failures of therapeutic trials. An obvious target emerging from our basic science is RT that could be pharmacologically inhibited. Critically, RT targets can potentially be accessed now, through multiple, existing Food and Drug Administration (FDA)-approved RT inhibitors that, in some cases, have demonstrated decades of post-approval efficacy and safety in both HIV and Hepatitis B patients. For example, Zidovudine is an orthosteric nucleoside analog RT inhibitor (NRTI), approved by the FDA for HIV treatment in 1987. This drug, along with many other FDA-approved NRTIs, can and should be tested through controlled clinical trials on SAD subjects as well as those in high-risk categories *before* overt AD manifests: Down syndrome,[21] familial AD,[4,5] and ApoE4 homozygotes.[22]

Remarkably, epidemiological support for this approach already exists: only one instance of verified AD in an HIV patient has been reported in the peer-reviewed literature,[23] despite a projected AD patient burden of thousands in the HIV-positive population, and unpublished assessments of medical insurer databases of aged HIV and Hepatitis B patients clearly support the use of NRTIs through a reduced number of AD patients. Careful analyses of national and international patient databases could provide additional information on the relationship between NRTI use and AD in patient sub-groups of different ethnicities. As important, AD patients today may not have access to any other clinical candidate that shows the long-term efficacy and, especially, long-term safety of the NRTIs that are available now. Considering the complete absence of disease-modifying therapies in AD, the decades of experience and acceptable human safety with NRTIs, and the legal precedent of using even *experimental* agents in untreatable diseases through the FDA's Expanded Access program[24,25] and the Right to Try Act,[26,27] there is a strong argument for off-label use of FDA-approved NRTIs in AD patient care, supported by the science of SGR.

Concluding Remarks

The presence of SGR, acting to produce *APP* gencDNAs in ND neurons, reflects its possible physiological role in normal neurons, while its dysregulation in AD offers new mechanistic insights, as well as therapeutic strategies, to interrupt the disease process. It may also be possible to access pathogenic gencDNAs themselves by targeted anti-sense or genome editing to repair the brain's cellular blueprint. Extrapolation to non-neuronal cells and non-brain cells that are long-lived may further expand the implications of this new phenomenon. The central importance of RTs in SGR suggests that other brain diseases may be understood and treated through SGR. Combined with the existence of FDA-approved RT inhibitors showing positive effects in retrospective analyses, effective first-generation therapies for AD (and possibly other disorders) may now be at hand for AD patients with no other options, and controlled clinical trials assessing efficacy and safety can establish desirable agents, optimal dosing, and appropriate patient selection, to identify and treat patients who might benefit from RT inhibitor therapy. Agents targeting specific brain RTs and/or recombined genes could emerge in the future.

Author Contributions
M-HL and JC prepared the manuscript.

ORCID iD
Jerold Chun 🔗 https://orcid.org/0000-0003-3964-0921

REFERENCES

1. Rehen SK, McConnell MJ, Kaushal D, Kingsbury MA, Yang AH, Chun J. Chromosomal variation in neurons of the developing and adult mammalian nervous system. *Proc Natl Acad Sci U S A*. 2001;98:13361–13366.
2. Rohrback S, Siddoway B, Liu CS, Chun J. Genomic mosaicism in the developing and adult brain. *Dev Neurobiol*. 2018;78:1026–1048.
3. Lee MH, Siddoway B, Kaeser GE, et al. Somatic APP gene recombination in Alzheimer's disease and normal neurons. *Nature*. 2018;563:639–645.
4. Hooli BV, Kovacs-Vajna ZM, Mullin K, et al. Rare autosomal copy number variations in early-onset familial Alzheimer's disease. *Mol Psychiatry*. 2014;19:676–681.
5. Hooli BV, Mohapatra G, Mattheisen M, et al. Role of common and rare APP DNA sequence variants in Alzheimer disease. *Neurology*. 2012;78:1250–1257.
6. Rovelet-Lecrux A, Hannequin D, Raux G, et al. APP locus duplication causes autosomal dominant early-onset Alzheimer disease with cerebral amyloid angiopathy. *Nat Genet*. 2006;38:24–26.
7. Bushman DM, Kaeser GE, Siddoway B, et al. Genomic mosaicism with increased amyloid precursor protein (APP) gene copy number in single neurons from sporadic Alzheimer's disease brains. *Elife*. 2015;4:05116.
8. Sciamanna I, De Luca C, Spadafora C. The reverse transcriptase encoded by LINE-1 retrotransposons in the genesis, progression, and therapy of cancer. *Front Chem*. 2016;4:6.
9. Tokuyama M, Kong Y, Song E, Jayewickreme T, Kang I, Iwasaki A. ERVmap analysis reveals genome-wide transcription of human endogenous retroviruses. *Proc Natl Acad Sci U S A*. 2018;115:12565–12572.
10. Rodic N, Burns K. Long interspersed element-1 (LINE-1): passenger or driver in human neoplasms. *PLoS Genet*. 2013;9:e1003402.
11. Lodato MA, Rodin RE, Bohrson CL, et al. Aging and neurodegeneration are associated with increased mutations in single human neurons. *Science*. 2018;359:555–559.
12. Westra JW, Barral S, Chun J. A reevaluation of tetraploidy in the Alzheimer's disease brain. *Neurodegener Dis*. 2009;6:221–229.
13. Westra JW, Rivera RR, Bushman DM, et al. Neuronal DNA content variation (DCV) with regional and individual differences in the human brain. *J Comp Neurol*. 2010;518:3981–4000.
14. Sorrells SF, Paredes MF, Cebrian-Silla A, et al. Human hippocampal neurogenesis drops sharply in children to undetectable levels in adults. *Nature*. 2018;555:377–381.
15. Kempermann G, Gage FH, Aigner L, et al. Human adult neurogenesis: evidence and remaining questions. *Cell Stem Cell*. 2018;23:25–30.
16. Ming GL, Song H. Adult neurogenesis in the mammalian brain: significant answers and significant questions. *Neuron*. 2011;70:687–702.
17. Teixeira J, Silva T, Andrade PB, Borges F. Alzheimer's disease and antioxidant therapy: how long how far? *Curr Med Chem*. 2013;20:2939–2952.
18. Lindahl T. Instability and decay of the primary structure of DNA. *Nature*. 1993;362:709–715.

19. Woodbine L, Brunton H, Goodarzi AA, Shibata A, Jeggo PA. Endogenously induced DNA double strand breaks arise in heterochromatic DNA regions and require ataxia telangiectasia mutated and Artemis for their repair. *Nucleic Acids Res*. 2011;39:6986–6997.

20. Madabhushi R, Gao F, Pfenning AR, et al. Activity-induced DNA breaks govern the expression of neuronal early-response genes. *Cell*. 2015;161: 1592–1605.

21. Wiseman FK, Al-Janabi T, Hardy J, et al. A genetic cause of Alzheimer disease: mechanistic insights from Down syndrome. *Nat Rev Neurosci*. 2015;16: 564–574.

22. Huang YA, Zhou B, Wernig M, Sudhof TC. ApoE2, ApoE3, and ApoE4 differentially stimulate APP transcription and abeta secretion. *Cell*. 2017;168:427. e421–441.e421.

23. Turner RS, Chadwick M, Horton WA, Simon GL, Jiang X, Esposito G. An individual with human immunodeficiency virus, dementia, and central nervous system amyloid deposition. *Alzheimers Dement (Amst)*. 2016;4:1–5.

24. Puthumana J, Miller JE, Kim J, Ross J. Availability of investigational medicines through the US Food and Drug Administration's expanded access and compassionate use programs. *JAMA Netw Open*. 2018;1:e180283.

25. Fountzilas E, Said R, Tsimberidou AM. Expanded access to investigational drugs: balancing patient safety with potential therapeutic benefits. *Expert Opin Investig Drugs*. 2018;27:155–162.

26. Traynor K. Federal right-to-try law aims to broaden access to investigational drugs. *Am J Health Syst Pharm*. 2018;75:1085–1086.

27. Brown B, Ortiz C, Dube K. Assessment of the Right-to-Try Law: the pros and the cons. *J Nucl Med*. 2018;59:1492–1493.