

Long-read sequencing reveals common structural variants as potential drivers of genetic risk for neurodegenerative diseases

Alex N Salazar¹ | Niccoló Tesi¹ | Daniel Alvarez Sirvent¹ | Lydian Knoop¹ |
Yaran Zhang¹ | Sven J van der Lee¹ | Sanduni Wijesekera¹ | Jana Krizova¹ |
Wiesje M. van der Flier² | Yolande A.L. Pijnenburg³ | Marcel JT Reinders⁴ |
Marc Hulsman¹ | Henne Holstege¹

¹Genomics of Neurodegenerative Diseases and Aging, Human Genetics, Vrije Universiteit Amsterdam, Amsterdam UMC, Amsterdam, Netherlands

²Alzheimer Center Amsterdam, Neurology, Vrije Universiteit Amsterdam, Amsterdam UMC location VUmc, Amsterdam, Netherlands

³Alzheimer Center Amsterdam, Neurology, Vrije Universiteit Amsterdam, Amsterdam UMC, Amsterdam, Netherlands

⁴Delft Bioinformatics Lab, Delft University of Technology, Delft, Delft, Netherlands

Correspondence

Niccoló Tesi, Genomics of Neurodegenerative Diseases and Aging, Human Genetics, Vrije Universiteit Amsterdam, Amsterdam UMC, Amsterdam, Netherlands.

Email: n.tesi@amsterdamumc.nl

Abstract

Background: Genome-Wide Association Studies (GWAS) have identified 86 SNPs associated with Alzheimer's disease (AD). GWAS-SNPs are markers of genetic variation in linkage disequilibrium (LD), which may drive the association with AD. One major class of genetic variation are Structural Variants (SVs), which can regulate transcription and translation of nearby genes. Here, we explored the interplay between large SVs (>50 bp) and AD-associated SNPs.

Method: We performed long-read whole-genome sequencing of 214 individuals representing the extreme ends of the cognitive spectrum: N = 93 AD patients (age 67.2±8.5), and N = 121 cognitively healthy centenarians (age 101.2±1.8). We identified SVs using *sniffles2* genome-wide, further characterised them through (local) *de novo* assembly, and annotated tandem repeats (TR) and transposable elements (TE). Next, we estimated LD between SVs and 86 AD-associated GWAS-SNPs. Finally, we compared SNP frequencies and SV sizes between AD cases and centenarians using logistic regression models.

Result: Across all 214 individuals, we found >27,404 SVs (>50 bp), with a Minor Allele Frequency (MAF) ≥ 5%. Most SVs were TRs (63%), followed by TEs (37%). We found that 37 AD GWAS-SNPs paired with 90 SVs in low to strong LD ($R^2 > 0.1$), 56 of which involved SNP-TR pairs (62%) and 34 involved SNP-TE pairs (38%). In six AD loci, the SVs associated more strongly with AD risk than the leading GWAS-SNP ($p < 0.05$). This includes *ADAM10*, *CTSH*, *SLC2A4RG*, *CD2AP*, *SPI1*, and *IDUA*, which were encompassed by complex haplotypes harbouring multiple TRs and TEs.

Conclusion: Our findings provide the first support for SVs as candidate-drivers of the association between respective GWAS loci with AD risk. Hence, SVs may lead to increased effect sizes.

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2024 The Alzheimer's Association. *Alzheimer's & Dementia* published by Wiley Periodicals LLC on behalf of Alzheimer's Association.