

GENETICS

Long-Read Sequencing Reveals Ancestral intragenic APOE Haplotypes with Distinct Roles in Alzheimer's Disease

Agustin Ruiz^{1,2,3,4} | Pablo García-González^{1,3} | Raquel Puerta¹ | Amanda Cano³ | Clàudia Olivé¹ | Marta Marquíé^{1,3} | Sergi Valero³ | Maitee Rosende-Roca¹ | Pilar Sanz⁵ | Montserrat Alegret³ | Frederic Brosseron⁶ | Pamela Martino-Adami⁷ | Itziar de Rojas^{3,5} | Michael T. Heneka⁸ | Alfredo Ramirez⁷ | Arcadi Navarro⁹ | María Eugenia Sáez¹⁰ | Lluís Tárraga^{1,3} | José E. Cavazos^{11,12} | Mercè Boada^{1,3} | Victoria Fernández⁵ | Alfredo Cabrera Socorro^{13,14}

¹Ace Alzheimer Center Barcelona – International University of Catalunya (UIC), Barcelona, Spain

²Glenn Biggs Institute for Alzheimer's & Neurodegenerative Diseases, University of Texas Health Science Center, San Antonio, TX, USA

³CIBERNED, Network Center for Biomedical Research in Neurodegenerative Diseases, National Institute of Health Carlos III, Madrid, Spain

⁴Department of Microbiology, Immunology and Molecular Genetics, University of Texas Health Science Center, San Antonio, TX, USA

⁵Ace Alzheimer Center Barcelona, Barcelona, Spain

⁶German Center for Neurodegenerative Diseases (DZNE), Venusberg-Campus 1, 53127, Bonn, Germany

⁷Division of Neurogenetics and Molecular Psychiatry, Department of Psychiatry and Psychotherapy, Faculty of Medicine and University Hospital Cologne, University of Cologne, Cologne, Germany

⁸Luxembourg Centre for Systems Biomedicine (LCSB), University of Luxembourg, Luxembourg, Luxembourg

⁹BarcelonaBeta Brain Research Center (BBRC), Barcelona, Spain

¹⁰CAEBI. Centro Andaluz de Estudios Bioinformáticos, Sevilla, Spain

¹¹Glenn Biggs Institute, San Antonio, TX, USA

Abstract

Background: The apolipoprotein E (APOE) $\epsilon 4$ allele remains the strongest genetic risk factor for late-onset Alzheimer's disease (AD), yet the marked variability in its pathogenicity suggests underlying genetic complexity. Historically, efforts to resolve the intragenic architecture of APOE have been hampered by the limitations of conventional genotyping and short-read sequencing, as well as the presence of homoplasy in common intragenic markers—misleading similarities arising from convergent variants.

Objective: We leveraged Oxford Nanopore Technology (ONT) to phase intragenic APOE variants, resolve homoplasy, and examine the impact of phased haplotypes on cerebrospinal fluid (CSF) APOE protein levels and AD progression.

Methods: Using long-read sequencing in a Spanish memory clinic cohort ($n = 1,267$), we reconstructed full-length 4 kb APOE haplotypes, identifying 59 unique configurations grouped into five major haplogroups. Common intragenic variants defined ancestral $\epsilon 4$ (4A, 4B) and $\epsilon 3$ (3A, 3B) haplogroups. These were analyzed for associations with CSF APOE levels (Olink platform) and progression from mild cognitive impairment (MCI) to dementia using adjusted Cox regression models.

Results: ONT sequencing successfully resolved homoplasy between the APOE promoter region—particularly at rs405509—and the canonical protein isoforms, uncovering common but functionally distinct $\epsilon 3A/B$ and $\epsilon 4A/B$ intragenic sub-haplotypes with independent biological effects. Carriers of the $\epsilon 4A$ haplotype exhibited significantly lower CSF APOE protein levels ($p = 0.004$), whereas the $\epsilon 3B$

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.

© 2025 The Alzheimer's Association. *Alzheimer's & Dementia* published by Wiley Periodicals LLC on behalf of Alzheimer's Association.

¹²Neurology and Physiology UT Health San Antonio, San Antonio, TX, USA

¹³Janssen Research & Development, A Division of Janssen Pharmaceutica, Beerse, Belgium, Beerse, Belgium

¹⁴Johnson & Johnson Innovative Medicine, Beerse, Belgium

Correspondence

Agustin Ruiz, Ace Alzheimer Center Barcelona – International University of Catalunya (UIC), Barcelona, Spain.
Email: ruiza5@uthscsa.edu

haplotype was associated with elevated CSF APOE protein levels ($p = 0.025$). Notably, both haplotypes were linked to a slower progression from MCI to AD, independent of APOE genotype, age, sex and core CSF biomarkers.

Conclusion: This study redefines the human APOE $\epsilon 3$ and $\epsilon 4$ alleles as genetically heterogeneous entities. Using ONT long-read sequencing, we achieved high-resolution mapping of intragenic haplotypic structure and regulatory variation previously obscured by conventional approaches. This enabled the identification of ancestral haplotypes with distinct functional profiles and potential relevance to Alzheimer's disease pathogenesis. These findings highlight the importance of incorporating haplotype-level resolution into Alzheimer's risk assessment, therapeutic targeting, and precision medicine strategies.