

GENETICS

Genomic Frontiers in Alzheimer's Research: A Primer on the Alzheimer's Disease Sequencing Project and its AI/ML Opportunities

Emily Greenfest-Allen¹ | Yuk Yee Leung¹ | Wan-Ping Lee¹ | Amanda Kuzma¹ |
Otto Valladares¹ | William S Bush² | Eden R. Martin³ | Adam C. Naj^{4,5} |
Gerald D. Schellenberg⁶ | Li-San Wang¹

¹Penn Neurodegeneration Genomics Center, Department of Pathology and Laboratory Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA

²Department of Population and Quantitative Health Sciences, Institute for Computational Biology, Case Western Reserve University, Cleveland, OH, USA

³Dr. John T. Macdonald Foundation Department of Human Genetics, University of Miami Miller School of Medicine, Miami, FL, USA

⁴University of Pennsylvania, Perelman School of Medicine, Philadelphia, PA, USA

⁵Penn Neurodegeneration Genomics Center, University of Pennsylvania, Perelman School of Medicine, Philadelphia, PA, USA

⁶Penn Neurodegeneration Genomics Center, Department of Pathology and Laboratory Medicine, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA, USA

Correspondence

Emily Greenfest-Allen, Penn Neurodegeneration Genomics Center, Department of Pathology and Laboratory Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA.
Email: allenem@pennmedicine.upenn.edu

Abstract

Background: In this introductory talk, we embark on a journey through the genomic frontiers of Alzheimer's disease (AD) research, illuminating the wide range of opportunities the pioneering Alzheimer's Disease Sequencing Project (ADSP) offers to AI/ML researchers. From predictive modeling to identifying novel biomarkers and therapeutic targets, we will explore the vast landscape of possibilities for ADSP data-driven discoveries that can help shape the future of Alzheimer's research and precision medicine.

Method: ADSP integrates numerous projects that collectively unravel the intricate genetic landscape of AD with the primary objective of advancing precision medicine for the millions affected globally by this devastating disease. Working toward the goal of sequencing and analyzing up to 150,000 complete genomes and associated clinical and functional data in the next five years, ADSP has amassed an unprecedented wealth of genomic data from diverse populations, providing a comprehensive and holistic understanding of the genetic underpinnings of AD.

Result: This presentation serves as a primer, exploring various components of the ADSP and discussing the unprecedented resources it presents AI/ML researchers: (1) Diversity Initiative: The ADSP places a paramount emphasis on diversity, ensuring the inclusion of a wide range of populations in its genomic dataset. The current (release 5) dataset includes whole genomes from 57,302 unique participants, including 6,875 with African ancestry, 5,523 with Asian, and 15,390 Hispanic/Latino individuals. (2) Phenotype Harmonization: Harmonizing phenotypic data across diverse cohorts is a critical aspect of the ADSP, facilitating meaningful comparisons and analyses. (3) Open Access Initiative: unrestricted access to harmonized summary statistics, curated

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.

genetic associations, and functional annotations enables generation of task- and domain-specific knowledge bases. (4) Functional Genomics: Moving beyond genetic variations, the ADSP incorporates functional genomics to discern the biological mechanisms at play.

Conclusion: The potential of AI/ML within the ADSP serves to emphasize the need for collaborative initiatives between AI/ML researchers and the broader Alzheimer's research community. The synergy between these fields holds the key to unlocking breakthroughs that can translate genomic insights into tangible clinical advancements.