

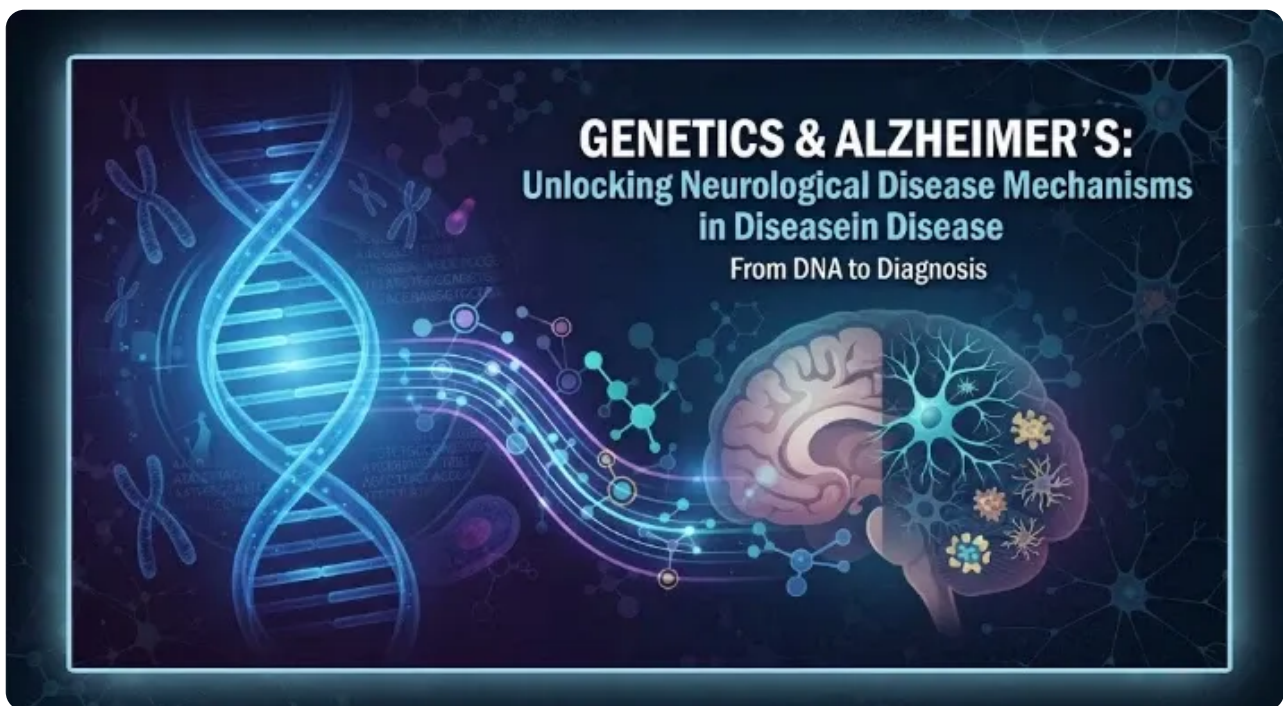
Human Genetics Mechanisms through Alzheimer's Disease in Neurological Disease

Sadia Tasmeeem

May 3, 2026

Abstract

The article reviews Alzheimer's disease genetics as a network of pathways rather than a collection of isolated risk loci. APOE is discussed as the strongest common genetic factor, with effects on lipid transport, amyloid biology, glial responses, vascular function, and ancestry-dependent risk. Additional loci such as CR1, CLU, PICALM, and ABCA7 connect disease susceptibility to complement, endocytosis, lipid handling, and innate immunity. The article stresses the importance of multi-ancestry cohorts, founder populations, rare variants, and studies of cognitive resilience. Genetic associations should be linked to specific cell types and disease stages through single-nucleus, chromatin, proteomic, imaging, and longitudinal biomarker data. CRISPR-edited neurons, astrocytes, and microglia can then test causality. Polygenic risk scores require careful calibration and portability assessment. The central argument is that inherited variation becomes most informative when followed from statistical association to cell type, molecular pathway, disease trajectory, and clinical phenotype.



Genetics of Alzheimer's disease has shifted from a single dominant risk locus to a network of pathways that include lipid transport, immune regulation, endocytosis, amyloid processing, and synaptic biology and vascular function. It was that expansion that transformed the concept of genetic risk. A variant is not normally used as a deterministic switch. Most common variants modify the probability by biological mechanisms that act over decades and which are much more complex and diverse than the calculation of probability.

APOE is still the most robust common genetic risk factor for late onset Alzheimer's disease. The epsilon four allele is a risk factor and is associated with earlier onset of disease and the epsilon two allele may be a protective variant in many populations. APOE has roles in lipid transport, amyloid aggregation and clearance, glial responses and vascular biology. It also has differential effects on ancestry groups; genetic risk should not be understood without considering population background.

Previous APOE studies in African American populations had revealed differences in magnitude for the associations compared to many European ancestry samples. This rose some doubts about the transferability of one risk estimate from one population to another. The observed genetic effects may be affected by local ancestry, environmental exposures, interacting variants, and social determinants. So, there is no reason to believe that the diverse cohort design is just an equity question, it is also part of mechanism discovery.

Beyond APOE, there was a wider landscape to expand beyond, with genome wide

association studies. Loci identified with CR1, CLU, and PICALM linked Alzheimer disease to complement biology, lipid handling, and endocytosis. Subsequent work that identified other regions led to the formulation of a polygenetic architecture of many small effects. These loci are a clue to a pathway; the variant is usually close to multiple genes and needs to be functionally characterized before it is securely attributed to a pathway.

The example of population specific discovery can be found in ABCA7. ABCA7 variants had high association with Alzheimer's in African Americans. ABCA7 is involved in lipid transport and phagocytosis, and is associated with a genetic risk, while already connected to pathways through APOE and innate immunity. The outcome illustrated the power of wider ancestry representation to uncover loci that can have significant relevance, but that is not apparent in other data sets.

Power was enhanced through large scale genetic studies by incorporation of clinically diagnosed cases, proxy phenotypes and population biobanks. These designs found other susceptibility regions and reinforced pathway enrichment of immune, lipid, and endosomal pathways. Discovery is enhanced with larger sample size but mechanisms can be blurred by phenotypic heterogeneity. A proxy case provided by family history is not the same as biomarker confirmed disease and is therefore not considered as a full replacement.

Rare variants further complicate matters. The effects of common variant studies account for only part of inherited susceptibility and rare coding changes may have a more significant biological effect in selected families or populations. Some of the founder communities are particularly useful as they bring a higher frequency of rare variants and make aspects of mapping easier. To study genetic architecture that may be challenging in more heterogeneous cohorts, genetic founder population studies, such as the Amish, have been undertaken.

Resilience is also informed by genetics. There are people who are not cognitively impaired even as they get older or have a significant neuropathological load. Research on cognitive preservation can identify variants that are protective of cognitive maintenance rather than those that are risk factors for disease. Protective biology may be of therapeutic relevance as a guide to pathways that allow function in the face of pathological stress. Research on susceptibility and resilience is useful in prevention.

To interpret a mechanism, variants need to be associated with cells and tissues. Single nucleus transcriptomics, chromatin accessibility, expression quantitative trait loci and proteomic information can help determine which cells harbor a genetic effect. A microglial expressing risk locus indicates a distinct causal pathway to a risk locus in

neuronal or vascular cells. Functional perturbation can be preceded by candidate gene narrowing due to colocalization and fine mapping.

Then causality is tested with functional models. Engineered risk and protective alleles can be introduced into iPSC derived neurons, astrocytes and microglia. The use of CRISPR editing enables the comparison of matched genetic backgrounds. Organoid and coculture systems can be used to study the relationship between cell types. There are a number of animal models for testing effects on amyloid, tau, inflammation, synaptic function, or cognition, with care taken when extrapolating results between species.

Another link between genotype and disease trajectory are the longitudinal biomarkers. Genetic risk can be assessed using any age at amyloid accumulation, tau spread, neurodegeneration, cognitive decline, or vascular change. These analyses may identify a variant that is involved in early stages of pathogenesis or a variant that impacts later stages of pathogenesis. The impact of a gene on the diagnosis of a disease may be different from the impact of a gene when it is involved in the course of a disease.

While the polygenetic risk scores represent many variants, one of the main drawbacks is the lack of portability. The accuracy of scores generated from European ancestry data is diminished in other population groups due to differences in allele frequencies, linkage disequilibrium, and effect size. Multiancestry or ancestry specific models can enhance performance. Calibration is required for clinical use, incremental value over age and family history, and communication is important as genetic probability is not a diagnosis.

The best approach is to relate human genetics to biological networks. Loci of the newer type (APOE, ABCA7, CR1, CLU, PICALM) do not function as individual switches. They are all related to lipid metabolism, immune function, endosomal trafficking, and cellular stress. Pathway nodes identified by network based interpretation can be those affected by multiple independent genetic signals, making it more robust to the conclusion that this is a mechanism that is important in populations.

However, due to the fact that the associated locus may harbor numerous correlated variants, fine mapping is required. Using statistical methods, the credible set can be narrowed and using functional annotations variants in enhancers, promoters or coding regions can be prioritised. The use of an automatic nearest gene approach for assigning loci should be avoided. The regulatory interactions between two genes that are not necessarily nearby can be displayed in chromatin contact maps, and if there is cell-specific expression information, it can be used to identify which candidate gene is affected by the risk variant.

Another restriction that must be exercised is the gene by gene interaction. It is obvious that there is interaction between biological pathways, and the statistical identification of epistasis will be very difficult unless the sample size is very large, since there are a lot of different possible combinations. Mechanistically informed tests can narrow that field to a set of genes in a common pathway, e.g. lipid handling or complement signaling. Replication is still required prior to an interaction becoming a causal model.

Clinical diagnosis is not always sufficient to provide a precise phenotype, and this can also be achieved by a neuropathological diagnosis. Amyloid, tau, vascular injury, Lewy body pathology and TDP forty three pathology can be analyzed separately from genetic variants. This breakdown is what could account for differences in the strength of the associations between a locus and dementia in different cohorts of individuals with varying mixes of pathology. Genetics may become more acute if the endpoint is a biological process rather than a general syndrome.

Evidence that is convergent should be considered more important for therapeutic interpretation. A gene that is supported by association, cell specific expression, protein change, effects of the longitudinal biomarkers, and functional perturbation has more weight than a gene supported by only statistical proximity. Convergence also aids in determination of direction. A target may need to be inhibited in one pathway, but activated in another, and genetics may sometimes suggest the direction in which a target is likely to behave as a protective target.

Replication between phenotype definitions can also help to elucidate mechanism. A locus supported by a clinically diagnosed disease, a biomarker confirmed disease, and a neuropathology has more evidence than a locus based on one diagnostic scheme. The genetic signal is consistent across definitions, indicating that it represents a fundamental part of Alzheimer biology, not ascertainment practice.

Genetics of Alzheimer's disease is thus a double source of evidence: experimental natural experiments generated by the random occurrence of inherited variation, and maps of biological systems repeatedly implicated in various cohorts. Best mechanistic studies involve multiple populations, specific phenotyping, time-dependent biomarkers, and functional validation. The utility of human genetics is maximized when a statistical association is followed from cell type to molecular pathway to disease stage and clinical phenotype. That's a sequence that makes inherited risk a testable explanation of neurological disease.