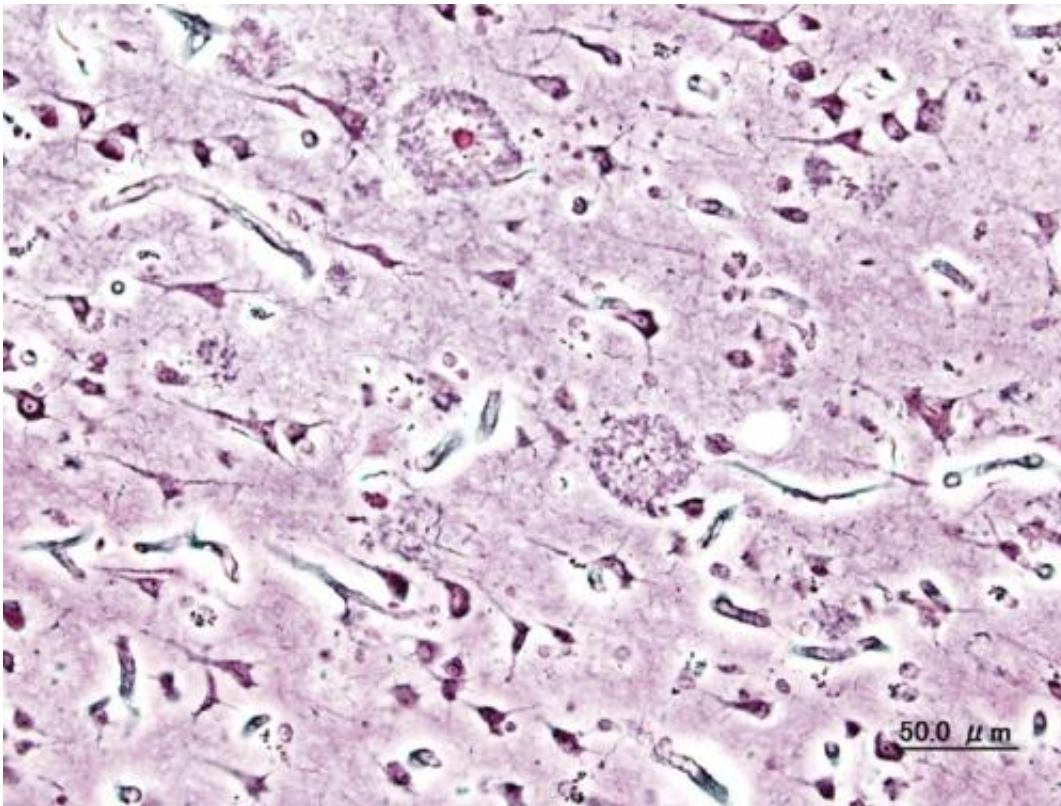


# Researchers use Blue Waters supercomputer to understand gene expression in the brain

February 4 2015, by Aaron Dubrow

---



Microscopic image of senile plaques seen in the cerebral cortex of a person with Alzheimer's disease of presenile onset. Credit: KGH, via Wikimedia Commons

The release of the film, "Still Alice," in September 2014 shone a much-needed light on Alzheimer's disease, a debilitating neurological disease that affects a growing number of Americans each year.

More than 5.2 million people in the U.S. are currently living with Alzheimer's. One out of nine Americans over 65 has Alzheimer's, and one out of three over 85 has the disease. For those over 65, it is the fifth leading cause of death.

There are several drugs on the market that can provide relief from Alzheimer's symptoms, but none stop the development of disease, in part because the root causes of Alzheimer's are still unclear.

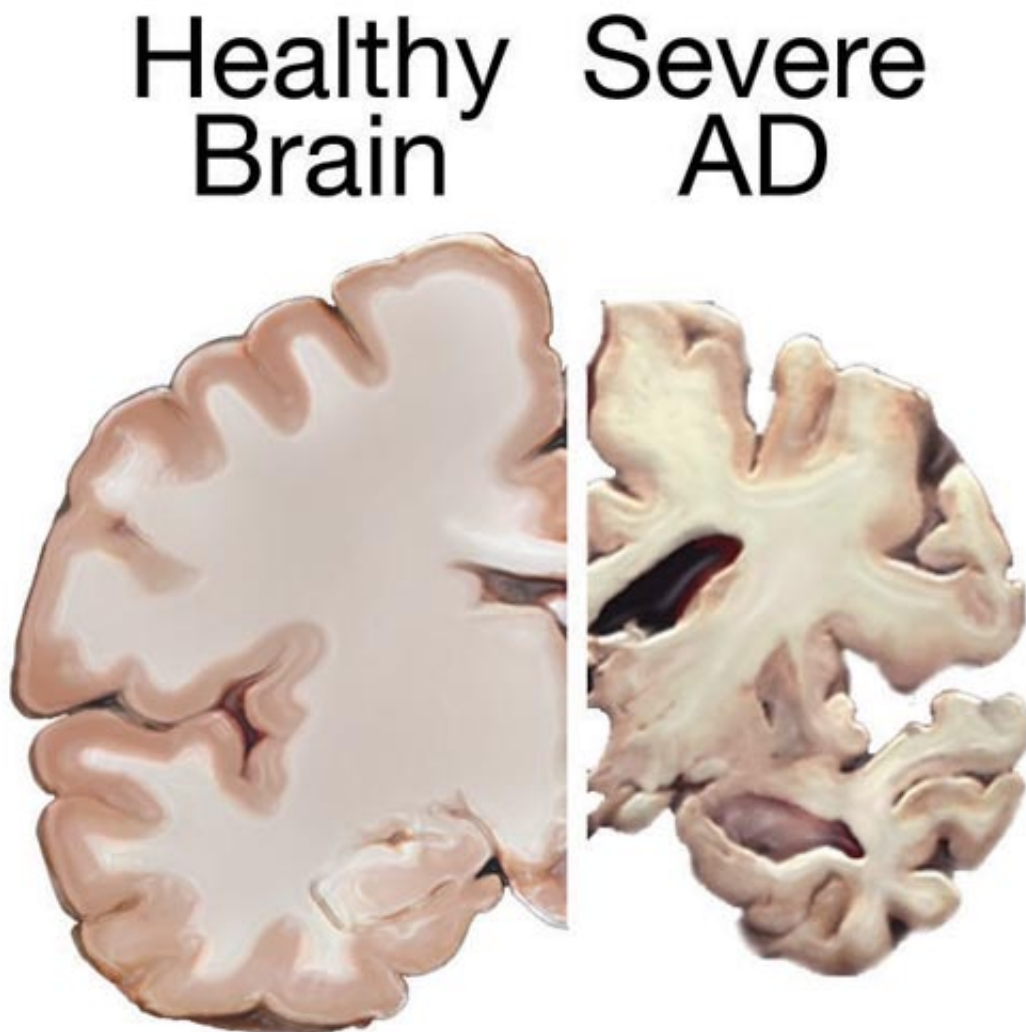
"We're interested in studying the genetics of Alzheimer's disease," said Mariet Allen, a post-doctoral fellow at the Mayo Clinic in Florida. "Can we identify genetic risk factors and improve our understanding of the biological pathways and cellular mechanisms that can play a role in the disease process?"

Allen is part of a team of researchers from the Mayo Clinic who are using Blue Waters, one of the most powerful supercomputers in the world, to decode the complicated language of genetic pathways in the brain. In doing so, they hope to provide insights into what genes and proteins are malfunctioning in the brain, causing amyloid beta plaques, tau protein tangles and brain atrophy due to neuronal cell loss—the telltale signs of the disease—and how these genes can be detected and addressed.

In the case of late onset Alzheimer's disease (LOAD), it is estimated that as much as 80 percent of risk is due to genetic factors. In recent years, researchers discovered 20 common genetic loci, in addition to the well-known APOE gene, that are found to increase or decrease risk for the disease. (Loci are specific locations of a gene, DNA sequence, or position on a chromosome.) These loci do not necessarily have a causal connection to the disease, but they provide useful information about high-risk patients.

Despite all that doctors have learned in recent years about the genetic basis of Alzheimer's, according to Allen, a substantial knowledge gap still exists. It has been estimated that likely less than 40 percent of genetic risk for LOAD can be explained by known loci. Furthermore, it is not always clear which are the affected genes at these known loci.

In other words, scientists have a long way to go to get a full picture of which genes are involved in processes related to the disease and how they interact.



A healthy brain compared to a brain suffering from Alzheimer's disease. Credit:

National Institutes of Health

The Mayo team and their colleagues had been very successful in the past in finding [genetic risk factors](#) using a method that matched individual differences in the DNA code—single-nucleotide polymorphisms or SNPs, to phenotypes—the outward appearances of the disease. In particular, the Mayo team focused on identifying SNPs that influence expression of genes in the brain. However, they now hypothesize that the single SNP method may be too simplistic to find all [genetic factors](#), and is likely not an accurate reflection of the complex biological interactions that take place in an organism.

For that reason, the Mayo researchers have recently turned their attention to investigating the brain using genetic interaction (epistasis) studies. Such studies allow researchers to understand the effects of pairs of gene changes on a given phenotype and can uncover additional genetic variants that influence gene expression and disease.

The process involves the analysis of billions of DNA base pairs (the familiar C, G, A and T) to find statistically significant correlations. Importantly, the search is not to discover simple one-to-one connections, since these have largely been found, but to study the interaction effects of pairs of DNA sequence variations.

Solving a problem of this size and complexity requires a huge amount of computational processing time, so the researchers turned to the Blue Waters supercomputer at the National Center for Supercomputing Applications (NCSA).

Supported by the National Science Foundation and the University of Illinois at Urbana-Champaign, Blue Waters allows scientists and

engineers across the country to tackle a wide range of challenging problems using massive computing and data processing power. From predicting the behavior of complex biological systems to simulating the evolution of the cosmos, Blue Waters assists researchers whose computing problems are at a scale or complexity that cannot be reasonably approached using any other method.

Allen and her colleagues used Blue Waters to rapidly advance their Alzheimer's epistasis study through NCSA's Private Sector Program, which lets teams outside of academia access the system.

Instead of requiring as much as a year or more of processing on a single workstation or university cluster, the research team was able to do each analysis on Blue Waters in less than two days.



Researchers from the Mayo Clinic used Blue Waters, one of the most powerful supercomputers in the world, to understand the genetic basis of Alzheimer's disease. Credit: NCSA

The researchers conducted three sets of analysis to investigate brain gene expression levels in a group of individuals without Alzheimer's, a group of individuals with Alzheimer's and then a combined analysis of both groups together. To date, these analyses have been completed for the almost 14,000 genes expressed in the majority of the brain samples studied.

Through their work with collaborators at NCSA and the University of Illinois at Urbana-Champaign (including Victor Jongeneel and Liudmila Mainzer), the Mayo team overcame many of the challenges that a project of this scope presented.

"The analysis of epistatic effects in large studies, such as ours, requires powerful computational resources and would not be possible without the unique computing capabilities of Blue Waters," wrote project lead Nilufer Ertekin-Taner from the Mayo Clinic.

"The Mayo Clinic project is emblematic of the type of problem that is beginning to emerge in computational medicine," said Irene Qualters, division director of Advanced Computing Infrastructure at NSF.

"Through engagement with the Blue Waters project, researchers at Mayo have demonstrated the potential of new analytic approaches in addressing the challenges of a daunting medical frontier."

The team reported on their progress at the Blue Waters Symposium in May 2014. Allen and her colleagues are currently processing and filtering the results so they can be analyzed.

"Recent studies by our collaborators and others have shown that both the risk for late onset Alzheimer's disease and gene expression are likely influenced by epistasis. However little is known about the effect of

genetic interactions on brain gene expression specifically and how this might influence risk for neurological diseases such as LOAD," said Allen. "The goal of our study is to address this knowledge gap; something we have been uniquely positioned to do using our existing data and the resources available on Blue Waters."

**More information:** Ridge PG, Mukherjee S, Crane PK, Kauwe JSK, Alzheimer's Disease Genetics Consortium (2013) Alzheimer's Disease: Analyzing the Missing Heritability. *PLoS ONE* 8(11): e79771. [DOI: 10.1371/journal.pone.0079771](https://doi.org/10.1371/journal.pone.0079771)

Alzheimer's Disease Facts and Figures:  
[www.alz.org/downloads/facts\\_figures\\_2013.pdf](http://www.alz.org/downloads/facts_figures_2013.pdf)

Provided by National Science Foundation

Citation: Researchers use Blue Waters supercomputer to understand gene expression in the brain (2015, February 4) retrieved 21 May 2024 from <https://medicalxpress.com/news/2015-02-blue-supercomputer-gene-brain.html>

<p>This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.</p>
--