NIH Opens First Batch of AD Genomic Data

The National Institutes of Health (NIH) announced on December 2 that researchers can now access the first batch of genome sequence data from the Alzheimer's Disease (AD) Sequencing Project (ADSP), one of the first initiatives undertaken in an intensified national program of research to prevent or effectively treat AD. The first data release includes whole genome sequence (WGS) data from 410 individuals in 89 families. Researchers deposited completed WGS data on 61 families and have deposited WGS data on parts of the remaining 28 families. The remaining data will be deposited in the near future.

"Providing raw DNA sequence data to a wide range of researchers proves a powerful crowd-sourced way to find genomic changes that put us at increased risk for this devastating disease," said NIH Director Francis S. Collins, MD, PhD, who announced the start of the project in February 2012. "The ADSP is designed to identify genetic risks for late-onset of AD, but it could also discover versions of genes that protect us. These insights could lead to a new era in prevention and treatment. "

The National Alzheimer's Project Act became law in 2011 in recognition of the need to do more to combat the disease. The law called for upgrading research efforts by the public and private sectors, as well as expanding access to and improving clinical and long-term care. One of the first actions taken by NIH under the Alzheimer's Act was the allocation of additional funding for a series of studies, including this genome sequencing effort. To carry out the ADSP, 2 NIH institutes, the National Human Genome Research Institute (NHGRI) and the National Institute on Aging (NIA), formed a collaboration to manage patient samples and genome sequencing. NHGRI has devoted $25 million in sequencing capacity at its 3 flagship centers: the Genome Institute at the Washington University School of Medicine (St. Louis, MO); the Human Genome Sequencing Center at Baylor College of Medicine (Houston, TX); and the Eli and Edythe L. Broad Institute of Harvard and the Massachusetts Institute of Technology (Cambridge).

The sequencing centers will produce WGS data from 582 subjects from 111 families in the Family Based Study, a collection of several different groups of patient cohorts, including: the NIH Late Onset of Alzheimer's Disease study, the National Cell Repository for Alzheimer's Disease, Alzheimer's Disease Genetics Consortium, and the Cohorts for Heart and Aging Research in Genomic Epidemiology, in addition to patient collections provided by several universities. The 3 sequencing centers also have begun whole exome sequencing on an additional 11,000 individuals, including 6,000 AD-affected individuals and 5,000 controls. Whole exome sequencing determines the order of genomic letters for the ~21,000 genes in the human genome, a considerably smaller dataset than the 3 billion letters examined in WGS but tightly focused on protein-producing genes.

The current deposit of genomic data includes phenotypes, such as the AD symptoms of affected individuals, as well as family "trees" and other demographic information. The rapid release of this community resource follows NIH standard rules of data sharing, including requirements for consent and institutional review board approval. Prepublication release of the raw sequence data also follows the Bermuda Principles and the Ft. Lauderdale Large Scale Biological Sequencing Projects accord of 2003.

"These studies have been designed to have enough statistical power to discover both risk alleles and protective alleles," said Adam Felsenfeld, PhD, director of the NHGRI Genome Sequencing Program. "The analysis of this genomic data is just getting started, and we are looking forward to what we might learn."

The National Alzheimer's Project Act ordered the establishment of a National Plan to Address Alzheimer's Disease. The plan established 5 primary goals. The first—to prevent and effectively treat AD by 2025—specifically recognizes genetics research as key to achieving the plan's goals, and a number of ongoing genetics projects are underway, including the NHGRI/NIA sequencing effort (www.nia.nih.gov/alzheimers/publication/2011-2012-alzheimers-disease-progress-report/). Researchers can access the recently deposited WGS data at www.ncbi.nlm.nih.gov/gap or at the National Institute on Aging Genetics of Alzheimer's Disease Data Storage Site (www.niagads.org/).

National Institutes of Health
Bethesda, MD