The International Brain and Behavior Consortium Genomics Gateway

Poster Abstract

David B. Rhee
Department of Genetics
Albert Einstein College of Medicine
1300 Morris Park Avenue
Bronx, NY 10461
david.rhee@einstein.yu.edu

Kevin Shieh
Department of Genetics
Albert Einstein College of Medicine
1300 Morris Park Avenue
Bronx, NY 10461
kevin.shieh@med.einstein.yu.edu

Pilib Ó Broin
Department of Genetics
Albert Einstein College of Medicine
1300 Morris Park Avenue
Bronx, NY 10461
pilib.obroin@einstein.yu.edu

Raquel G. Gur
Departments of Psychiatry, Neurology & Radiology
University of Pennsylvania
Perelman School of Medicine
Philadelphia, PA 19104
raquel@upenn.edu

Aaron Golden
Department of Genetics
Albert Einstein College of Medicine
1300 Morris Park Avenue
Bronx, NY 10461
aaron.golden@einstein.yu.edu

ABSTRACT

The International Consortium on Brain and Behavior in 22q11.2 Deletion Syndrome (22q11DS) is a recently funded collaboration of 22 institutions worldwide, whose aim is to combine genomic with neuropsychiatric and neurobehavioral paradigms to advance the understanding of the pathogenesis of schizophrenia (SZ) and related phenotypes for patients with 22q11.2 deletion syndrome. 22q11DS is caused by a deletion of ~ 3 Mb in extent localized to one copy of chromosome 22, and its current incidence is 1 per 4000 live births, making it one of the most common genetic disorders - individuals with 22q11DS have a number of physical (cardiovascular, cranio-facial) and cognitive (learning difficulties) clinical features in common, although there is a large variation in severity and extent. Patients with 22q11DS face however a substantial risk ($\sim$25-30%) of developing SZ in adolescence or beyond, with illness presentation and course similar to SZ in the general population ($\sim$1%). As part of its work program, the Consortium will genotype 1000 patients with 22q11DS recruited worldwide as part of this study using SNP arrays, and in addition conduct whole genome sequencing (WGS) on 600 individuals so as to uncover genetic variation that may contribute to the heterogeneity of neuropsychiatric and neurobehavioral phenotypes of SZ and psychosis. Sequencing will involve 30$\times$ coverage using Illumina HiSeq2500 sequencing at the Hudson Alpha Institute for Biotechnology, with the resulting alignment (using the BWA aligner) and variant calling (using a variant caller called PEMapper) performed at Emory University. The resulting variant files will be sent to the Albert Einstein College of Medicine who will host the Consortium's genomics database system. All clinical and associated phenotypic data will be stored and managed at the University of Pennsylvania. In this poster we will detail how the resulting variant files will be integrated into a MySQL based genomics database, that will implement the Ensembl database schema, and will be made available to the Consortium as a Science Gateway infrastructure courtesy of XSEDE. Adoption of the same underlying data storage/management schema as used by the 1000 Genomes consortium would similarly allow visualization of specific genomic loci via the popular Ensembl (and other) genome browser and also support complex bioinformatics interactions with both the Consortium's variant data as well as the full set of other curated human genomics datasets using Ensembl's powerful API. This latter aspect will be of particular relevance to the various research groups within the Consortium who are scheduled to process the collated variant data. Hosting of the genomics database as an XSEDE Gateway resource provides the Consortium with a secure and controlled web-portal and terminal access to these sensitive resources. We will outline the current operational status of the Gateway using the first batch of genotyped and sequenced participant samples and describe how access to consortia members and the wider community will be facilitated in collaboration with the clinical database repository hosted by the University of Pennsylvannia. This poster represents an ideal opportunity to allow community feedback as regards the provisioning of future resources and functionalities on the Consortium's Genomics Gateway from interested members of the XSEDE community.

The International Brain and Behavior Consortium acknowledges the support of the National Institute of Mental Health through award 1U01MH101719.