NCRAD

May 2016

Conference Calls

Thank you to everyone who participated in our NCRAD-ADC conference calls this year. We spoke with the Center Director and/or Clinical Core Director at each Center along with additional staff. We enjoy having the opportunity to speak with each center and learn about ways we can help facilitate sample collection and discuss new and ongoing initiatives. We hope you find the calls informative and please do not hesitate to contact us with additional questions.

More Genotype Data Posted

A variety of genetic data is generated from DNA samples shared with NCRAD by the ADCs. NCRAD obtains *APOE* for all ADC samples matched to NACC UDS data. NACC posted *APOE* data in February and April. To obtain these data, visit the NACC website: https://www.alz.washington.edu/adgc.html, select your Center, and choose the option to download *APOE* data from NCRAD. As always, please compare these data with any internal *APOE* genotype data you have generated. Notify NCRAD if you find any discrepancies.

The Alzheimer Disease Genetics Consortium (ADGC) continues to generate GWAS and exome chip data for a subset of ADC-contributed samples. These data can be found on the same page at NACC.

The Alzheimer Disease Sequencing Project (ADSP) generated whole exome sequencing data from over 3,200 ADC samples. To access data for subjects seen at your center, investigators must submit a Data Access Request (DAR) to dbGaP for review. (http://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000572.v1.p1) Information for mapping ADSP IDs to the original patient IDs from the contributing ADCs is available from NACC.

Genomic Data Sharing Policy Documentation

A huge THANK YOU to those who have already completed this updated Genomic Data Sharing (GDS) form!

The Alzheimer's Disease Sequencing Project is continuing to plan for the next phase of sequencing. In order for your Centers' samples to be included in this phase, the NIH Genomic Data Sharing Policy (https://gds.nih.gov/index.html) documentation must be updated. This documentation certifies that the samples you have provided to NCRAD are allowed to be shared and how they can be shared.

NCRAD has been working with every ADC to update the required paperwork for the Genomic Data Sharing Policy to ensure ADC samples can be included. Please note, that this document was updated in the fall of 2015. If you completed the documentation in the spring of 2015 or earlier, the form will need to be updated to the latest version.

The ADSP replication phase is already in the process of selecting samples. We appreciate your help in providing the required documentation as quickly as possible. If you are unsure if you have completed this or not, please contact Kelley Faber at kelfaber@iu.edu.

A Central Repository with DNA available to match the rich dataset collected for all subjects seen in the ADCs is a very valuable resource for the field of AD research. We hope you will continue to support this effort!

Please contact us with any questions or concerns about NCRAD at 800-526-2839/317-274-7360, by email at kelfaber@iu.edu or visit our web-site: www.ncrad.org Thanks!!

NIAGADS The National Institute on Aging Genetics of Alzheimer's Disease Data Storage Site

ADSP Version 7 Release

On April 12, 2016, the ADSP released its seventh version of sequencing data for public distribution through dbGaP (phs000572.v7.p4).

The **seventh ADSP data release** includes:

(1) WES SNV Consensus VCF and PLINK files (10,913 total subjects, of which 3,256 come from ADCs)
(2) WGS SNV Consensus VCF files (578 total subjects)
(3) WES Indel Concordant PLINK files(10,913 total subjects, of which 3,256 come from ADCs)
(4) WGS Indel Concordant PLINK files (578 total subjects)

Please use the release notes provided by dbGaP to obtain detailed information about study release updates. We anticipate the consensus Indel VCF files to be released in early summer 2016.

ADSP Data Available through dbGaP:

	ADSP - Whole Genome Sequencing	ADSP - Whole Exome Sequencing	Comments
DNA-Seq (BAM)	n=578	n=10913	Sequence data available (plus n=38 replications w/out genotype data)
Concordant SNV Genotypes (PLINK format)	N/A	n=10913	QC'ed genotypes that are concordant between the Atlas (Baylor's) and GATK (Broad's) calling pipelines (a subset of the consensus genotype set)
Consensus Genotypes (PLINK and VCF format)	n=578	n=10913	QC'ed genotypes that are concordant between Atlas and GATK pipelines as well as those that that were called uniquely by Atlas or GATK
Concordant Indel Genotypes (PLINK format)	n=578	n=10913	QC'ed genotypes that are concordant between the Atlas and GATK calling pipelines
Phenotype Data	n=4735	n=10913	Data of n=53 phenotype variables available (plus administrative data), including APOE genotype. WGS phenotypes include data of connecting family members.

Exome array data available through NACC

The ADGC has generated exome array data for >11,000 ADC samples. NIAGADS has provided these data to NACC and the data is available for download by each ADC center. More information about the array used can be found on the Illumina website.

ADSP Quality Control Workflow

The ADSP Quality Control (QC) workflow was developed by the ADSP QC Working Group and was applied to all ADSP Discovery Phase WGS and WES data. The workflow takes as input project-level VCF files from SNV calling performed using GATK-Haplotypecaller (Broad Institute) and Altas2 (Baylor College of Medicine) on the same sample set; performs thorough variant-level and sample-level QC on each set of SNV calls; and then compares genotypes at variants present in both call sets to create a single set of VCF files containing genotypes concordant between callers. This workflow produces both QCed VCFs and annotation of the VCF files including variant-level and sample-level metrics and QC exclusion information. The workflow is available at: https://www.niagads.org/adsp-qc-workflow

Genomics DB Version 2.0 Now Available

The NIAGADS Genomics DB v2.0 is now available at https://www.niagads.org/genomics.

New in Version 2.0:

- Enhanced search interface and improved presentation of gene and SNP information makes it easier than ever to identify AD-relevant sequence features.
- 72 new datasets are availble to search, including both NIAGADS GWAS Summary Statistics and functional genomics datasets from ENCODE and FANTOM5.
- New tools have been added to facilitate research, such as pathway or functional enrichment analysis and co-location searches.

Email rcweibel@upenn.edu to be added to the Genomics DB mailing list. For questions about the Genomics Database, email genomicsdb@niagads.org.