

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](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE68814)).

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 available 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.