RELEASE NOTES- AMP-AD SV-xQTL

Data Release Date: April 24, 2023 **Dataset Version:** ng00118.v1

Release Information:
The first release (April 24, 2023) includes SV genotypes in VCF file format, phenotype data, and SV-xQTL association summary statistics. Subjects in ROS/MAP not consented for individual SV calls were removed from the individual level data submission but are still part of the aggregate association summary statistics.

Dataset Description:
Structural variants (SVs) were discovered in 1,760 donors by running a combination of seven different tools to capture the main classes of variation, including deletions (DEL), duplications (DUP), insertions (INS), inversions (INV), mobile element insertions (MEI), and complex rearrangements (CPX). We mapped associations of 25,421 SVs with MAF ≥ 0.01 in the ROS/MAP cohorts to four different molecular phenotypes in the DLPFC. These molecular phenotypes were measured for a partially overlapping set of samples and included gene expression for 15,582 genes (n=456), 110,092 splicing junctions proportions measured by “percent spliced in” values (PSI) (n=505), histone acetylation (H3K9ac) peaks (n=571), and proteomic data for 7,960 proteins (n=272). We refer to these analyses as SV-xQTL, in which we map differences in measurements of each molecular phenotype associated with specific SV's. Therefore, each SV-xQTL is an SV-phenotype pair (i.e., SV-eQTL, SV-sQTL, SV-haQTL, or SV-pQTL). All phenotype measurements were adjusted prior to associations to account for known (e.g., sex and ancestry principal components) and unknown covariates, determined either with PEER (probabilistic estimation of expression residuals) or PCA (principal component analysis), and the allele alternative to the genome of reference was considered as effect allele. This identified 3,191 SV-eQTL, 2,866 SV-sQTL, 399 SV-pQTL, and 1,454 SV-haQTL (FDR < 0.05).

SV calls were generated using data from the AMP-AD Whole Genome Sequencing Harmonization Study made available through the AD Knowledge Portal.

File Manifest: https://st1.niagads.org/portal/download-public/NG00118.v1/fm

Subject Consents:
Sequenced subjects in this dataset belong to the following consent levels as indicated by the submitting study IRBs:

Consent Level*	# Subjects
DS-ND-IRB-PUB-NPU	715
HMB-IRB-PUB	305
GRU-IRB-PUB	429
Total	1369

*Consent level definitions can be found on the Data Use Limitations page.

Dataset Accession Numbers Available in ng00118.v1:

Type	Description	Accession
Dataset	AMP-AD Structural Variation – Genotyping Data and Summary Statistics	ng00118
Study	Integrating whole-genome sequencing with multi-omic data reveals the impact of structural variants on gene regulation in the human brain - Vialle et al. 2022	sa000028
Sampleset	AMP-AD WGS – SV Calls	snd10041
Fileset	AMP-AD SV Genotyping Data	fsa000041
Fileset	AMP-AD SV-xQTL Association Summary Statistics	fsa000042