NIAGADS Update: Spring 2023

Li-San Wang, PhD

University of Pennsylvania

COI: Nothing to Disclose

NIAGADS

NIAGADS At a Glance

Data Sharing

96 datasets

>172,701 samples

474 Unique Requests

218 Institutional Certifications

Userbase

>250 articles cite U24

>3,167 verified users

ADSP

62 cohorts

24 current grant subawards

Exomes

20,503

Genomes

~350% increase from 2021 to 2023

(16,905 in 2021 to ~60,000 genomes in 2023)

ADSP Support

14 workgroups /

5 consortia

309 members

49 institutions

5,730 documents

16 in-person/virtual meetings

1,668 conference calls

NIAGADS

ADSP WGS Data Release

ADRCs can get WGS/WES data from their own samples from us!

Send an email to niagads@pennmedicine.upenn.edu to initiate.

Release 4 Preview

19,456 new genomes

released in October 2022

Release includes

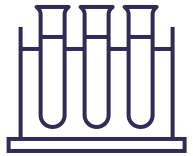
- Sequencing read alignments in CRAM (compressed BAM) format
- genomic Variant Call Format (gVCF) files
- Phenotypes in ADSP format, basic diagnosis, age, sex, self-reported race/ethnicity

Joint genotype calls of

36,361 total genomes

- including R3 in VCF format (pVCF)
- 40 Cohorts
- 20 Research Use Limitation Levels (Consents) from > 60 GDS Forms

ADRC Samples in ADSP



- **4,590 exomes** (**22%** of total ADSP WES)
- **11,195 genomes** (**30%** of total ADSP WGS)

ADSP-Phenotype Harmonization Consortium (PHC) Phase 0 Release

- ADSP Phenotype Harmonization Consortium (ADSP-PHC) released their first set of harmonized phenotypes
- READMEs, data dictionaries, and harmonized data files are available for the following domains:
  - Cognition
  - Fluid biomarker
  - Neuropathology

Cohort	Cognition	Biomarker	Neuropath
ACT	1,337	0	0
ADNI	1,566	1,165	0
KGAD	0	64	0
MAP- Rush	639	0	538
MARS	48	0	11
NACC	10,488	805	4,649
NIA-LOAD	0	2	262
ROS	583	0	532
Total	**14,661**	**2,036**	**5,992**

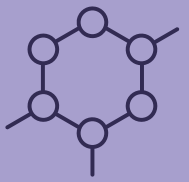
PIs: Tim Hohman / Vanderbilt U
Michael Cuccaro / U Miami
Arthur Toga / USC

NIAGADS

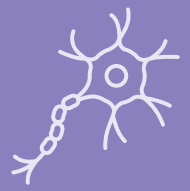
ADSP Spring 2023 R4 QC Release

Coming in Q2 2023

QC'd genotypes



Individual level structural variant calls



Phenotypes

- updates
- Instructional **quick guide**
- Script to combine phenotype files for a standard AD case/control analysis

Identity by descent (IBD), Principal Component Analysis Coordinates, and other additional information that can be used in genetic association analysis

NIAGADS

ADSP R5 WGS Release

Coming in Late 2023

- CRAMs, gVCFs, joint genotype called pVCF
- SV individual level calls

~23,128 new genomes,
totaling **59,489 genomes** joint genotype called in R5

12 new cohorts added

PHC Phase 1 release will include:

- vascular risk factors
- neuroimaging
- additional cohorts

NIAGADS

ADRC GWAS Data Release - ADGC

ADC1–7

Released on NIAGADS

ADC8–12

Being prepared to release on NIAGADS

- Imputation using Haplotype Reference Consortium (HRC) panel (~33K genomes, 39M variants)

ADC14–15

Undergoing QC

- TOPMed Imputation (~97K genomes, 308M variants)

Totaling GWAS SNP Array Data for 31,115 ADRC Samples

ADGC is updating Imputation for ADC1-12 using TOPMed panel, data under review

NIAGADS

ADRC Genetic and Genomic Data Collections

NIAGADS is supporting the development of customized webpages with an in-depth summary of ADRC center studies and the data sets available.

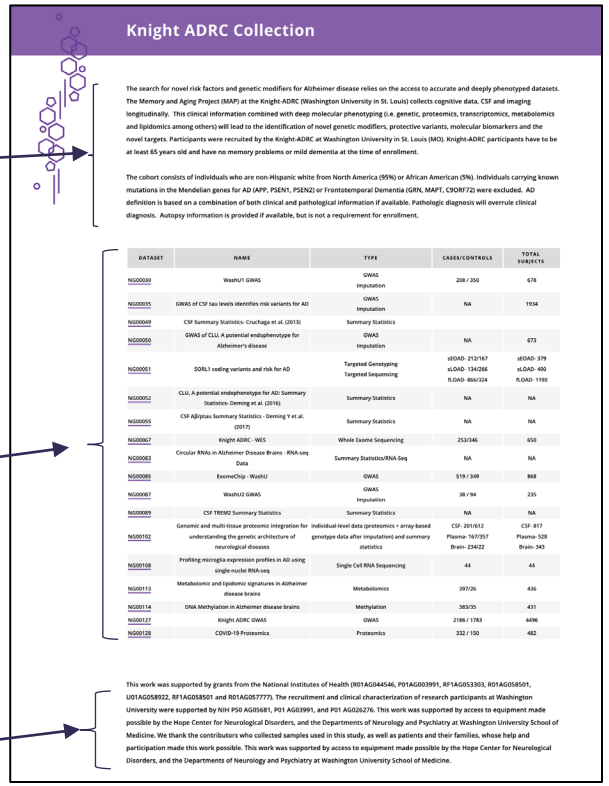
E.g. Knight ADRC Data Collection (with help from Carlos Cruchaga)

Short summary about the study, data, and cohort

Overview of the data sets:

Knight ADRC consists of 18 datasets inclusive of:
- GWAS
- RNA-seq (bulk/single cell) methylation
- Proteomics
- Summary Statistics

Funding sources and acknowledgements



Knight ADRC Collection

The search for novel risk factors and genetic modifiers for Alzheimer disease relies on the access to accurate and deeply phenotyped datasets. The Memory and Aging Project (MAP) at the Knight-ADRC (Washington University in St. Louis) collects cognitive data, CSF and imaging longitudinally. This clinical information combined with deep molecular phenotyping (i.e. genetic, proteomics, transcriptomics, metabolomics and lipidomics among others) will lead to the identification of novel genetic modifiers, protective variants, molecular biomarkers and the novel targets. Participants were recruited by the Knight-ADRC at Washington University in St. Louis (MO). Knight-ADRC participants have to be at least 65 years old and have no memory problems or mild dementia at the time of enrollment.

The cohort consists of individuals who are non-Hispanic white from North America (95%) or African American (5%). Individuals carrying known mutations in the Mendelian genes for AD (APP, PSEN1, PSEN2) or Frontotemporal Dementia (GRN, MAPT, C9ORF72) were excluded. AD definition is based on a combination of both clinical and pathological information if available. Pathologic diagnosis will overrule clinical diagnosis. Autopsy information is provided if available, but is not a requirement for enrollment.

DATASET	NAME	TYPE	CASES/CONTROLS	TOTAL SUBJECTS
NG00030	WashU1 GWAS	GWAS Imputation	208 / 350	678
NG00035	GWAS of CSF tau levels identifies risk variants for AD	GWAS Imputation	NA	1934
NG00049	CSF Summary Statistics- Cruchaga et al. (2013)	Summary Statistics		
NG00050	GWAS of CLU, A potential endophenotype for Alzheimer's disease	GWAS Imputation	NA	673
NG00051	SORL1 coding variants and risk for AD	Targeted Genotyping Targeted Sequencing	sEOAD- 212/167 sLOAD- 134/266 fLOAD- 866/324	sEOAD- 379 sLOAD- 400 fLOAD- 1190
NG00052	CLU, A potential endophenotype for AD: Summary Statistics- Deming et al. (2016)	Summary Statistics	NA	NA
NG00055	CSF Aβ/ptau Summary Statistics - Deming Y et al. (2017)	Summary Statistics	NA	NA
NG00067	Knight ADRC - WES	Whole Exome Sequencing	253/346	650
NG00083	Circular RNAs in Alzheimer Disease Brains - RNA-seq Data	Summary Statistics/RNA-Seq	NA	NA
NG00085	ExomeChip - WashU	GWAS	519 / 349	868
NG00087	WashU2 GWAS	GWAS Imputation	38 / 94	235
NG00089	CSF TREM2 Summary Statistics	Summary Statistics	NA	NA
NG00102	Genomic and multi-tissue proteomic integration for understanding the genetic architecture of neurological diseases	Individual-level data (proteomics + array-based genotype data after imputation) and summary statistics	CSF- 201/612 Plasma- 167/357 Brain- 234/22	CSF- 817 Plasma- 528 Brain- 345
NG00108	Profiling microglia expression profiles in AD using single-nuclei RNA-seq	Single Cell RNA Sequencing	44	44
NG00113	Metabolomic and lipidomic signatures in Alzheimer disease brains	Metabolomics	397/26	436
NG00114	DNA Methylation in Alzheimer disease brains	Methylation	383/35	431
NG00127	Knight ADRC GWAS	GWAS	2186 / 1783	4496
NG00128	COVID-19 Proteomics	Proteomics	332 / 150	482

This work was supported by grants from the National Institutes of Health (R01AG044546, P01AG003991, RF1AG053303, R01AG058501, U01AG058922, RF1AG058501 and R01AG057777). The recruitment and clinical characterization of research participants at Washington University were supported by NIH P50 AG05681, P01 AG03991, and P01 AG026276. This work was supported by access to equipment made possible by the Hope Center for Neurological Disorders, and the Departments of Neurology and Psychiatry at Washington University School of Medicine. We thank the contributors who collected samples used in this study, as well as patients and their families, whose help and participation made this work possible. This work was supported by access to equipment made possible by the Hope Center for Neurological Disorders, and the Departments of Neurology and Psychiatry at Washington University School of Medicine.

Work in progress on pages for Stanford ADRC and UAB/HudsonAlpha

If centers are interested, please contact NIAGADS@pennmedicine.upenn.edu to being the development process

See the page in action at: www.niagads.org/knight-adrc-collection

NIAGADS

ADSP Functional Genomics Consortium

PIs
Phil De Jager (Columbia)
Carlos Cruchaga (WashU)
Gao Wang (Columbia)
Fanny Leung (NIAGADS)

- Raw data going to Sage Bionetworks AMP-AD Portal
- Collaboration between Sage and NIAGADS for data integration

xQTL analysis – aimed release second half of 2023

2023 Pilot Data Free at a Glance		
Cohort	**Tissue / Cell**	**Omics QTL**
ROSMAP	Brain tissues (DLPFC, PCC, AC), various DLPFC cells, microglia, monocyte	haQTL (H3K9ac), meQTL, eQTL, sQTL, pQTL, metaQTL
Knight ADRC	Brain, CSF, plasma	meQTL, eQTL, sQTL, pQTL, metaQTL
STARNET	Myeloid (macrophage, monocyte)	eQTL
DIAN	Blood derived	meQTL
EFIGA	CSF and plasma	pQTL, metaQTL
MAGENTA	Blood derived	eQTL, sQTL
Mount Sinai Brain Bank	Brain	meQTL, eQTL, sQTL, pQTL
PART Working Group	Brain	meQTL

NIAGADS

Supporting Data Use

User Support

NIAGADS is always looking to improve user experience. Over the next year we are planning to introduce several new resources:

- Office Hours
- Video Tutorials
- NIAGADS Roadshow

Glad to talk at your ADRC and meet colleagues!

Awareness

Expanding NIAGADS reach through increase investigator awareness will be accomplished through informational outreach sessions at AAIC and ASHG in 2023

- Exhibit Hall Booth
- In Booth Presentations
- ADSP Session at AAIC

NIAGADS

Increasing Diversity in AD Genetics Data: ADSP Follow-Up Sequencing (FUS)

GCAD Release Date and Funding/Sequencing Status	Primary Ancestry/Ethnicity	Number of Alzheimer Disease Cases	Number of Controls	Other including MCI, ADRD, and PSP	Totals
GCAD Release 5** - Funded and scheduled for sequencing in 2023 - Data release planned for 2023	Black/African American	2,635	5,038	810	8,483
	Asian	2,336	5,096	35	7,467
	Latino/Hispanic	4,053	7,652	2,348	14,053
	Non-Hispanic White	8,565	11,642	5,447	25,654
	Total	17,589	29,428	8,640	55,657

NIAGADS

Asian Cohort for Alzheimer's Disease

ACAD

U19 (Pending): 5,000 participants/5 years in US and Canada

- Chinese / Korean / Vietnamese Ancestry in US and Canada

Collaborations with ADRCs / Investigators from Boston U, Columbia, Mount Sinai, Indiana U, NYU, Penn, UCI, UCSF, Stanford, USC, UCSD

Culturally appropriate recruitment and assessment

- Data collection packet developed **based on UDS with NACC input**

Protocols, SOPs, Consent, Translation procedure, QA and Training material, REDCap schema and data manuals

DNA for genetics and plasma for blood-based biomarkers

- **NCRAD hosts samples and generates biomarker data**

We look forward to collaborations with ADRCs!