Accelerating Medicines Partnership®
Parkinson’s Disease

Debra Babcock, NINDS
Pablo Sardi, Sanofi
AMP® PD

Creating a platform with harmonized multi-omic data resources to support the identification of progression biomarkers & improve clinical trial design

- Providing a platform to identify prognostic/disease progression biomarkers for PD (Stage 1)
- Enabling deep molecular characterization and longitudinal clinical profiling of PD (Stage 2)
- Identifying biomarkers (Stage 3)

AMP® PD

Stage 1  Stage 2  Stage 3  NCE

0 – 18 months 12 – 48 months 36 – 72 months

- Launch of AMP PD knowledge platform amp-pd.org
- Strategic database and tool expansion
- Initiation of data analyses
- Final data integration
- Biomarker-based data competition
- Data analysis results

We are here

Feb ‘18

Feb ‘24
AMP PD Knowledge Platform

In 2019, the AMP PD Knowledge Platform went live, with a public-facing website for access registration, data storage on the Google Cloud and data access and analysis through the Terra Platform.

In 2021, the AMP PD harmonized cohorts were federated with the Global Parkinson’s Genetics Program (GP2), which is genotyping 150,000 subjects with PD worldwide. All cohorts can be accessed through the AMP PD Knowledge Platform, providing a seamless experience through a collaborative cloud-based resource and computing environment, providing a one-stop-shop for Parkinson’s Disease researchers via:

- A single data use agreement that provides access to unified AMP PD and complementary GP2 Cohorts data.
- A consistent use of Terra and Google cloud agnostic-based computing environments for all data sets.
Current data on the AMP® PD Knowledge Platform

Eight unified cohorts (harmonized, longitudinal)
1. BioFIND
2. Harvard Biomarkers Study (HBS)
3. Parkinson's disease Biomarkers Program (PDBP)
4. Parkinson's Progression Marker Initiative (PPMI)
5. Lewy body dementia case-control cohort (LBD)
6. LRRK2 Cohort Consortium (LCC)
7. Study of Isradipine as a Disease Modifying Agent in Subjects With Early Parkinson Disease, Phase 3 (STEADY-PD3)
8. Study of Urate Elevation in Parkinson’s Disease, Phase 3 (SURE-PD3)
9. Federated access to Global Parkinson’s Genetics Program (GP2)

Data Types:
1. Clinical
2. Genomics
3. Transcriptomics
4. Proteomics (Targeted and Untargeted)
5. scRNA Seq from Postmortem Brain tissue
Additional Data to be available on the AMP-PD platform

- Extracellular Vesicle RNA seq data (2400 samples)
- Transcriptomics data from the Harvard Biomarker study cohort (2250 samples)
Recent Biomarkers identified using AMP® PD Data

“Inflammatory Blood Biomarkers Are Associated with Long-Term Clinical Disease Severity in Parkinson’s Disease”. Hepp DH et. al., Int J Mol Sci. 2023 Oct; 24(19): 14915

<table>
<thead>
<tr>
<th>Differentiates PD vs Control</th>
<th>Higher Levels Predict Worse Symptoms</th>
</tr>
</thead>
<tbody>
<tr>
<td>IL-12B</td>
<td>CCL23</td>
</tr>
<tr>
<td>OPG</td>
<td>CCL25</td>
</tr>
<tr>
<td>CXCL11</td>
<td>TNFRSF9</td>
</tr>
<tr>
<td>CSF-1</td>
<td>TGF-α</td>
</tr>
<tr>
<td></td>
<td>VEGFA</td>
</tr>
</tbody>
</table>


Higher Expression Predicts MCI Progression

| ** IGLC1                      |
| IGLV3-21                     |
| IGLV1-44                     |
| CARD-17                      |
| ZFP91-CNTF                   |
Resource for exploring nominated genes and proteins from AMP PD data

https://target-explorer.amp-pd.org/
### Summary of Key Accomplishments

The AMP PD program is built on Parkinson’s Disease-related study data that has been collected, harmonized, and made publicly available to identify new biomarkers and develop new PD treatments.

**Outcomes:**

<table>
<thead>
<tr>
<th>checked</th>
<th>HARMONIZED DATA FROM 8 COHORTS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Harmonized data from eight cohorts, including ~10,800 participants, and generated omics data from a majority of their biosamples collected under similar protocols</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>checked</th>
<th>LEVERAGED RESOURCES</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Overcame challenges associated with resources, time, cost, and data availability is supported by the collective efforts of multiple studies</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>checked</th>
<th>PROMOTED DATA ACCESSIBILITY</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Numerous groups (80+ publications) have utilized data from AMP PD to discern potential targets and biomarkers for Parkinson’s Disease, as well as critical findings about the disease pathobiology</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>checked</th>
<th>REGISTERED MULTIPLE USERS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>The AMP PD data platform has 1000+ registered users</td>
</tr>
</tbody>
</table>
Phase 2: AMP® Parkinson’s Disease and Related Disorders (PDRD)

Specific Goal: To Identify and validate proof of mechanism biomarkers tied to Parkinson’s and other α-synucleinopathies.

- Aim 1: Longitudinal Multi-omic Molecular Profiling
- Aim 2: Discover/Validate PD and PDRD Biomarkers
- Aim 3: Multiscale Analysis Framework

Active fundraising stage to initiate in the second quarter of 2024.
For AMP® PDRD Partnership Questions, Please contact

Sri Pullagura  
Program Manager, Neuroscience  
Translational Science Partnerships  
spullagura@fnih.org

Heidi Blythe  
Director of Development  
hblythe@fnih.org
Thank you!
Established a cloud-native data portal with data access through the Terra platform (knowledge platform)
AMP PD Knowledge Portal

2018 -> Present

Activation -> Outcomes

Matt Bookman (Verily Life Sciences)
Accelerating Parkinson’s Research via AMP PD

**Objective**

To empower scientific collaborations and enable the discovery of rich Parkinsons data to accelerate biomarker discovery and the advancement of Parkinson’s disease therapies.

**Outcome**

The AMP PD Knowledge Portal harmonizes multi-omics (WGS, RNA sequencing data) with clinical data from thousands of PD patients across studies to provide the longitudinal view needed to inform more robust biomarkers and clinical trials.
Overarching Implementation Goal

Activate the Partnership ➔ Build the Platform ➔ Activate the Community
Activate the AMP PD Partnership
Initial framing (Jan 2018)

- Objective
  - Build a knowledge portal to discover and validate biomarkers for Parkinson's Disease
  - Launch within 2 years (by Jan 2020)

- Contributors
  - Small directly contracted team
    - (Shared) FNIH Project Director
    - (Half) Data Coordinator
  - Large number of invested parties across the partnership
    - Partners
    - Connected science / data science organizations
Collaboration: Modern science is a team sport

If you want to go fast, go alone.
If you want to go far, go together.

- Proverb of unknown origin
Identify system components

- Data
  - It's the most important thing

- Portal
  - It's the way to communicate with everyone and is independent of the data and tools

- Data Explorers
  - "Custom" tools for this project's data

- Workbench
  - General purpose platform for biomedical research
System Components

Portal
Data Explorers
Researcher Workbench

Data
Activate the workgroups/workstreams

- Get started, identify goals, identify resource/responsibility gaps, identify decision making process
- (Short term) fill gaps with current team
- (Long term) identify resourcing needs
  - Contracted building and maintenance of the portal (Sapient)
  - Contracted data harmonization (Rancho Biosciences)
  - Hired project manager (@ FNIH)
  - Hired cloud operator (@ Technome)
  - Focused partnership's resources on data, science, explorers, and workbench
Data Working Group and Subgroups

Data Use Policies Subgroup:
- Deb Babcock (NINDS) – Co-Chair
- Rosa Canet-Aviles (FNHI) – Co-Chair
- Vanessa Arredondo (Moff)
- Matt Bookman (Vently)
- Alyssa Reimer (Moff)
- Clements Scherzer (Harvard)
- Lubo Smolensky (Moff)
- Marg Sutherland (NINDS)

Data Working Group:
- Marg Sutherland (NINDS) – Co-Chair
- Matt Bookman (Vently) – Co-Chair
- Vanessa Arredondo (Moff)
- Deb Babcock (NINDS)
- Shamere Bivans (Celgene)
- Jeff Biju (NCI)
- Rosa Canet-Aviles (FNHI)
- David Craig (USC)
- Todd Fisher (Sano)
- Mark Fraser (Moff)
- Alex Guttenidge (GSK)
- Gehrad Krimin (Sano)

Proteomics Subgroup:
- Rosa Canet-Aviles (FNHI)
- Bradford Casey (Moff)
- Andy Chastain (Celgene)
- Peter Cicherman (Vently)
- Samantha Hutton (Moff)
- Barry Landin (Data Coordinator)
- Pablo Sardi (Sano)
- Thomas Snyder (Vently)
- Marg Sutherland (NINDS)
- Paulina Wolff (Sano)

Transcriptomics Subgroup:
- Kendal Jansen (Tgen) – Co-Chair
- Marg Sutherland (NINDS) – Co-Chair
- Shamere Bivans (Celgene)
- Matt Bookman (Vently)
- Bradford Casey (Moff)
- Rosa Canet-Aviles (FNHI)
- David Craig (USC)
- Alex Guttenidge (GSK)
- Dinosh Kumar (Sano)
- Barry Landin (Data Coordinator)
- Shawn Levy (Hudson Alpha)
- Michael Nagle (Pfizer)
- Jerrod Schwartz (Vently)
- Thomas Snyder (Vently)
- Christine Swanson-Fischer (NINDS)
- Ivo Violich (USC)

WGS Subgroup:
- Andy Singleton (NIH/LNG) – Chair
- Shamere Bivans (Celgene)
- Matt Bookman (Vently)
- Rosa Canet-Aviles (FNHI)
- Bradford Casey (Moff)
- Matt Edwards (Vently)
- Mark Fraser (Moff)
- Dinesh Kumar (Sano)
- Barry Landin (Data Coordinator)
- David Pullford (GSK)
- Shini Shankara (Sano)
- Thomas Snyder (Vently)
- Marg Sutherland (NINDS)
- Mike Nalls (NIA)
- Jenna Hernandez (NIA)
- Hampton Leonard (NIA)
- Hirohito Iwaki (NIA)
- Boris Scholze (NIA)
- Willy Nojopramo (Vently)

Portal Subgroup:
- Matt Bookman (Vently) – Co-Chair
- Luba Smolensky (Moff) – Co-Chair
- Ainsley Azzo (Saphient)
- Deb Babcock (NINDS)
- Rosa Canet-Aviles (FNHI)
- Elaine Heitzman (Saphient)
- Barry Landin (Data Coordinator)
- Jon Meyrick (Saphient)
- Young Stewart-Frederick (Saphient)
- Marg Sutherland (NINDS)
- Carl Wonders (NINDS)

Clinical Data Harmonization Subgroup:
- Alyssa Reimer (Moff) – Co-Chair
- Deb Babcock (NINDS) – Co-Chair
- Marg Sutherland (NINDS)
- Matt Bookman (Vently)
- Rosa Canet-Aviles (FNHI)
- Lindsey Gaiser (Celgene)
- Barry Landin (Data Coordinator)
- Michal Nagle (Pfizer)
- Clements Scherzer (Harvard)

Managed and coordinated through FNHI
Components, Owners, Contributors
Activate data working groups for each subtype

- Attendees
  - Scientists, data scientists, delivery team
- Always asking: what would researchers expect?
  - Are we innovating or using community standards?
  - Provide data in native/standard file formats
  - **Add** aggregate, normalized, or transformed versions of data
  - **Add** tables for tabular data
- Oversight of Solutions Architect
  - Allow for modality idiosyncrasies while avoiding *arbitrary* inconsistency
Release 1 (2019)
Architecture Implemented
Access Request Process

1. Complete AMP PD Access Form
2. Coordinate with AMP PD ACT
3. Access AMP PD Data

- Access Control managed AMP PD (ACT + Technome)
- "Granted Access" = data + tools + code

- Data (Google Cloud)
- Tools (Data Explorers, Jupyter, RStudio)
- Examples (Workspaces)
Researcher Access - Data and Tools
AMP PD Today
Complimentary Datasets
AMP PD Researcher Resources

- Explorers
**Exploration**

**Summary Data Dashboard**

Total Case vs Control Participants

<table>
<thead>
<tr>
<th>Category</th>
<th>Case</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total PD Case participants</td>
<td>3,484</td>
<td></td>
</tr>
<tr>
<td>Total PD Control participants</td>
<td>4,193</td>
<td></td>
</tr>
<tr>
<td>Total other cases participants</td>
<td>2,298</td>
<td></td>
</tr>
</tbody>
</table>

**Biological Sex Within Cohort**

**Most Common Participant Diagnoses**

- Internal PD
- External PD
- Movement disorder
- Neuropathy
- Leukoencephalopathy
- Cognitive deficit
- Depression
- Fatigue
- Sleep
- Fatigue
- AD
- Depression
- PD

**Explore Parkinson’s Disease Genes**

The AMP PD Target Explorer is your go-to resource for exploring nominated genes and proteins from AMP PD data that might be implicated in Parkinson’s Disease. Explore targets from genomics, transcriptomics and proteomics data, contributed by the AMP PD community. Together, we can validate targets and move Parkinson’s Disease research forward.

**Public**


https://target-explorer.amp-pd.org/
Exploration

Cohort Dashboard

Total Global Data

<table>
<thead>
<tr>
<th>Locations</th>
<th>Cohorts</th>
</tr>
</thead>
<tbody>
<tr>
<td>59</td>
<td>169</td>
</tr>
</tbody>
</table>

- 179,142 Samples Expected
- 44,831 Samples Completed

https://gp2.org/cohort-dashboard-advanced/

GP2 Monogenic Resource Map

https://gp2.org/monogenic-resource-map/
Exploration

Tier 1

https://clinical-data-explorer.amp-pd.org/
Exploration

https://data-explorer.amp-pd.org/
Exploration

Tier 2

Community Contributed

https://amp-rna.uc.r.appspot.com
AMP PD Researcher Resources

- Workspaces
Acclerating Medicines Partnership Parkinson's Disease (AMP PD)

AMP PD (Accelerating Medicine Partnership in Parkinson's Disease) was launched in 2018 as a public/private partnership intended to build a database with deep molecular characterization and longitudinal clinical profiling of PD patient data and biosamples with the goal of identifying and validating diagnostic, prognostic and/or disease progression biomarkers for PD.

A critical component of this partnership is broad sharing of the AMP PD data and analyses with the biomedical community to advance research in PD. AMP PD utilizes well characterized cohorts with existing biosamples and clinical data that were collected under comparable protocols and using common data elements.

A list of participating organizations and contributing studies can be found below.

Data

Data Overview

The current release of AMP PD is release 4.0 and of GP2, its federated dataset, is 4.0. See below for release history.

AMP PD Release 4.0 contains

- 10,908 participants
- 10,490 whole genome sequence (WGS) samples
- 8,461 transcriptomics (RNAseq) samples
- 621 untargeted proteomics participants at various timepoints, including raw and mzML sample level data
- 413 targeted proteomics participants at various timepoints from Release 3.0 in a combined, normalized release

https://app.terra.bio/#workspaces/amp-pd-public/AMP-PD-In-Terra
## Workspaces

### AMP PD - Getting Started Tier 1 - Clinical Access - Version 4

#### Data
AMP PD data is stored on Google Cloud Platform.

#### Google Cloud Storage
Raw files are stored in Google Cloud Storage buckets:
- `gs/amp-pd-data`: Clinical data files and other participant metadata.

#### Google BigQuery
Tabular data is stored in Google BigQuery for SQL query access:

### Notebooks

When you are ready to run the notebooks in your cloned instance, be sure to load the data from the Jupyter notebooks menu by selecting `Cell > All Output` to clear output or by selecting `Cell > Run All`.

#### Getting started

**Python 3**
- Py3 - Clinical - load a CSV from Cloud Storage
- Py3 - Clinical - load a table from BigQuery

**R**
- R - Clinical - load a CSV from Cloud Storage
- R - Clinical - load a table from BigQuery

---

**Jupyter**
- Py3 - Clinical - load a CSV from Cloud Storage.ipynb
- Py3 - Clinical - load a table from BigQuery.ipynb
- Py3 - Saving Notebook Results.ipynb
- R - Clinical - load a CSV from Cloud Storage.ipynb
- R - Clinical - load a table from BigQuery.ipynb
- R - Saving Notebook Results.ipynb
## Tier 2 - Clinical and Omics Access

<table>
<thead>
<tr>
<th>Topic</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Getting Started Tier 2 - Clinical and Omics Access</td>
<td>Get started with AMP PD clinical, RNASeq, and WGS data. Tour the layout of files in Cloud Storage and tables in BigQuery. Load data into a Python 3 or R notebook for analysis and visualization.</td>
</tr>
<tr>
<td>Proteomics Getting Started</td>
<td>Get started using Terra notebooks to analyze AMP PD proteomics data.</td>
</tr>
<tr>
<td>Proteomics QC and Analysis</td>
<td>Access QC and analysis notebooks for AMP PD.</td>
</tr>
<tr>
<td>RNASeq Release Workflows</td>
<td>Learn how to run workflows on AMP PD RNASeq data in Terra.</td>
</tr>
<tr>
<td>RNASeq QC and PCA</td>
<td>Understand AMP PD RNASeq data, data quality, and analysis tools available.</td>
</tr>
<tr>
<td>Polygenic Risk Scores</td>
<td>Calculate Polygenic Risk Scores (PRS), a way to determine risk of developing disease, based on the number of genetic changes related to the disease.</td>
</tr>
</tbody>
</table>

https://app.terra.bio/#workspaces/amp-pd-public/AMP-PD-In-Terra
## Workspaces

### Community-Provided Workspaces

<table>
<thead>
<tr>
<th>Workspace</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simulated Proteomics Data Set Workspace</td>
<td>Work with a simulated proteomics data set. No AMP PD access required! Note, a billing project is necessary.</td>
</tr>
<tr>
<td>Accelerating Parkinson's Diagnosis using Multi-omics and Artificial Intelligence</td>
<td>This workspace created by Dr. Ruifeng Hu for a NINDS funded U01 grant. Within the workspace, there are scripts for data pre-processing, differentially expressed genes analysis and Parkinson's Disease diagnosis classifier models will be introduced.</td>
</tr>
<tr>
<td>AMP PD Tandem Repeat Expansion Pipeline Workflows</td>
<td>Workspace, developed by Bharati Jadhav at the Icahn School of Medicine at Mount Sinai, describes the pipeline to identify short tandem repeat expansions using the STretch and GangSTR workflows on the Terra platform. These workflows are designed for Whole Genome Sequencing data starting from mapped cram files.</td>
</tr>
<tr>
<td>Target Explorer Differential Protein Expression</td>
<td>See how differential protein lists were generated for AMP PD's newest tool: Target Explorer.</td>
</tr>
<tr>
<td>Target Explorer Transcriptomics Differential Expression</td>
<td>See how differential gene expression lists were generated for AMP PD's newest tool: Target Explorer.</td>
</tr>
</tbody>
</table>

AMP PD – Researcher Engagement
Researcher Engagement - Webinars

- 2019 AMP PD Webinar (NINDS, Technome, Broad)
- 2022 Variant Aware Alignment on AMP PD RNASeq data
  - Ramiya Sivakumar (USC)
- 2022 APOE status across AMP PD subjects and cohorts
  - Victoria Dardov (Technome)
- 2022 Extracellular RNA expression analysis from CSF and Plasma
  - Elizabeth Hutchens (TGen)
- 2022 Predictive Models of PD Risk
  - Ruifeng Hu (Brigham and Women's Hospital, Harvard Medical School)
- 2022 Determining Polygenic Risk Score
  - Mary Makarious (NIA)
- 2022 International HUPO (Cedars-Sinai, NINDS, Sanofi, Technome, Verily)
Researcher Engagement - Hackathons / Competitions

• 2021 GP2/IPDGC Hackathon
• 2023 GP2 Hackathon
• 2023 Proteomics Kaggle Competition
Researcher Engagement – AMP PD Newsletter

- 5 AMP PD Newsletters since July 2022
Researchers and Outcomes

- Total: ~1,000
  - Tier 1: ~700
  - Tier 2: ~300

- Publications
  - AMP PD Publication Archive
    - 84 papers
  - Google Scholar: "AMP PD" "Parkinson" "Accelerating"
    - 164 results

- Nominated Targets (per the Target Explorer)
  - 13,000+ targets
  - 14 target lists
  - 6 teams
Thank-you
AMP-PD: Single-nucleus analysis in PD

AMP Symposium
February 5th, 2024

Panos Roussos, M.D., Ph.D.

Friedman Brain Institute
Department of Genetics and Genomic Sciences
Department of Psychiatry
James J. Peters VA Medical Center

Center for Disease Neurogenomics
Single-nucleus transcriptome profiling across multiple brain regions in Parkinson's Disease

Goal:

• perform snRNA-seq and WGS in postmortem brain tissue from 100 patients with PD and unaffected controls.

• perform sequencing in four five separate brain regions, including some regions known to be affected by PD pathology and at least one region known to be relatively unaffected.

• Omics and clinical data will be broadly shared through the AMP-PD Knowledge Portal.
Brain regions based on PD progression

Braak Stages of PD

Region related to AMP-AD and psychAD
- Dorsolateral prefrontal cortex

Cognition Emotion
- Primary motor cortex

Motor signs
- Globus pallidus interna

Non-motor signs
- Dorsal motor nucleus of the Xth nerve

Stage 6
Stage 5
Stage 4
Stage 3
Stage 2
Stage 1

Unaffected region
- Primary visual cortex

snRNAseq was performed in each dissection. We estimated 4,000 nuclei per region and donor.

Final dataset would include 2M single nucleus profiles

We performed RIN and DNA extraction for WGS
Cohort’s characteristics

Brain bank composition

- Harvard: 24%
- MSSM: 30%
- Udall: 21%
- UMBEB: 25%

Ethnicity

- White: 89%
- Black: 9%
- Unknown: 2%
- American Indian or Alaska Native: 8%

Sex

- Female: 62%
- Male: 38%

PD diagnosis

- PD or Parkinsonism: 75%
- Control: 25%

Age by diagnosis

- Control: 40-60 years
- PD or Parkinsonism: 60-80 years

Age by sex

- Male: 40-60 years
- Female: 60-80 years

Braak staging by brain bank

- Harvard: Red
- MSSM: Orange
- Udall: Green
- UMBEB: Blue

Diagnosis by brain bank

- Control: Orange
- PD or Parkinsonism: Red
Single nuclei RNA-seq data generation pipeline

- Brain dissections
- Tissue homogenization
- Sample hashing
- Isolation of nuclei by FANS
- Nuclei pooling
- Barcoded beads
- Multiplexed single nuclei RNA-seq

(n=100 individuals)
**snRNA-seq QC: Donors / samples**

<table>
<thead>
<tr>
<th>Sample-centered</th>
<th>Sample count</th>
<th>Note</th>
</tr>
</thead>
<tbody>
<tr>
<td>Received from brain banks</td>
<td>500</td>
<td></td>
</tr>
<tr>
<td>Sent for sequencing</td>
<td>494</td>
<td>6 samples not sent due to poor wet-lab QC metrics</td>
</tr>
<tr>
<td>Passing sample-level QC</td>
<td>444</td>
<td>50 samples removed due to:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Low number of cells</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Low genotype match with expected genotype</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- High genotype match with multiple genotypes (contamination)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Donor-centered</th>
<th>Donor count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Donors with samples from 5 brain regions</td>
<td>79</td>
</tr>
<tr>
<td>4 brain regions</td>
<td>8</td>
</tr>
<tr>
<td>3 brain regions</td>
<td>5</td>
</tr>
<tr>
<td>2 brain regions</td>
<td>3</td>
</tr>
<tr>
<td>1 brain regions</td>
<td>2</td>
</tr>
<tr>
<td>0 brain regions</td>
<td>3</td>
</tr>
</tbody>
</table>

Avg number of cells per removed / kept samples: 1,446 / 5,908
snRNA-seq QC: By brain region

**Number of samples per brain region**

- DMN: 89
- GPI: 89
- PMc: 91
- PFC: 90
- PVC: 84

**Average cell count per brain region**

- DMN: 4607
- GPI: 5334
- PMc: 6248
- PFC: 6761
- PVC: 6613

**Total cell count per brain region**

- DMN: 410108
- GPI: 474762
- PMc: 568609
- PFC: 608563
- PVC: 555510
snRNA-seq QC: Cells

Cell filtering
- UMI: $1,500 \leq n \leq 110,000$
- Gene counts: $1.1k \leq n \leq 12.5k$
- Max fraction of chrM reads: 2%

released data

Doublet detection

Colored by cell type clusters

Colored by brain regions
SN Brain Data | Release 4 Data Summary

- Clinical Data
  - **GCS**: gs://amp-pd-data/releases/2023_v4release_1027/clinical/
  - **BigQuery**: amp-pd-research:2023_v4release_1027

- WGS Data (GCS)
  - **CRAM**: gs://amp-pd-genomics/samples/wgs-BR-DSNWGS/gatk/cram/
  - **VCF**: gs://amp-pd-genomics/samples/wgs-BR-DSNWGS/gatk/vcf/

- RNAseq Data (GCS)
  - **Individual H5AD**:  
    gs://amp-pd-transcriptomics/samples/rnaseq-BR-RSNSR/starsolo/quantification
  - **Merged H5AD**:  
    gs://amp-pd-transcriptomics/releases/2023_v4release_1027/rnaseq-BR-RSNSR/starsolo/quantification

All data are uploaded and ready for release (slide provided by Technome team)
psychAD: powerful dataset to study interindividual variation across disease, NPS, pathology, cognitive status, aging and genetic architecture
Building tools for analysis of population scale single cell/nucleus omics with complex design

**Compositional analysis**
- Linear mixed models
- Precision weights
- Computational scaling

**Differential expression**
- Linear mixed models
- Precision weights
- Pseudobulk aggregation
- Computational scaling

---

Hoffman et al (2023) Efficient differential expression analysis of large-scale single cell transcriptomics data using dreamlet.


---

Center for Disease Neurogenomics
PD vs Controls in PsychAD (33/126 samples)
PD vs Controls in PsychAD (33/126 samples)
Transcriptomics in AMP-PD

David W. Craig, Ph.D.
Feb 5th, 2024
City of Hope

Presenting Work Of

Ivo Violich  Ramiyapriya Sivakumar  Kayla Xu  Kendall Jensen, Ph.D.  Elizabeth Hutchens, Ph.D.  Eric Alsop, Ph.D.

Mark Cookson, Ph.D.  Andy Singleton, Ph.D.  J. Raphaal Gibbs, Ph.D.  Mark Nalls, Ph.D.

Karen Crawford, Ph.D.  Art Toga, Ph.D.
Transcriptomics in AMP-PD

Support the *use* and *analysis* of AMP PD data, with a specific focus on integrative visualization and analysis of RNA transcriptomic data with the existing genomic data.

- Overview
- Resources
- Samples + QC
- Findings
AMP-PD Transcriptomics Cohort: Random Primed Transcriptomics on PaxGene Isolated Whole Blood

WGS and RNA-Seq On > 3,000 Individuals

PPMI

M0  M6  M12  M18  M24  M36

WGS and RNA-Seq On > 3,000 Individuals

PDBP

WGS Indiv

PDBP 1445 D10418 C157

BioFIND RNA

Clinical by Case-Control

Graph Split By: Study  Colored By: Diagnosis

Counts

PPMI  RNA

BioFIND RNA

Color – Diagnosis

- Corticobasal Degeneration
- Dementia With Lewy Bodies
- Essential Tremor
- Fahn’s Syndrome
- Idiopathic PD
- Parkinson’s Disease
- Parkinsonism
- Progressive Supranuclear Palsy
- Progressive Supranuclear Palsy
- Progressive Supranuclear Palsy
AMP-PD Transcriptomics Cohort: Random Primed Transcriptomics on PaxGene Isolated Whole Blood

Data & Methods

Overview

Quality Metrics

Cohort Analysis

Methods Overview

AMP-PD Transcriptomics project Github workflows (e.g. STAR alignment, quantification with Salmon, and quality control assessment).

About AMP PD Transcriptomics Resource

Study information, including finding essential information for analysis, links to downloadable interactive tables, Terra notebooks, Github repositories, and descriptions of methods/analyses.

Github Repository

Overview of library preparation & protocol, workflows, sequencing, and links to data dictionaries and overview of standard pipeline for quality control.

Graph Building

Graphical Sample Explorer

- Sample counts based on visit, study, case/control status, and genetic status
- High QC are samples that were within 2 standard deviations of median usable bases
- Overall 1,749 of 2,356 were defined as High QC
- Factor rows are sample numbers at each visit, where M6 is Month 6, M6 is Month 6, and so forth

Dynamic graphing (Click on Wi panel to change axis)
- Change Panel graph layout
- X-axis, Y-axis, coloring
- Automatic Graphing (Diagrams to tables, scatter, histograms, based on data type, 2D graph types available)

Numbers: Dynamic Table Building

<table>
<thead>
<tr>
<th>QC/QC Level</th>
<th>Enroll:Study</th>
<th>Enroll:Diagnosis</th>
<th>Enroll:Case Control</th>
<th>Case</th>
<th>Control</th>
<th>Other</th>
<th>Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BioFIND</td>
<td></td>
<td></td>
<td></td>
<td>92</td>
<td>92</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PDNP</td>
<td></td>
<td></td>
<td></td>
<td>70</td>
<td>70</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other Neurological Disorder(s)</td>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corticobasal Degeneration</td>
<td></td>
<td></td>
<td></td>
<td>6</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dementia With Lewy Bodies</td>
<td></td>
<td></td>
<td></td>
<td>5</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Essential Tremor</td>
<td></td>
<td></td>
<td></td>
<td>33</td>
<td>33</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multiple System Atrophy</td>
<td></td>
<td></td>
<td></td>
<td>50</td>
<td>50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No PD Nor Other Neurological Disorder</td>
<td></td>
<td></td>
<td></td>
<td>509</td>
<td>509</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Olivopontocerebellar Atrophy</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parkinson's Disease</td>
<td></td>
<td></td>
<td></td>
<td>1.846</td>
<td>1.846</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Essential Tremor</td>
<td></td>
<td></td>
<td></td>
<td>8</td>
<td>8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Idiopathic PD</td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Progressive Supranuclear Palsy</td>
<td></td>
<td></td>
<td></td>
<td>46</td>
<td>46</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parkinson's Disease</td>
<td></td>
<td></td>
<td></td>
<td>209</td>
<td>209</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PDNP</td>
<td></td>
<td></td>
<td></td>
<td>2,363</td>
<td>2,363</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other Neurological Disorder(s)</td>
<td></td>
<td></td>
<td></td>
<td>31</td>
<td>31</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prodromal motor PD</td>
<td></td>
<td></td>
<td></td>
<td>11</td>
<td>11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prodromal non-motor PD</td>
<td></td>
<td></td>
<td></td>
<td>206</td>
<td>206</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BioFIND</td>
<td></td>
<td></td>
<td></td>
<td>22</td>
<td>22</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Idiopathic PD</td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other Neurological Disorder(s)</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corticobasal Degeneration</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Essential Tremor</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multiple System Atrophy</td>
<td></td>
<td></td>
<td></td>
<td>16</td>
<td>16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No PD Nor Other Neurological Disorder</td>
<td></td>
<td></td>
<td></td>
<td>203</td>
<td>203</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parkinson's Disease</td>
<td></td>
<td></td>
<td></td>
<td>33</td>
<td>33</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Essential Tremor</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parkinson's Disease</td>
<td></td>
<td></td>
<td></td>
<td>22</td>
<td>22</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Idiopathic PD</td>
<td></td>
<td></td>
<td></td>
<td>326</td>
<td>326</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PDNP</td>
<td></td>
<td></td>
<td></td>
<td>189</td>
<td>189</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other Neurological Disorder(s)</td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prodromal motor PD</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prodromal non-motor PD</td>
<td></td>
<td></td>
<td></td>
<td>26</td>
<td>26</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Totals</td>
<td></td>
<td></td>
<td></td>
<td>5,042</td>
<td>2,836</td>
<td>478</td>
<td>8,356</td>
</tr>
</tbody>
</table>
**Workflow & QC**

### Initial Workflow

- Participants enrolled from 33 clinic sites.
- PA.xgene whole blood
  - Baseline and months 6, 12, 24 and 36
- RNA isolation + storage
  - In lysis buffer, 
- Transcripts library preparation and sequencing
  - FASTQs

### PCA

**PCA by PC1 vs. PC2**

- Graph Split By: Sex
- Colored By: Study Arm

**PC Correlates to Laboratory, Clinical, and Biological Variables**

- R-squared Correlation by Variable vs. PC
  - Colored By: Val

### Variance Analysis by Covariates

- AVS metric analysis
  - 1. ANOVA (analysis 1-15)
  - 2. PCA (assortative set points)
- Disease or genetic classification
- Differential expression analysis
AMP-PD Transcriptomics Cohort: Random Primed Transcriptomics on PaxGene Isolated Whole Blood

**Complex Dynamic Graphing**
- Fraction Variance Explained in Gene RNA
- Variable: Gene
- Color By Variance: 0.005, 0.025, 0.045
- Vet by Neutrophil % (RNA) vs. RNA Counts
- Color By Status: Case, Control

**Numbers: Dynamic Table Building**
- Genomic Browser: chr4:89,697,537 - 89,849,737 (152kbp)
- Binned Coverage
- Binned Junctions
- Whole Blood eQTLs
- Whole Blood splice QTLs

**Locus Plots**
Findings: Neutrophils are enriched in PD

**PPMI Diff. Expression**

**PDBP Diff. Expression**

**Lab Cell Counts**

**Lab Cell Counts By Different PD Types**

**DE Prodromal / Idiopathic PD**
Findings: SNCA is differentially expressed when accounting for estimated Neutrophils

- **Multiple Models Tested**
  - LM model w/ 18 Blood Atlas Informed genes
**Differential Splicing**

- Difficulty Observing Psi due to pre-spliced mRNA
- *KANSL1* shows some evidence for differential splicing

![Differential Splicing Diagram](image_url)

**Genetrack**
- Intron, exons clearly evident

**Coverage Track**
- Showing read depth, Mismatches

**Split in alignments inform intron splicing**
- Blue is an inferred junction

**Pre-spliced mRNA**
- Intronic reads
Interactive Resource

AMP-PD Data (Secure)

Cloud Identity-Aware Proxy

AMP Tier 1
AMP Tier 2

GROUP MANAGEMENT

GROUPS

Create a New Group

GROUP NAME

Group Email
Role

CP2_Tier1
CP2_Tier1@foundinpd.org
Member

amp-pd-clinical-access
amp-pd-clinical-access@foundinpd.org
Member

amp-pd-casper-ma
amp-pd-casper-ma@foundinpd.org
Member

amp-pd-wg-ma-qc
amp-pd-wg-ma-qc@foundinpd.org
Member

FOUNDIN-PD
https://foundinpd.org
- **Interactive Resource For Exploring Transcriptomics Analysis**
  - Methods/Resources/Protocols
  - Data Explorers
  - Gene Explorers

- **Neutrophils are enriched in PD**
  - Proportion neutrophils increases over time in prodromal and with age

- **SNCA is differentially expressed controlling for neutrophils**

**Acknowledgements**

Ivo Violich  Ramiyapriya Sivakumar  Kayla Xu  Kendall Jensen, Ph.D.  Elizabeth Hutchens, Ph.D.  Eric Alsop, Ph.D.

Mark Cookson, Ph.D.
Andy Singleton, Ph.D.
J. Rapheal Gibbs, Ph.D.
Mark Nalls, Ph.D.

Karen Crawford, Ph.D.
Art Toga, Ph.D.
Targeted proteomics for biomarker discovery in Parkinson’s Disease

An AbbVie : AMP-PD collaboration

Aparna Vasanthakumar
Ameya Kulkarni

February 6, 2024
Study contributions & acknowledgments

Abbvie
Jan Stoehr
Jeff Waring
Maria Quinton
Alvis Hu
Elina Regan
Marieta Hyde
Tifani Anton
Julian Sefrin
Roland Heym
Jennifer Mollon
Lynne Reuter

FNIH
Eline Appelman
Stacey Adam
Sri Pullagura

Technome
Victoria Dardov
Barry Landin

University Medicine Gottingen
Brit Mollenhauer
Mohammad Dakna
The goal of Olink analyses is to identify biomarkers for clinical use

- Proteomics (OLINK)
  Discover potential novel biomarkers
- Develop specific assays for biomarkers discovered by OLINK
- Develop and validate clinical assays for novel biomarkers
- Incorporate novel biomarker in clinical trials
Olink Explore Workflow: Proximity Extension Assay with Next Gen Seq readout

Panels available
- **Olink Explore 1536** (4x 384-plex kits)
- **Olink Explore 3072** (8 x 384-plex kits)
- **Olink Explore HT** (5343 proteins)

**Immunoassay:** Pair of oligonucleotide-labeled antibodies bind to target protein

**Extension:** In close proximity, probes hybridize with each other and are pre-amplified.

**Amplification and Reading:** Newly created dsDNA further amplified by PCR and read out by NGS or qPCR.

Results in 36 hours; Hands-on time: <5 hours
Proposal for Olink Analysis at AbbVie

• Main objectives:
  • Interrogate a longitudinal sample set which will include matched plasma and CSF to complement earlier efforts. Samples will come from:
    1. Prodromal PD
    2. PD Dx<0.16 yr
    3. Healthy control

Additional **bridging samples** will be included to account for any potential batches and to compare to previously analyzed data: 16-20 CSF, 16-20 plasma from previous PPMI runs
Increasing the Sample Number Increases the Power of Detecting Significant Differences between the Groups

- Power calculation recommends n=70/group to detect a significant difference at power of 0.94.
- The current proposal would put the number of unique individuals per group at ~130

<table>
<thead>
<tr>
<th>Cohort</th>
<th>numbers (total)</th>
<th>TOTAL Plasma</th>
<th>TOTAL CSF</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PPMI prodromal</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>non-converter</td>
<td>42</td>
<td>122</td>
<td>78</td>
</tr>
<tr>
<td>converter</td>
<td>25</td>
<td>258</td>
<td>90</td>
</tr>
<tr>
<td><strong>PPMI PD Dx&lt;0.16 years</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>non-converter</td>
<td>44</td>
<td>131</td>
<td>111</td>
</tr>
<tr>
<td>converter</td>
<td>22</td>
<td>77</td>
<td>9</td>
</tr>
<tr>
<td><strong>PARS</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>non-converter</td>
<td>56</td>
<td>167</td>
<td>63</td>
</tr>
<tr>
<td>converter</td>
<td>22</td>
<td>77</td>
<td>9</td>
</tr>
<tr>
<td><strong>DeNoPa</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>non-converter</td>
<td>45</td>
<td>63</td>
<td>27</td>
</tr>
<tr>
<td>converter</td>
<td>15</td>
<td>50</td>
<td>10</td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Information provided by FNIH as part of AMP-PD proposal
Research Questions

1. Differential Protein Abundance between PD & HC- at each time point and longitudinal
2. Proteomic changes with PD progression- defined as a function of MDS-UPDRS and time
3. Matched Plasma vs CSF comparisons
4. Prodromal cohort- converters vs non-converters- in collaboration with Brit Mollenhauer (DeNoPa)
5. Predictive capability of baseline protein values for progressive changes in MDS-UPDRS
6. Targeted comparison of protein abundance between carriers and non-carriers for PD-relevant genes
7. Validation of target differences in conditioned medium from iPSC-derived dopaminergic model

Datasets: Olink 1536 Explore Panel in matched plasma and CSF from AMP-PD (PPMI/PDBP) cohort
Bridging Normalization

- Recommended for integration and comparison between two separate projects
- 39 samples in CSF and plasma are common between the preview and AbbVie datasets
- Bridging with subset normalization used to calculate median of the paired differences between the bridge samples
- This adjustment factor used to add to the NPX values of remaining samples
Principal Components Analysis (CSF shown as an example)

- All models adjusted for age at baseline, sex and project
- Progression models are adjusted for sex and project
Proteomic differences between Parkinson’s Disease patients and Healthy Controls

Robust regression (limma); Covariates: Age at baseline, Sex, Dataset

<table>
<thead>
<tr>
<th>Time</th>
<th>N individuals per group</th>
<th>Number of DEP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PD</td>
<td>HC</td>
</tr>
<tr>
<td>Baseline</td>
<td>139</td>
<td>170</td>
</tr>
<tr>
<td>Month 6</td>
<td>40</td>
<td>50</td>
</tr>
<tr>
<td>Month 12</td>
<td>83</td>
<td>80</td>
</tr>
<tr>
<td>Month 18</td>
<td>14</td>
<td>15</td>
</tr>
<tr>
<td>Month 24</td>
<td>120</td>
<td>125</td>
</tr>
<tr>
<td>Month 36</td>
<td>87</td>
<td>81</td>
</tr>
<tr>
<td>Month 48</td>
<td>58</td>
<td>101</td>
</tr>
<tr>
<td>Month 60</td>
<td>17</td>
<td>31</td>
</tr>
</tbody>
</table>

Plasma PD vs Ctrl (Year 4)
109 proteins (105 ↑; 4 ↓)

CSF PD vs Ctrl (Year 2)
28 proteins (3 ↑; 25 ↓)
Proteomic changes with PD progression (MDS-UPDRS)

Plasma- PD progression (MDS-UPDRS)
836 proteins (807 ↑; 29 ↓)

CSF- PD progression (MDS-UPDRS)
199 proteins (191 ↑; 8 ↓)

Linear Mixed Model (limma- duplicatecorrelation)
Fixed effects: Sex and dataset; Random effect: Subject ID
Tissue-specific enrichment of PD progression associated genes

19 brain-specific proteins enriched in the plasma (GTEx - Tissue Enrichment)

9 PD-progression plasma proteins enriched in the brain (GTEx)

\[ \rho = 0.33; p = 2e-5 \]
Drug-induced peripheral changes vs disease-biomarkers of PD

• DDC (Dopa-decarboxylase) or AADC (Aromatic L-amino acid decarboxylase) identified as the most significant differentially expressed protein in plasma at all time points
  o DDC is a metabolite responsible for synthesis of dopamine and serotonin via decarboxylation of L-dopa
  o PD patients are treated with peripheral inhibitors of DDC to prevent premature peripheral conversion of L-dopa to dopamine before reaching the CNS

• Other consistent DEP include
  o HPGDS (hematopoietic prostaglandin D synthase) in plasma
  o LPO (Lactoperoxidase) in CSF could be a direct effect of L-dopa and carbidopa

• For progression, MDS-UPDRS III measurements from drug naïve patients were used wherever available.

Rutledge et al, 2022 medRxiv
Protein-RNA comparisons for PD progression associated genes

Plasma (Total = 50; 26 genes in Q1)

CSF (Total = 10; 6 genes in Q1)
Next steps- validation, integration and expansion to the prodromal cohort

**Validation**
- Validation of the top differentially abundant proteins in independent datasets
- Comparisons with the secreted proteome of iPSC-derived dopaminergic neurons

**Prodromal Cohort**
- Additional samples analyzed for prodromal motor and prodromal non-motor patients, some of whom convert to PD
- Integration with DeNoPa cohort (Brit Mollenhauer)

**Additional Analysis**
- Integration with WGS and RNA-Seq available on the same set of individuals
- Building a classifier (logistic regression/RF) between diagnoses
- Expanding to other relevant clinical metrics (components of MDS-UPDRS; MMSE; H&Y)
- Comparing effect sizes in drug naïve vs treated patients
abbvie
Prior Omics Work in AMP-PD

- >10,500 unique participants
- Eight unified cohorts: BioFIND, HBS, LBD, LCC, PDBP, PPMI, STEADY-PD3, SURE-PD3
CSF Data QC- AbbVie D02

Panel | QC Warning | Assay Warning
--- | --- | ---
Cardiometabolic | 252596 | 5704 | 257600 | 700
Inflammation | 252257 | 5343 | 256900 | 700
Neurology | 248412 | 8488 | 256900 | 0
Oncology | 248042 | 9558 | 252700 | 4900
Plasma Data QC

More variability in plasma NPX values than CSF (notice the y-axis)

<table>
<thead>
<tr>
<th>Panel</th>
<th>QC Warning</th>
<th>Assay Warning</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pass</td>
<td>Warn</td>
</tr>
<tr>
<td>Cardiometabolic</td>
<td>338413</td>
<td>9185</td>
</tr>
<tr>
<td>Inflammation</td>
<td>338411</td>
<td>8245</td>
</tr>
<tr>
<td>Neurology</td>
<td>330793</td>
<td>14921</td>
</tr>
<tr>
<td>Oncology</td>
<td>324208</td>
<td>20564</td>
</tr>
</tbody>
</table>
Low abundance proteins when NPX < LOD (All proteins are retained)

CSF (576 proteins with >50% missing frequency)

Plasma (77 proteins with >50% missing frequency)
# AMP-PD CSF Samples Results Summary

## 1. Project information

<table>
<thead>
<tr>
<th>No. of samples</th>
<th>No. of plates</th>
<th>Normalization method</th>
</tr>
</thead>
<tbody>
<tr>
<td>684</td>
<td>8</td>
<td>Intensity normalization</td>
</tr>
</tbody>
</table>

## 2.1 QC summary

<table>
<thead>
<tr>
<th>Olink Panel</th>
<th>Samples passed QC</th>
<th>Samples passed QC (%)</th>
<th>Datapoints passed QC</th>
<th>Datapoints passed QC (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Explore 384 Cardiometabolic</td>
<td>626 / 684</td>
<td>92</td>
<td>246692 / 252396</td>
<td>98</td>
</tr>
<tr>
<td>Explore 384 Inflammation</td>
<td>629 / 684</td>
<td>92</td>
<td>246369 / 251712</td>
<td>98</td>
</tr>
<tr>
<td>Explore 384 Neurology</td>
<td>605 / 684</td>
<td>88</td>
<td>242540 / 251028</td>
<td>97</td>
</tr>
<tr>
<td>Explore 384 Oncology</td>
<td>598 / 684</td>
<td>87</td>
<td>242154 / 251712</td>
<td>96</td>
</tr>
</tbody>
</table>

## 2.2.1 Average %CV

<table>
<thead>
<tr>
<th>Olink Panel</th>
<th>Intra-assay %CV</th>
<th>Inter-assay %CV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Explore 384 Cardiometabolic</td>
<td>8</td>
<td>24</td>
</tr>
<tr>
<td>Explore 384 Inflammation</td>
<td>9</td>
<td>24</td>
</tr>
<tr>
<td>Explore 384 Neurology</td>
<td>10</td>
<td>22</td>
</tr>
<tr>
<td>Explore 384 Oncology</td>
<td>11</td>
<td>21</td>
</tr>
</tbody>
</table>
CSF Samples Cardiometabolic Z-score
# AMP-PD Plasma Samples Results Summary

## 1. Project information

<table>
<thead>
<tr>
<th>No. of samples</th>
<th>No. of plates</th>
<th>Normalization method</th>
</tr>
</thead>
<tbody>
<tr>
<td>920</td>
<td>11</td>
<td>Intensity normalization</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Olink Panel</th>
<th>Samples passed QC</th>
<th>Samples passed QC (%)</th>
<th>Datapoints passed QC</th>
<th>Datapoints passed QC (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Explore 384 Cardiometabolic</td>
<td>830 / 920</td>
<td>90</td>
<td>330295 / 339480</td>
<td>97</td>
</tr>
<tr>
<td>Explore 384 Inflammation</td>
<td>838 / 920</td>
<td>91</td>
<td>330315 / 338560</td>
<td>98</td>
</tr>
<tr>
<td>Explore 384 Neurology</td>
<td>807 / 920</td>
<td>88</td>
<td>322719 / 337640</td>
<td>96</td>
</tr>
<tr>
<td>Explore 384 Oncology</td>
<td>774 / 920</td>
<td>84</td>
<td>316156 / 336720</td>
<td>94</td>
</tr>
</tbody>
</table>

## 2.2.1 Average %CV

<table>
<thead>
<tr>
<th>Olink Panel</th>
<th>Intra-assay %CV</th>
<th>Inter-assay %CV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Explore 384 Cardiometabolic</td>
<td>9</td>
<td>18</td>
</tr>
<tr>
<td>Explore 384 Inflammation</td>
<td>8</td>
<td>19</td>
</tr>
<tr>
<td>Explore 384 Neurology</td>
<td>10</td>
<td>22</td>
</tr>
<tr>
<td>Explore 384 Oncology</td>
<td>9</td>
<td>20</td>
</tr>
</tbody>
</table>
Plasma Samples Cardiometabolic Z-score
Exploring Genetic Links to Parkinson's Disease in African and African Admixed Populations: A Genome-Wide Analysis

Mary B Makarious
February 5th 2024
AMP-PD: FNIH AMP Symposium
About the
Global Parkinson’s Genetics Program
GP2 is generating data from studies across the world

Data from over 24,000 samples now available with 40,000 more in the pipeline
Nearly 60 countries represented, and over 160,000 samples promised

1. Go to: https://amp-pd.org/
2. Sign DUA
3. Verify email address
4. Wait ~2 weeks for GP2 access
5. You now have access to AMP-PD and GP2 data at https://terra.bio/
Cohort Enrollment

- Compliance (consent form review)
- GP2 agreements (includes MTA's, data-sharing agreements, publication policy and more)
- Preparing of sample manifests, DNA and clinical data for GP2 processing

GP2 Processing

- Clinical data review and harmonization
- DNA processing and generating genotyping, WGS or long-read WGS
- Data processing and harmonization
- Providing data back to cohort provider

Data Sharing

- GP2 summary statistics
- GP2 data browsers
- Data sharing via AMP-PD (single sign-on)
- Online training
Study Cohorts


Genetic Studies and Findings

Leveraging the African and African Admixed Populations
Ancestral diversity in PD genetics has been limited thus far.

There is both a social and scientific imperative to understanding the genetic basis of PD across global populations.

Examination of previously understudied ancestral groups, even with a modest sample, could lead to novel and striking findings.

This has implications for therapeutic development and deployment.

The Nigerian Parkinson Disease Research Network, BLAAC PD, NIH & 23andMe teamed up to conduct this collaborative study

Team
- 24 Nigerian institutions
- 4 BLAAC PD sites
- Other studies recruited by GP2
- 78 authors (46 from African institutions)

Training to URP trainees was provided virtually and in-person

Data Overview

African cohorts
- **GP2 AFR:** 693 cases and 1,009 controls
- **IPDGC Africa Neurochip:** 326 cases and 306 controls
- **Total:** 1,019 cases and 1,315 controls

African admixed cohorts
- **GP2 AAC:** 179 cases and 1,135 controls
- **23andMe:** 288 cases and 193,985 controls
- **Total:** 467 cases and 195,093 controls

**Total:** 1,486 cases and 196,408 controls

---

**Identification of genetic risk loci and causal insights associated with Parkinson's disease in African and African admixed populations: a genome-wide association study**

Mie Ritzig, PhD ♦ | Sara Bandres-Ciga, PhD ♦ | Mary B Makarious, BSc ♦ | Oludamilola Omotola Ojo, MD ♦ | Peter Wilder, BSc | Oladimi Victoria Abiodun, FWACP | Kristin S Levine, MS | Prof Sani Atta Abubakar, MBBS ♦ | Charles Okoro Acholu, MBBS | Dan Vitala, MS | Olaleye Akinmoludan Adeniji, MBBS | Osigwe Paul Agabi, MBBS ♦ | Matthew J Koretsky, BSc ♦ | Uchechi Agajan, MBBS ♦ | Deborah A Hall, MD, PhD ♦ | Prof Rufus Oyolua Akiyemi, PhD ♦ | Tao Xie, MD, PhD ♦ | Mohammed Wulgo Ali, MBBS | Ejaaz A Shamim, MD | Ifeanyi Amadi-Okwu, FACP ♦ | Mahesh Padmanaban, MD ♦ | Oyewotu Michael Ariigbodi, MBBS ♦ | David G Standaert, MD, PhD ♦ | Abiodun Hamad, FWACP ♦ | Marissa N Dean, MD ♦ | Cyriel Oshobami Ereneh, MBBS ♦ | Inas Elsayed, PhD ♦ | Temitope Hannah Faramobi, MBBS ♦ | Olaitan Okunuyo, PhD | Michael Bimbole Fawole, MSc ♦ | Kimberley J Billing, PhD ♦ | Prof Frank Akwashabe Mhamid, MBCHB ♦ | Pilar Alvarez Jerez, BSc ♦ | Emmanuel Uzodimma Iwuzo, FACP ♦ | Breanea Baker, BSc ♦ | Prof Morenikeji Adeyoyin Kosoko, MBBS ♦ | Laksh Malik, MFS ♦ | Paul Osemeke Nwari, MBBS ♦ | Kemi Okafor Dada, MD ♦ | Ernest Owundu Nwosu, FACP ♦ | Abigail Miano-Burkhardt, BSc ♦ | Prof Yakub Willberforce Nyamadzwi, MBBS ♦ | Zih-Hua Fang, PhD ♦ | Prof Tahanaya Olubogbo Obiabo, MBCHB ♦ | Jillian H Kluss, PhD ♦ | Olaniye Adedoyin Odeniyi, MBBS ♦ | Dena G Hernandez, PhD ♦ | Francis Ehiobuon Odafe, MBBS ♦ | Nahid Tayebi, PhD ♦ | Prof Francis Ibi Ojini, MSc ♦ | Ellen Sidransky, MD ♦ | Gerald Awele Onweegbuze, MBBS ♦ | Andrea M D’Souza, BSc ♦ | Godwin Osasugu Osagih, MBBS ♦ | Bahaafa Berho, BSc ♦ | Nosakhare Osemwegie, MBBS ♦ | Xylena Reed, PhD ♦ | Prof Ojopiti Oluwose Oshinaike, FWACP ♦ | Hampton L Leonard, MS ♦ | Folajimi Morenikeji Ohubogun, MBCHB ♦ | Chelsea X Alvarez, MS ♦ | Shyngele Imewar Okhakhire, MBBS ♦ | Simon Isuchukwu Ozoemena, FACP ♦ | Sarah Chiabiri Samuel, MBBS ♦ | Funmilola Oluwole Taiwo, MBCHB ♦ | Prof Kolawole Wasu Wahab, MD ♦ | Yusuf Agboola Zafar, MSc ♦ | Hiroaka Iwaki, MD ♦ | Jonggeol Jeffrey Kim, BA ♦ | Prof Huw R Morris, PhD FRCP ♦ | Prof John Hardy, PhD ♦ | Mike A Nalls, PhD ♦ | Karl Hellborg, PhD ♦ | Lucy Norcliffe-Kaufmann, PhD ♦ | Nigeria Parkinson Disease Research Network ♦ | International Parkinson’s Disease Genomics Consortium Africa ♦ | Black and African American Connections to Parkinson’s Disease Study Group ♦ | the 23andMe Research Team ♦ | Cornelis Blauwendraad, PhD ♦ | Prof Henry Houdens, MD ♦ | Andrew Singleton, PhD ♦ | Prof Nijede Okunle Okudahoro, MD ♦ on behalf of the Global Parkinson’s Genetics Program ♦

Show less • Show footnotes

---

https://github.com/GP2code/GP2-AFR-AAC-metaGWAS
Workflow

1. Data Collection
   - Recruitment
   - Genotyping
   - Imputation and Quality Control
   - Ancestry Predictions

2. Genome-Wide Association Studies
   - GWAS per Independent Cohort
   - African ONLY Meta-GWAS
   - African Admixed ONLY Meta-GWAS

3. Follow-up Analyses
   - ADMIXTURE and Haplotype Analysis
   - Runs of Homozygosity
   - Polygenic Risk Scores
   - Short- and Long-Read Sequencing
   - Glucocerebrosidase Activity
   - eQTL Analysis

Additional Information:
- African: 1,015 cases, 1,296 controls
- African-Admixed: 473 cases, 195,134 controls
African and African admixed GWAS Meta-Analysis

- **rs3115534**, intronic *GBA1*
- \( P = 2.397 \times 10^{-14} \)
  - \( \text{OR} = 1.57 \)
  - \( \text{BETA} = 0.50 \)
  - \( \text{SE} = 0.07 \)
- \( N = 1,488 \) cases, \( 196,430 \) controls
- \( \lambda = 1.01 \)
- \( I^2 = 36.5 \)

**Datasets:**
- GP2 African + GP2 African admixed
- IPDGC Africa Neurochip + 23andMe African admixed
Meta Analysis - 90 European Risk Loci

Interpretation

- Color → Direction (blue = +positive; orange = -negative)
- Shape → Significance; p<0.05 within ancestry (O = yes; X = no)
- Size → Absolute beta effect (small = small effect; large = large effect)
- Missing → Not found in that ancestry’s summary statistics
**Meta Analysis - Results**

### GP2 African data

<table>
<thead>
<tr>
<th>CHR</th>
<th>SNP</th>
<th>A1</th>
<th>A2</th>
<th>MAF</th>
<th>MAF_A</th>
<th>MAF_U</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>rs3115534</td>
<td>G</td>
<td>T</td>
<td>0.2552</td>
<td>0.3362</td>
<td>0.1988</td>
</tr>
</tbody>
</table>

**CASES CONTROLS Times**

- **HMZ (G/G)**
  - CASES: 130/952 (13.7%)
  - CONTROLS: 49/1230 (4.0%)
  - Times: 3.42

- **HET**
  - CASES: 398/952 (41.8%)
  - CONTROLS: 435/1230 (35.4%)
  - Times: 1.18

### GP2 African admixed data

<table>
<thead>
<tr>
<th>CHR</th>
<th>SNP</th>
<th>A1</th>
<th>A2</th>
<th>MAF</th>
<th>MAF_A</th>
<th>MAF_U</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>rs3115534</td>
<td>G</td>
<td>T</td>
<td>0.1485</td>
<td>0.2268</td>
<td>0.136</td>
</tr>
</tbody>
</table>

**CASES CONTROLS Times**

- **HMZ (G/G)**
  - CASES: 11/183 (6%)
  - CONTROLS: 18/1140 (1.6%)
  - Times: 3.81

- **HET**
  - CASES: 61/183 (33%)
  - CONTROLS: 274/1140 (24.0%)
  - Times: 1.39

### 23andMe African admixed

**CASES CONTROLS Times**

- **HMZ (G/G)**
  - CASES: 10/288 (3.47%)
  - CONTROLS: 3537/193985 (1.82%)
  - Times: 1.92

- **HET**
  - CASES: 85/288 (29.5%)
  - CONTROLS: 44967/193985 (23.2%)
  - Times: 1.27
rs3115534 Analyses

- **Effect size: rs3115534 as a risk allele**
  - In African ancestry individuals
    - ~13% of cases are homozygous → ~3.5x as likely to get PD
    - ~40% of cases are carriers → ~1.2x as likely to get PD

- **Age-at-onset: rs3115534 vs. AAO**
  - n=793 cases
  - P = 5.29E-4, beta = -2.004, SE = 0.57
  - ~2-3 years earlier disease development per risk allele

- **Linear Regressions vs. % Genomic Admixture**
  - rs3115534-G positively correlated with % African ancestry → suggesting African founder effect
  - P = 0.01, beta = 0.064, SE = 0.024
What is the functional effect underlying this signal?

- WGS did not identify any coding variant
- LRS did not identify any structural variant
- Fine-mapping prioritized rs3115534 based on PP (71.4%)
- Is the GWAS signal a Quantitative Trait Locus?
  - eQTL, mQTL, pQTL, sQTL, caQTL…?
GBA1 - rs3115534 is an expression QTL in Blood

Gene expression in African Americans, Puerto Ricans and Mexican Americans reveals ancestry-specific patterns of genetic architecture

Linda Kachuri, Angel C. Y. Mak, Donglei Hu, Celeste Eng, Scott Huntsman, Jennifer R. Elhawary, Namrata Gupta, Stacey Gabriel, Shujie Xiao, Kevin L. Keys, Akinwumi Oni-Orisan, José R. Rodríguez-Santana, Michael A. LeNoir, Luisa N. Borrell, Noah A. Zaitlen, L. Keoki Williams, Christopher R. Gignoux, Esteban González Burchard & Elad Ziv

Nature Genetics 55, 952–963 (2023) | Cite this article

Gene ID: ENSG00000177628.16

<table>
<thead>
<tr>
<th>rsID</th>
<th>Samples</th>
<th>maf</th>
<th>pval</th>
<th>slope</th>
<th>slope_se</th>
<th>qval</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs3115534</td>
<td>209</td>
<td>0.152576</td>
<td>9.99E-25</td>
<td>-0.238011</td>
<td>0.0223042</td>
<td>1.42E-19</td>
</tr>
</tbody>
</table>

LocusZoom comparing African American expression QTL data versus GP2 AFR-AAC GWAS summary statistics
• Collaborative efforts have made the recruitment of diverse samples possible
  • Accessible through a joint agreement with AMP-PD on Terra

• ~1500 PD cases of AFR/AAC ancestry → ancestry-specific SNP via a GWAS meta-analysis
  • Not previously identified or in LD with previously known GBA1 variants associated with PD risk
  • Higher population attributable risk compared to known GBA1 variants

• rs3115534 is the first intronic variant in GBA1 that has been identified by a GWAS
  • Highly correlated to % African ancestry → founder effect
  • Intronic variant → modification via gene expression
  • Gene expression → RNA-based therapeutics?