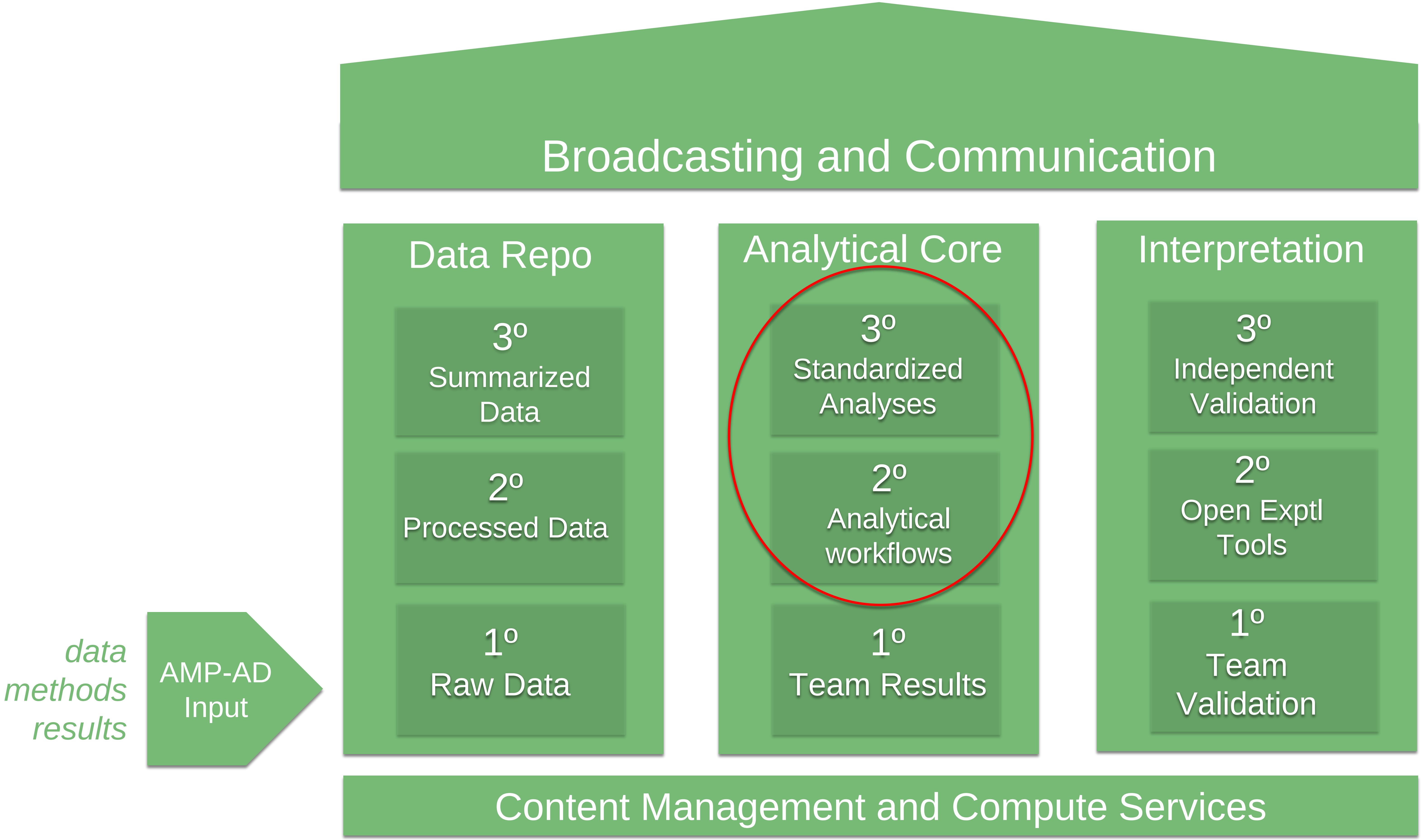
Identification of Disease Insights Through Consortium-Wide Meta-Analyses

Ben Logsdon, PhD
Director of Neurodegenerative Research, Sage Bionetworks
on behalf of the AMP-AD analytic working groups

SageBionetworks

powered by Synapse



Broadcasting and Communication

Data Repo

3^o
Summarized Data

2^o
Processed Data

1^o
Raw Data

Analytical Core

3^o
Standardized Analyses

2^o
Analytical workflows

1^o
Team Results

Interpretation

3^o
Independent Validation

2^o
Open Exptl Tools

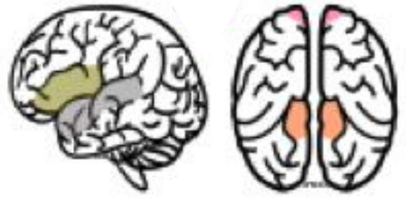
1^o
Team Validation

data
methods
results

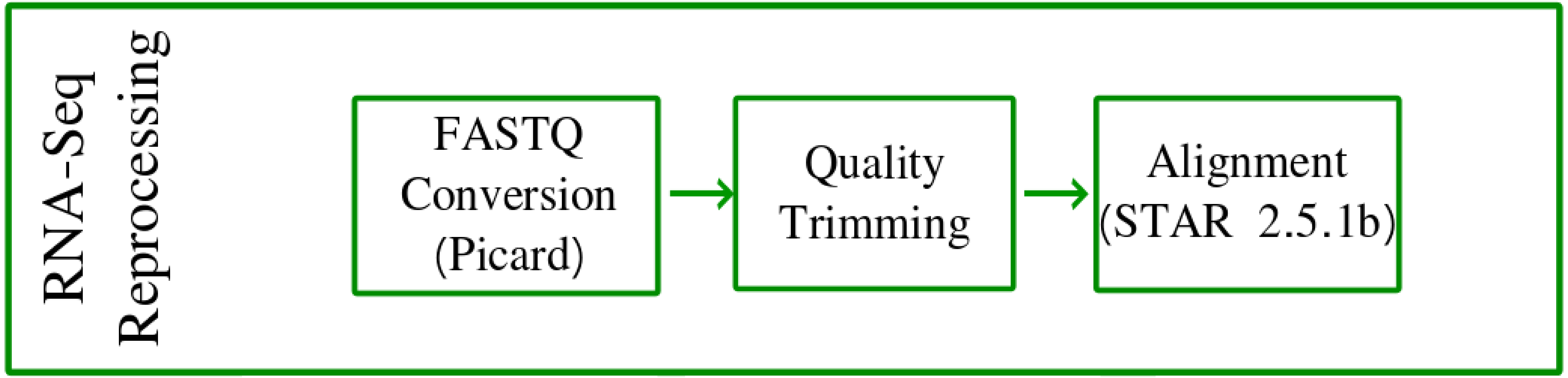
AMP-AD Input

Content Management and Compute Services

RNA Sequencing

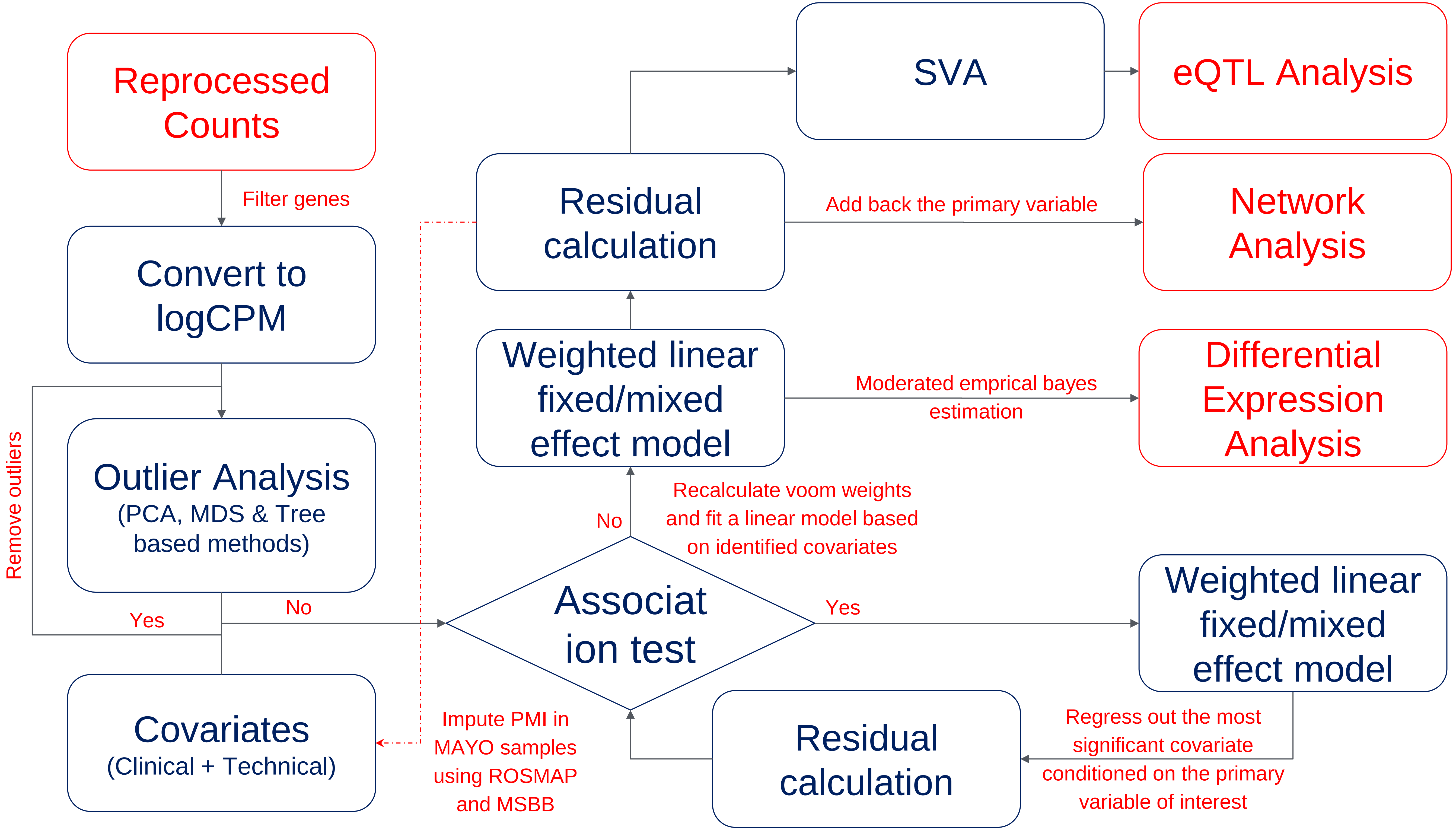
Source	**Brain Regions**	**Library Preparation**	**Read Length**	**Depth of Sequencing**
ROSMAP (n=700)	 Dorsolateral prefrontal cortex	Poly-A Selection	101 bp PE	50-100M per library
MSBB (n=300)	 Frontal Pole, Superiortemporal Gyrus, Parahipocampal Gyrus, Inferiorfrontal Gyrus	RiboZero	100 bp SE	120M per library
MAYO RNASEQ (n=270)	 Cerebellum Temporal Cortex	Poly-A Selection	100 bp PE	120M per library

RNAseq Processing



Gene level counts generated using the HTSeq algorithm

Covariate Analysis Workflow



Reprocessed Counts

Filter genes

Convert to logCPM

Outlier Analysis
(PCA, MDS & Tree based methods)

Remove outliers

Yes

No

Covariates
(Clinical + Technical)

Impute PMI in MAYO samples using ROSMAP and MSBB

Association test

No

Recalculate voom weights and fit a linear model based on identified covariates

Weighted linear fixed/mixed effect model

Moderated emprical bayes estimation

Differential Expression Analysis

Residual calculation

Add back the primary variable

Network Analysis

SVA

eQTL Analysis

Yes

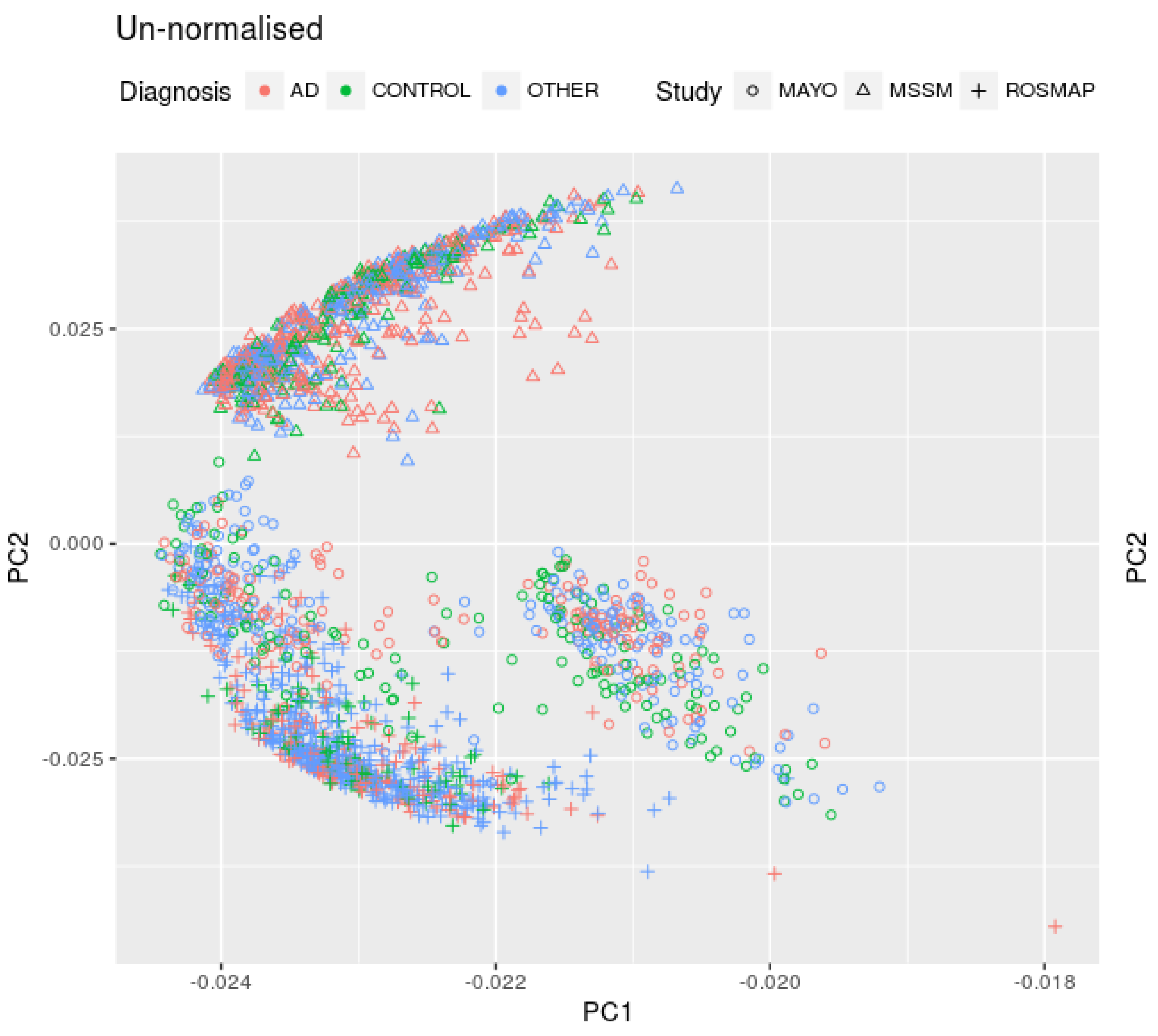
Weighted linear fixed/mixed effect model

Regress out the most significant covariate conditioned on the primary variable of interest

Residual calculation

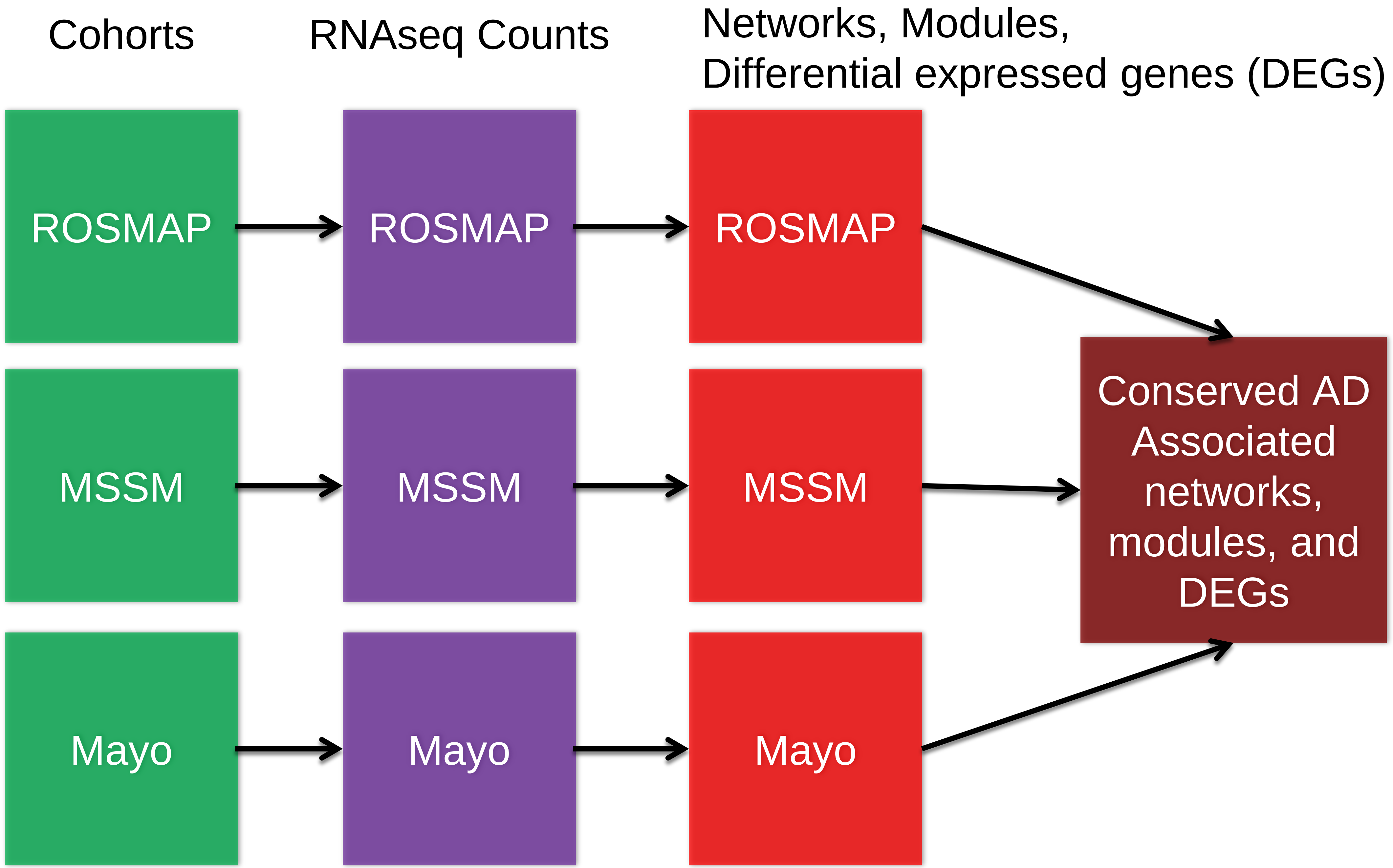
Thanneer Perumal, Sage Bionetworks

Making genomic data useful *across studies* requires *a lot* of *collaborative* work



Thanneer Perumal (Sage), Mingui Wang (MSSM), Cory Funk (ISB), Mariet Allen (Mayo), Kristen Dang (Sage), Chris Gaiteri (Rush), Phil Ebert (Eli Lilly), Yuhao Lin (Eli Lilly), Jishu Xu (Broad)

Identifying robust AD transcriptomic signatures



Cohorts

RNAseq Counts

Networks, Modules,
Differential expressed genes (DEGs)

ROSMAP → ROSMAP → ROSMAP

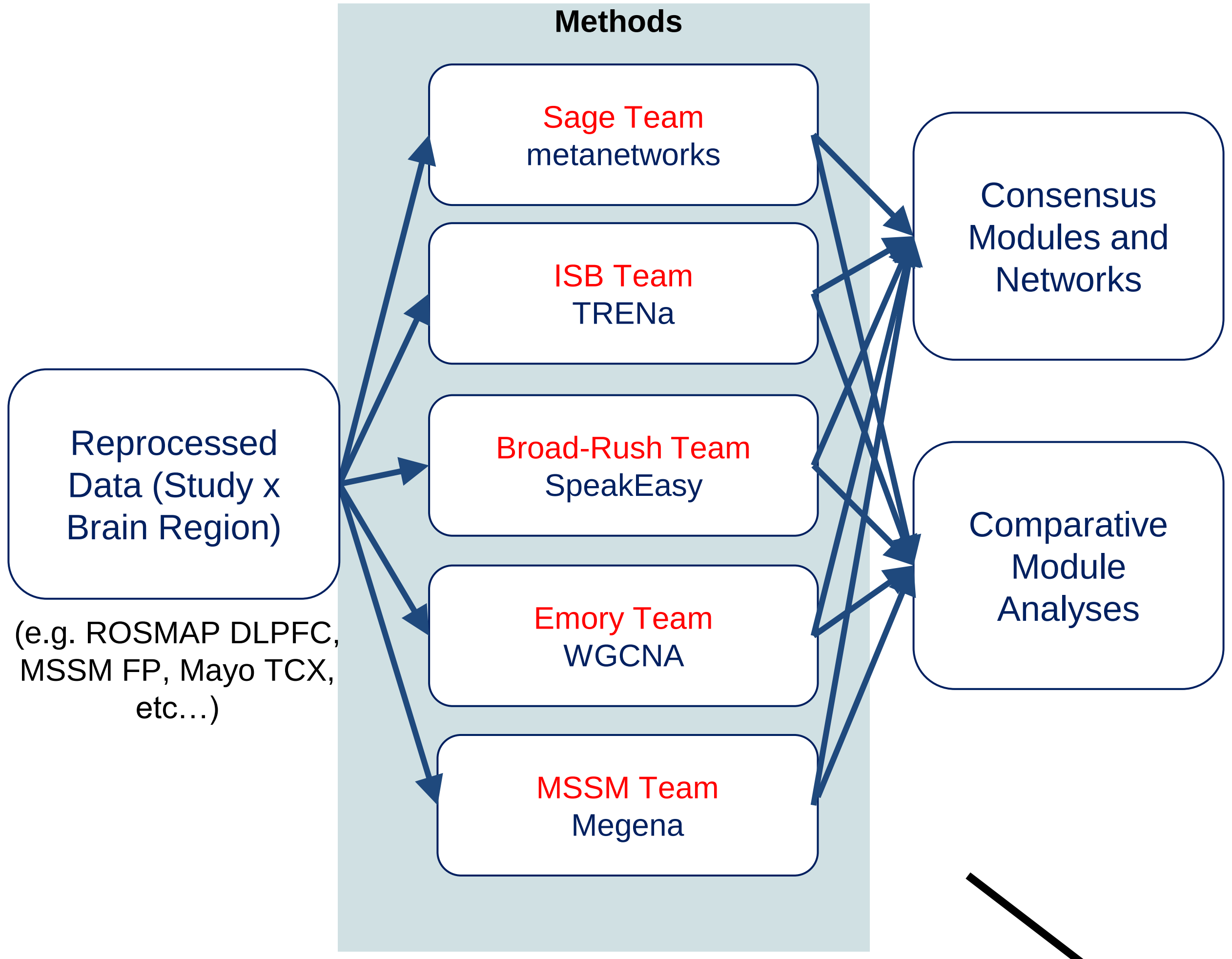
MSSM → MSSM → MSSM

Mayo → Mayo → Mayo

Conserved AD Associated networks, modules, and DEGs

Analysis strategy

Multi method coexpression network analysis


Methods

Reprocessed Data (Study x Brain Region)

(e.g. ROSMAP DLPFC, MSSM FP, Mayo TCX, etc…)

Sage Team
metanetworks

ISB Team
TRENa

Broad-Rush Team
SpeakEasy

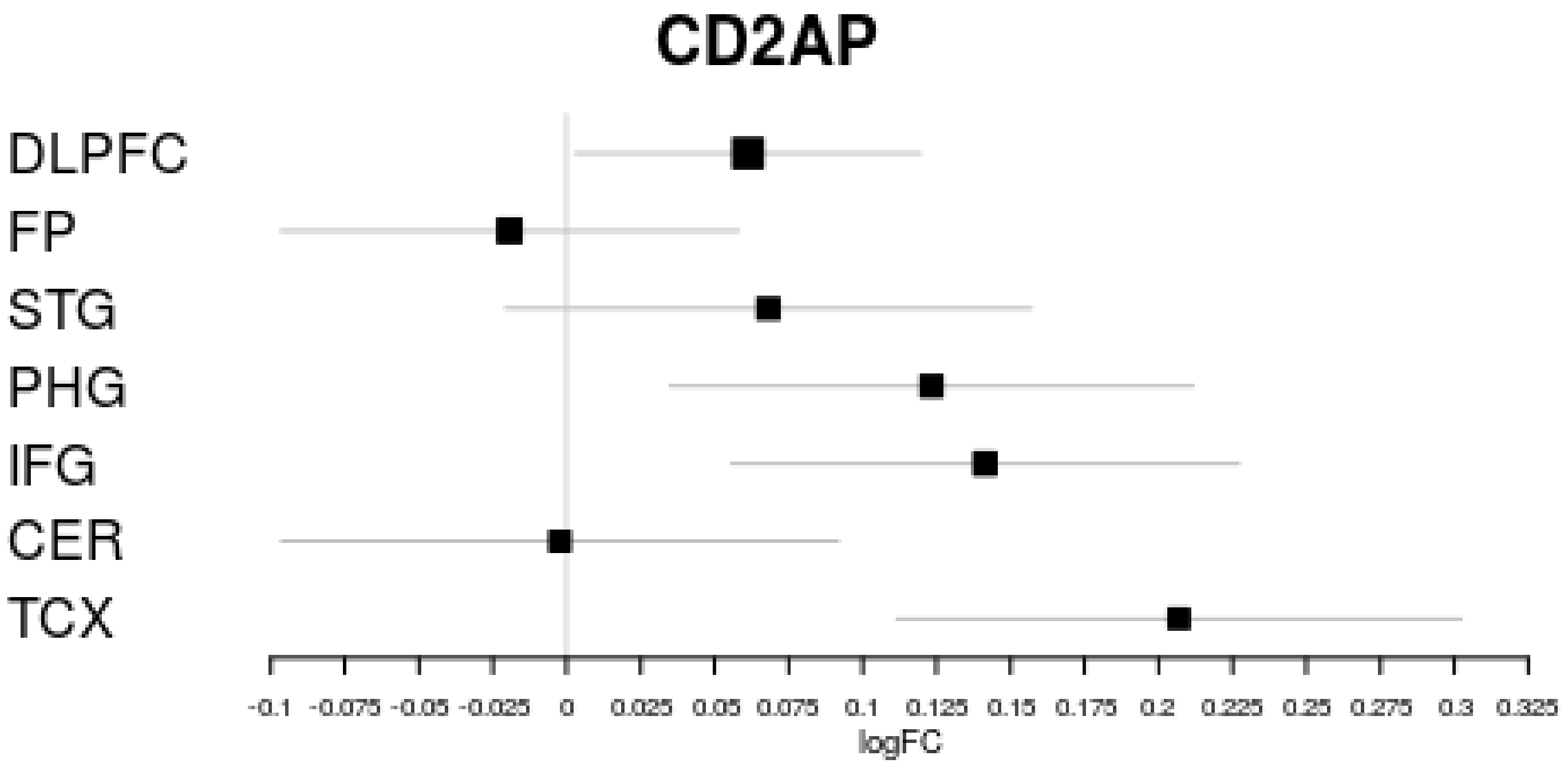
Emory Team
WGCNA

MSSM Team
Megena

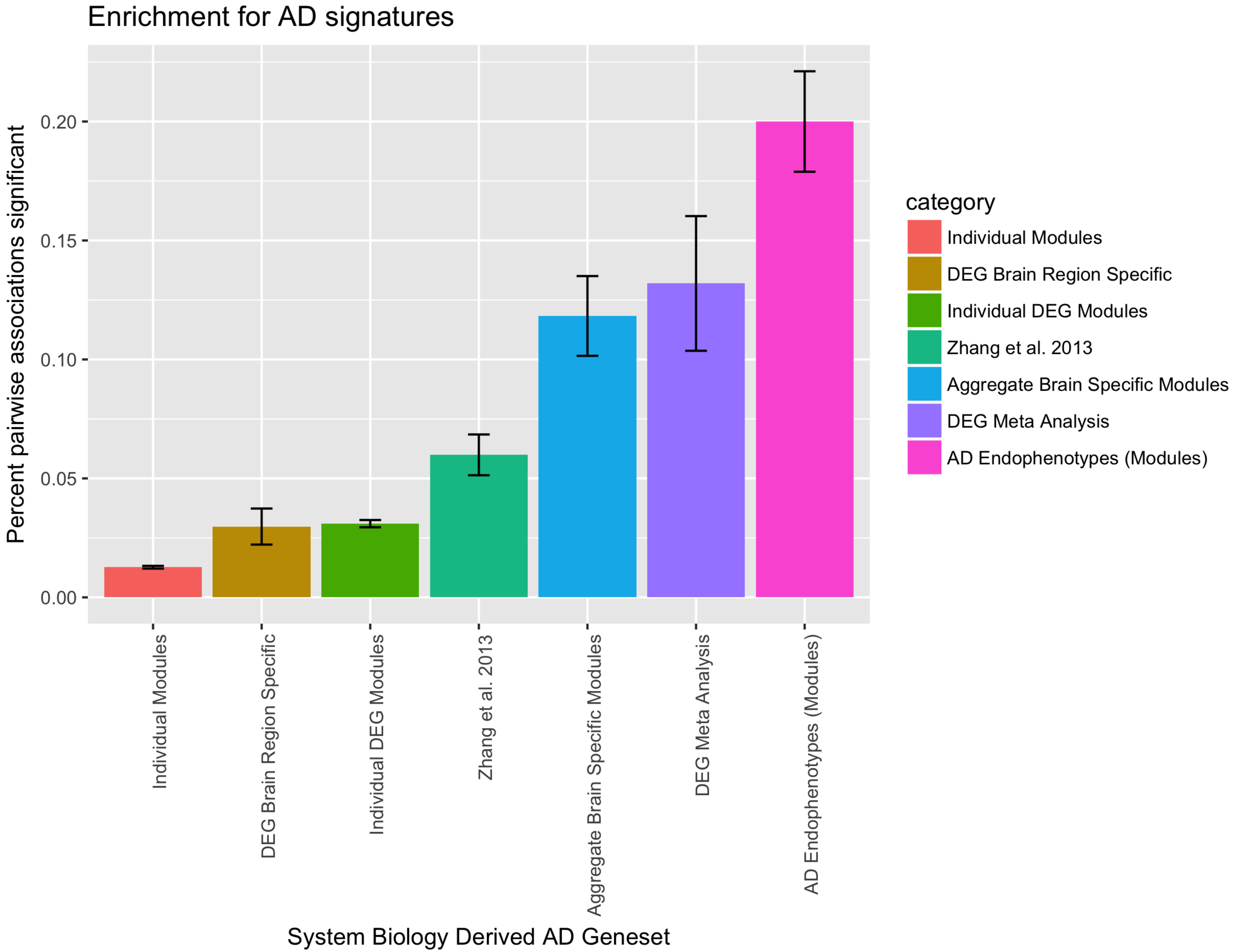
Consensus Modules and Networks

Comparative Module Analyses

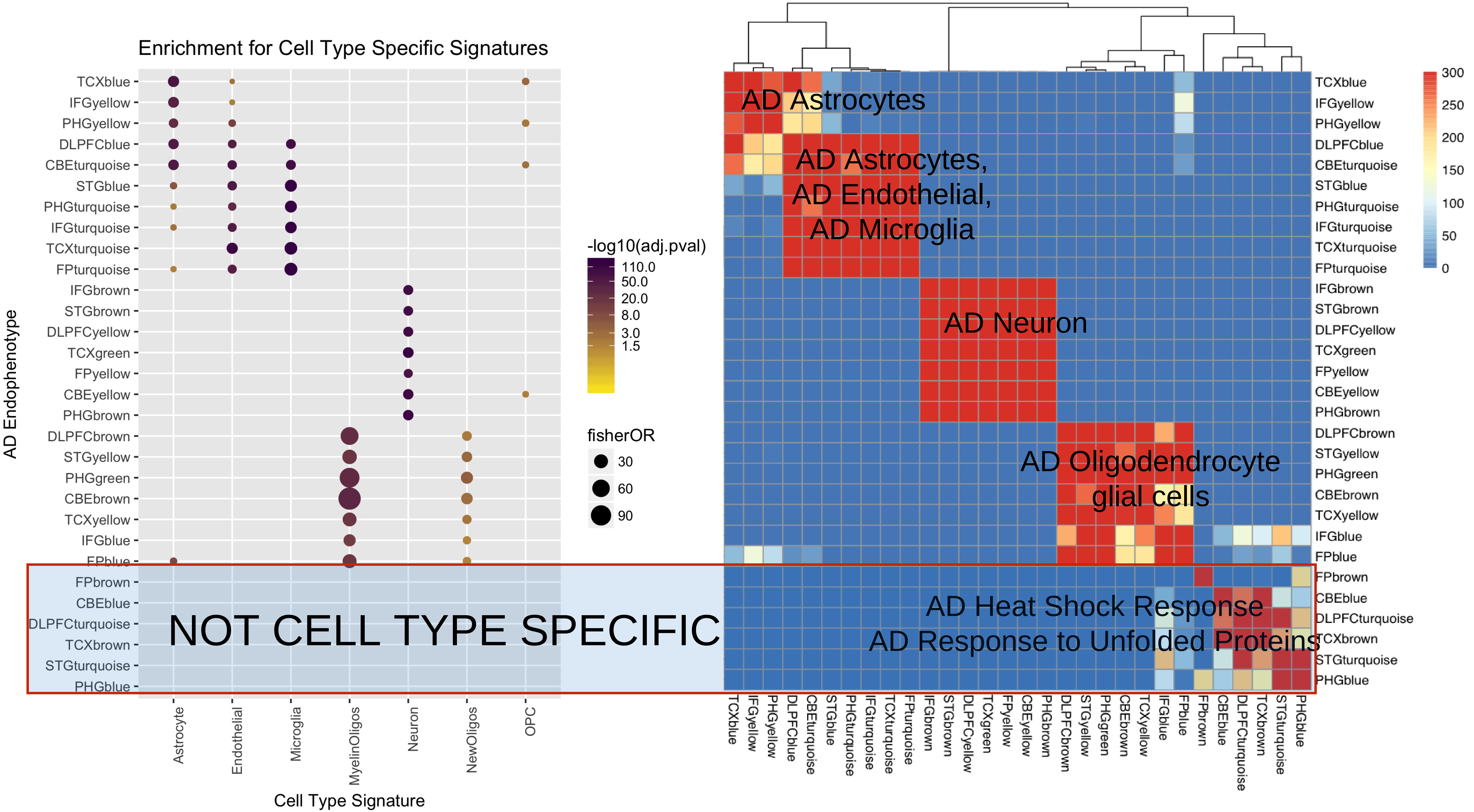
Differential Expression Meta Analysis



AD Transcriptomic Differentially Expressed coExpression Modules (DEEMs)


Enrichment for AD signatures
Percent pairwise associations significant
0.20
0.15
0.10
0.05
0.00
Individual Modules
DEG Brain Region Specific
Individual DEG Modules
Zhang et al. 2013
Aggregate Brain Specific Modules
DEG Meta Analysis
AD Endophenotypes (Modules)
System Biology Derived AD Geneset
category
Individual Modules
DEG Brain Region Specific
Individual DEG Modules
Zhang et al. 2013
Aggregate Brain Specific Modules
DEG Meta Analysis
AD Endophenotypes (Modules)

AD endophenotypes show strong cell type specificity



AD Astrocytes

AD Astrocytes,
AD Endothelial,
AD Microglia

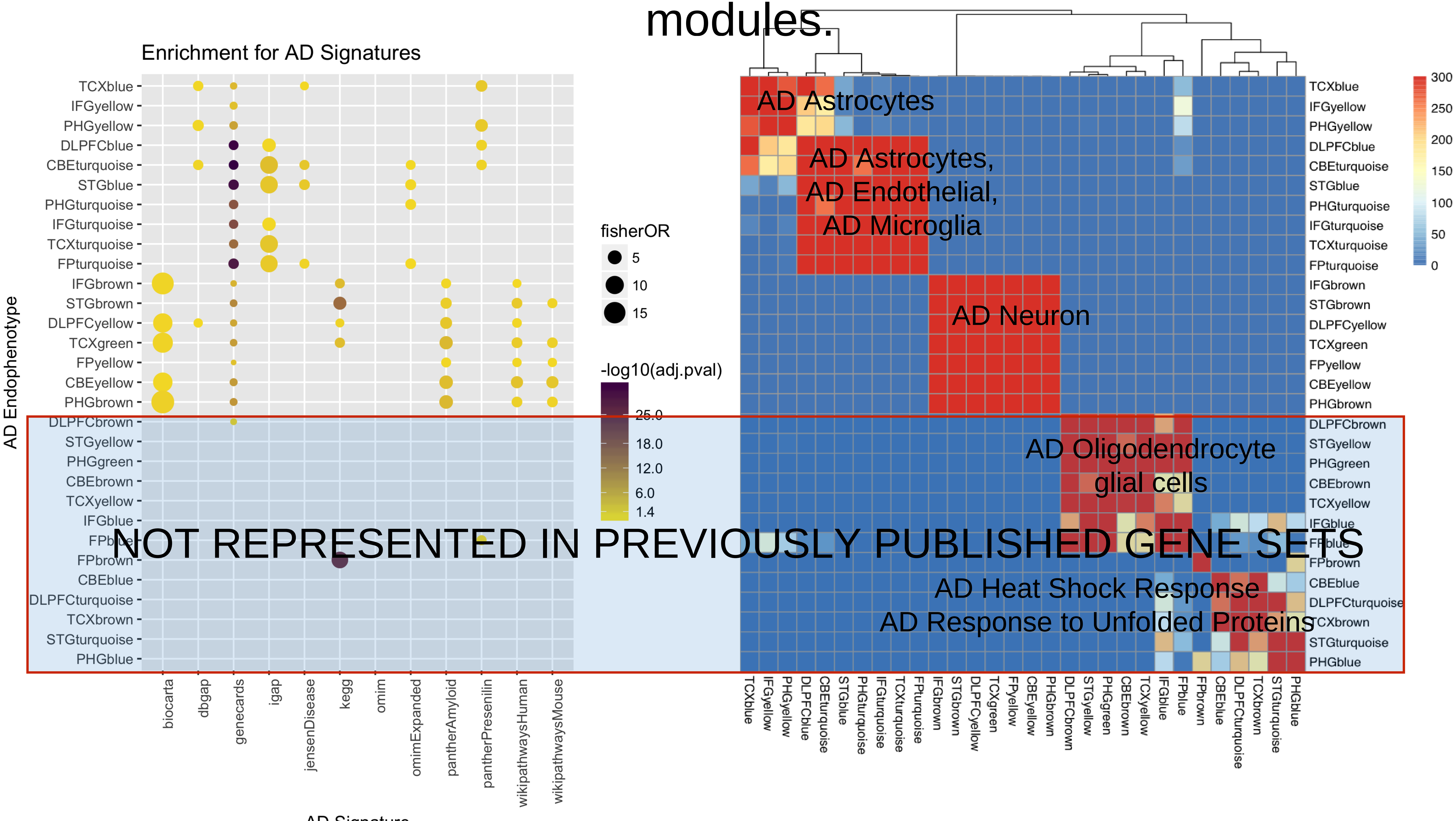
AD Neuron

AD Oligodendrocyte
glial cells

AD Heat Shock Response
AD Response to Unfolded Proteins

NOT CELL TYPE SPECIFIC

Known AD signatures localize in predominantly glial and neuronal modules.



AD Astrocytes

AD Astrocytes,
AD Endothelial,
AD Microglia

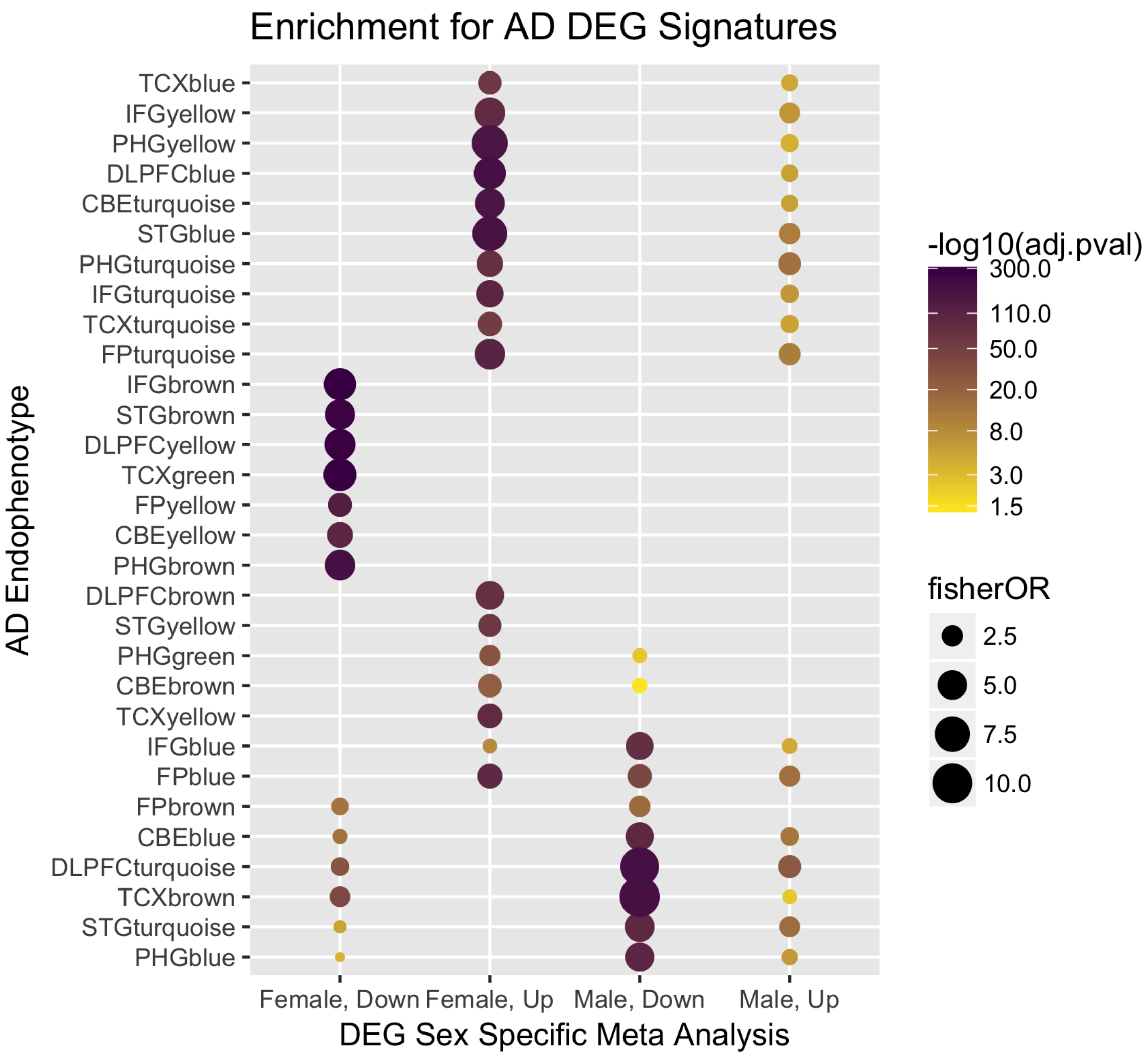
AD Neuron

AD Oligodendrocyte
glial cells

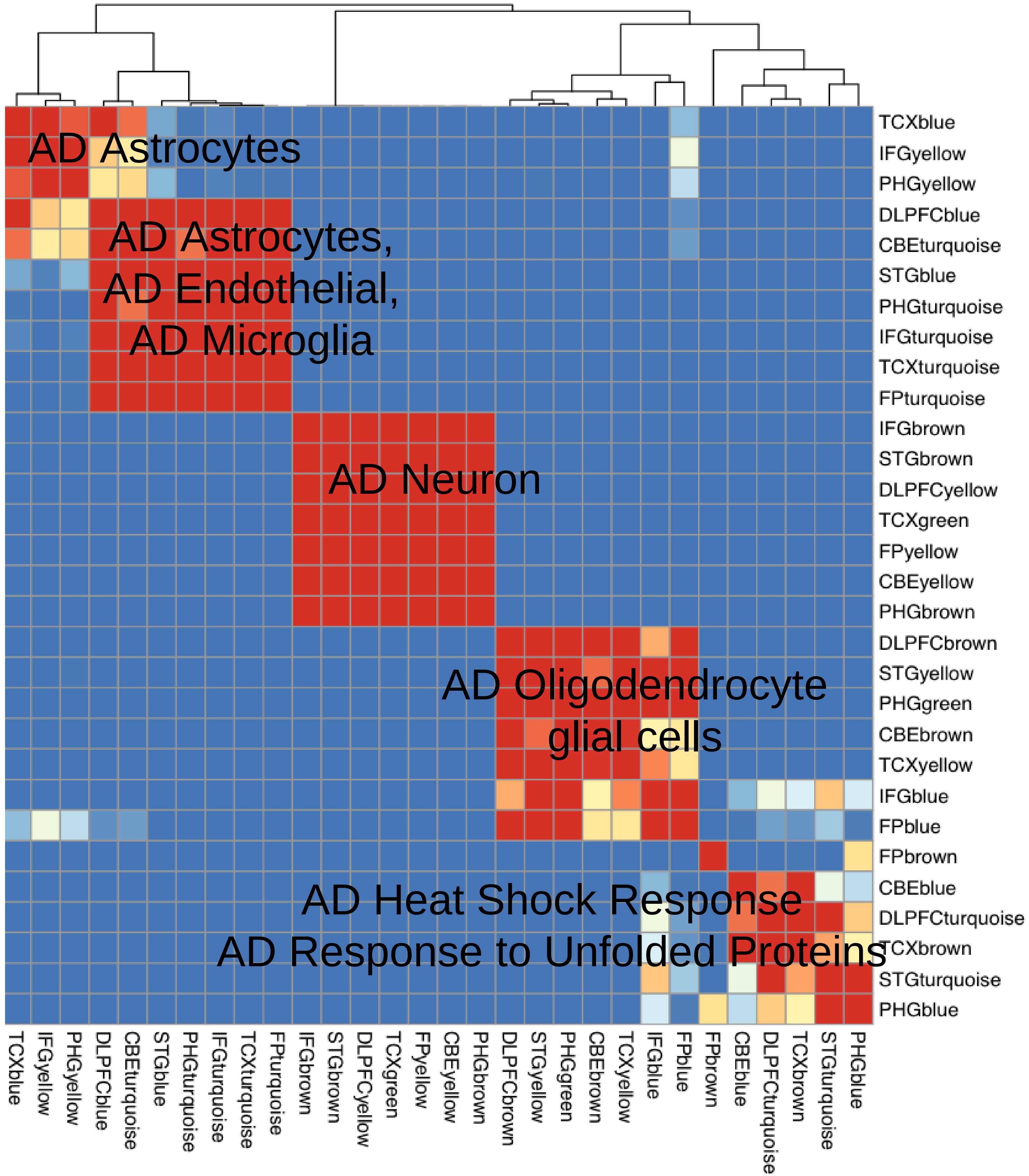
NOT REPRESENTED IN PREVIOUSLY PUBLISHED GENE SETS

AD Heat Shock Response
AD Response to Unfolded Proteins

AD endophenotypes show strong sex specificity



Enrichment for AD DEG Signatures



AD Astrocytes

AD Astrocytes,
AD Endothelial,
AD Microglia

AD Neuron

AD Oligodendrocyte
glial cells

AD Heat Shock Response
AD Response to Unfolded Proteins

Differential Expression: AD vs Control (Sex Specific)

Model:

Gx = BrainRegion.Diagnosis.Sex + Study + RIN + RIN^2 + PMI + AOD + Batch + PCT_RIBOSOMAL_BASES + PCT_INTRONIC_BASES + PCT_CODING_BASES + PCT_PF_READS_ALIGNED + PCT_INTERGENIC_BASES + (1|Individual_ID)

Source	**Region**	**# of Genes (at FDR of 0.05 and fold-change of 1.2)**			
		Female		**Male**	
		DOWN	**UP**	**DOWN**	**UP**
ROSMAP	DLPFC	325	352	92	30
MSBB	FP	4	0	24	20
	IFG	2	0	3	3
	PHG	587	361	198	121
	STG	6	1	1	1
MAYO	CER	2461	2199	1250	1151
	TCX	3060	2635	1719	1955

MUCH HIGHER

Thanneer Perumal

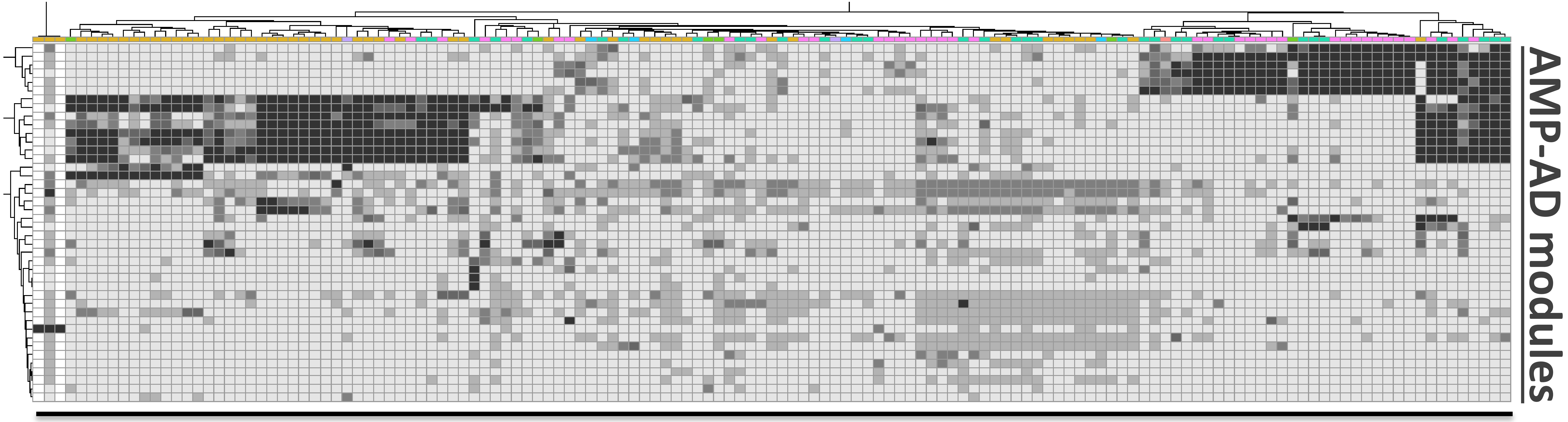
Genetics and Sex Interaction in AD

- We computed a *Genomic Risk Score* (GRS) based on the IGAP SNPs in ROSMAP from imputed genotype data

- We found significant (adjusted p-value: $1x10^{-4}$) sex by GRS interaction associations with DLPFCbrown (oligodendrocyte myelinating cells enriched module) eigengenes.

- We also found significant (p-value: 2.9x10-3) sex by GRS interaction association with Braak and CERAD score in ROSMAP.

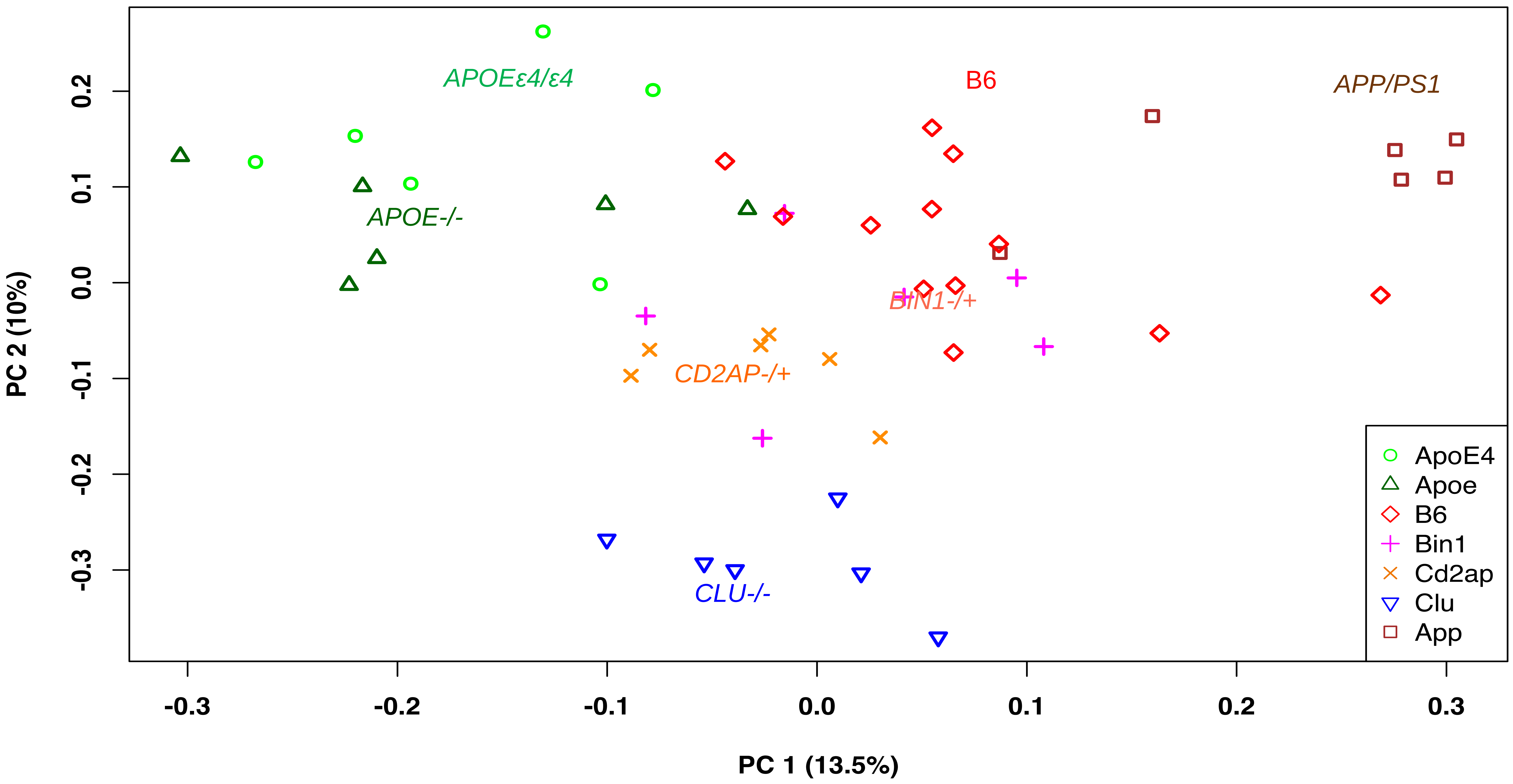
Enrichment in Human Modules



139 mouse model Differentially Expressed Genes (DEGs) in Brain

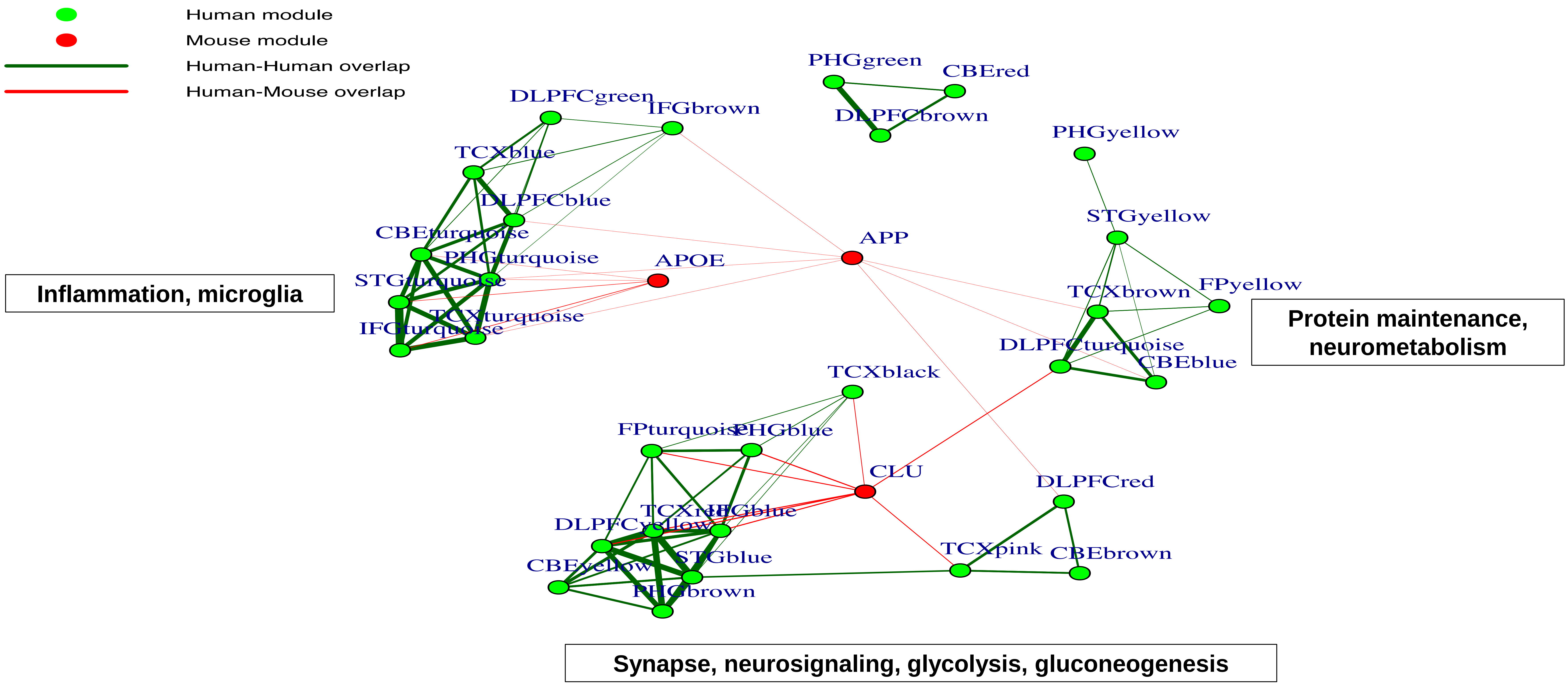
Joshua Shulman & Zhandong Liu,
Baylor College of Medicine
Cross-species working group

Mouse Knockout Models in MODEL-AD



Slide courtesy of Greg Carter

Human and mouse modules overlap



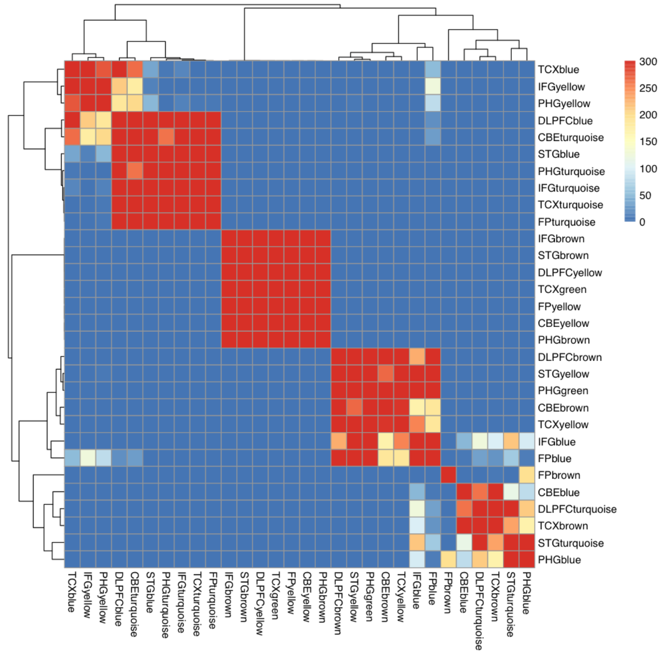
Inflammation, microglia

Protein maintenance, neurometabolism

Synapse, neurosignaling, glycolysis, gluconeogenesis

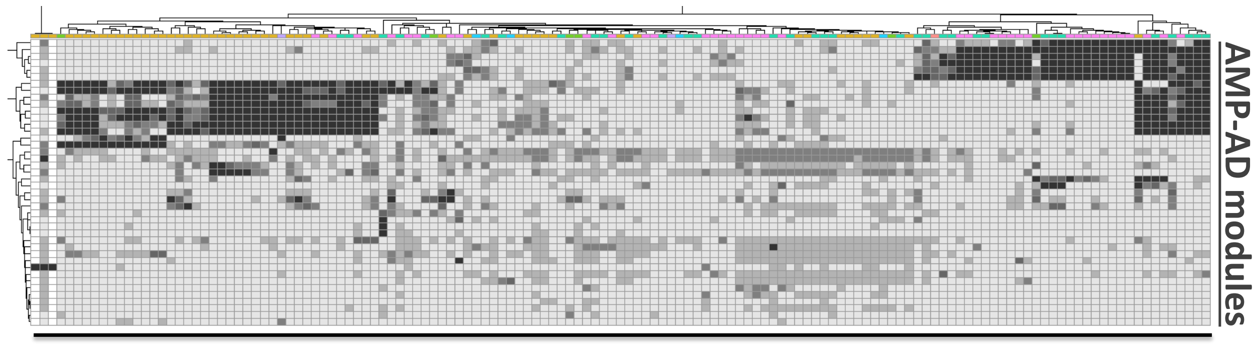
Slide courtesy of Greg Carter

- Disease Processes
- Patient Sub-Types
- Heterogeneity of disease



- Decompose disease Heterogeneity
- Identify missing models
- Better model patient specific signatures of disease

Enrichment in Human Modules



139 mouse model Differentially Expressed Genes (DEGs) in Brain

Questions?

- Also, we’re hiring! We have two positions open – a postdoc in computational neurodegenerative research and a research associate in genomic workflows.

- Please contact me – ben.logsdon@sagebionetworks.org - if interested

Acknowledgements

Sage Bionetworks
Phil Snyder
Yooree Chae
Solly Sieberts
Kristen Dang
Thanneer Perumal
Lara Mangravite

Columbia-Broad-Rush
Chris Gaiteri
Lei Yu
Jishu Xu
David Bennett
Phil De Jager

Mayo-ISB-UFL
Mariet Allen
Cory Funk
Minerva Carrasquilo
Nathan Price
Steve Younkin
Nilufer Ertekin-Taner

MSSM
Minghui Wang
Noam Beckmann
Sam Gandy
Vahram Haroutunian
Jun Zhu
Koichi Iijima
Scott Noggle
Bin Zhang
Eric Schadt

Emory
Eric Dammer
Duc Duong
James Lah
Ranjita Betarbet
Joshua Shulman
Vivek Swarup
Madhav Thambisetty
Tom Montine
Juan Troncoso
Daniel Geschwind
John Trojanowski
Allan Levey

Harvard-MIT
Tao Lu
Fan Gao
Alison Mungenast
Bruce Yankner
Li-Huei Tsai

AbbVie
Gyan Srivastava
Viswanath Devanarayan

Biogen
Ayla Ergun
Sally John

GSK
Hongkang Mei
Xiaoyan Zhong
Jill Richardson
Feng Yu
Guofeng Meng
Fiona Kelly

Duke
Jessie Tenenbaum
Rima Kaddurah Daouk

Lilly
David Airey
David Collier
Michael O’Neill
Hong Wang
Philip Ebert
Jeffrey Dage
Laura Nisenbaum

Columbia-SUNY
Jose Luchsinger
Adam Brickman
Herman Moreno
Sergio Angulo
Luna Buitrago

Mayo
Guojun Bu
Nilufer Ertekin-Taner
Mariet Allen
Minerva Carrasquillo
Takahisa Kanekyo

Emory (M2OVE-AD)
Ihab Hajjar
Arshed Quyyumi

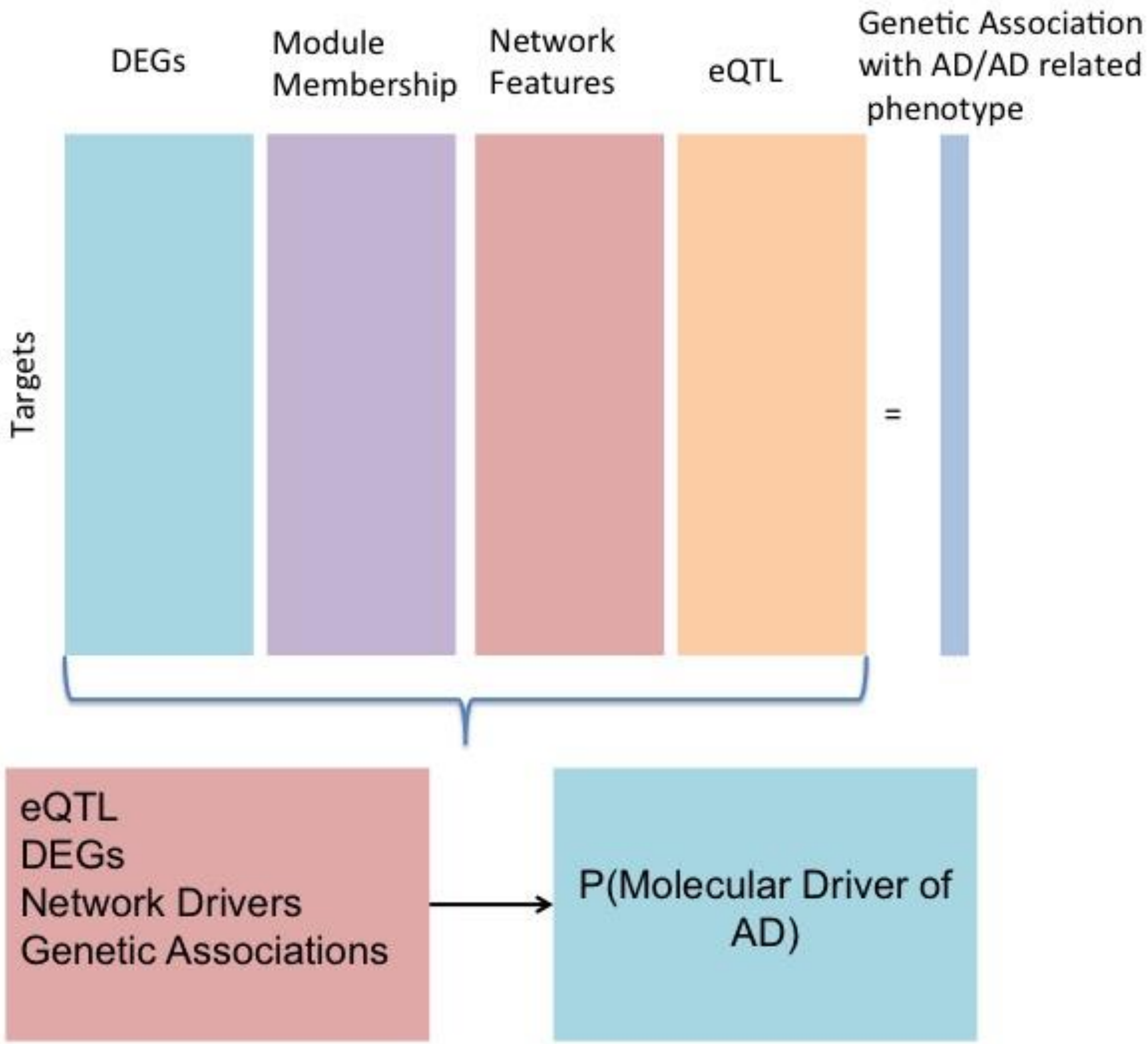
Duke -UPenn
Rima Kaddurah Daouk
Mitchel kling
Arshed Quyyumi

Harvard
Steven Greenberg
Brian Bacskai

IU-Jax-Sage (MODEL-AD)
Bruce Lamb
Greg Carter
Gareth Howell
Paul Territo
Mike Sasner
Stacey Rizo
Christoph Preuss
Asli Uyar

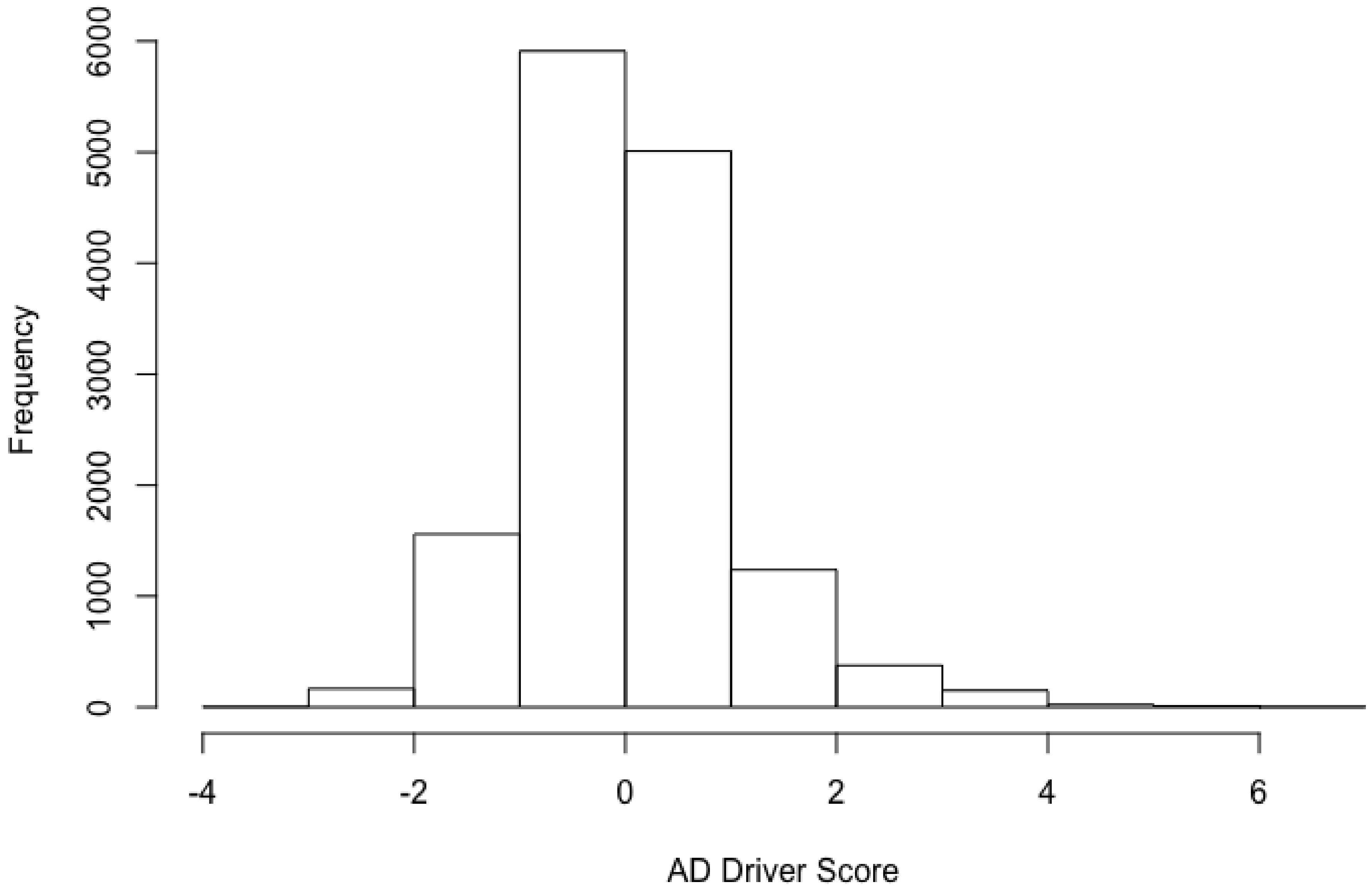
Funding
We would like to graciously
thank the NIA for their
financial support:
U01 AG046152, U01 AG046139,
U01 AG046161, U01 AG046170

Ranking system for Alzheimer’s disease drivers



Top Ranked Genes

gene	adDriverScore	external_gene_name
ENSG00000066336	6.894854674	SPI1
ENSG00000168918	6.539246527	INPP5D
ENSG00000120885	6.24964901	CLU
ENSG00000123338	5.924403628	NCKAP1L
ENSG00000197943	5.829814658	PLCG2
ENSG00000142347	5.732757074	MYO1F
ENSG00000105281	5.540961722	SLC1A5
ENSG00000183255	5.505479877	PTTG1IP
ENSG00000160255	5.497834261	ITGB2
ENSG00000095970	5.349427681	TREM2
ENSG00000160219	5.173553159	GAB3
ENSG00000108798	5.157684373	ABI3
ENSG00000128604	5.025162817	IRF5
ENSG00000019582	4.975204454	CD74
ENSG00000138964	4.875343453	PARVG
ENSG00000204257	4.865171894	HLA-DMA
ENSG00000134516	4.819653436	DOCK2
ENSG00000129667	4.721231461	RHBDF2
ENSG00000023892	4.50325394	DEF6



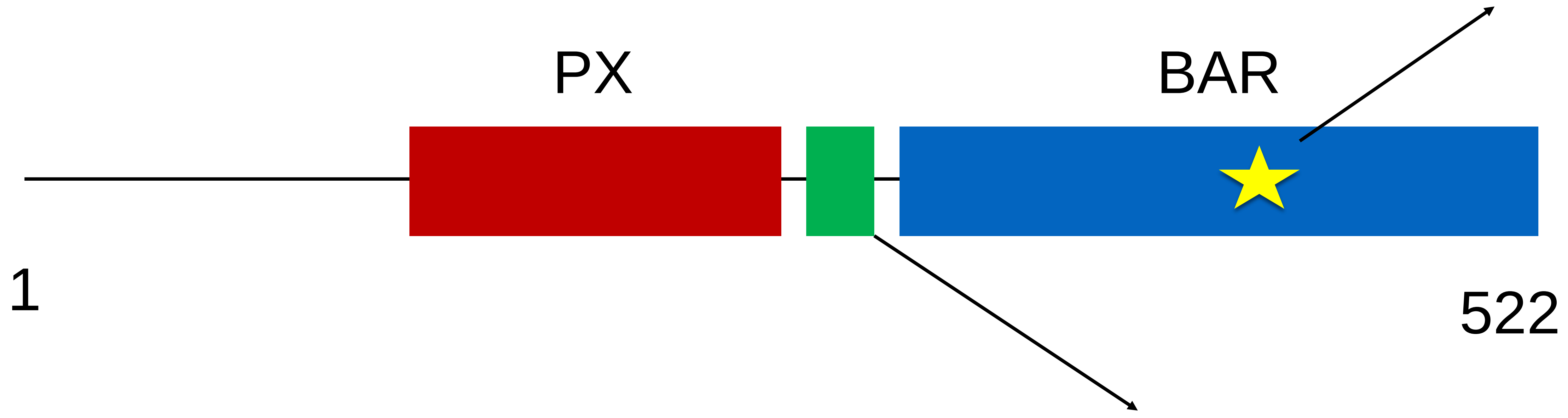
Re-prioritized genes based on score

- We select the top 1000 scored genes from the model to screen for variants with evidence in IGAP stage 1+2.

- We identify 667 variants of which 404 pass a Bonferroni correction ($P < 7 \times 10^{-5}$).

- We identify six new genomic regions associated with Alzheimer’s disease: *HBEGF*, *RABEP1*, *DNAH17*, *C1S*, *SNX1*, and *PFKFB3*.

SNX1

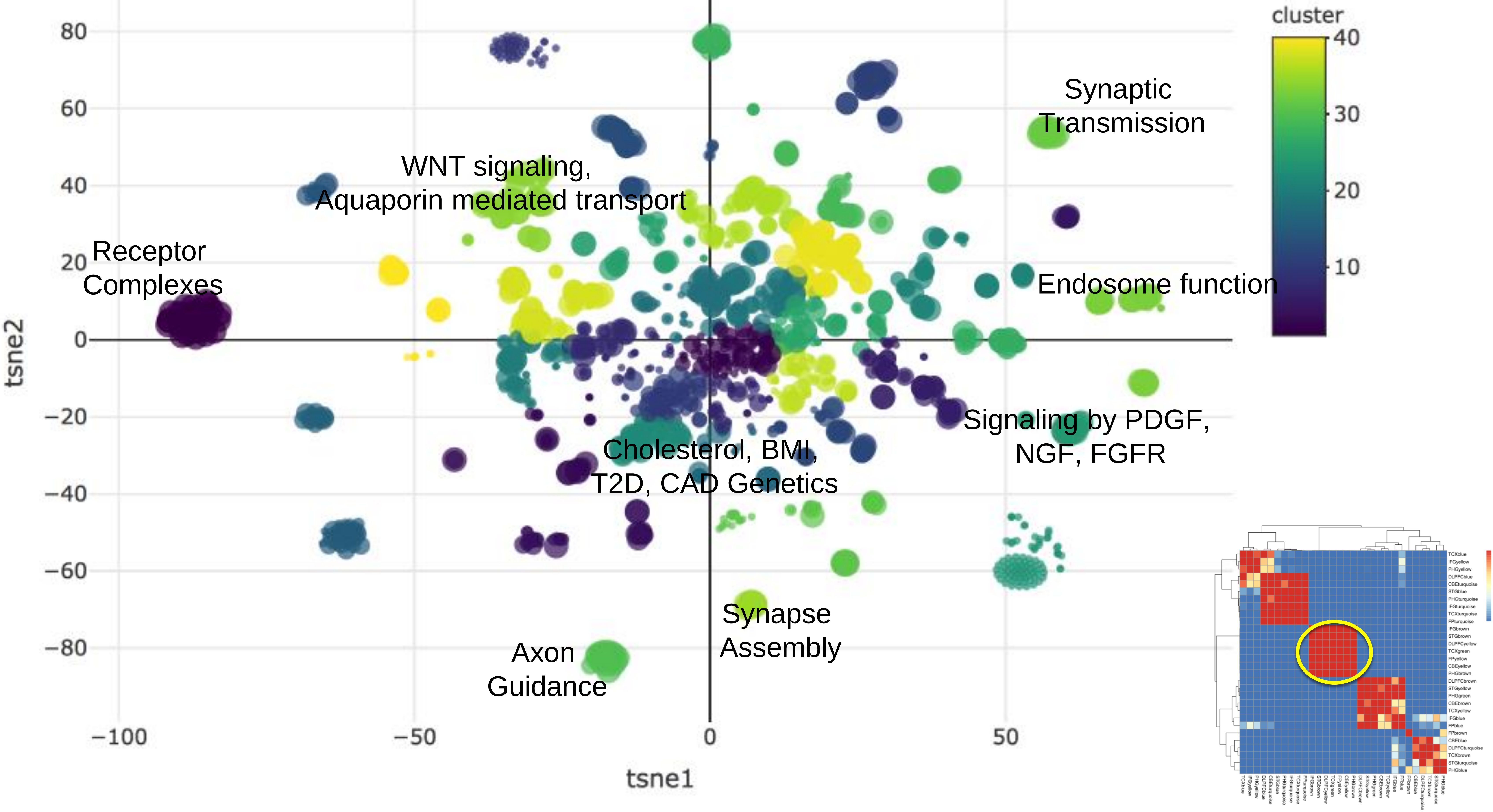


rs1802376 : p.D466N

Membrane binding amphipathic helix (PMID: 19816406)

rs1802376
IGAP P-value: 6x10^{-6}
ADSP exome P-value: 2.6x10^{-3}
CADD score: 28.3

Modules are rich in disease relevant pathway information



Synaptic Transmission

WNT signaling, Aquaporin mediated transport

Receptor Complexes

Endosome function

Signaling by PDGF, NGF, FGFR

Cholesterol, BMI, T2D, CAD Genetics

Synapse Assembly

Axon Guidance