From Genes to Multiomics, Deep Profiling to Better Understand Alzheimer's Disease

Victoria Fernandez, PhD
Assistant Professor of Psychiatry
Washington University School of Medicine in St. Louis
2023 Brain Aging Conference; March 10th

Washington University in St.Louis

FernandezLAB

NeuroGenomics &
Informatics Center
at Washington University

Overview

Intro
- review of Alzheimer disease genomics
  - ADAD, EOAD, LOAD
- Genetic methods
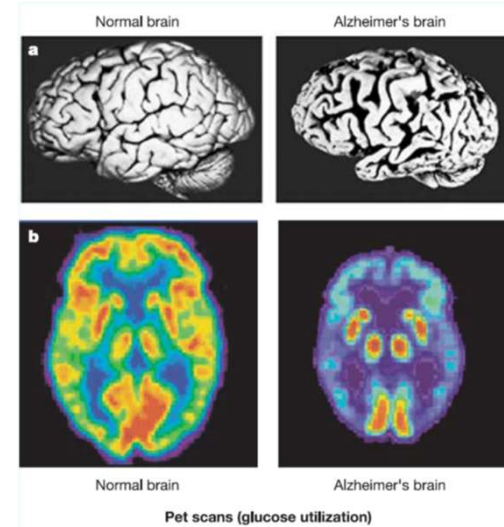- Omic methods
- Data integration

Genetic studies by us
- A large-scale Genome-Wide Association Study of Early Onset Alzheimer's Disease
- The Familial Alzheimer Sequencing (FASe) Project
- OMIC approaches to identify molecular contributors to AD

Alzheimer Disease (AD)

Clinically

- gradual onset and progression of memory impairment
- deficits in executive functioning, language, visuo-spatial abilities
- personality, behavior and self-care

Pathologically

- Reduction in volume and neuronal death
- Extracellular plaques of amyloid beta (**Aβ**)
- Intracellular tangles of hyperphosphorylated **Tau**
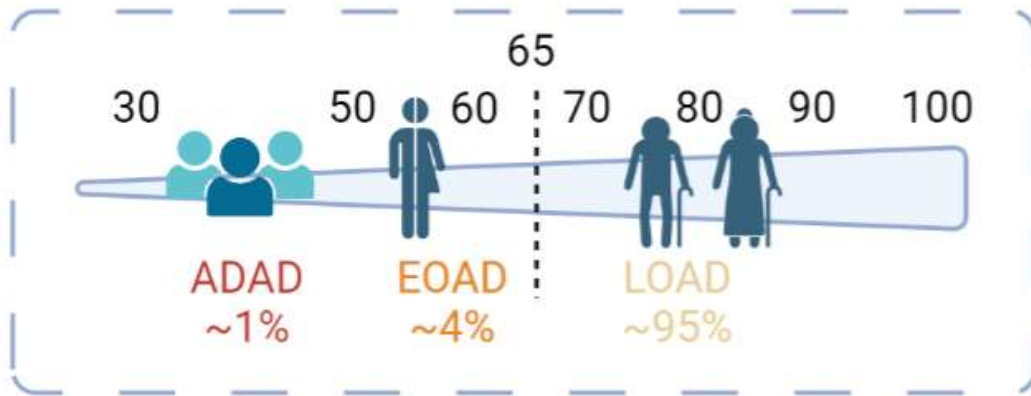

Mattson,2004,Nature, 430:631-639
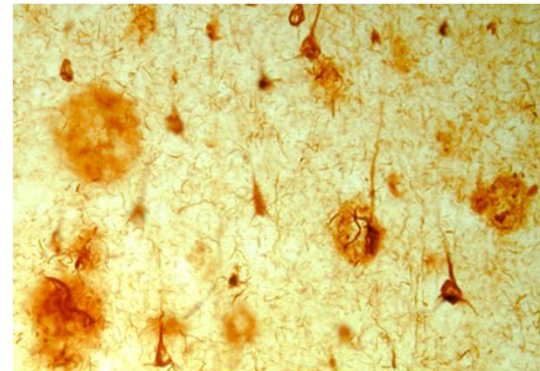
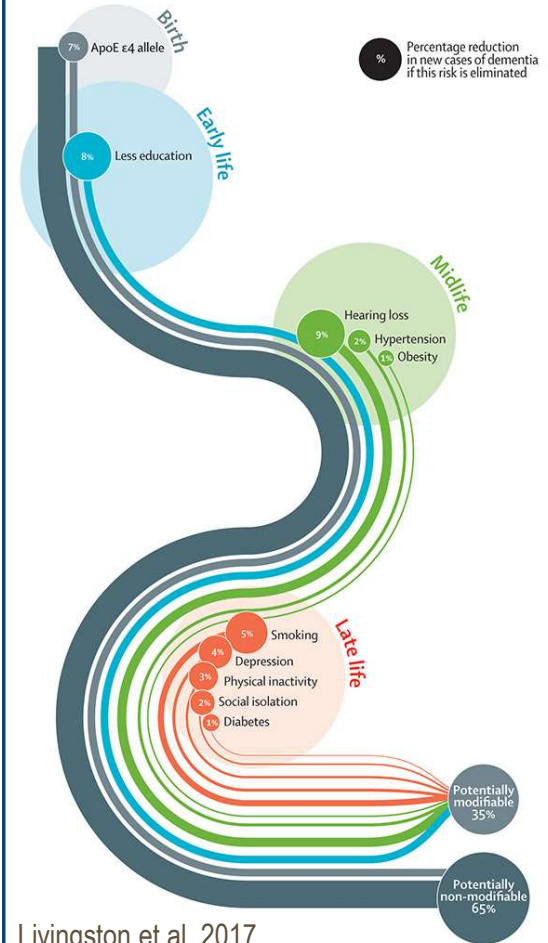# AD risk factors

Risk factors

- Age: risk doubles every 5 years after 60 years old
- 60-80 % is due to genetic causes

Gatz 2006, Arch Gen Psy 63:168-174

Risk factors for dementia

The Lancet Commission presents a new life-course model showing potentially modifiable, and non-modifiable, risk factors for dementia.
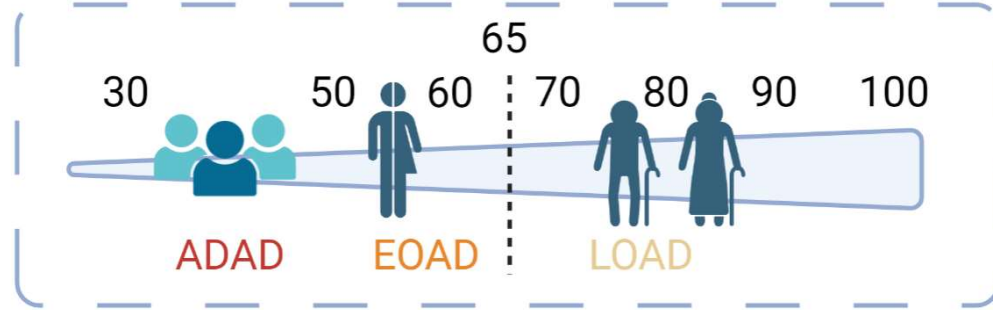
Livingston et al. 2017

ADAD

Segregation of a missense mutation in the amyloid precursor protein gene with familial Alzheimer's disease

Alison Goate*, Marie-Christine Chartier-Harlin*, Mike Mullan*, Jeremy Brown*, Fiona Crawford*, Liana Fidani*, Luis Giuffra†, Andrew Haynes‡, Nick Irving*, Louise James‡, Rebecca Mant||, Phillippa Newton*, Karen Rooke*, Penelope Roques*, Chris Talbot*, Margaret Pericak-Vance§, Allen Roses§, Robert Williamson*, Martin Rossor*, Mike Owen|| & John Hardy*¶

ADAD

PSEN1

PSEN2

65

30 50 60 70 80 90 100

ADAD EOAD LOAD
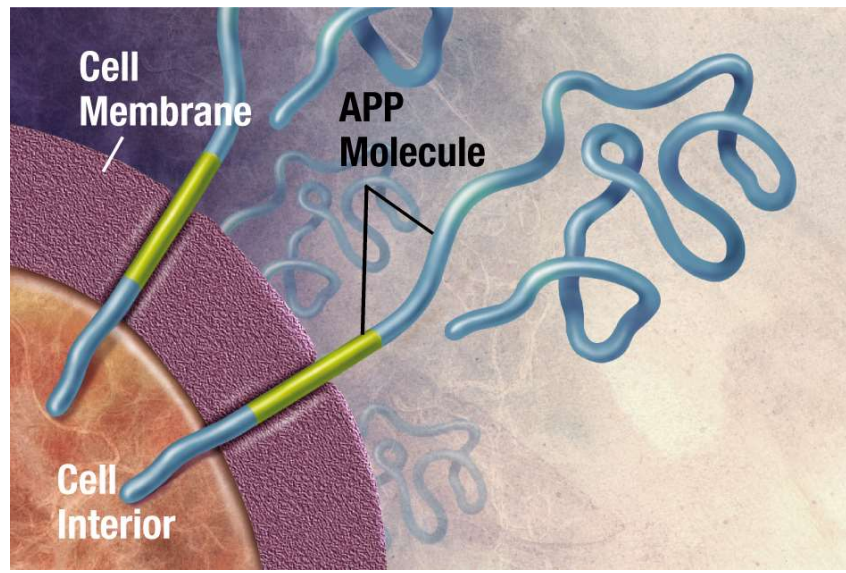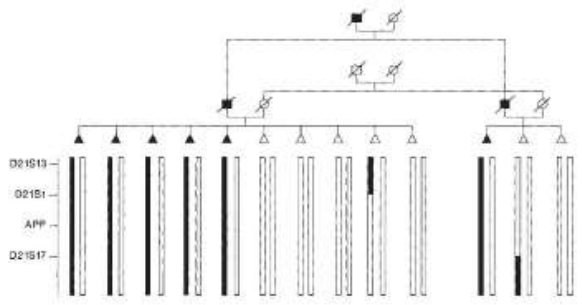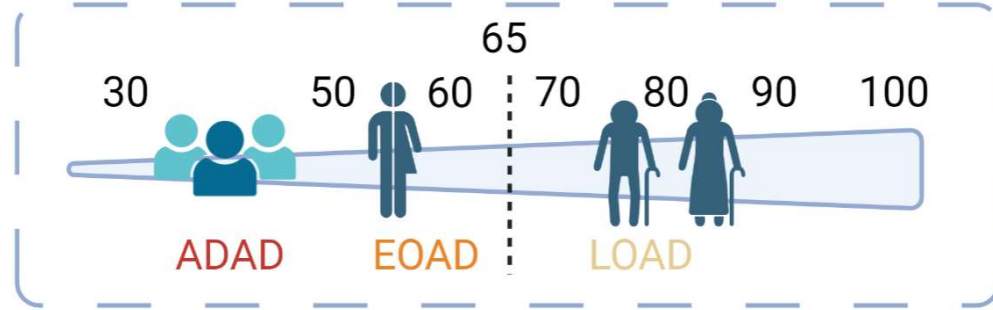
A variant of Alzheimer's disease with spastic paraparesis and unusual plaques due to deletion of exon 9 of presenilin 1

RICHARD CROOK[1], AULI VERKKONIEMI[2], JORDI PEREZ-TUR[1], NITIN MEHTA[1], MATT BAKER[1], HENRY HOULDEN[1], MATT FARRER[1], MIKE HUTTON[1], SARAH LINCOLN[1], JOHN HARDY[1], KATRINA GWINN[3], MIRJA SOMER[4], ANDERS PAETAU[5], HANNU KALIMO[6,7], RAIJA YLIKOSKI[7], MINNA POYHONEN[8] STEVE KUCERA[9] & MATTI HALTIA[5].
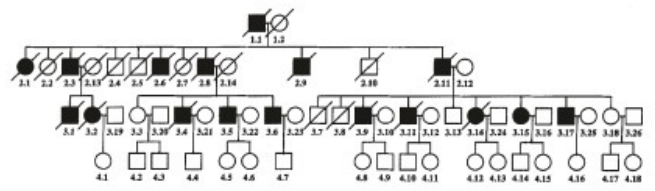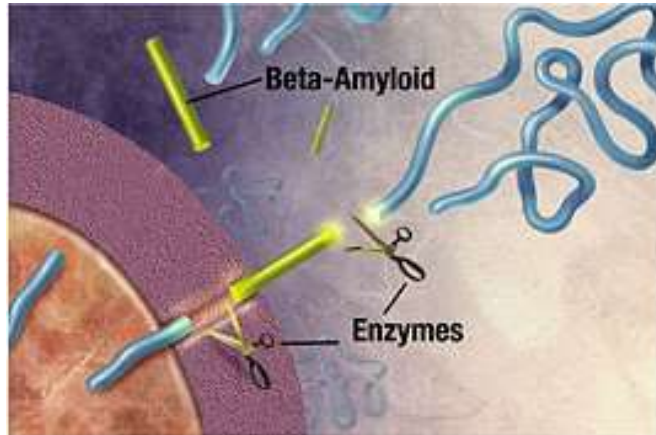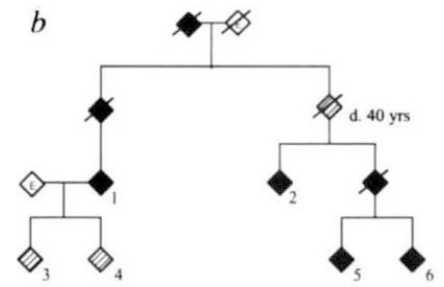
Fig. 1 Finnish pedigree showing 17 individuals affected by a variant of Alzheimer 's disease. Clinical information on 12 of these individuals (age of onset 45–57 years) in three generations is presented in Table 1.

Familial Alzheimer's disease in kindreds with missense mutations in a gene on chromosome 1 related to the Alzheimer's disease type 3 gene

E. I. Rogaev*, R. Sherrington*, E. A. Rogaeva*, G. Levesque*, M. Ikeda*, Y. Liang*, H. Chi*, C. Lin*, K. Holman*, T. Tsuda*, L. Mar†, S. Sorbi‡, B. Nacmias‡, S. Piacentini‡, L. Amaducci‡, I. Chumakov§, D. Cohen§, L. Lannfelt||, P. E. Fraser*, J. M. Rommens† & P. H. St George-Hyslop*¶
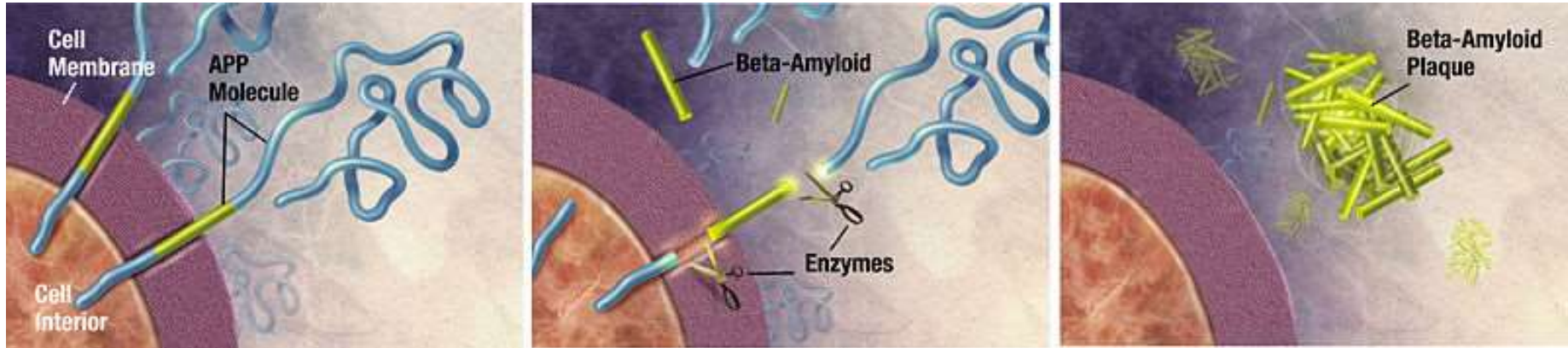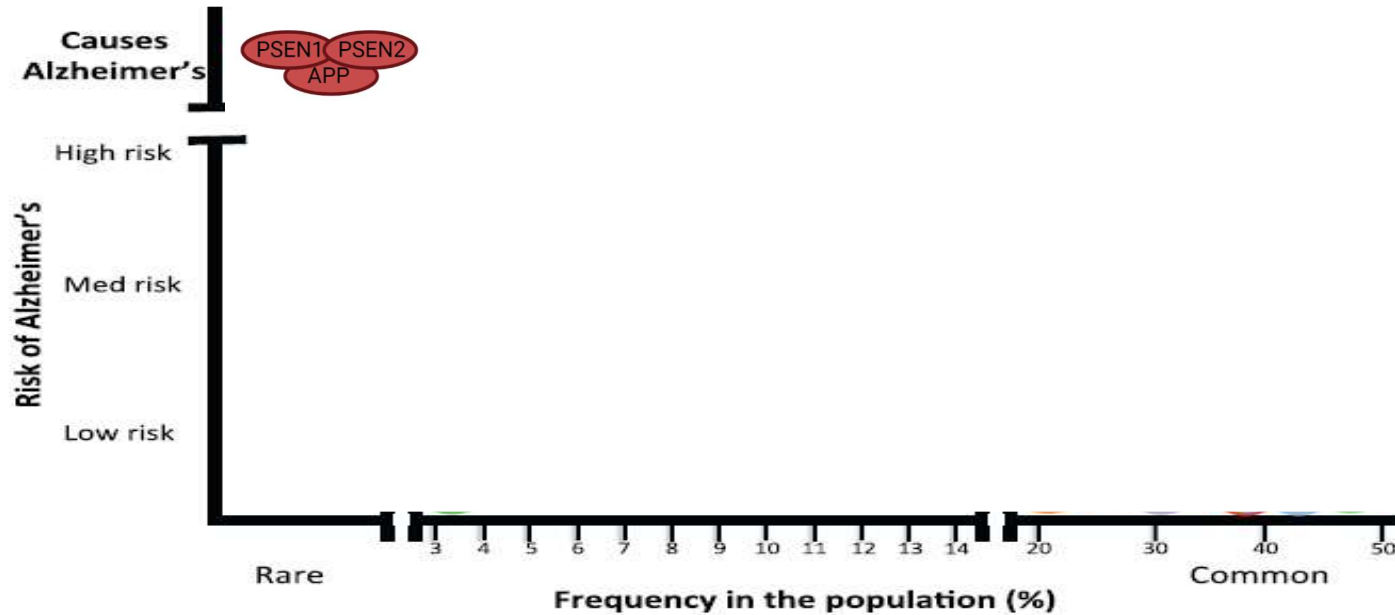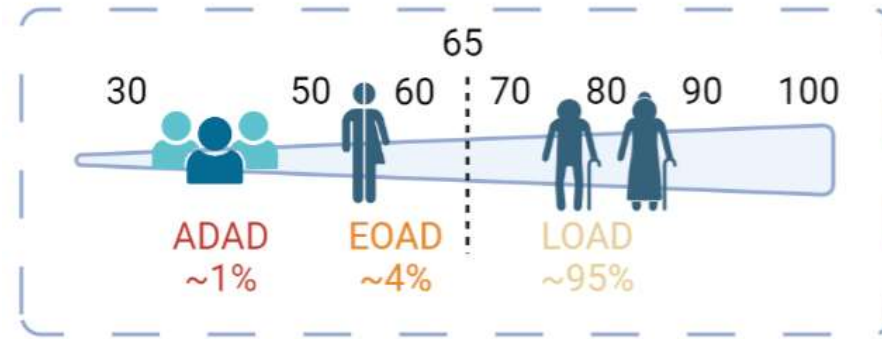
Beta-Amyloid

Enzymes

b

d. 40 yrs

ADAD

PSEN1   APP   PSEN2

APP mutations → increase β-secretase cleavage

PS1/PS2 mutations → increase γ-secretase cleavage

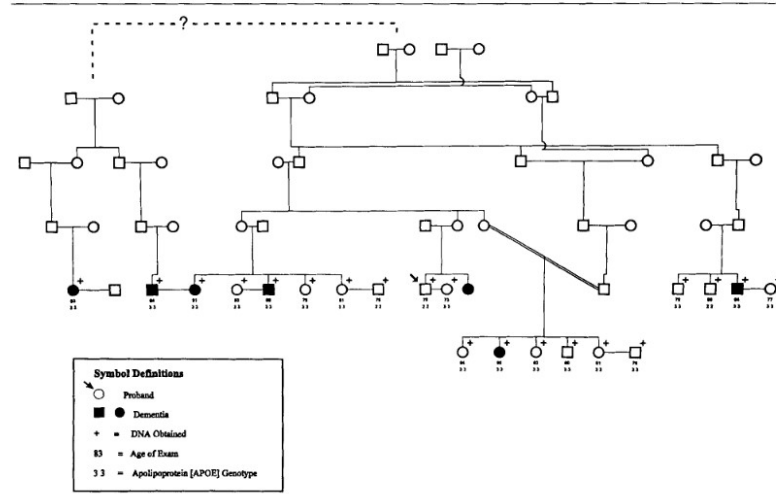# AD genetics

AD APOE

ORIGINAL ARTICLES

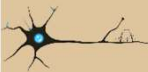# Alzheimer's Disease and Apolipoprotein E-4 Allele in an Amish Population

M. A. Pericak-Vance, PhD,* C. C. Johnson, PhD,† J. B. Rimmler, MS,* A. M. Saunders, PhD,*
L. C. Robinson, BS,* E. G. D'Hondt, BS,† C. E. Jackson, MD,† and J. L. Haines, PhD‡

Alzheimer's Disease (AD) is a complex genetic disorder with four loci already identified. Mutations in three of these, the amyloid precursor protein, presenilin I, and presenilin II, cause early-onset AD. The apolipoprotein E (APOE) gene contributes primarily to late-onset AD. The APOE-4 allele acts in a dose-related fashion to increase risk and decrease the age-of-onset distribution in AD. We examined the effect of APOE on AD in a previously unstudied Amish population that has a lower prevalence of dementia compared with other populations. We sampled a large inbred family with 6 late-onset AD members. We also genotyped 53 individuals from the general Amish population as controls for the APOE allele frequency estimates. The frequency of the APOE-4 allele in the Amish controls was 0.037 ± 0.02. This differed significantly compared with three independent sets of non-Amish white controls ($p < 2 \times 10^{-4}$, $p < 6 \times 10^{-5}$, and $p < 2 \times 10^{-6}$). In addition, all Amish AD-affected individuals had APOE 3/3 genotypes; no APOE X/4 or 4/4 individuals were observed. We suggest that the lower frequency of dementia in the Amish may be partially explained by the decreased frequency of the APOE-4 allele in this population, and that the inbred nature of this pedigree, with its strong clustering of cases contrasted against the lower frequency of dementia, indicates that additional genetic factors influence late-onset AD.
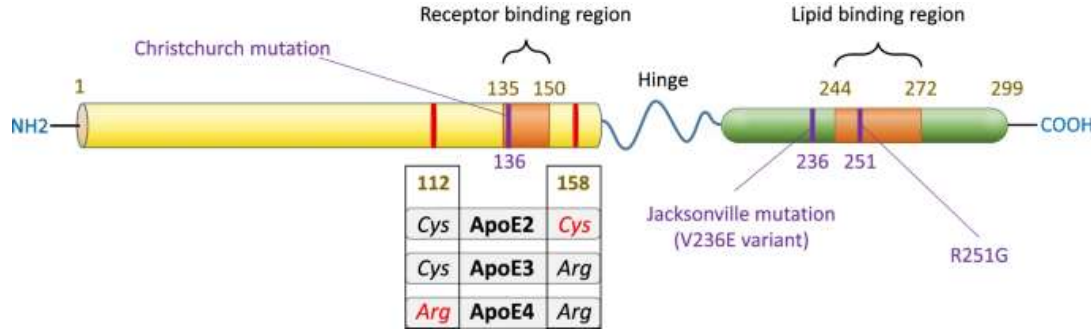
Pedigree of Amish family.

AD *APOE*

Christchurch mutation

Receptor binding region

Lipid binding region

1   135 150   Hinge   244   272   299

NH2   136   236 251   COOH

112		158
Cys	ApoE2	Cys
Cys	ApoE3	Arg
Arg	ApoE4	Arg

Jacksonville mutation
(V236E variant)

R251G

ASTROCYTE

APOE2 APOE3 APOE4 ABCA1

Lipid droplets   LXR   RXR

VLDLR   LRP1

HSPGs

AMPA NMDA LDLR

E2>E3>E4   E4?E3?E2   E4>E3>E2

ENDOSOME   Lipid oxidation   ROS

NEURON

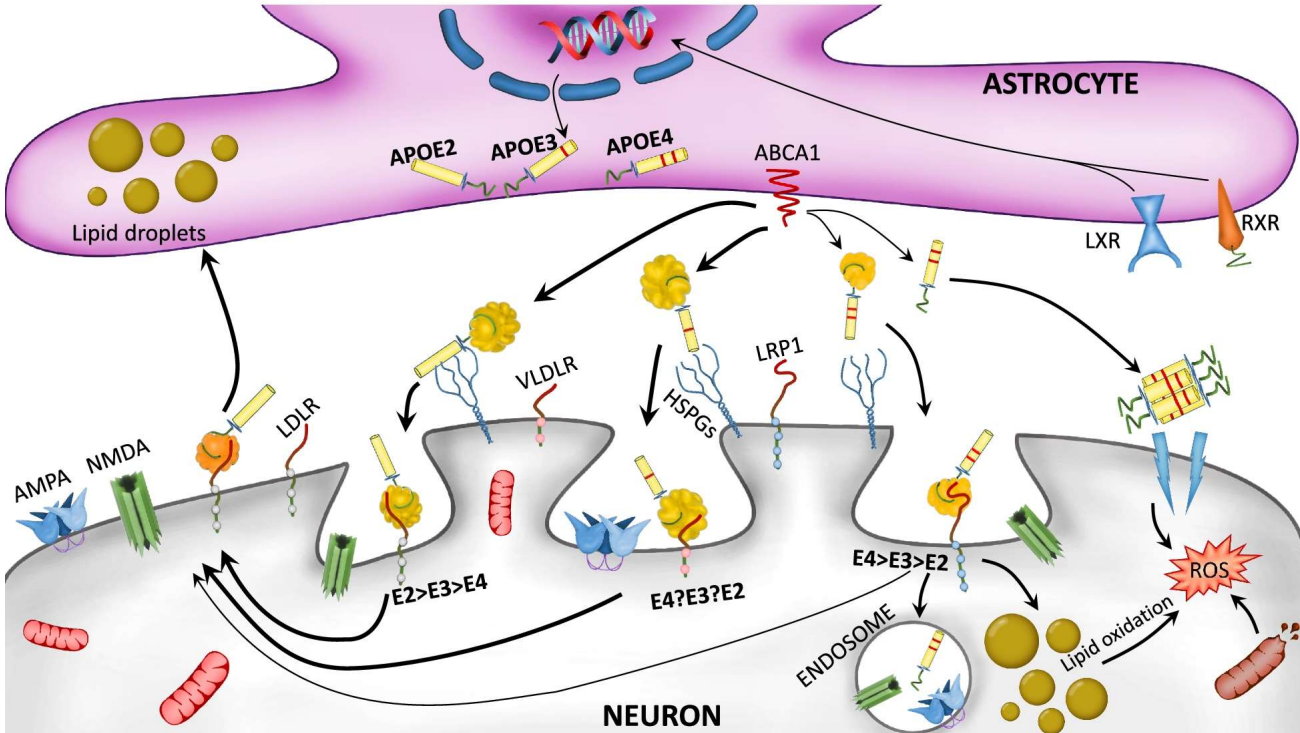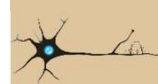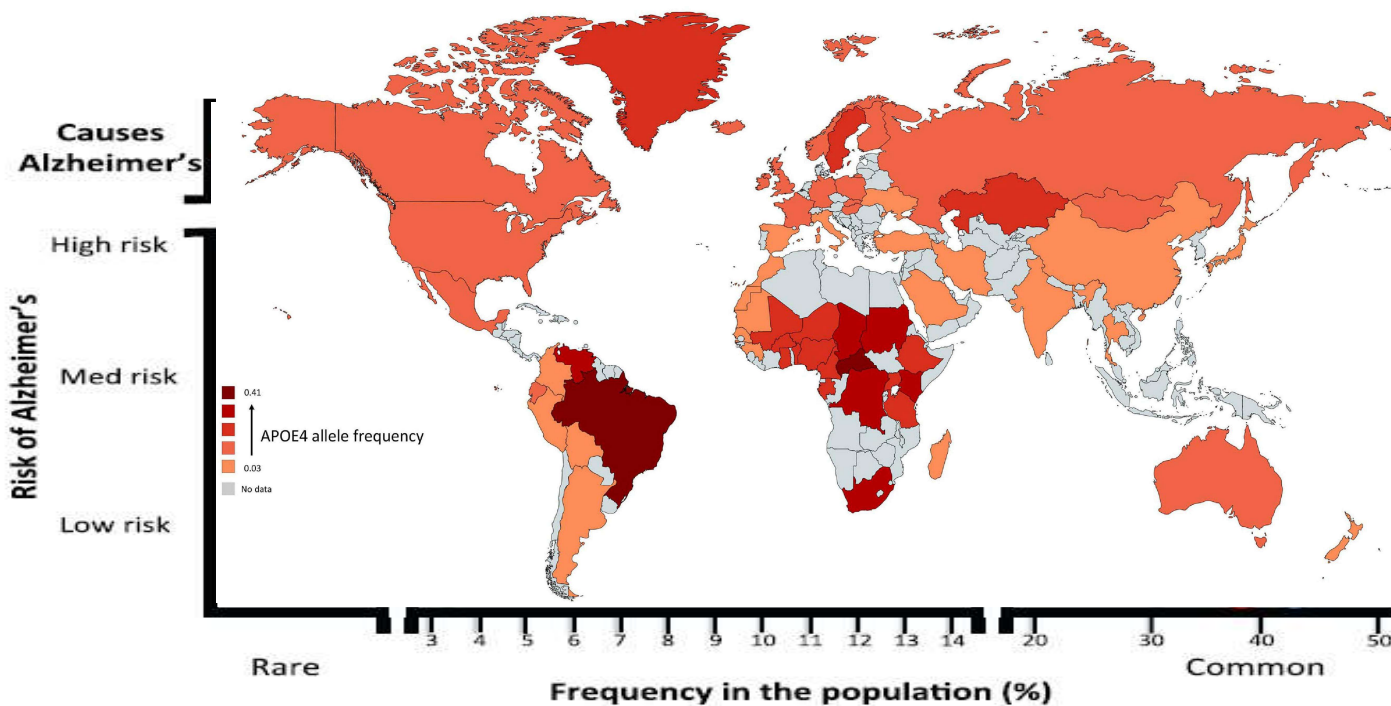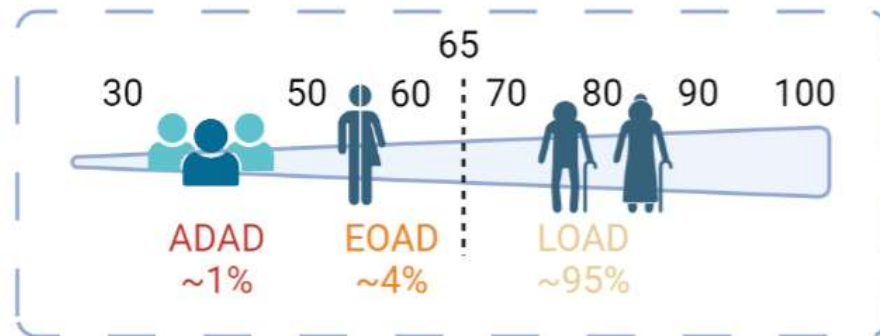## APOE ε4 in AD
- BBB disruption
- Synaptic impairment
- Protein clearance
- Protein aggregation
- Microglia activation
- Autophagy impairment

Neurogenomics & Informatics

AD genetics

Genome Wide Association Analysis

cases

CA = 35,247
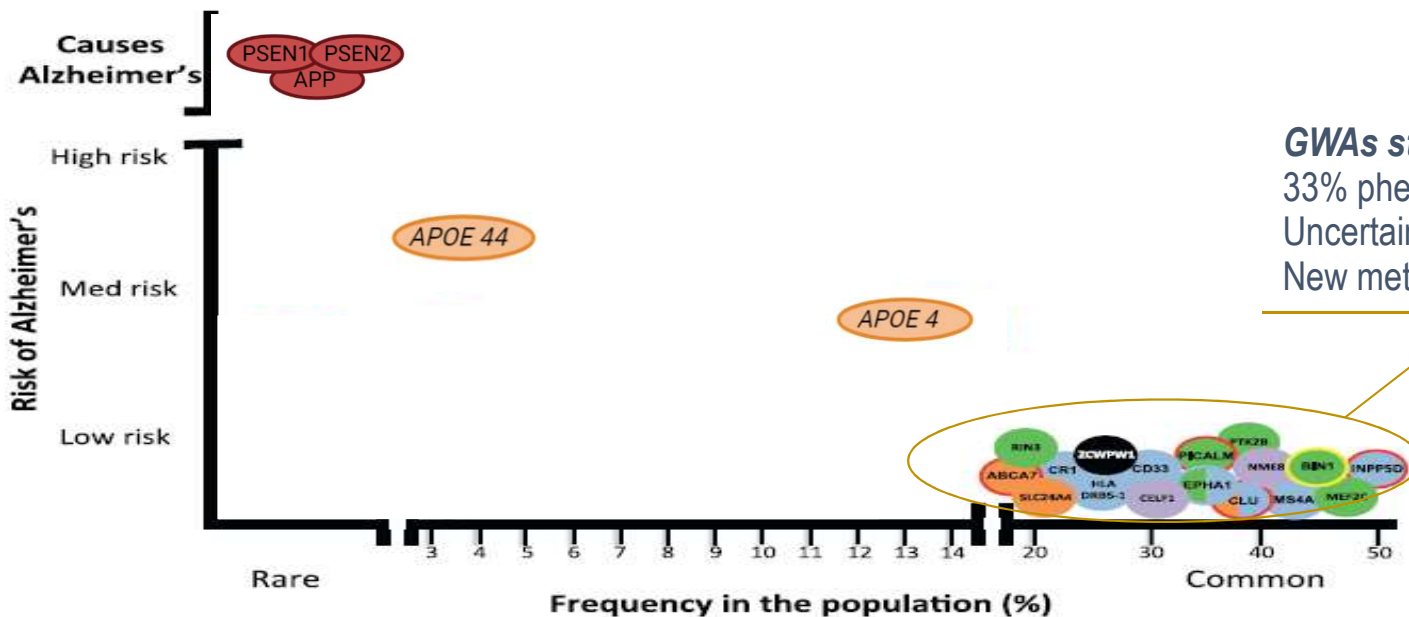CO = 59,163

controls

$APOE\ \mathcal{E}4$
$Freq_{CA}$ = 38.7 %
$Freq_{CO}$ = 13.7 %

$OR = Freq_{CA}/Freq_{CO}$ = 2.82
$P\ (\chi^2)$ = 2.9 × 10^{-45}

Kunkle et al., 2019, Nature Genetics, 51:414–430
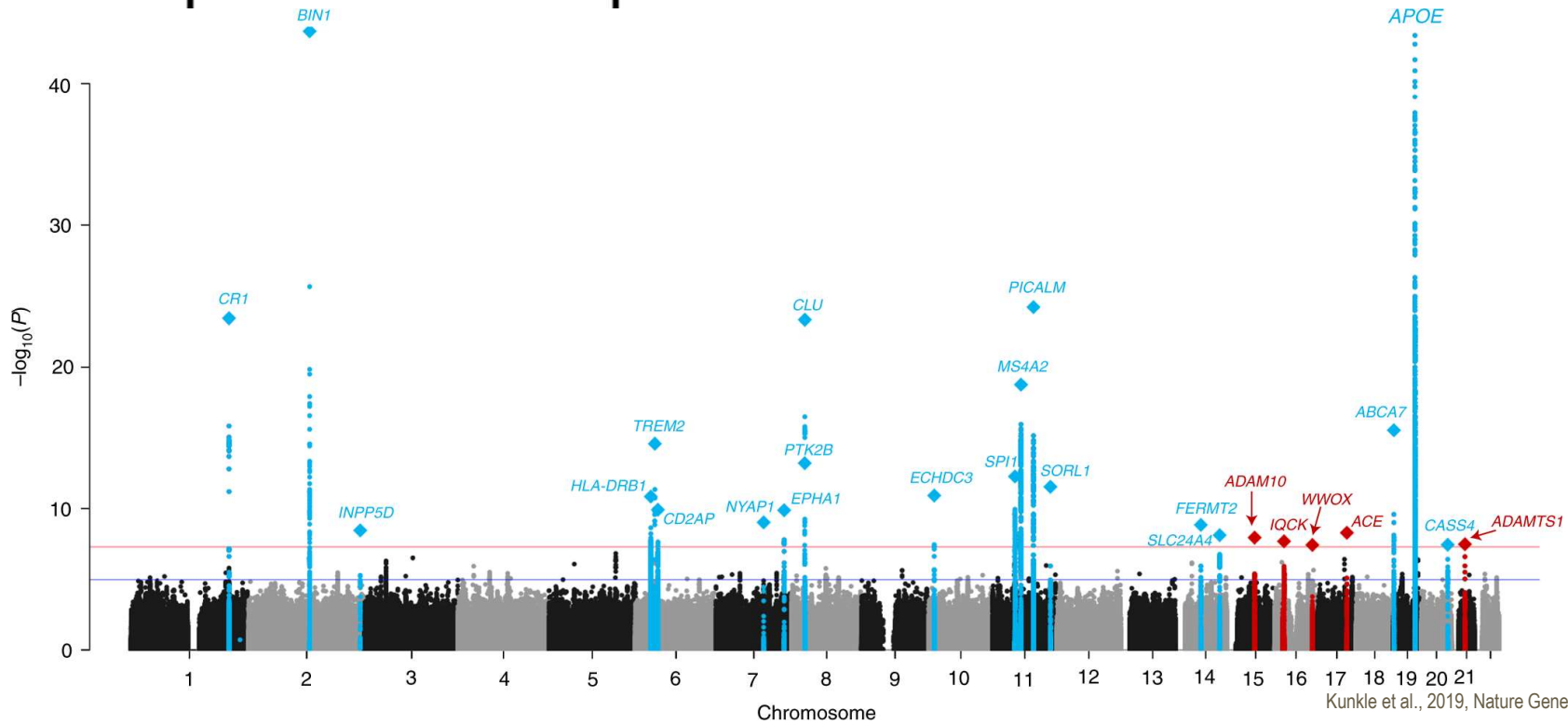
AD genetics
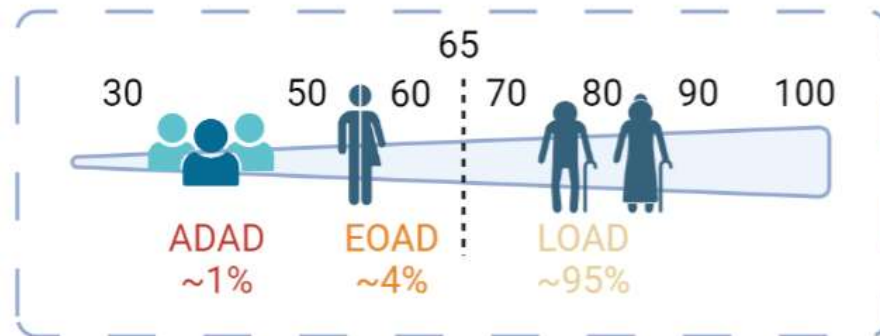
GWAs studies
33% phenotypic variance
Uncertain effect
New metabolic pathways

Neurogenomics & Informatics
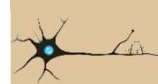
AD pathways
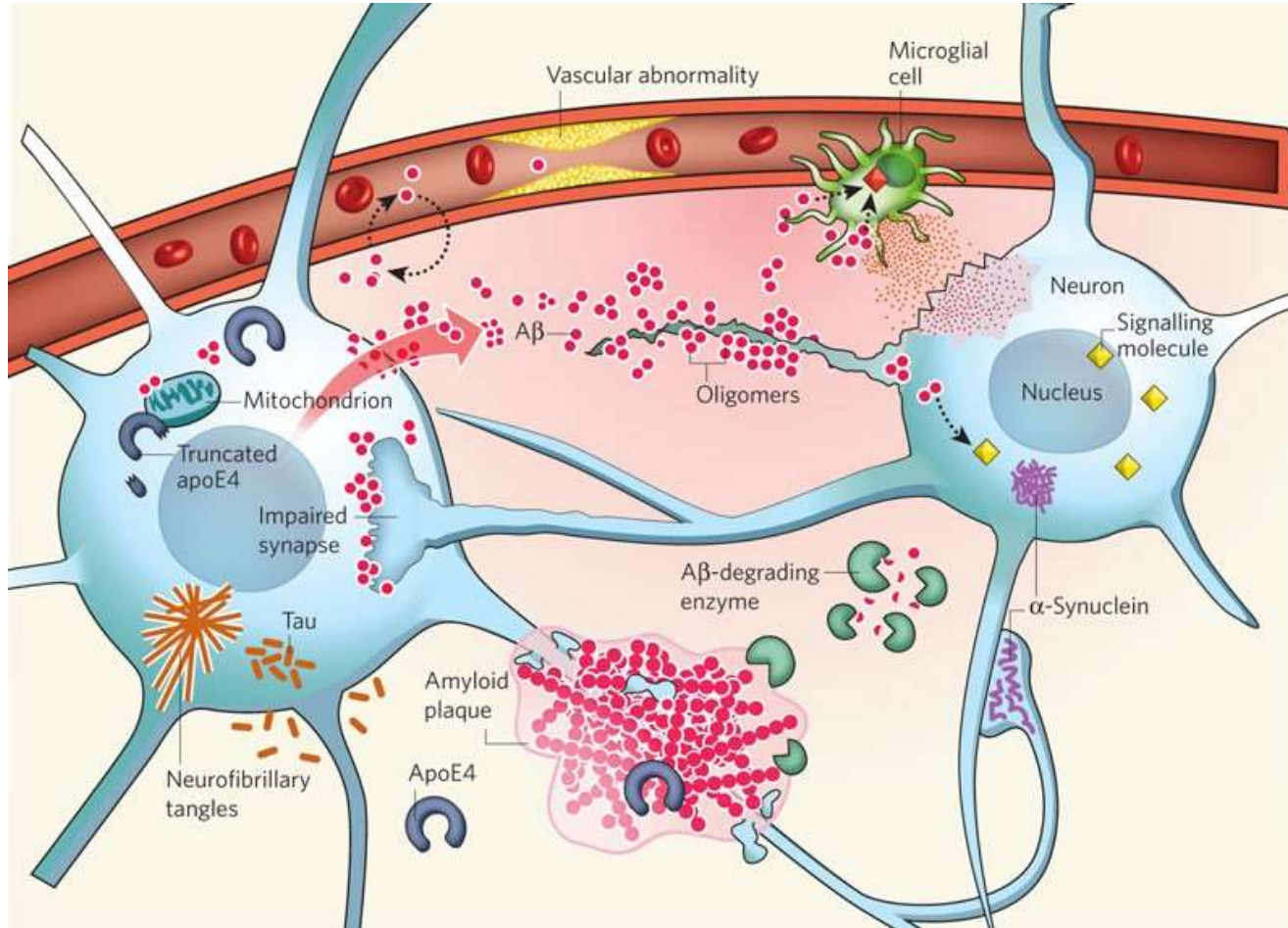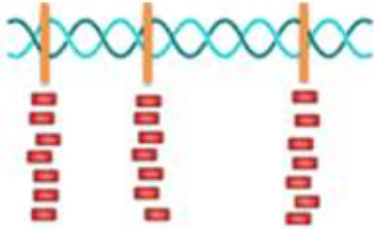

- Immune response
- Endocytosis
- Tau metabolism
- Cytoskeleton / Axon development

https://www.nature.com/articles/461895a

GWAs vs WES vs WGS

Targeted sequencing

- Sequencing region: specific regions (could be customized)
- Sequencing Depth: >500X
- Identify all kinds of variants including SNPs, INDELs in specific regions
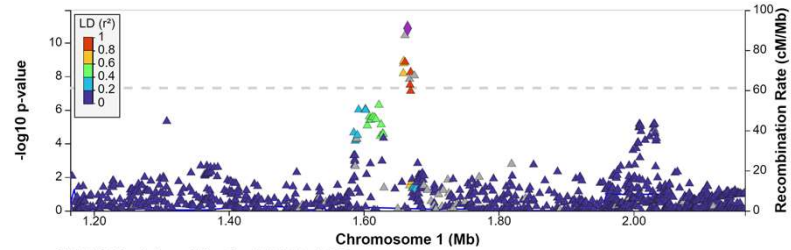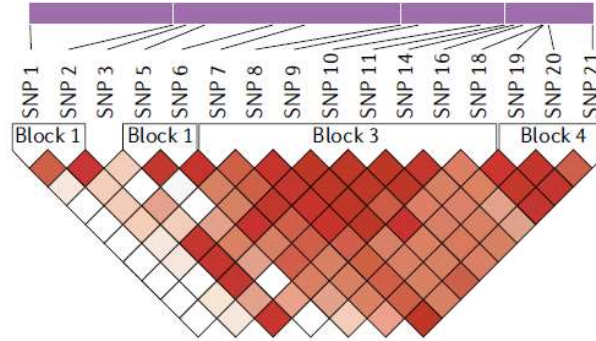- Most Cost effective
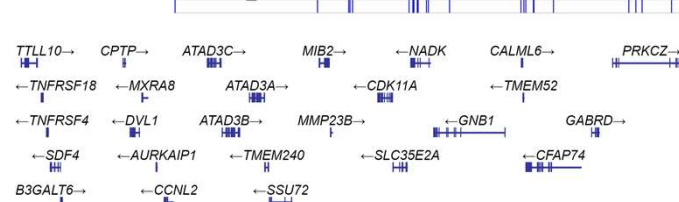
GWAs vs WES vs WGS

Targeted sequencing
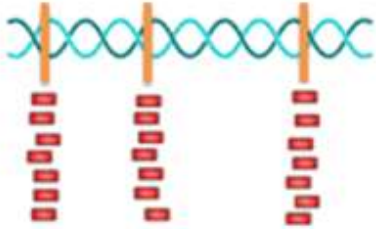
- Sequencing region: specific regions (could be customized)
- Sequencing Depth: >500X
- Identify all kinds of variants including SNPs, INDELs in specific regions
- Most Cost effective

Whole exome sequencing
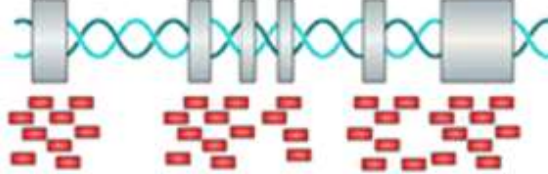
- Sequencing region: whole exome
- Sequencing Depth: >50X ~ 100X
- Identify all kinds of variants including SNPs, INDELs and SV in coding region.
- Cost effective

Whole genome sequencing
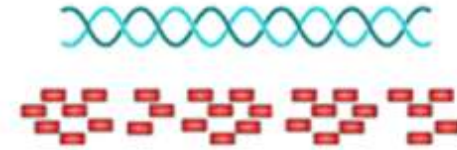
- Sequencing region: whole genome
- Sequencing Depth: >30X
- Covers everything — can identify all kinds of variants including SNPs, INDELs and SV.

Neurogenomics & Informatics

AD genetics

Karch & Goate, 2015, Biol Psych, 77:43-51
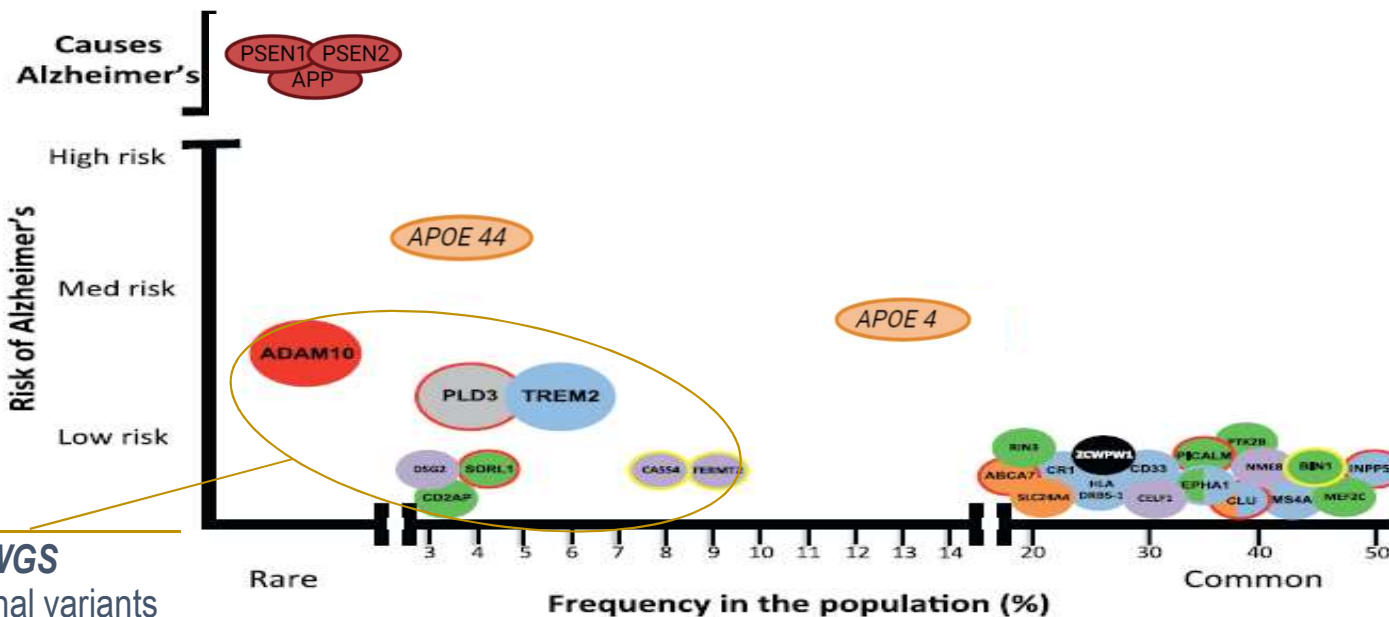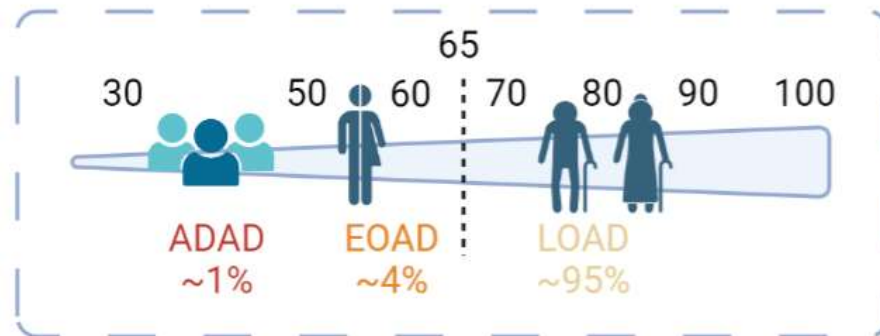
WES-WGS functional variants

AD beyond genetics
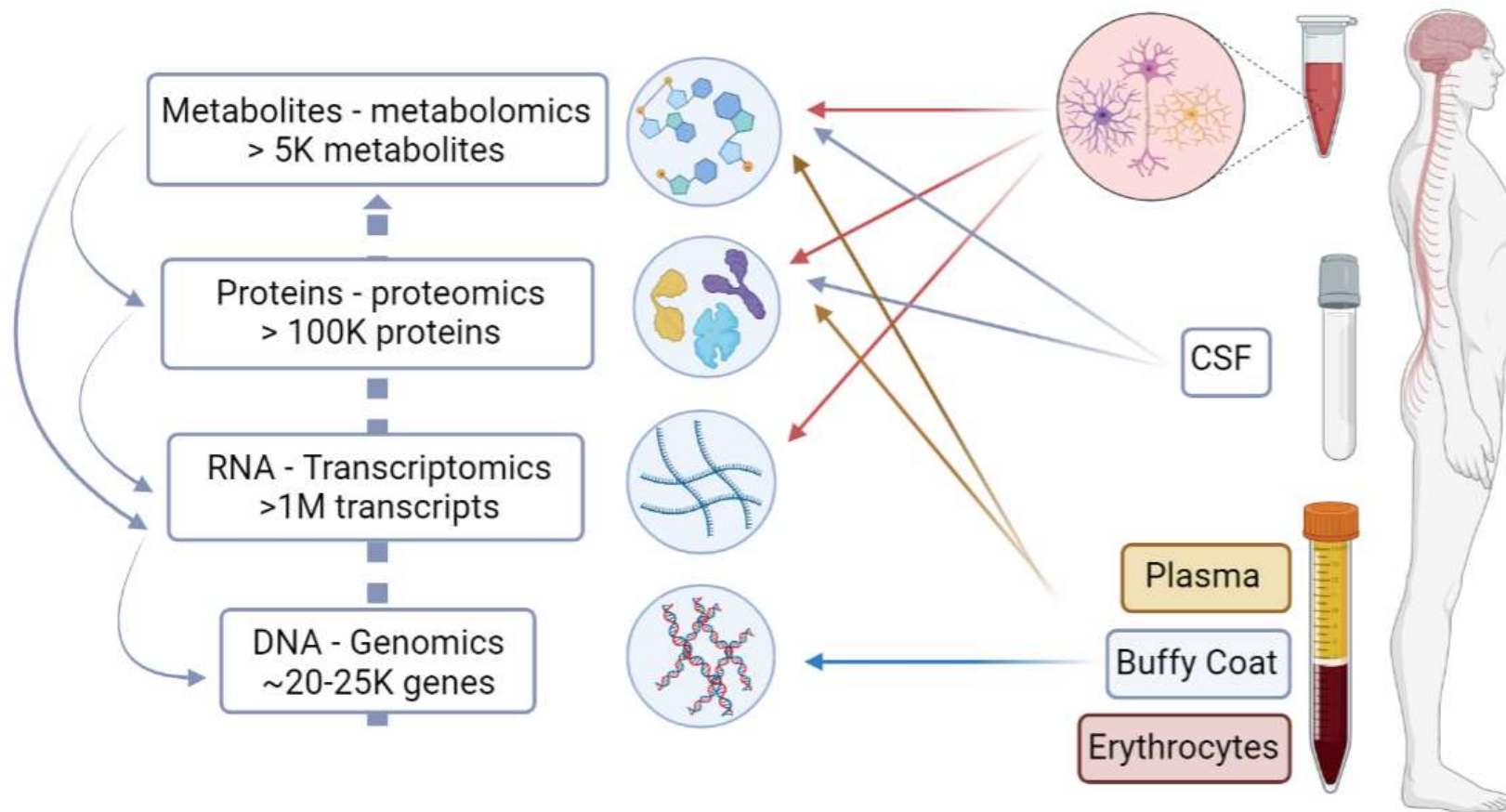
Data integration

Fig. 3: Illustration of functional follow-up of GWAS.

a What are the associated loci?
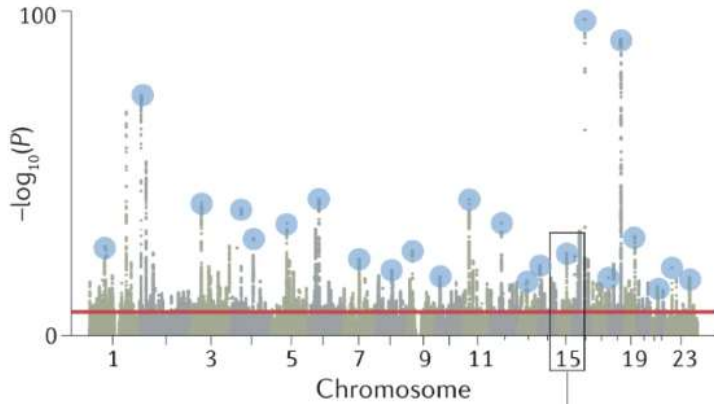
b What are the likely causal variants?
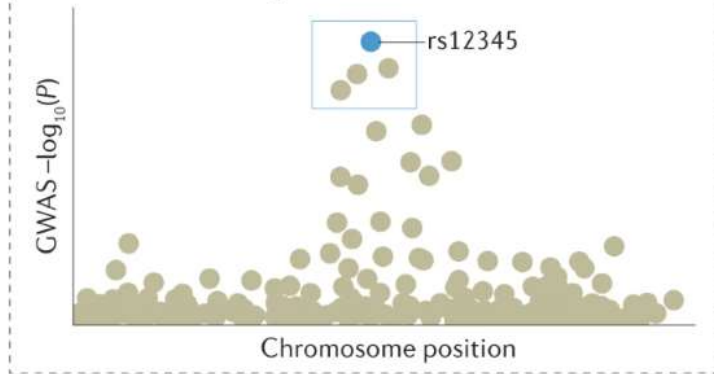
c What are the epigenomic effects of variants?
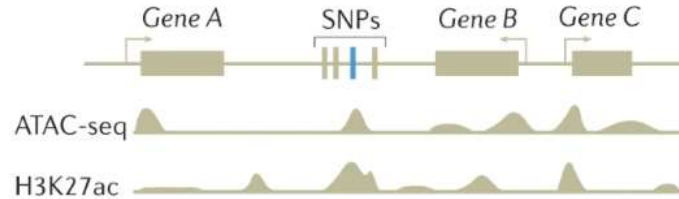
d What are the target genes in the locus?
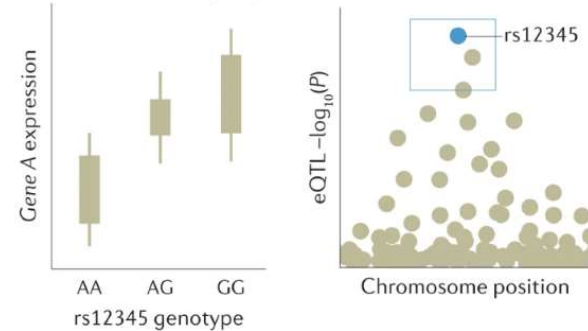
e What are the affected pathways?
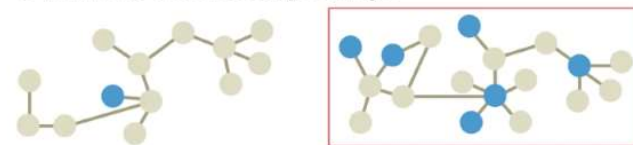
AD beyond genetics
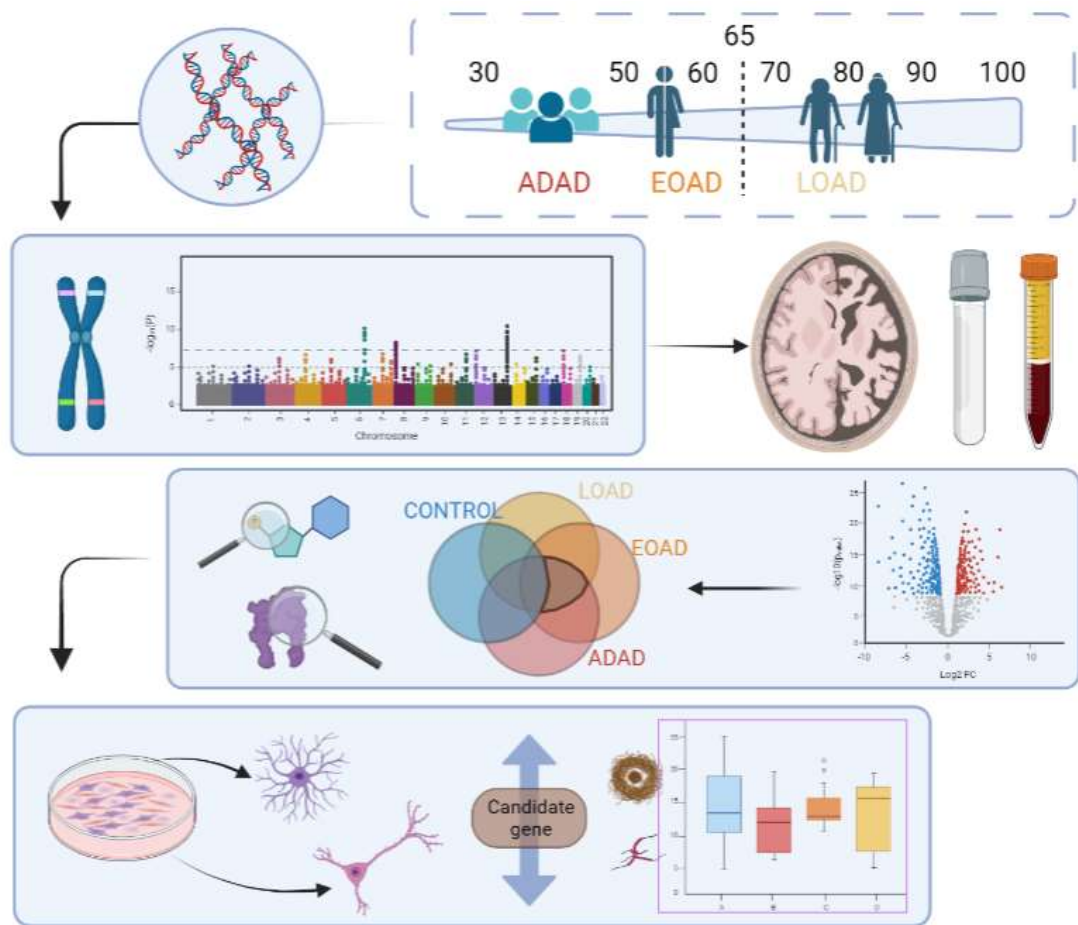
The Familial Alzheimer Sequencing (FASe) Project

A large-scale Genome-Wide Association Study of Early Onset Alzheimer's Disease

OMIC approaches to identify molecular contributors to AD

Datasets

KnightADRC **NIA-LOAD**

Family inclusion criteria
- > 3 affected members
- APOE ε4 does NOT segregate with disease
- Proband does not carry pathogenic mutations in Mendelian genes

Participant inclusion criteria
- Cases   - clinical dementia rating (CDR) > 0.5
  - AAO > 65 yo
- Controls   - CDR = 0
  - ALA > oldest affected AAO within family

	N	Age	%Fe	%APOE-ε4
CA	864	73	52%	71%
CO	426	86	61%	50%

Fam size	2	3	4	5	6	7	8	9	Total
N	7	75	105	58	26	19	2	4	296

Gene-based analysis

Rationale: Single variant test of rare variants have very low power for detecting association

Hypothesis: testing collective effect of a set or fare variants may increase the power

→ **gene-based analysis**

Control sequences

Case sequences

Nomination of novel candidate genes

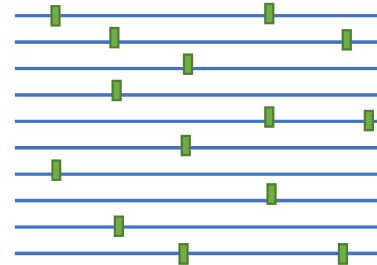|	Collapsing	Variance-component	TDT
	FarVAT-CMC	FarVAT-SKATO	RVGDT
gene	pval	pval	pval
CHRD2	0.007	7.37×10^{-7}	0.990
CLCN2	0.006	1.12×10^{-5}	1.000
NLRP9	2.81×10^{-4}	2.59×10^{-4}	0.998
PTK2B	1.23×10^{-4}	4.93×10^{-4}	1.000
HDLBP	0.021	1.22×10^{-4}	0.996
MAS1L	4.65×10^{-4}	4.23×10^{-4}	0.998
CPAMD8	6.91×10^{-5}	4.23×10^{-4}	9.99×10^{-4}

RESULTS:
- **PTK2B** is a previously reported gene
- **CPAMD8** is gene-wide suggestive by RVGDT, FarVAT-SKATO and FarVAT-CMC
  - 38 rare non-synonymous variants with different degrees of segregation in several families

CPAMD8 - 19:17039023:A:C

Perfect segregation / Imperfect segregation

AAO	--	75	74	--	85	87	89	--	--	67	65	60	65	--
ALA	78	86	69	91	92	91	93	66	68	72	77	68	74	71
APOE	33	33	43	32	43	32	43	43	33	44	43	33	33	43

Neurogenomics & Informatics

Fernandez et al. 2018
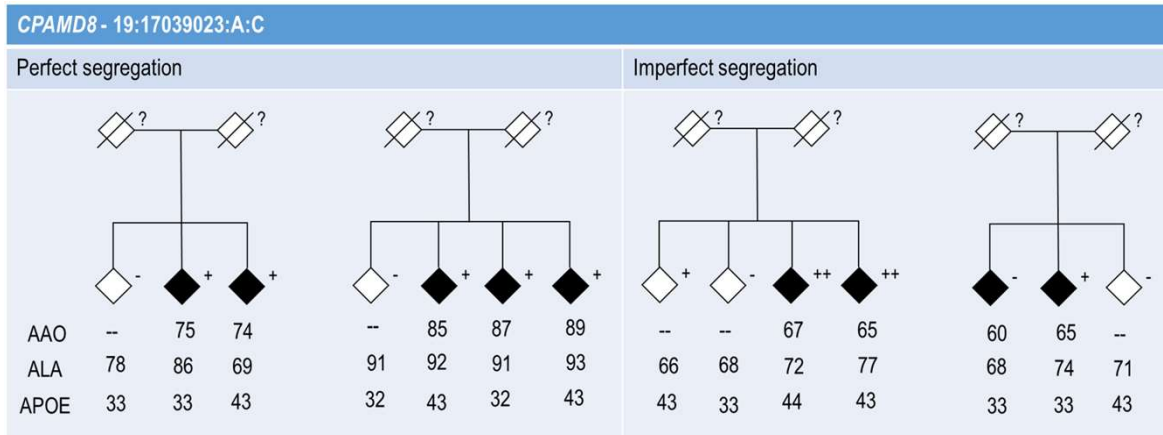
Higher expression of *CPAMD8* in cases than controls
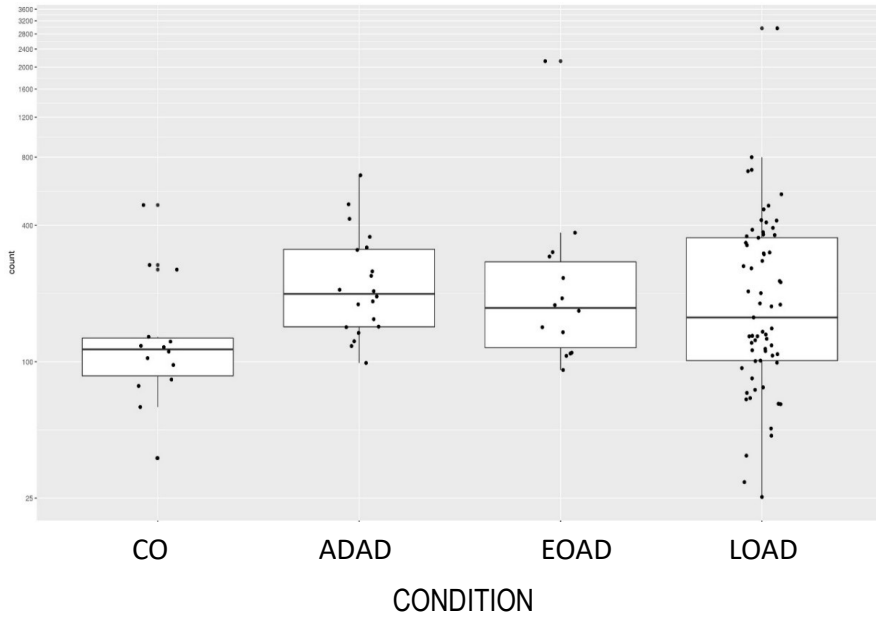
In-house transcriptomic data from 103 parietal tissue
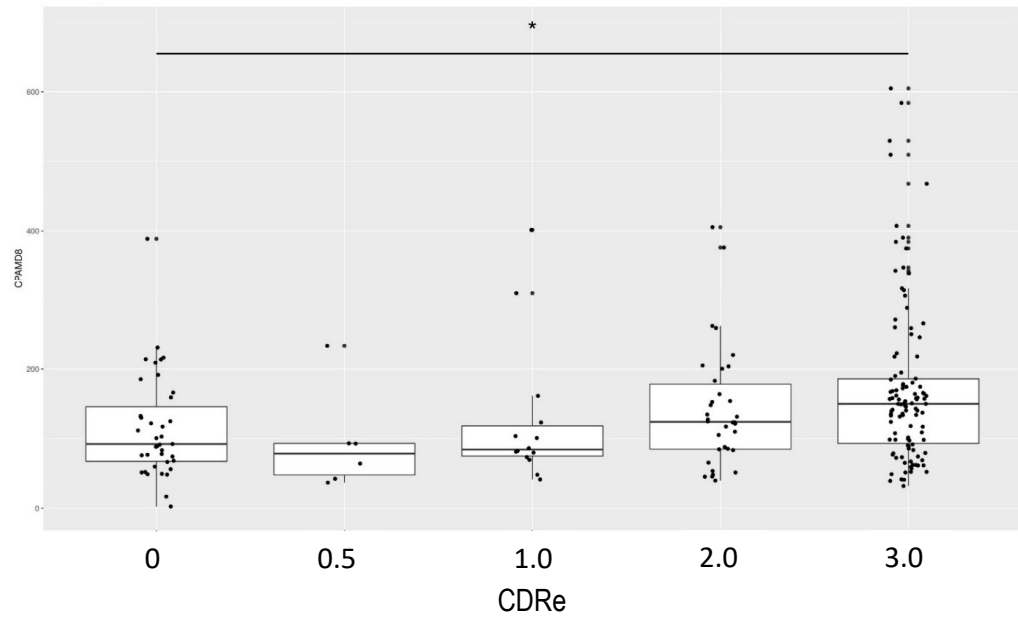
Transcriptomic data from MSBB, Inferior Frontal Gyrus (BM44)
CDRe (Kruskal-Wallis p=8.45×10-3)

CPAMD8 interacts with APP

CPAMD8 alters extracellular Aβ40, Aβ42

CPAMD8 alters APP processing
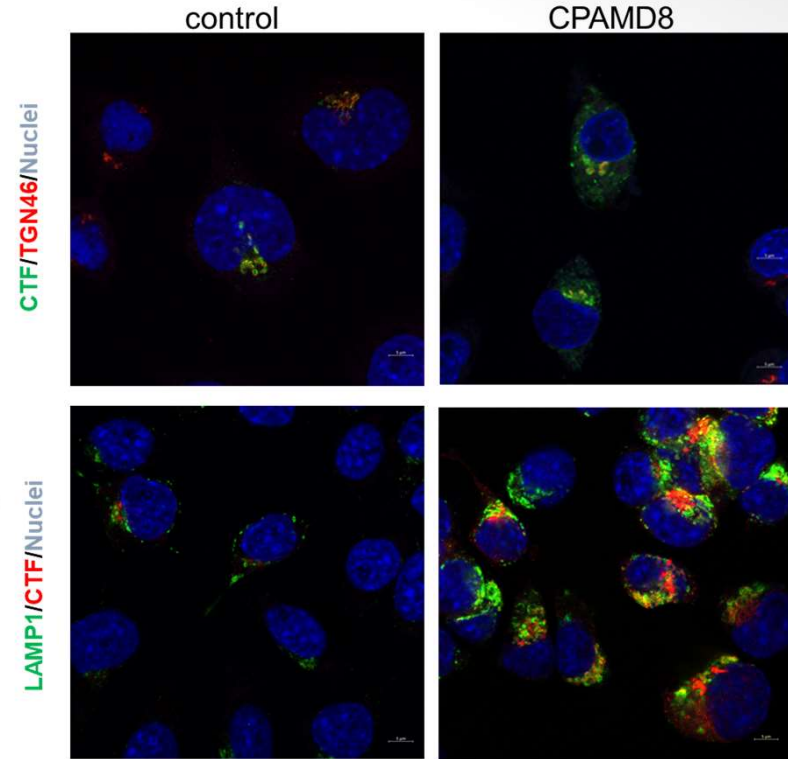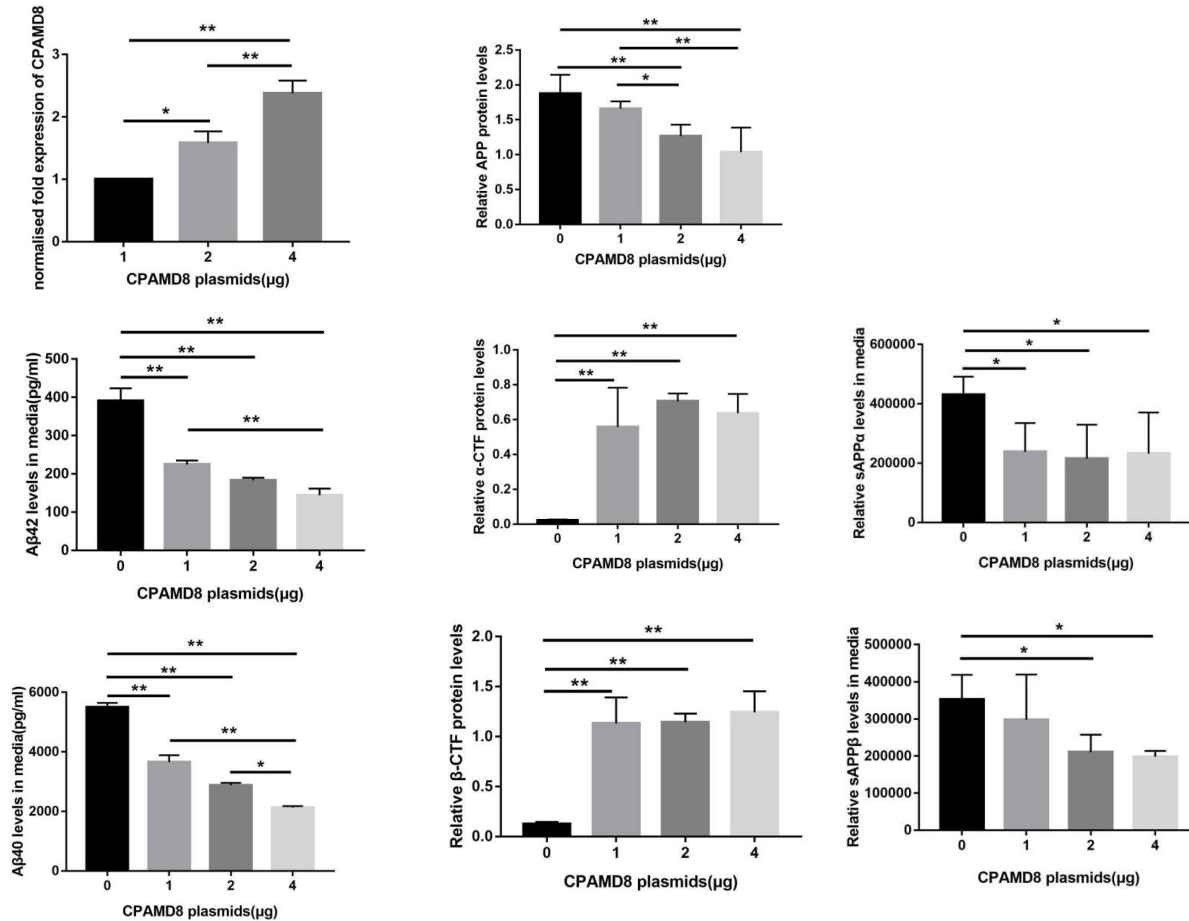
Co-localization of CPAMD8 and APP

CPAMD8 acts via the authopahgy-lysosome
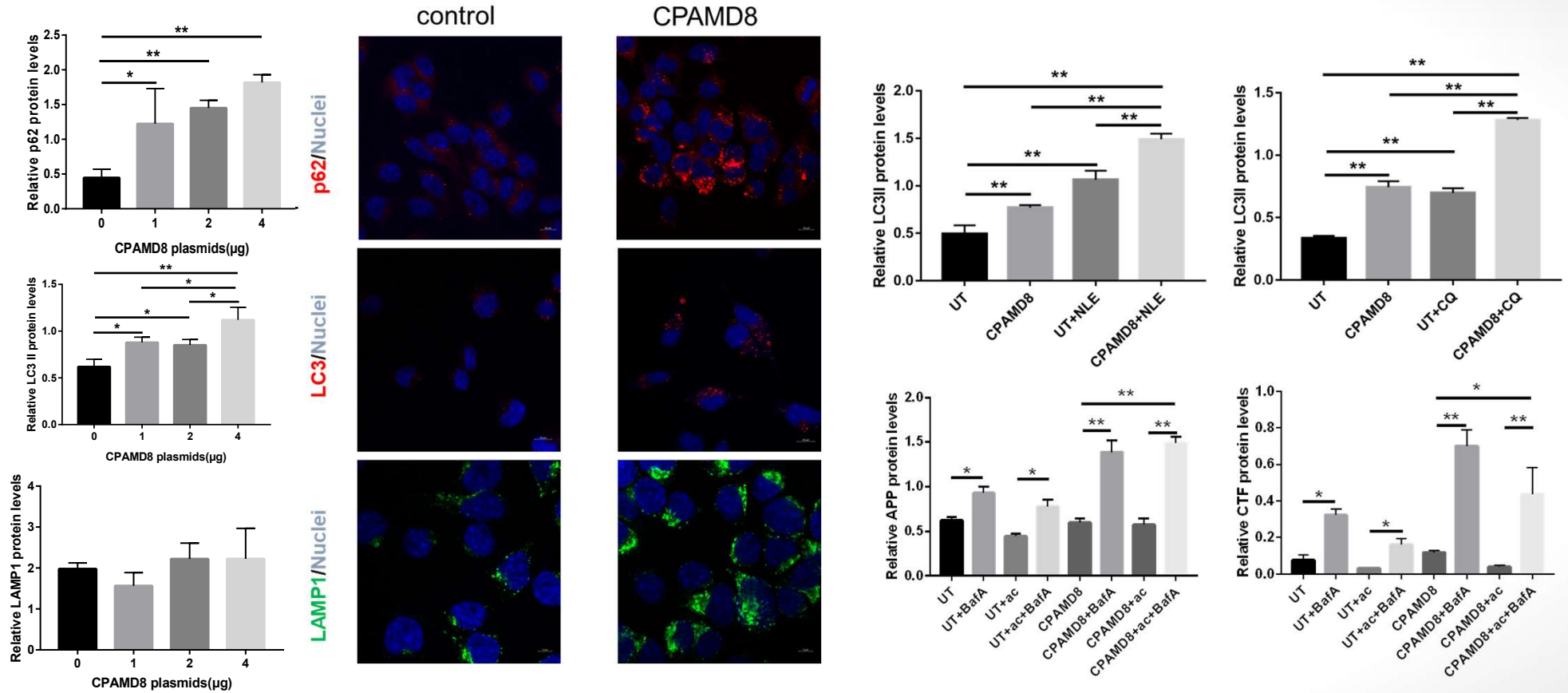
Increased CPAMD8 and lysosomal components

CONCLUSIONS

- Gene-based and segregation analysis suggests CPAMD8 influences AD risk
  - Found variants that segregate in several families

- Higher expression in cases compared to controls (based on status and CDR)

- *CPAMD8* alters APP processing through the regulation of the autophagy-lysosome pathway

Neurogenomics & Informatics

Multi-ethnic meta-analyses for Early Onset Alzheimer's Disease

Joseph Bradley,
PhD student

COLUMBIA
UNIVERSITY

Washington
University in St.Louis

UHealth

R01AG064614

Rationale

- About ~5% of Alzheimer disease cases have an early onset <65-70 yo (**EOAD**)
- Only 1% of EOAD is caused by mutations in APP, PSEN1 or PSEN2 (ADAD)
- Majority of EOAD heritability remains unexplained; most studies target LOAD

Aims

- Identify additional EOAD-associated variants through large-scale sequencing data
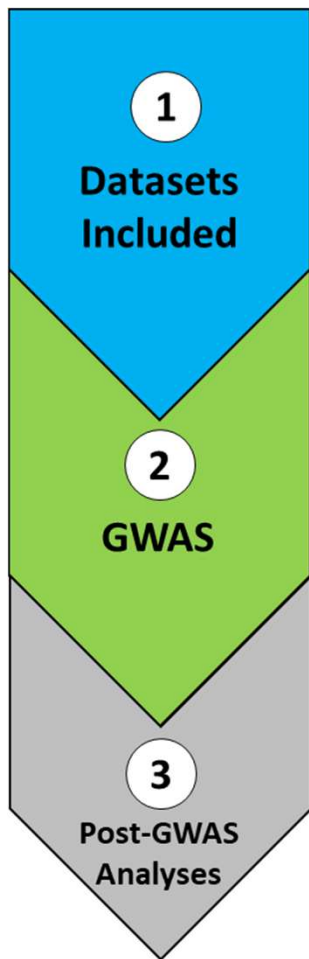- We used data from well characterized groups as well as publicly available data:
  - Knight-ADRC
  - GCAD

Analysis Workflow

Neurogenomics & Informatics

1 Datasets Included

Non-Hispanic White (NHW, Phenotype + Genotype) N=19,668

African American (AA, Phenotype + Genotype) N=4,445

Asian (AZN, Phenotype + Genotype) N=1,213

NHW (Summary Statistics) N=788,989
Bellenguez et al., 2022 (LOAD)

AA (Summary Statistics) N=7,970
Kunkle et al., 2021 (LOAD)

AZN (Summary Statistics) N=8,036
Shigemizu et al., 2021 (LOAD, Japanese)

2 GWAS

NHW-specific Joint analysis (Status ~ SNP + Sex + PC1-10)
GWAS loci: Chr1(×2), Chr2(×2), Chr4(×2), Chr6(×2), Chr7, Chr9, Chr10, Chr11(×2), Chr15, Chr19(×3), Chr20, Chr22

AA-specific Joint analysis (Status ~ SNP + Sex + PC1-10)
GWAS loci: Chr10, Chr12, Chr19, Chr21

AZN-specific Joint analysis (Status ~ SNP + Sex + PC1-10)

Multi-Ancestry meta-analysis N=25,326
(NHW=19,668, AA=4,445, AZN=1,213)
New Genome-wide significant (GWS) loci: Chr19
Fixed effects, sample size-weighted

3 Post-GWAS Analyses

Annotation, Gene based analysis, QTL mapping
Variant annotation MAGMA gene-based analysis eQTL mapping

Gene Prioritization
Summarize colocalization, Annotation, and gene-based analysis. Score to find causal gene in each locus

Trans-ethnic Fine-mapping
Trans-ethnic annotation of GWS SNPs from each ethnicity in each other ethnicity

EOAD LOAD Overlap
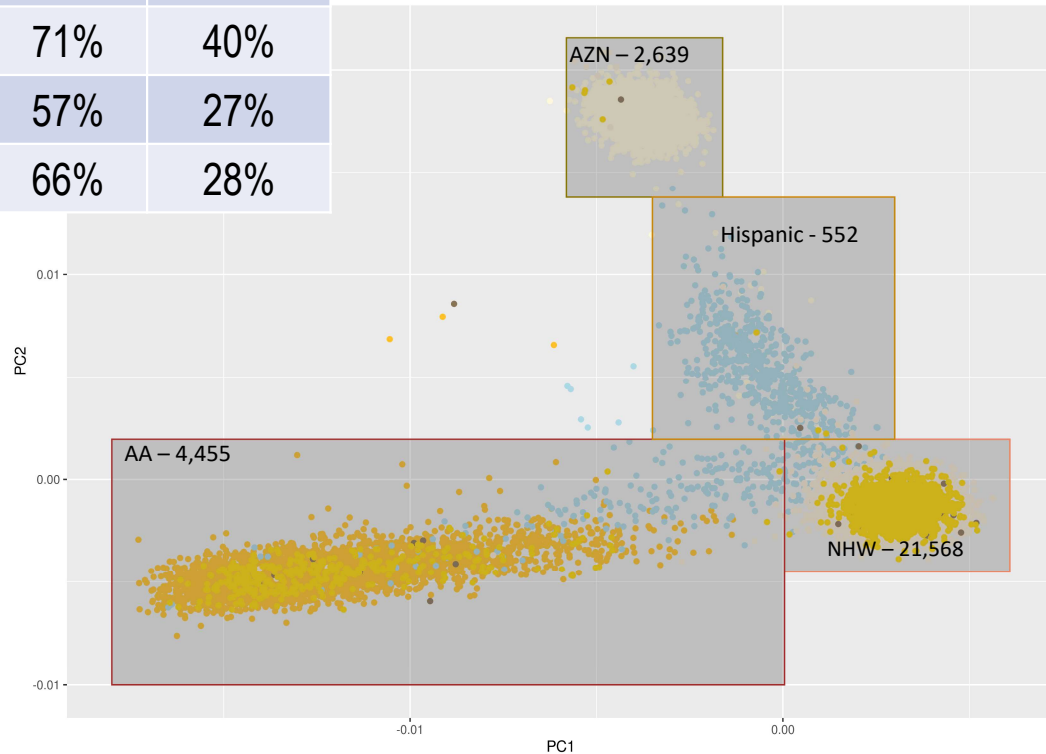Quantify genetic overlap of LOAD and EOAD with LDSC, PRS, effect size correlation, and top hits overlap

Pathway analysis
FUMA pathway analysis and mQTLs colocalization

Summary of demographics and analysis framework

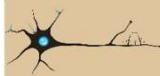|	# cases (CA)	# control (CO)	% Fem	% APOE
NHW	6,282	13,386	57%	26%
Afr Am	782	3,663	71%	40%
Asian	617	1,714	57%	27%
Hispanic	280	270	66%	28%

Non-Hispanic Whites
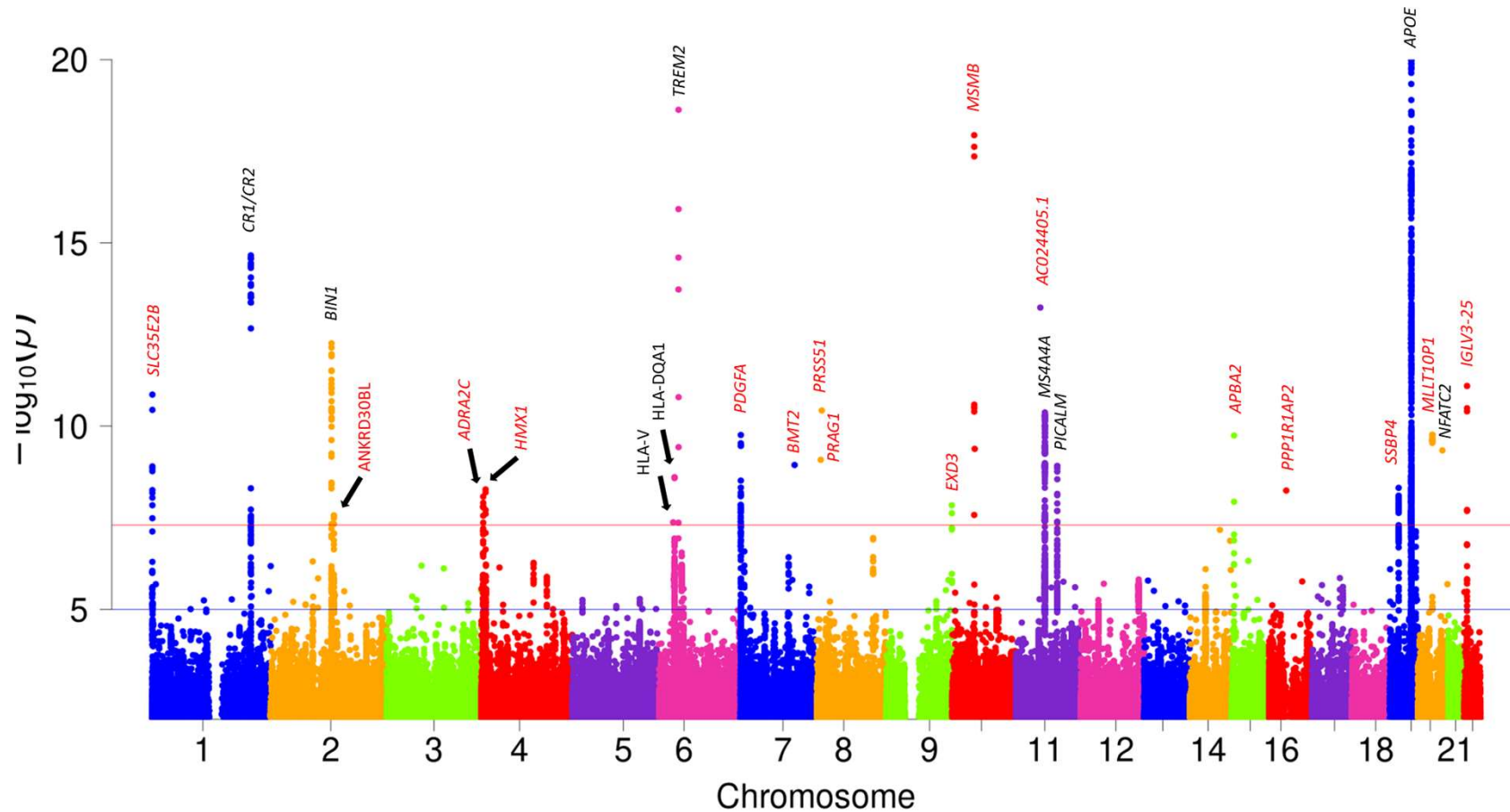
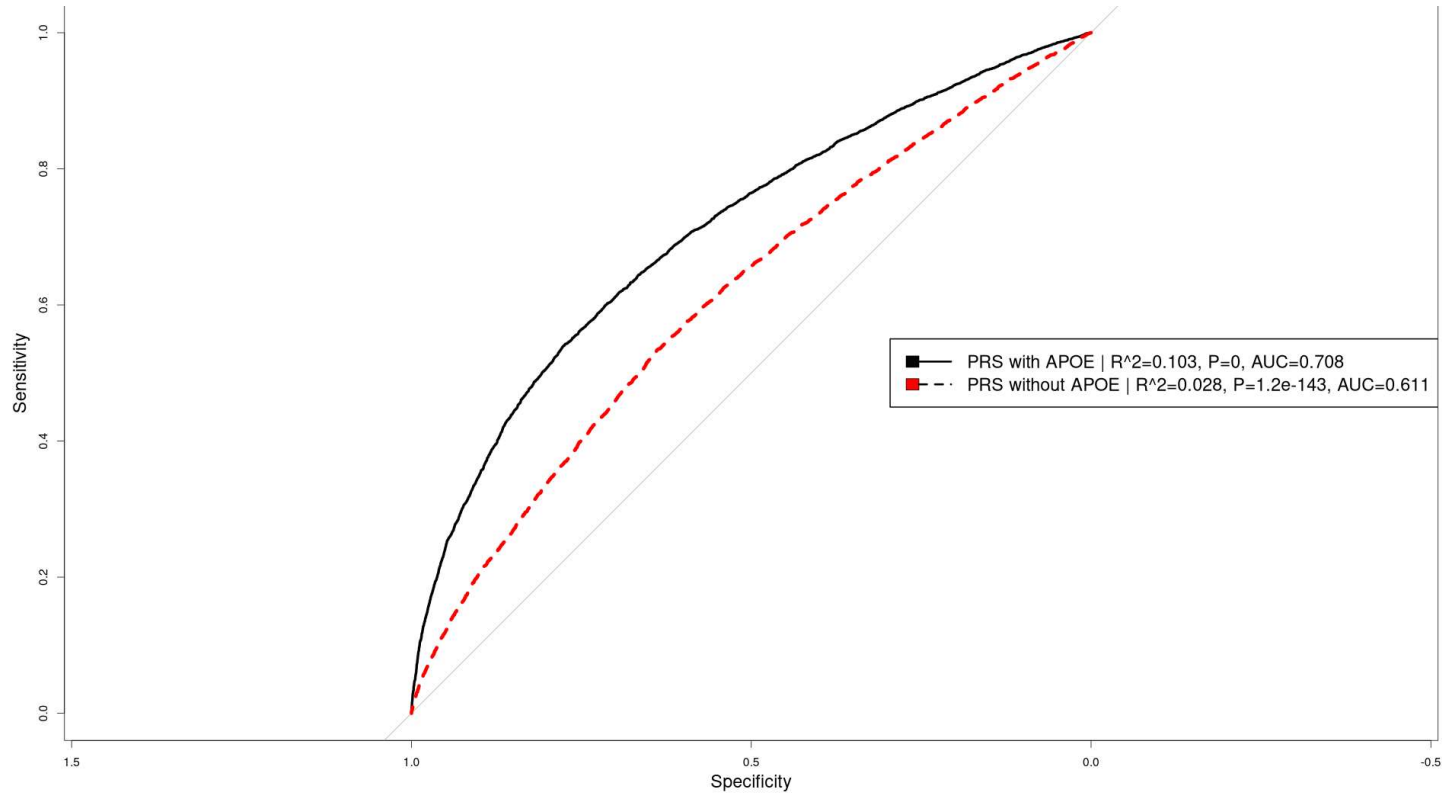NHW, n=19,668: λ=1.053 Covars: Sex, PC1-10 - Identified 25 total/16 novel Significant Loci

Novel

PRS for LOAD has limited prediction for EOAD
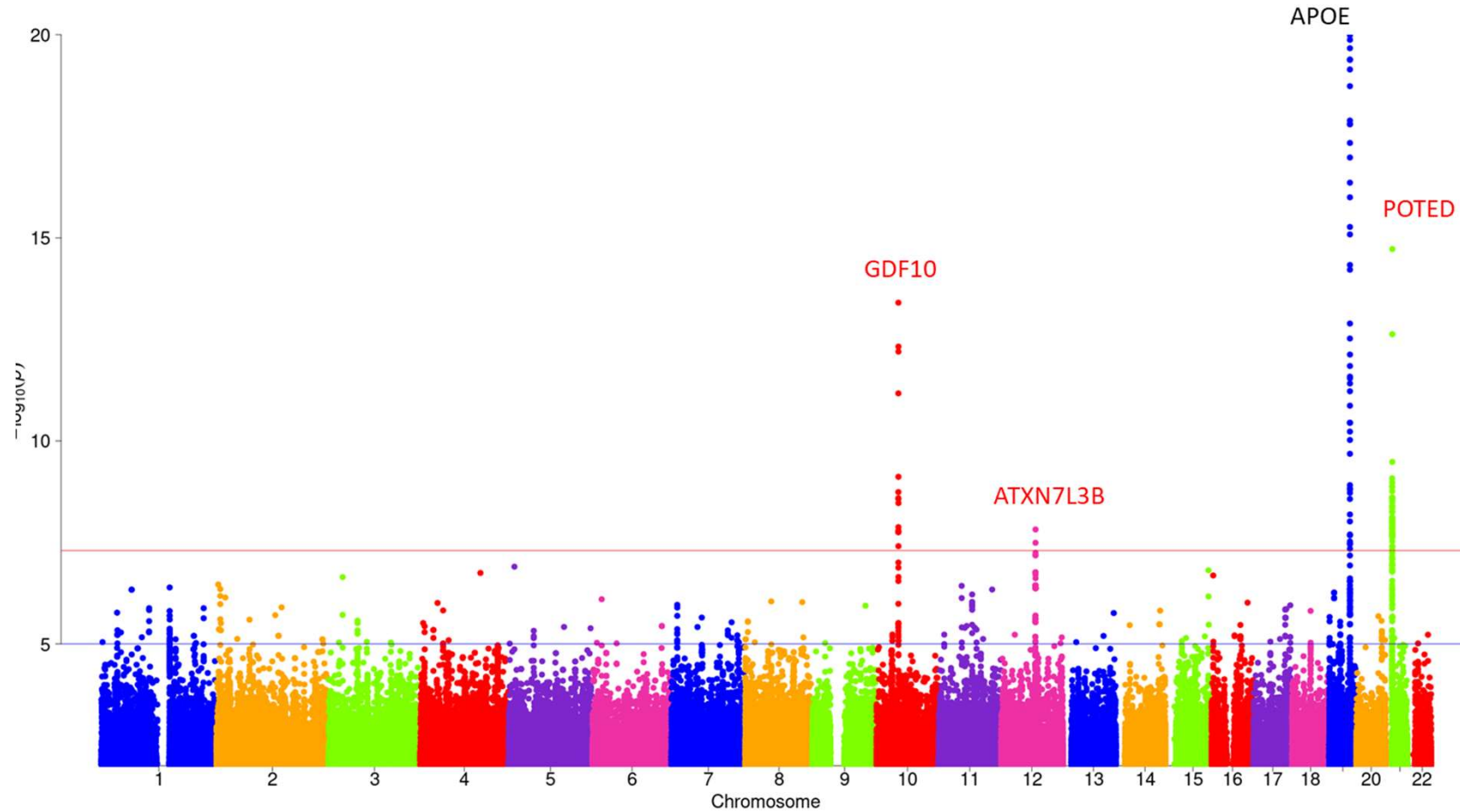
PRS with APOE | R^2=0.103, P=0, AUC=0.708
PRS without APOE | R^2=0.028, P=1.2e-143, AUC=0.611

African Americans

AA, n=4,445: λ=1.008 Covars: Sex, PC1-10

Novel

Asian

APOE

AZN, n=1,213: λ=1.008 Covars: Sex, PC1-10

Neurogenomics & Informatics
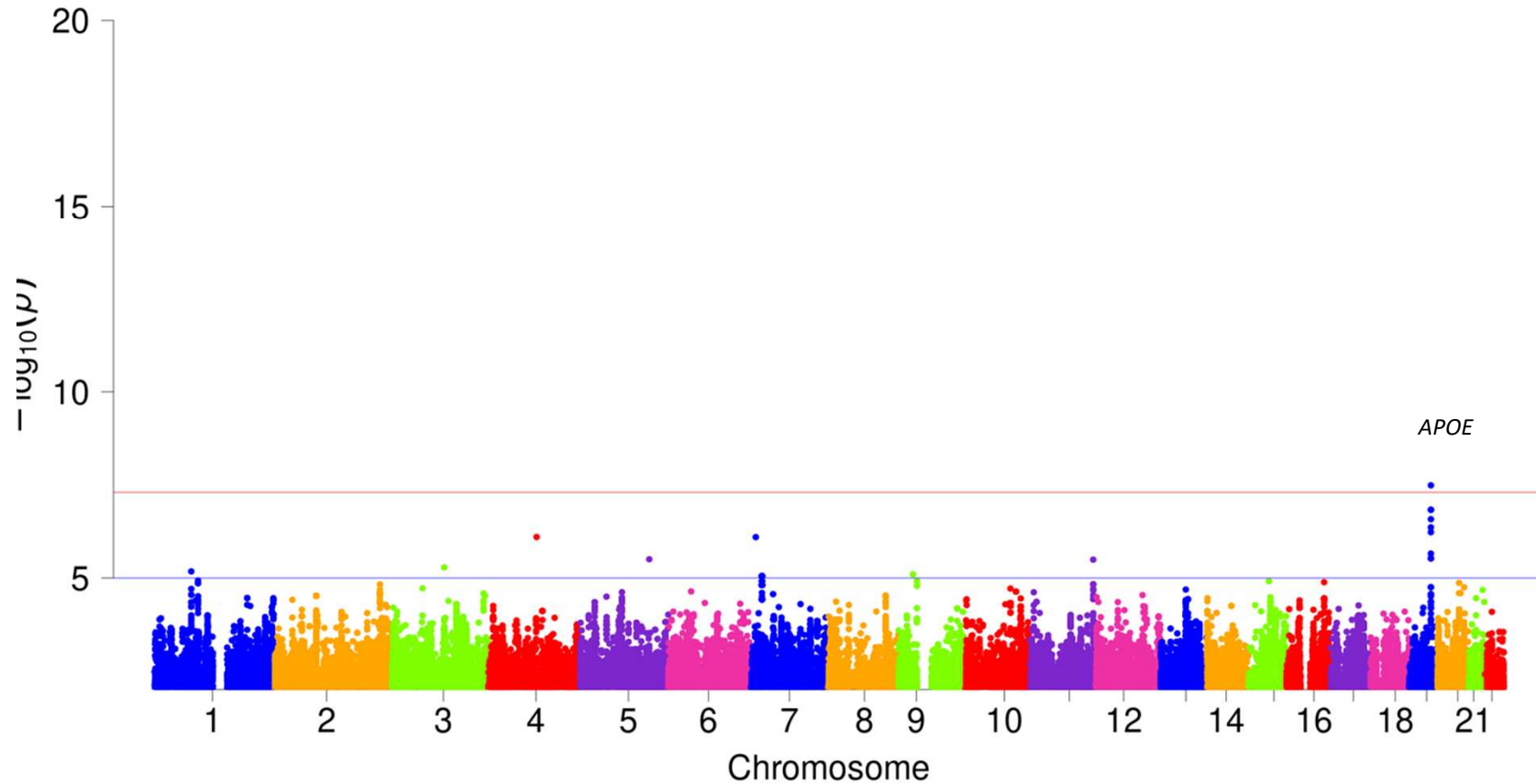
Meta-analysis – reveals two new hits
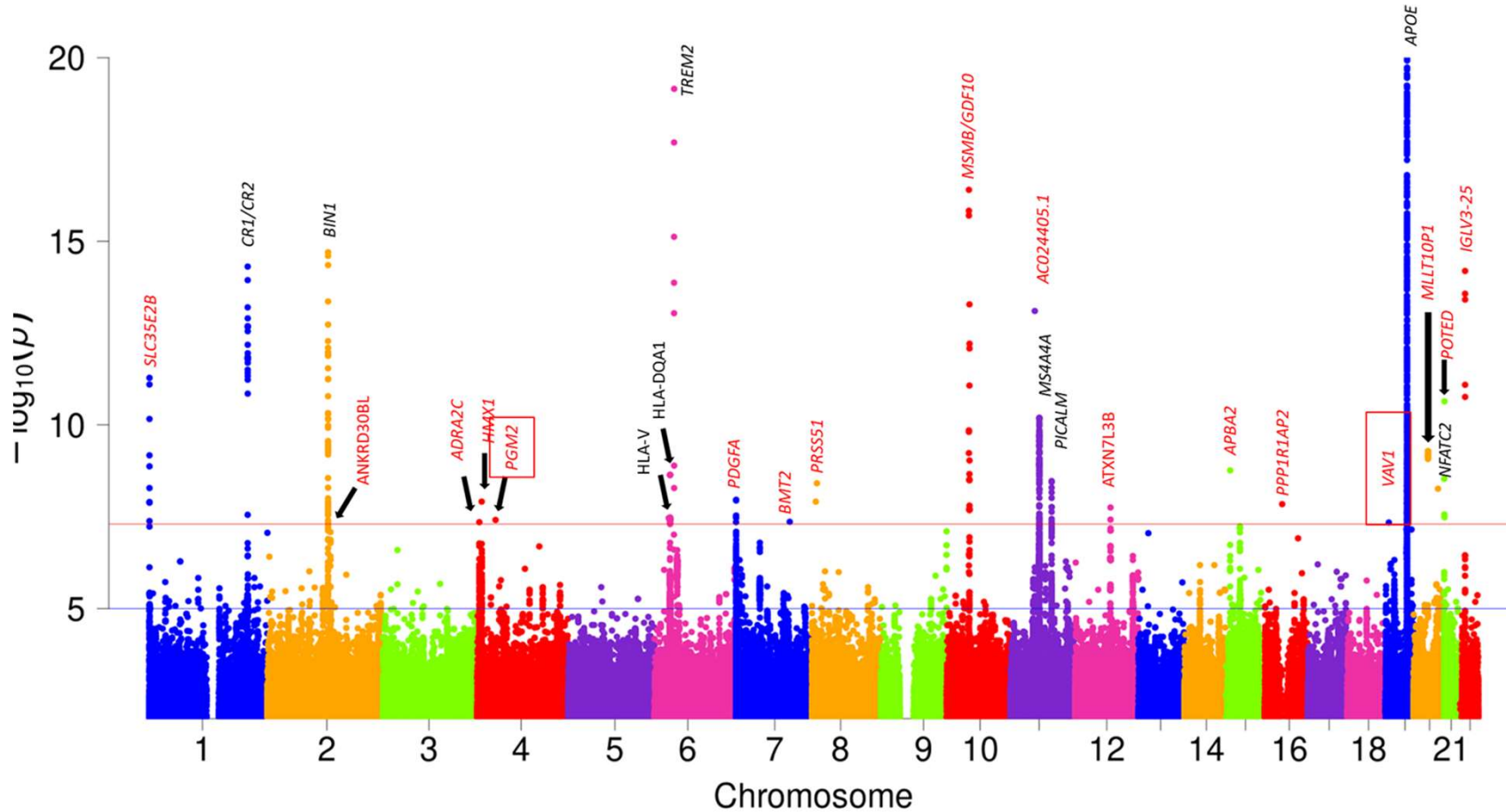
Multi-ethnic Meta-analysis, n=25,326 sample size weighted

Conclusions

We found 25 Significant loci in NHW

• Some replicate in LOAD but most are novel

• LDSC shows significant genetic overlap with LOAD, but unique signals and moderate AUC/R2 from PRS suggest there's still a lot of difference in genetic background

We find 4 significant loci in AA and one which are significant in Asian

• All AA loci except APOE are novel.

• No novel AZN loci

We are currently combining this information with transcriptomics and proteomics data to perform QTL mapping and annotation, to determine like and probable functional genes in each locus

OMIC approaches to identify molecular contributors to AD

Jessie Sanford
Bioinformatic Scientist

Fernandez LAB

Washington
University in St.Louis

Anirudh Sivaraman
Research Technician

39

Bulk Brain transcriptomic data

Condition	N	Age	AAD	PMI	% Fe	CDR				Braak			
						<0.5	1-2	3	NA	1-2	3-4	5-6	NA
ADAD	19	44	52	14.45	42%	0	4	4	11	0	0	15	4
EOAD	13	62	76	11.15	54%	0	1	12	0	0	0	10	3
LOAD	55	77	87	12.88	56%	5	20	30	0	4	9	35	7
AD	87	72	78	12.97	53%	5	25	46	11	4	9	60	14
HC	16	78	86	10.20	63%	15	1	0	0	14	2	0	0

* N= sample size; age = age at onset for AD and age at last assessment for HC; ADA=average age at death; CDR: average CDR at death; Braak: aveage braak; PMI: post mortem interval.

Generated transcriptomic data from 103 parietal tissue

Differential gene expression (DGE) analysis, controlling for PMI, TIN and gender, identified 57 genes [p<0.05 after FDR (padj)] that are differentially expressed between AD and HC.

DGE identifies 57 genes
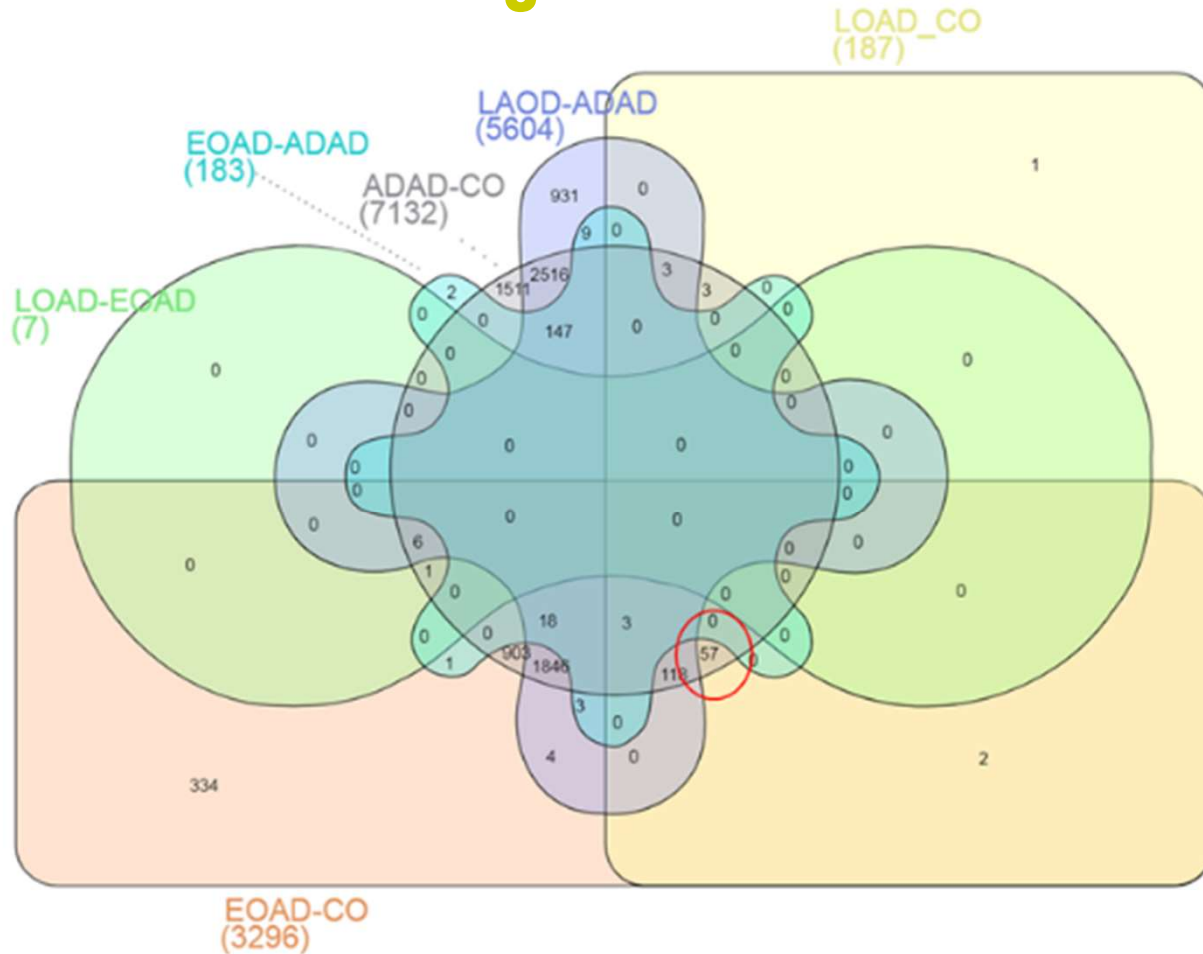
Venn diagram of differentially expressed genes across AD etiologies and against HC.

The 57 genes in the red circle are: PLXNA3,ARMCX5,TTC38,LSS,SMTN,NOL4L,PALM3,TMED1,LZTS2,AQP1,LIX1L, ANXA2R,SCLT1,GCM1,INPPL1,ADAMTS2,PFKFB4,GLT8D1,RASL12,BCORL1,S100A4,STAT1,TRIM33,SMG5,SLC44A5,ELN,NOL10,TSN,ANP32B,GDF1,SYNGR2, GAREM2,PDE7B,JUND,IQCH,NAGK,SLC12A9,GLI2,IFIT3,IGFBP5,PNPLA6,SIRT5, AZI2,KSR1,FAM84B,PDE10A,BSPRY,FIBCD1,TRIM65,EPS8L2,TPRA1,RHOF,TGFB3,**AGFG2**,CRAMP1,ZBTB12,TMC6

AGFG2 is under GWAs signal
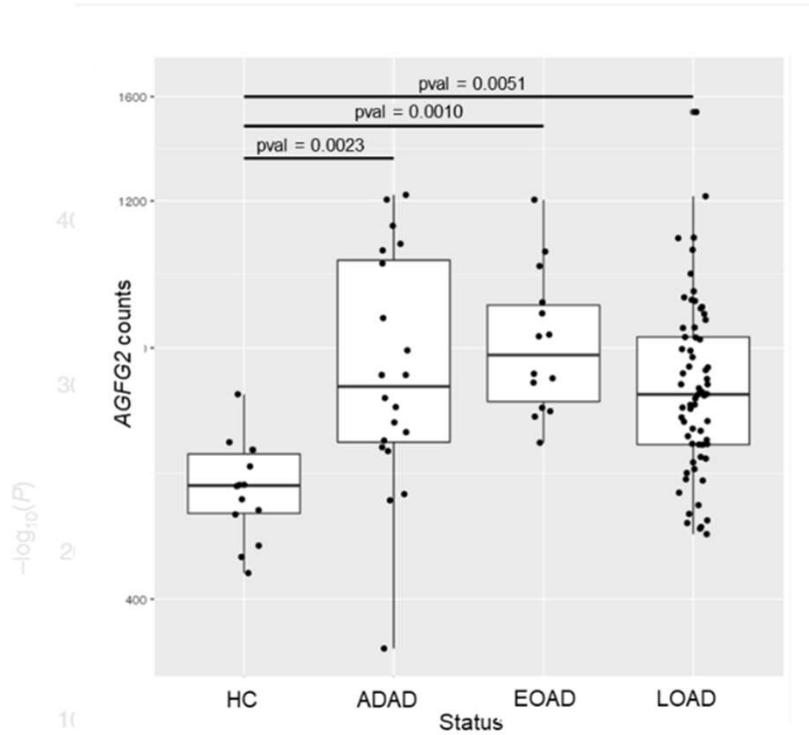
AGFG2 gene counts across HC and AD etiologies

Lambert et al. (2013)
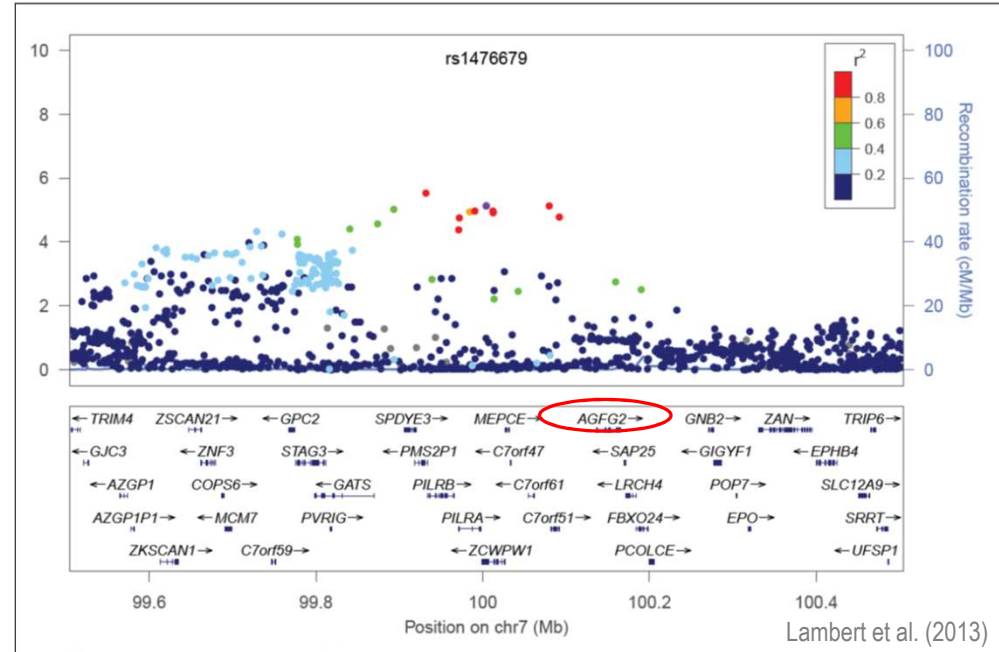
Replication in independent datasets

Table 1. Kruskal-Wallis test of *AGFG2* differential expression values in MSBB, Mayo and ROSMAP dataset per each brain area tissue available and the following phenotypes: CDRe (or CogDx), BraakTau, Status and Etiology.

Dataset	Brain Tissue	CDRe	BraakTau	Status	Etiology
Knight-ADRC (discovery)	Parietal	6.44×10^{-05}	5.21×10^{-03}	8.81×10^{-06}	4.11×10^{-05}
MSBB	BM10	5.02×10^{-05}	0.06	8.63×10^{-03}	0.07
	BM22	2.01×10^{-04}	0.003	9.98×10^{-03}	0.062
	BM36	6.39×10^{-06}	5.96×10^{-04}	7.74×10^{-04}	1.53×10^{-03}
	BM44	4.21×10^{-05}	5.09×10^{-03}	9.49×10^{-05}	1.48×10^{-04}
Mayo	Cerebellum	NA	NA	3.36×10^{-03}	1.10×10^{-02}
	Temporal Cortex	NA	NA	5.88×10^{-12}	4.77×10^{-11}
ROSMAP	DLPC	1.62×10^{-04}	0.052	8.34×10^{-05}	3.96×10^{-05}

Kinght-ADRC

CO AD

MSBB-BM36

CO PS-AD AD

Mayo - TCX

CO AD

ROSMAP

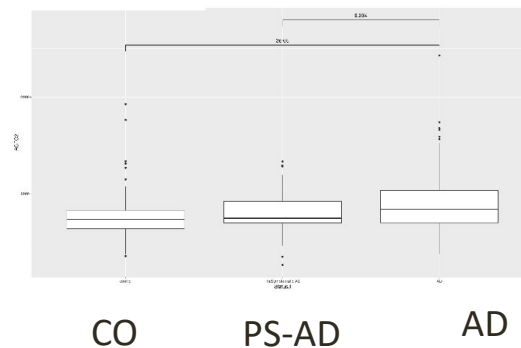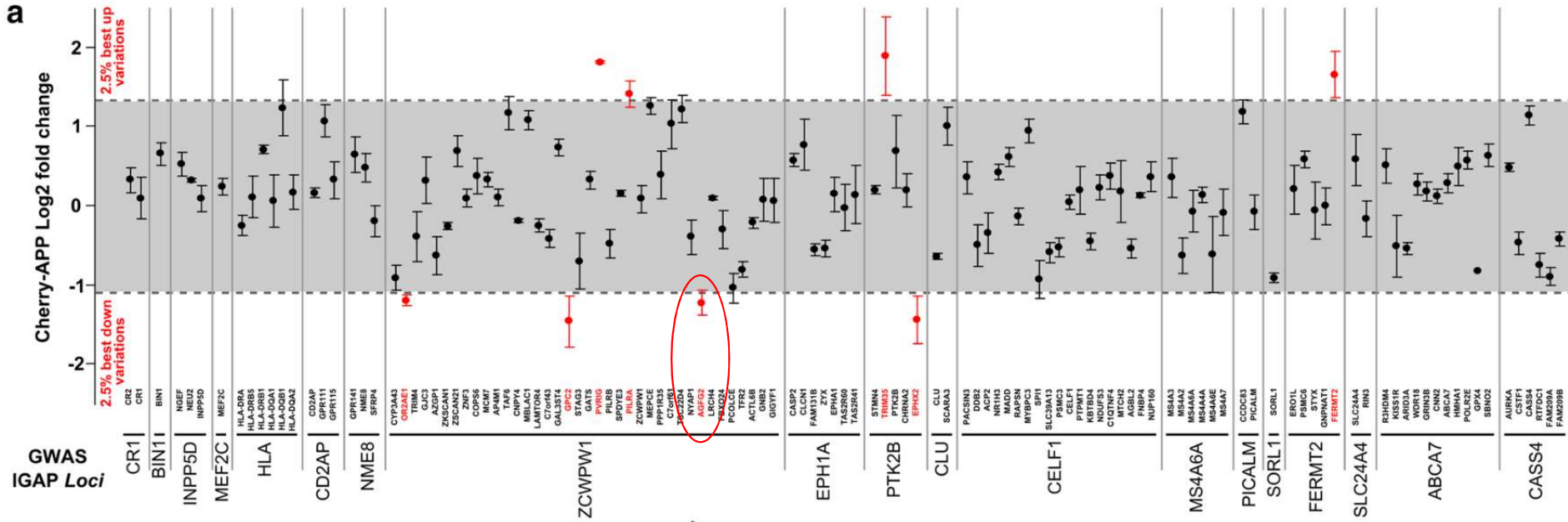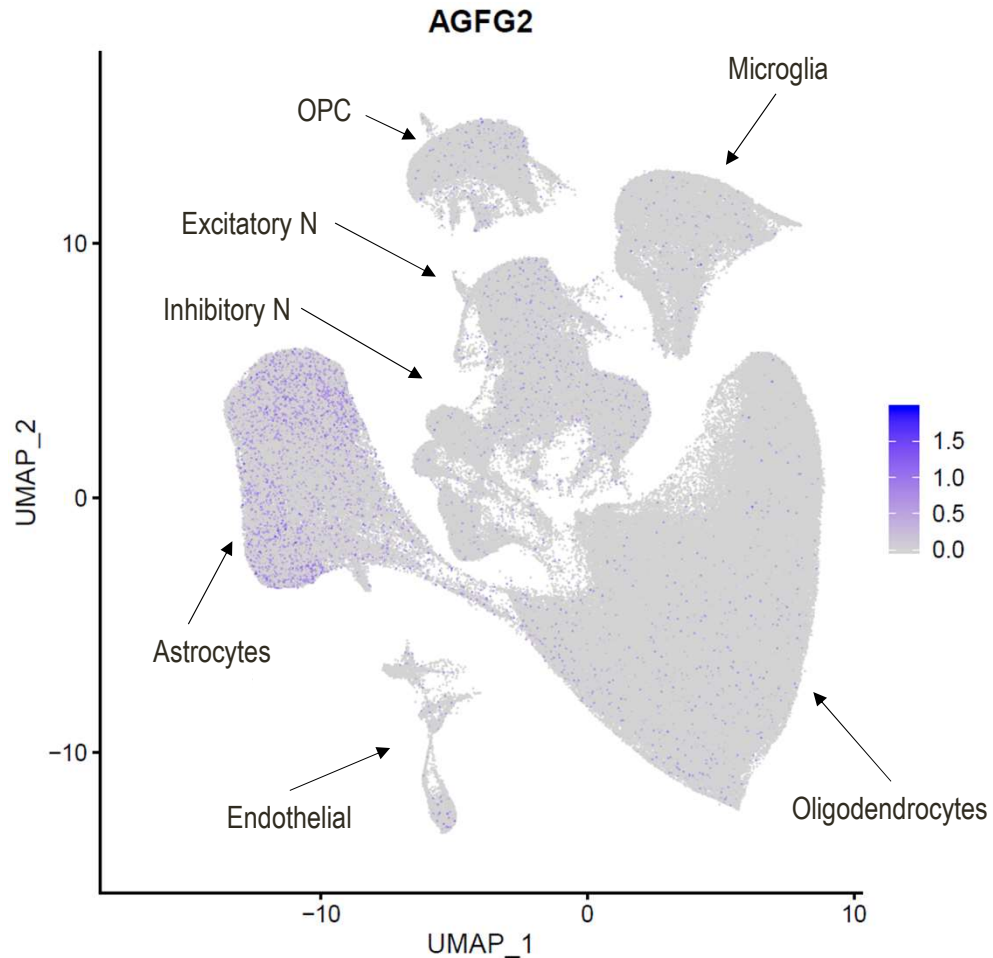CO PS-AD AD
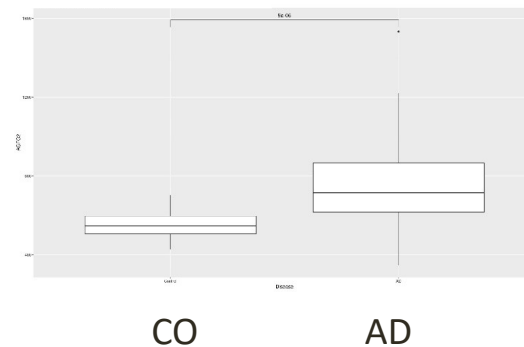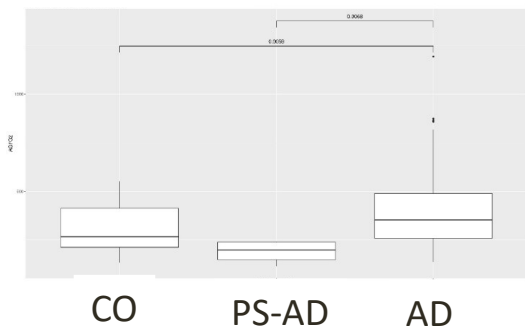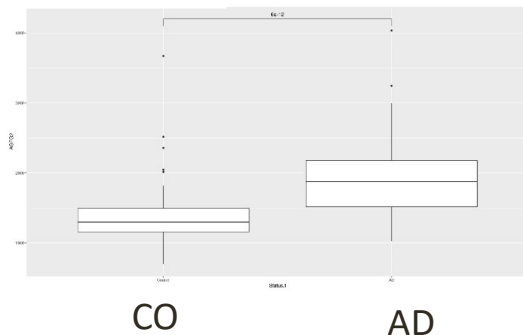
AGFG2 could alter APP processing

Performed siRNA on HEK293 cells stably over-expressing mCherry-APP695wt-YFP. *AGFG2* was among the top eight genes to dysregulate APP processing - Chapuis et al., Acta Neuropath, 2017

AGFG2 is an astrocyte gene

AGFG2

Single nuclei expression data from 69 brain donors

→ AD cases have more Aβ plaques
→ AD cases have higher *AGFG2* expression
→ *AGFG2* interferes with APP processing Chapuis et al.
→ Astrocytes are also and active producers of APP
(Liao, 2016, J. Neurosc, 36(5):1730-1746)

Hypothesis
α - higher expression of *AGFG2* could be participating into more release of APP to the media

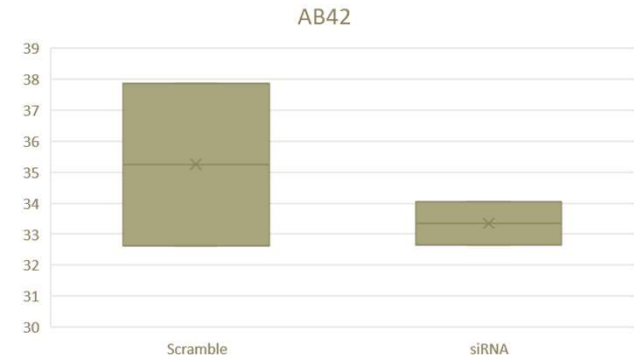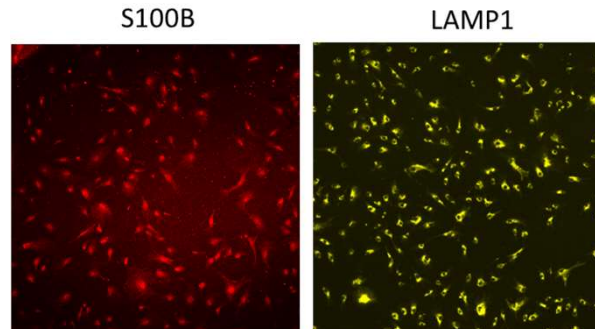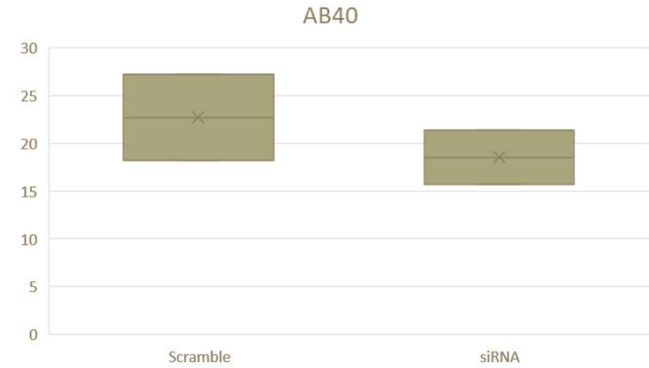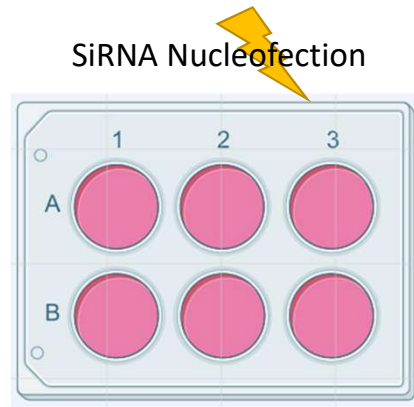# Changes in *AGFG2* expression alter Aβ – iPSC astrocytes
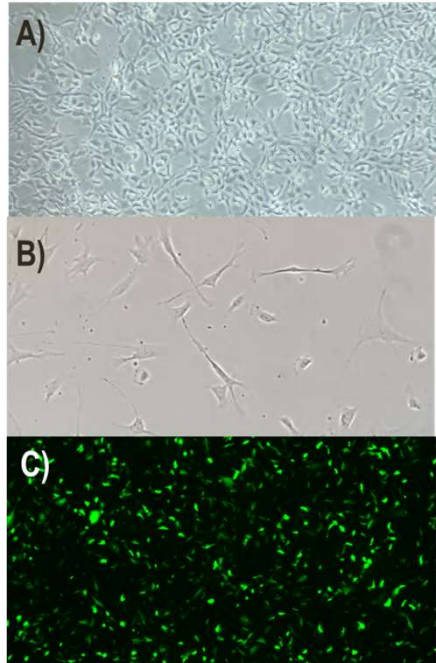
Figure. A) iPSC 20x; B) Astrocyte D30 20x
C) GFAP nucleofection

SiRNA Nucleofection

Scr

↓AGFG2

S100B

LAMP1

AB40

AB42

Summary

- Genomic analyses of clinically stratified participants
  - Multi-ethnic meta-analyses of EOAD
    - Meta-analyses reveals two novel loci: *PGM2, VAV1*
  - Family based approaches
    - Gene-based analyses in LOAD families identified *CPAMD8* as novel candidate gene

- Transcriptomic analysis of combined clinical categories
  - novel candidate gene AGFG2
  - is an astrocyte expressed gene
  - could be promoting higher release of APP

Neurogenomics & Informatics

47

Acknowledgements

Thanks to the members at the NGI, Dr. Cruchaga, and most importantly the families and participants of the Knight ADRC and ADGC