Dissecting the Genetic Basic of Complex Disease Using AI

Michael Snyder, PhD

Stanford University School of Medicine

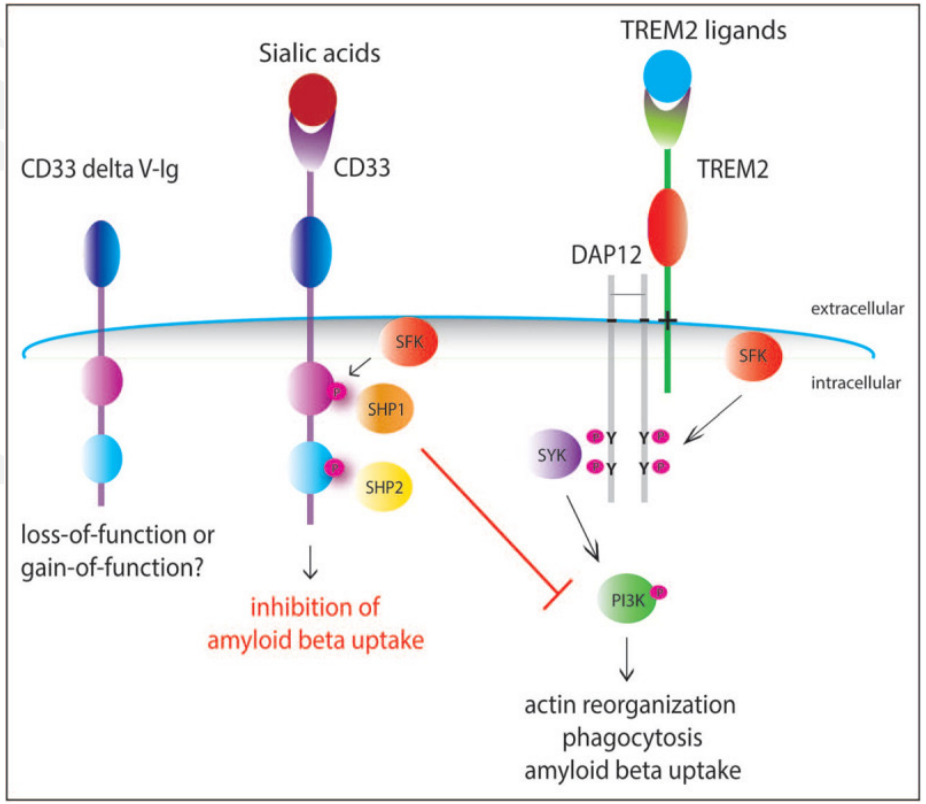
April 2026

Conflicts: Personalis, SensOmics, Qbio, January AI, Filtricine, Mirvie, Fodsel, Protos, Xthera, Marble Therapeutics, Exposomcs, Iollo, RTHM, Mitrix, Yuvan, Onza, Next Thought, Orange Street Ventures

Major Goals in Genetics and Genomics

1. Understand the Genetic Basic of Human Disease
    - Provides insights into biological pathways involved
    - e.g. Alzheimers and immune function e.g. *CD33, TREM2*

2. Predict Genetic Risk and Prognosis



Genome Sequencing – First 70 People

Twelve have important pathogenic mutations:

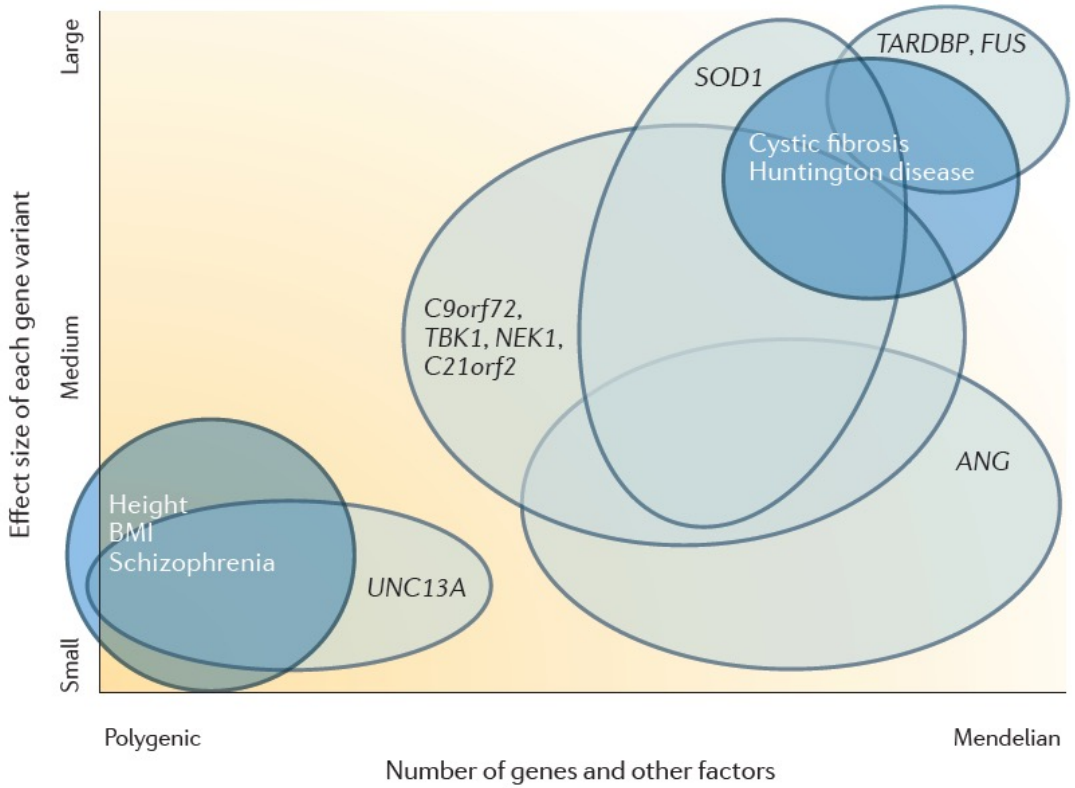
- SDHB (2X): high freq. of neuroendocrine tumors^
- APC (2X): Colon cancer
- BRCA1: Breast & ovarian cancer
- MUTYH: Colon cancer
- SLC7A9: Cystinuria
- RBM20: Dilated cardiomyopathy^
- CHEK2: Breast cancer
- PROC: Affects coagulation
- HNF1A: MODY mutation^
- ABCC8: Hyperinsulinemic hypoglycemia

Rego, … Snyder et al. Mol. Case Reports 2020

Genetics involves:

Single gene (Mendelian) MODY genes (e.g. HNF1a)

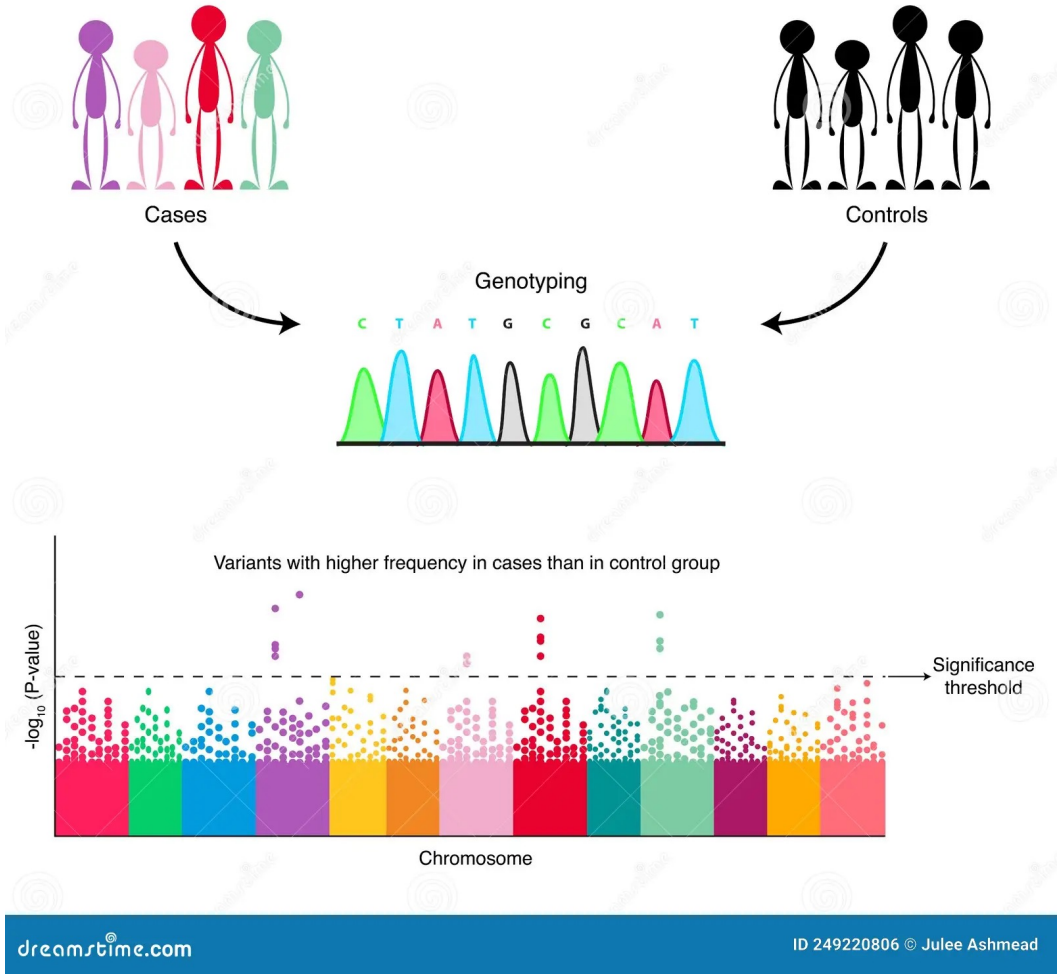
Complex disease: Thought to be due to many genetic changes of small effect



Many

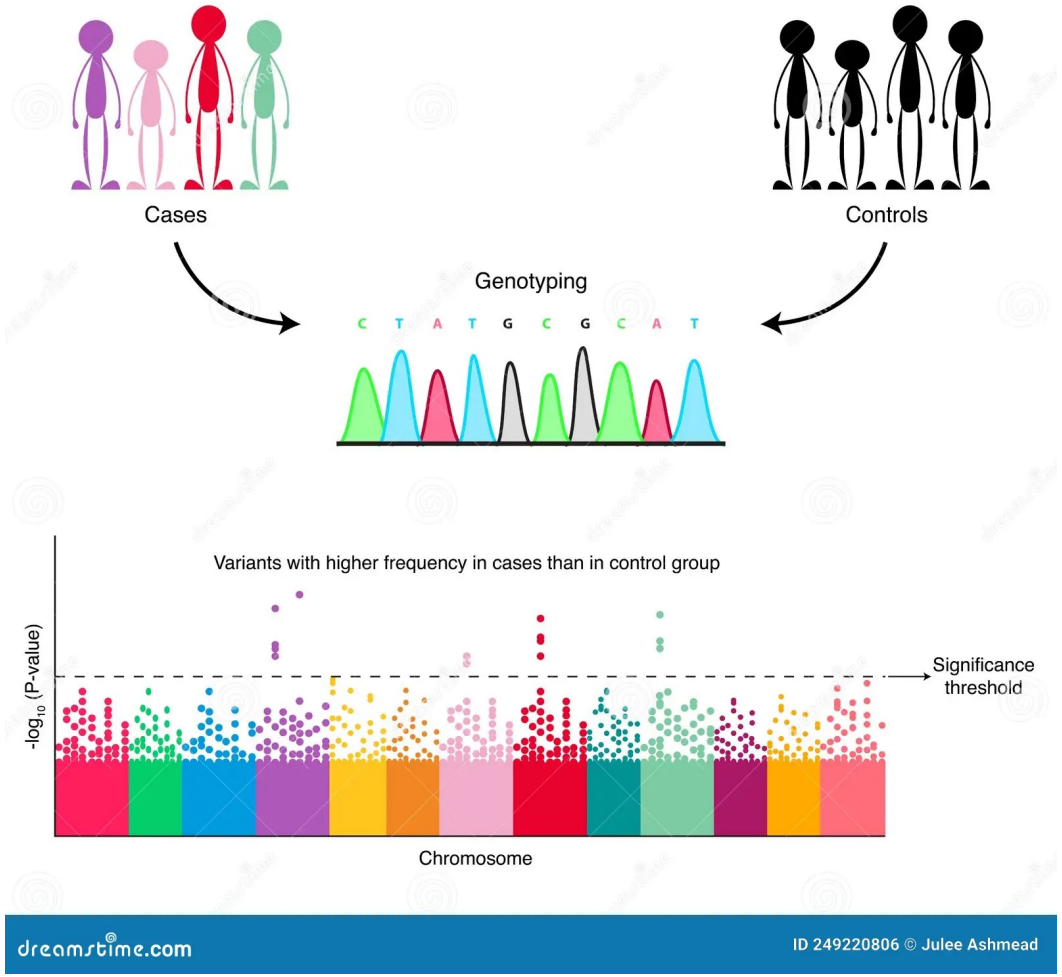
Single gene

Genome-Wide Association Studies (GWAS)



https://www.dreamstime.com/gwas-image249220806

Genome-Wide Association Studies (GWAS)



https://www.dreamstime.com/gwas-image249220806

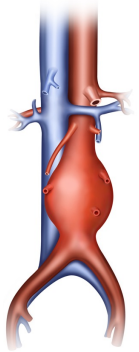
Polygenic Risk Score

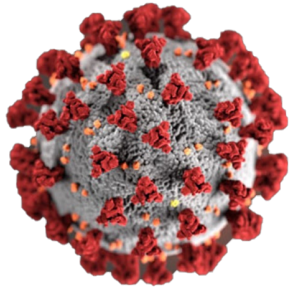
Disease	# SNPs
Coronary Artery Disease	6.6M
Atrial Fibrillation	6.7M
Type 2 Diabetes	6.9M
Inflammatory Bowel Disease	6.9M
Breast Cancer	5.2K

Khera et al. Nat Genet. 2018 50: 1219–1224

Three Approaches For Analyzing Complex Disease

1. Rare pathogenic gene variants
2. Common variants from GWAS signals
3. Noncoding rare variants





Li, Pan, Zhang, …Snyder Cell 2018

Zhang, Cooper-Knock … Snyder *Neuron* 2022

Zhang, Cooper-Knock et al , *Cell Systems 2022*

Abdominal Aortic Aneurysm (AAA)

Article

Cell

Decoding the Genomics of Abdominal Aortic Aneurysm

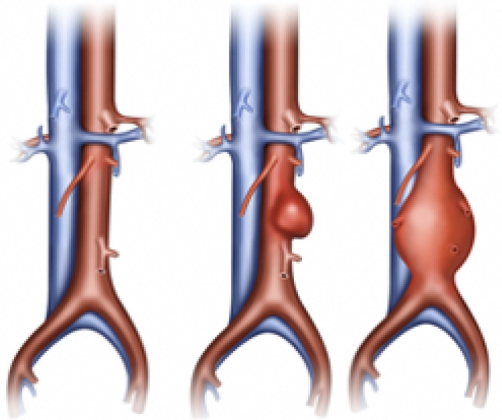
Jingjing Li,[1,2,6,7] Cuiping Pan,[1,3,6,7] Sai Zhang,[1,6,7] Joshua M. Spin,[3,4,6] Alicia Deng,[3,4,6] Lawrence L.K. Leung,[3,4,6] Ronald L. Dalman,[5,6] Philip S. Tsao,[3,4,6,*] and Michael Snyder[1,6,8,*]
[1]Department of Genetics, Center for Genomics and Personalized Medicine, Stanford University School of Medicine, Stanford, CA 94305, USA

Clinical Facts

- Asymptomatic at onset - fast growing
- 90% mortality rate upon rupture
- The 10th leading cause of death in US
- **No early screening tool**

Epidemiology

- **Heritability: 70%**
- Aged population >50 yo.
- Lifestyle matters
- High blood pressure
- Cholesterol etc



Identifying Genes Associated with AAA

268 AAA cases

133 AAA controls

50X

Whole genome deep sequencing

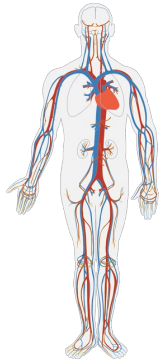
EHR mutation network gene

HEAL: a machine learning framework

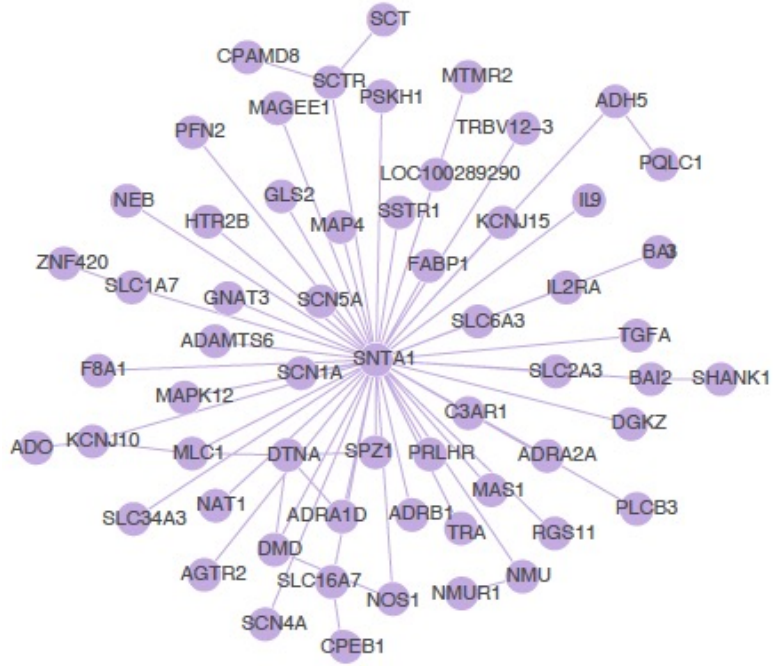
Call pathogenic mutations

HNRNPCL1 HLA-DQB2 CBWD3 PYGM KCNC3

Identify 60 AAA genes

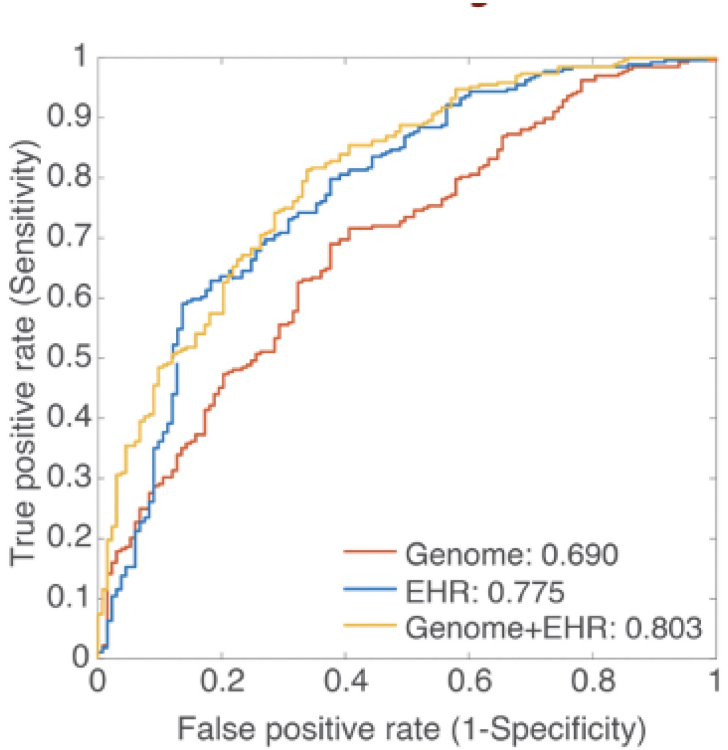


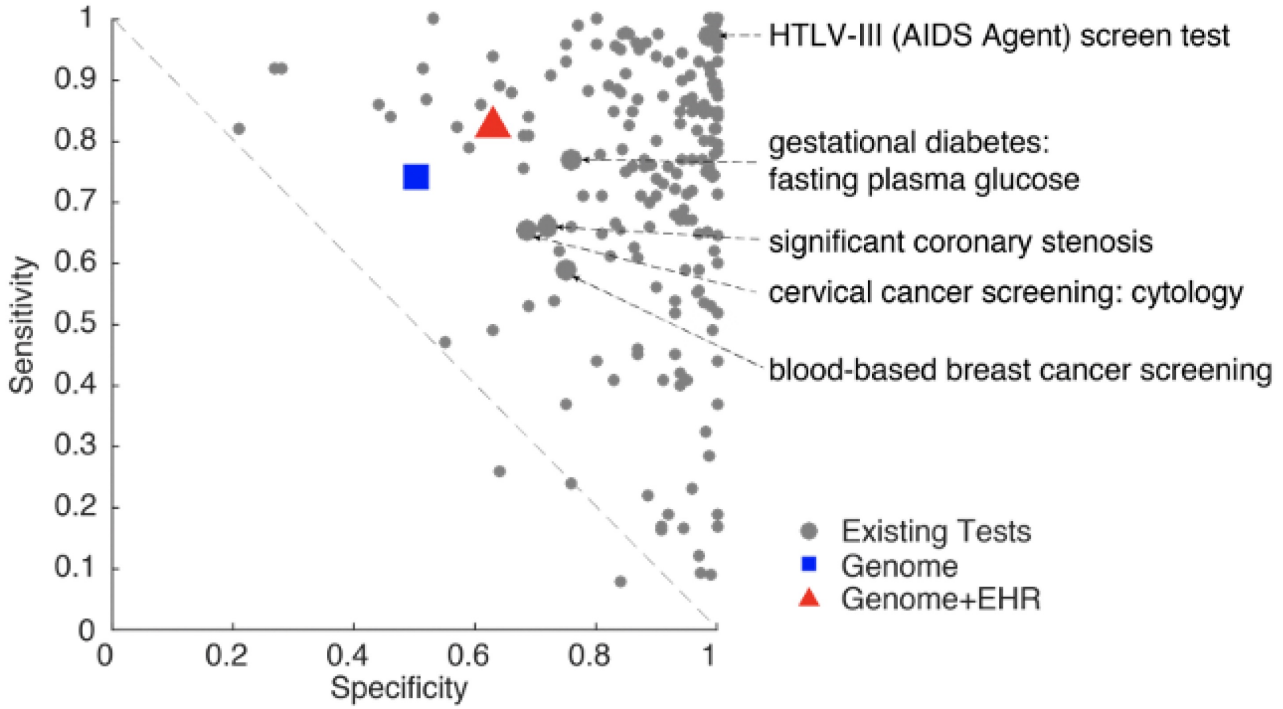
Blood Circulation, Blood Pressure, Cardiomyopathy



Jingjing LI, Cuiping Pan, Sai Zhang .. Phil Tsao, Cell 2018

HEAL accurately predicts AAA risk -> clinical utility





ME/CFS

- Affect 2-4% of US population
- Known to have a genetic component
  - SNP based: Chris Ponting 9.5%
- Few (7-10) genes identified

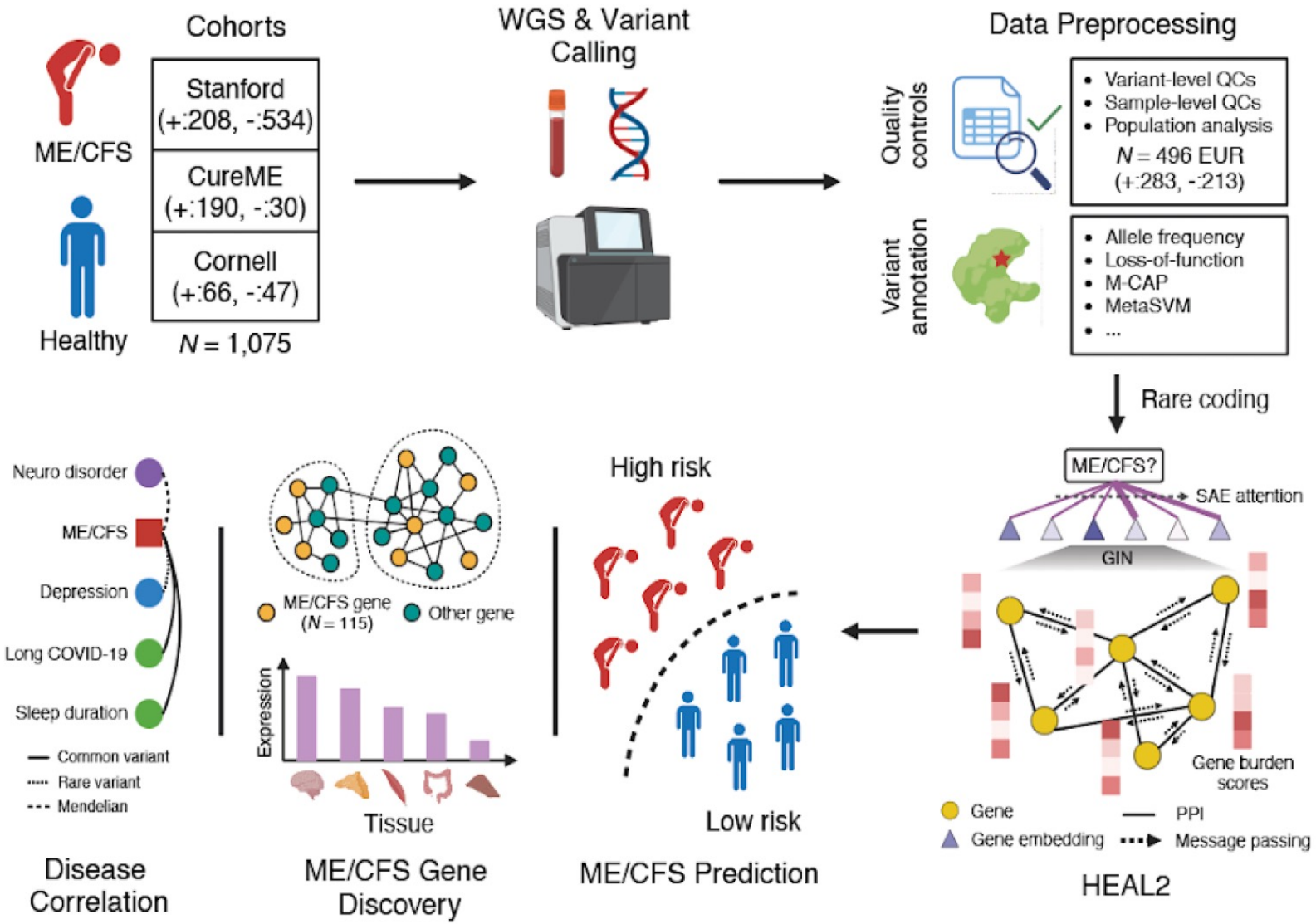
Cohorts

ME/CFS Disease Cohorts	Sequenced Samples
Maureen Hanson (Cornell)	N=113 (Cases-66, Controls-47)
Fereshteh (Stanford)	N=364 (Cases-208, Controls-156)
UK Biobank	N=220 (Cases-190, Controls-30)
Total	**697**

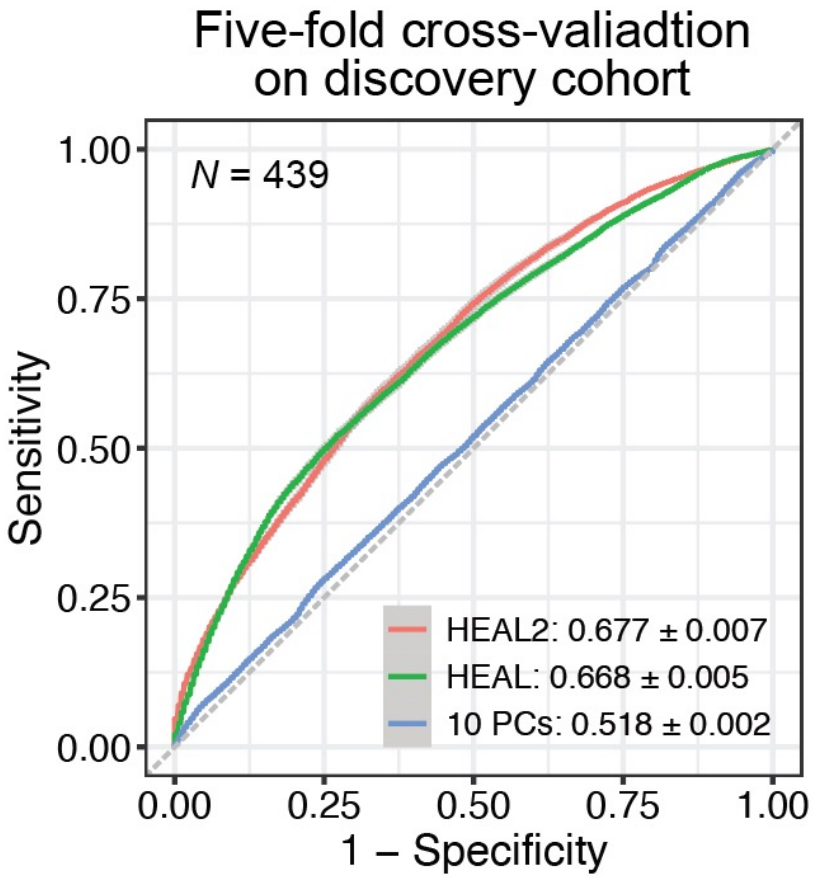
Controls	Samples
iPOP	110 (7 MECFS carriers/cases)
hPOP	268
Total	**378**

ME/CFS Study design

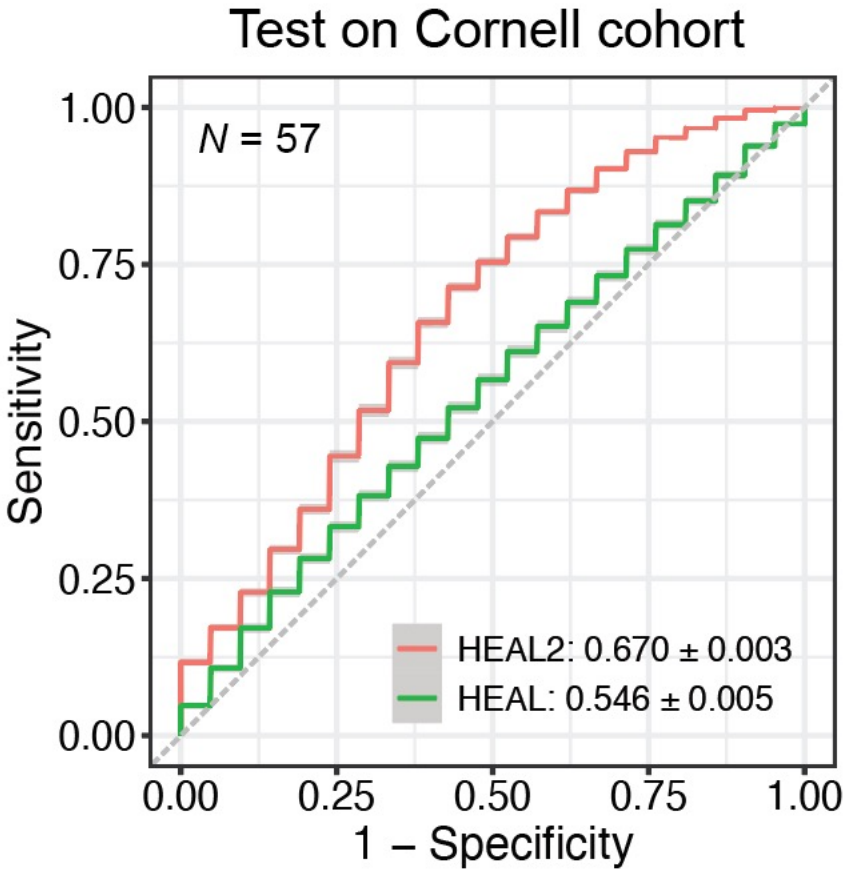
- **HEAL2** extends our previous model, HEAL, by aggregating the mutational burden across genes and utilizing a graph neural network to model gene-gene interactions, derived from known protein-protein interactions



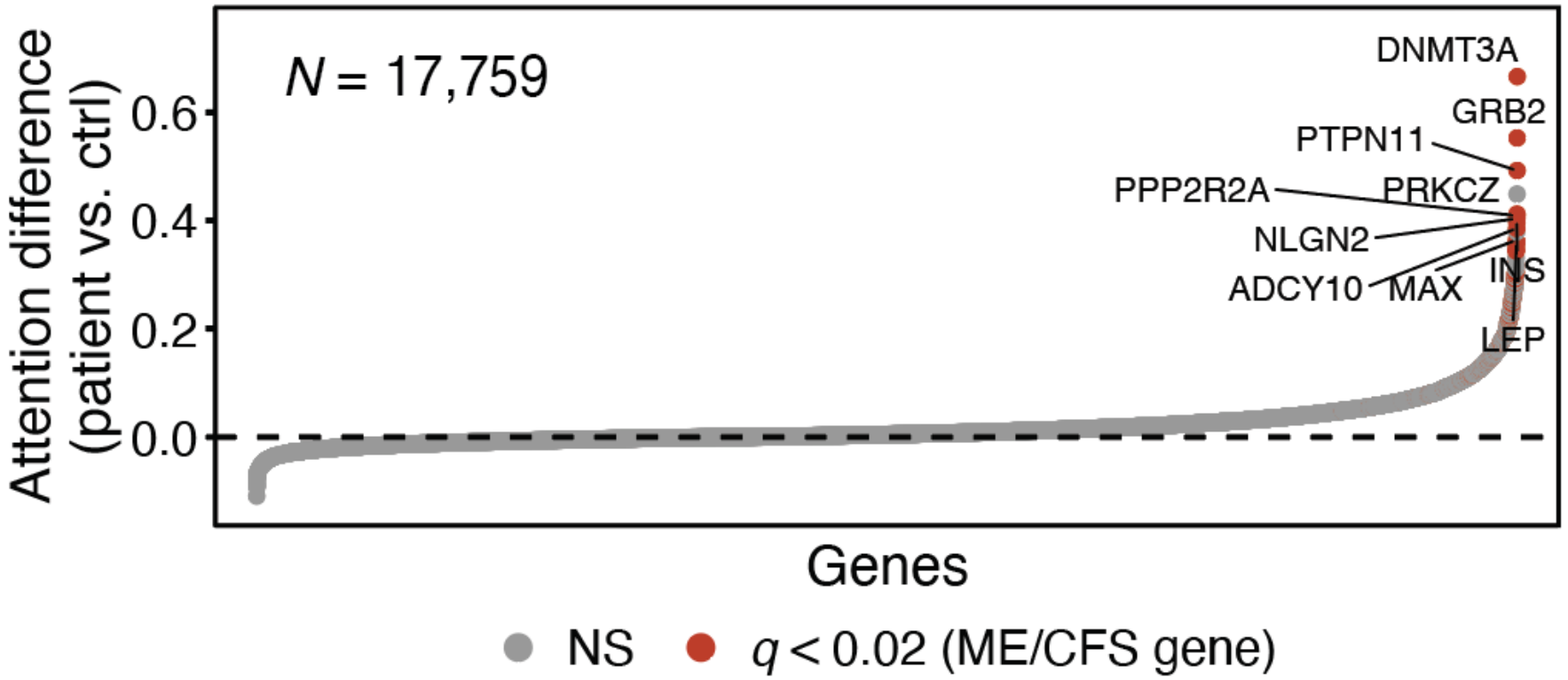
HEAL2 accurately predicts ME/CFS from rare variants – outperforming HEAL



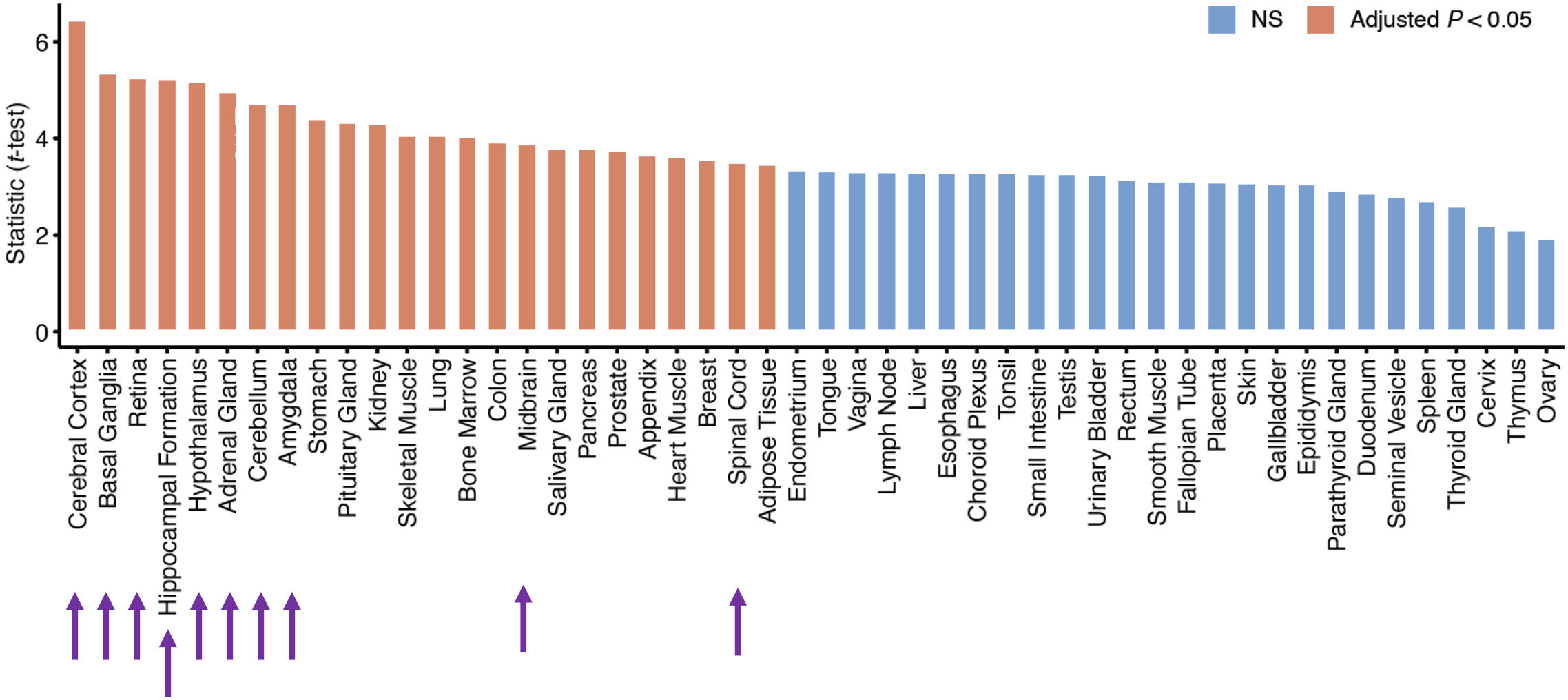
Discovery cohort:
Stanford + UK CureME



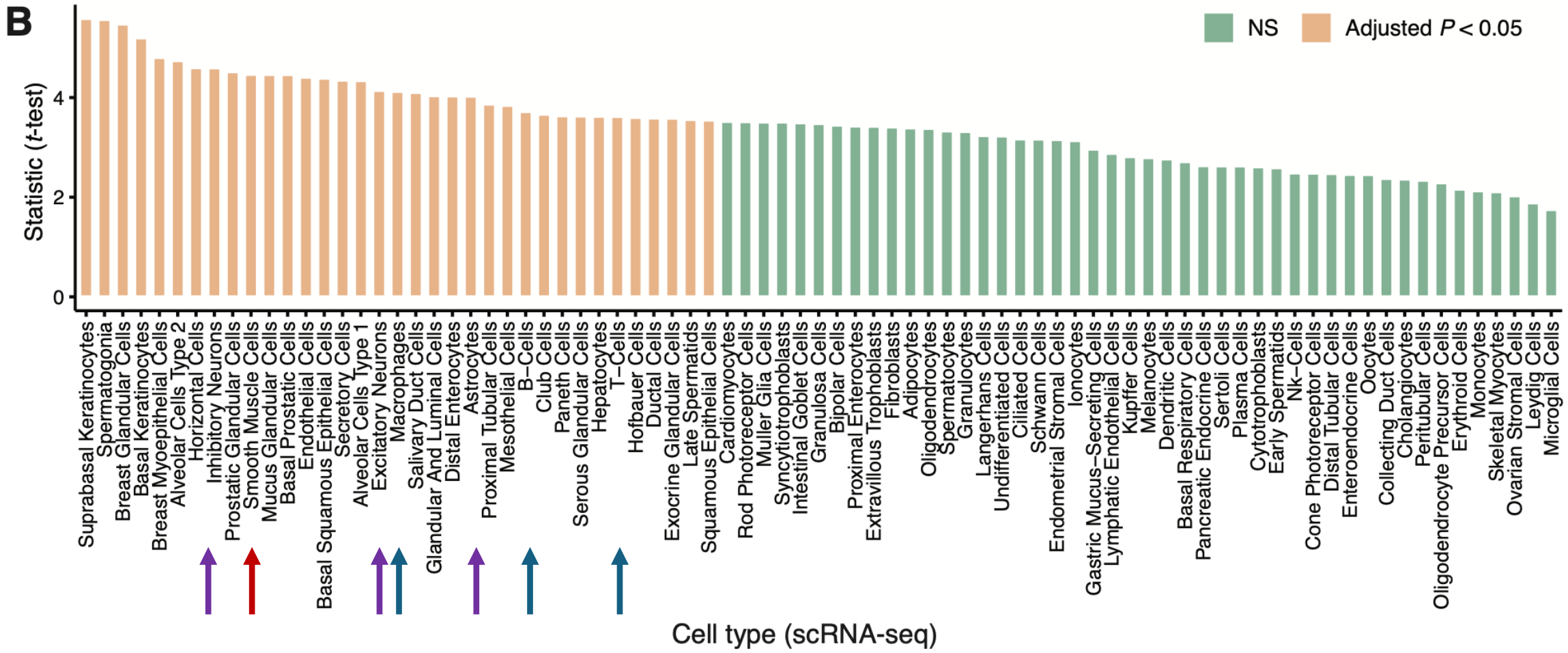
HEAL2 identified 115 ME/CFS genes distinguishing cases from controls



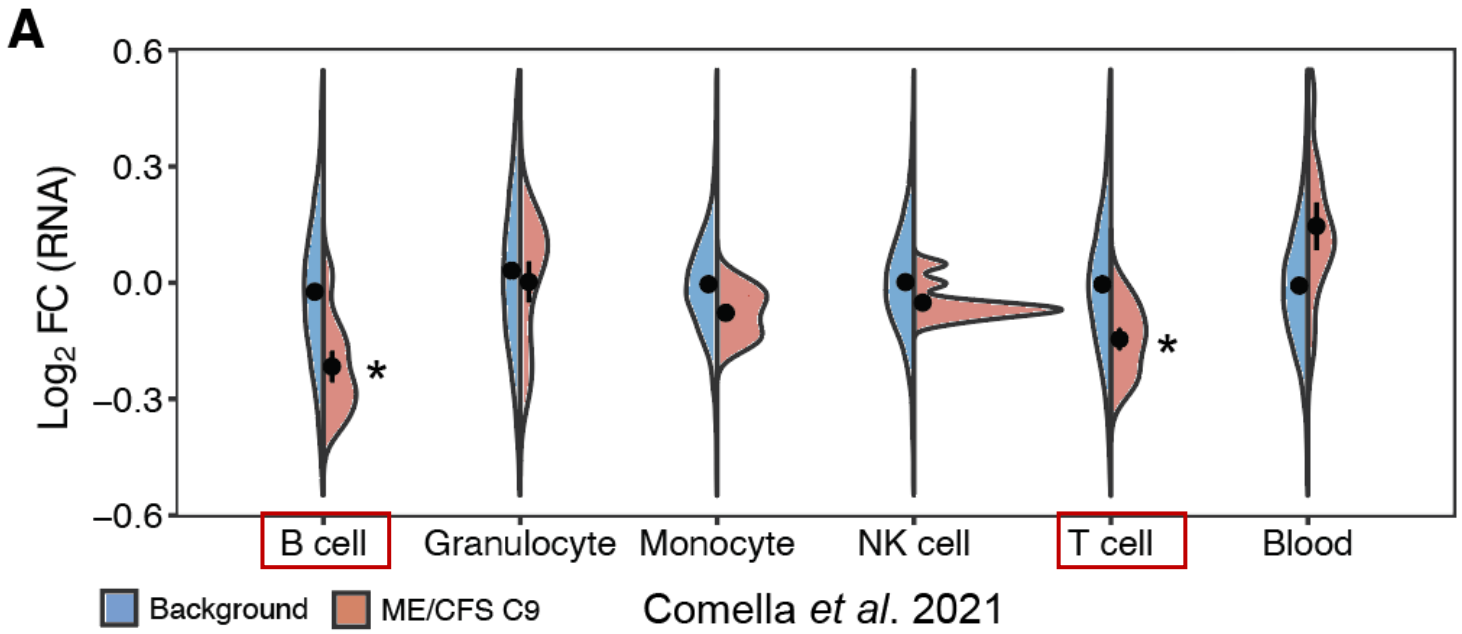
The identified ME/CFS genes are highly expressed in CNS

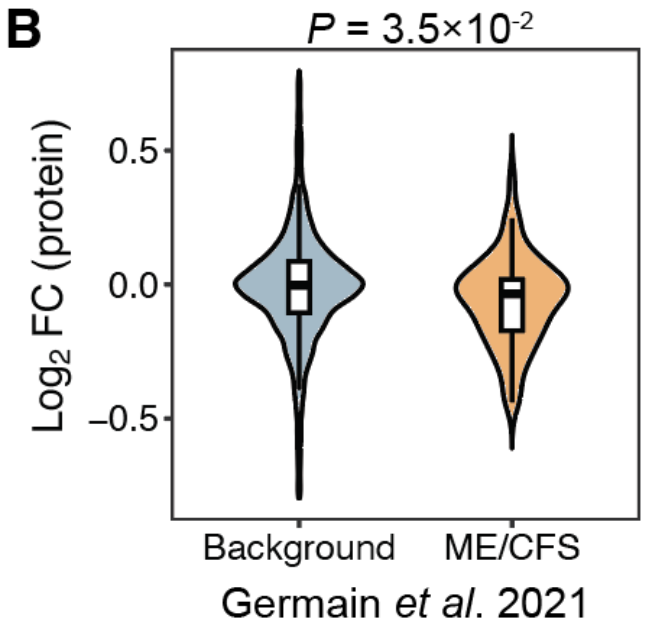


ME/CFS genes are highly expressed in neurons, smooth muscle cells, and immune cells



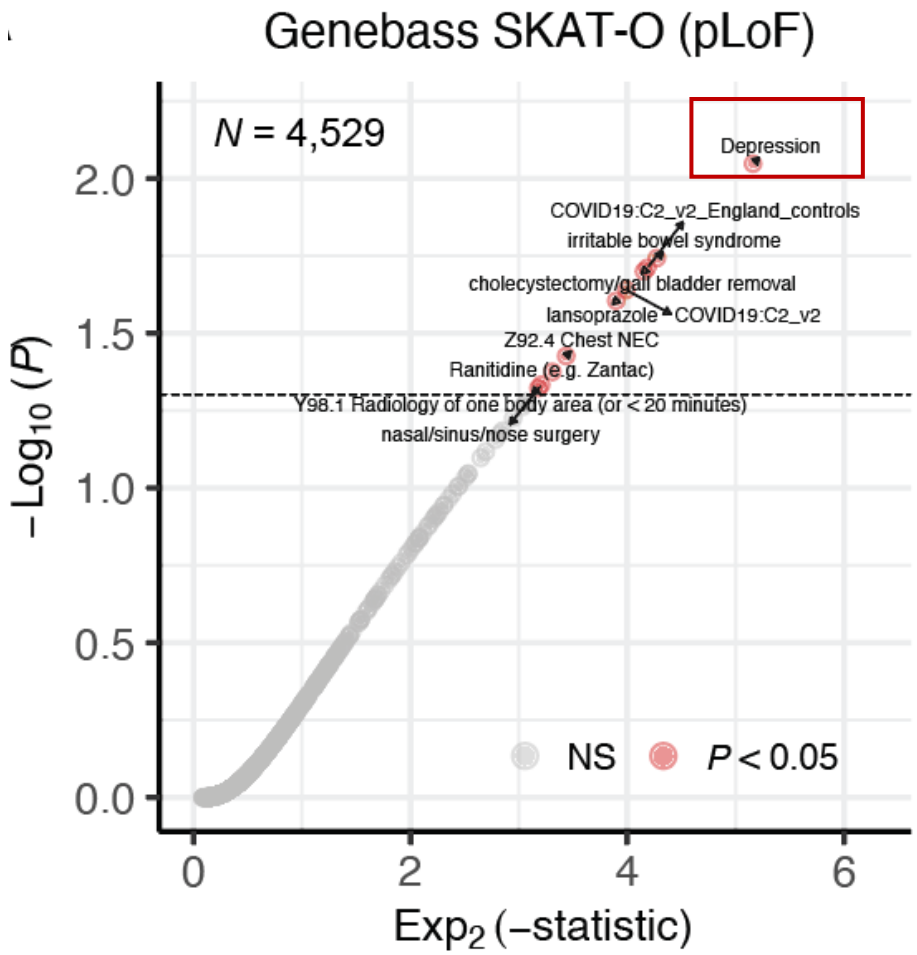
ME/CFS genes are down-regulated in ME/CFS patients

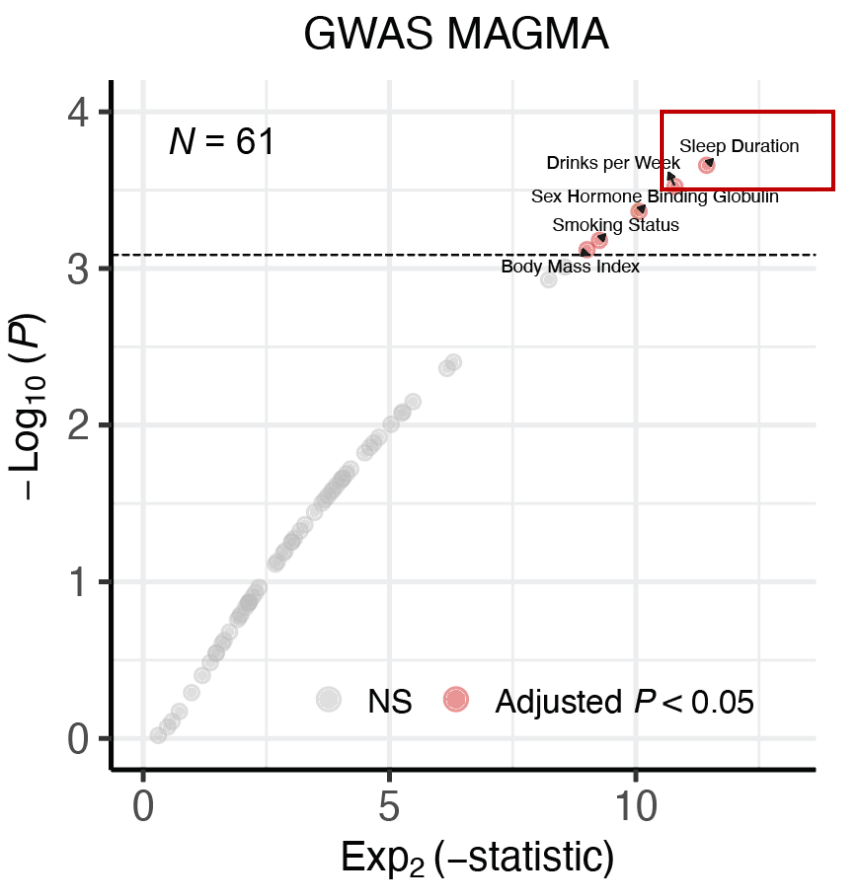




and down-regulated in patient B and T cells

ME/CFS genetically correlated with other diseases and traits





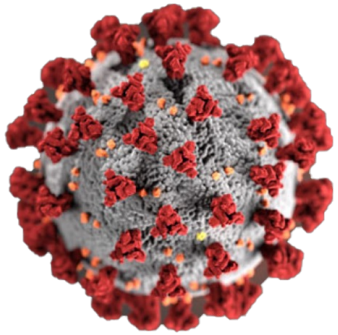
Sleep disturbance is a key symptom of ME/CFS

Genetics of Complex Disease Using Machine Learning & GWAS Information

Sai Zhang, Johnathan Cooper-Knock

1. ALS

2. COVID-19 Severity

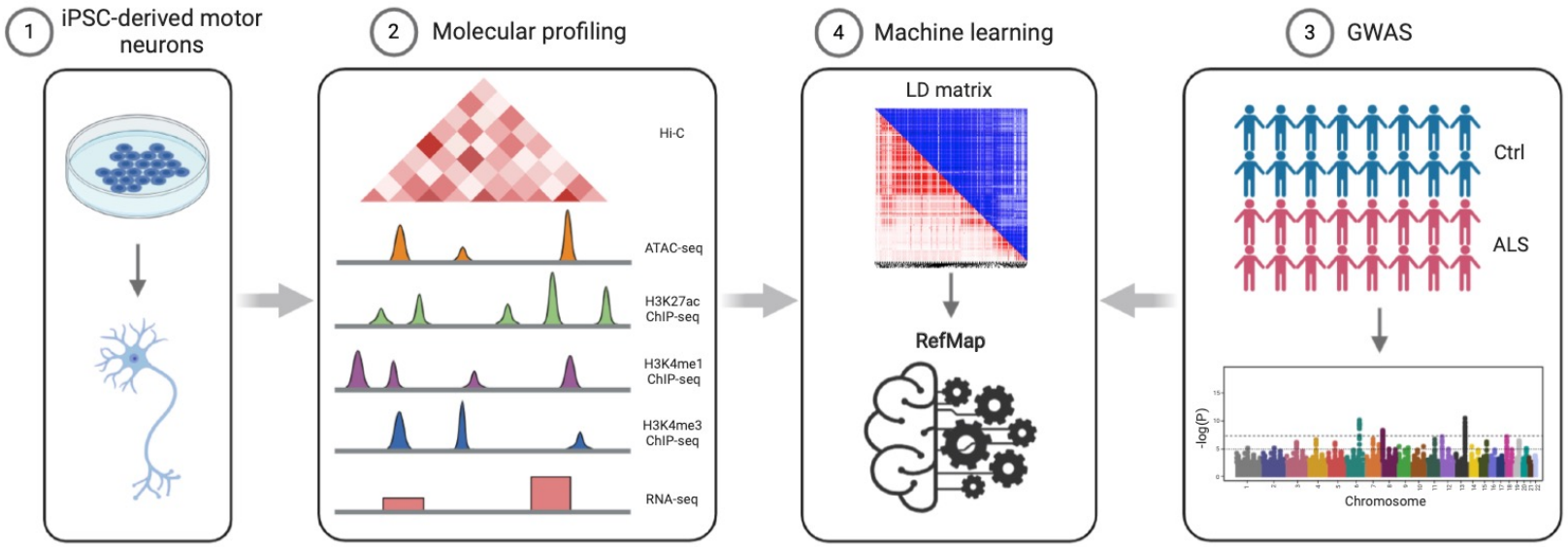


Zhang, Cooper-Knock .. Michael Snyder *Neuron* 2022

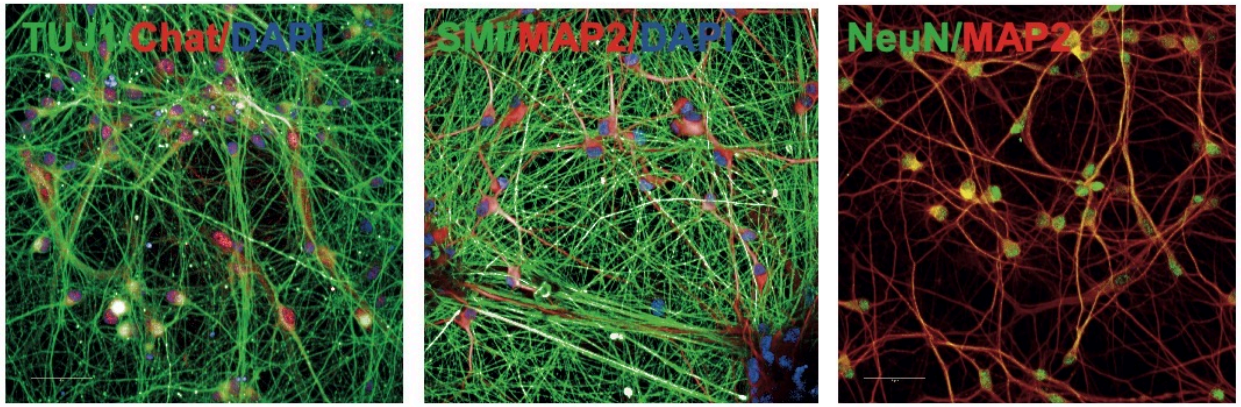
Zhang et al , *Cell Systems* 2022

Zhang et al , *BioRixv* 2024

Identification of ALS Genes Using Machine Learning (AI)

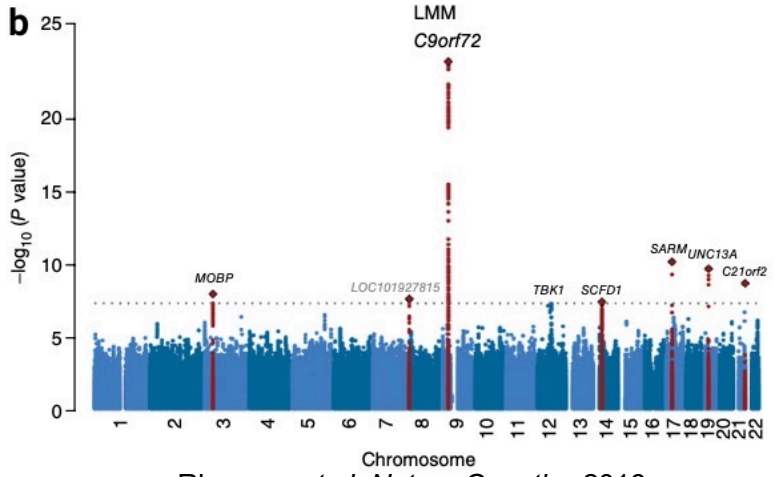


Profiling motor neurons for gene regulatory regions ATAC-Seq, etc



Zhang et al. *Neuron* 2022

GWAS: 12,577 cases and 23,475 controls



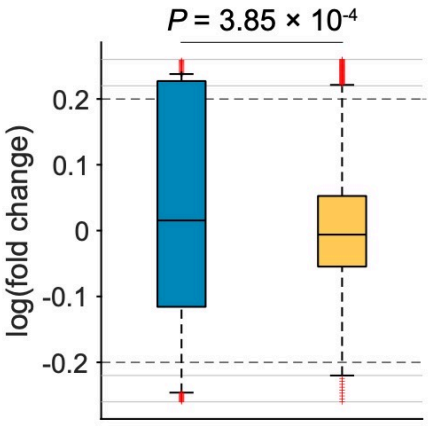
Rheenen *et al*. *Nature Genetics* 2016

RefMap identifies: 690 ALS genes (over the 7 known previously)
Accounts for 36% of heritability (over the 6% known previously)

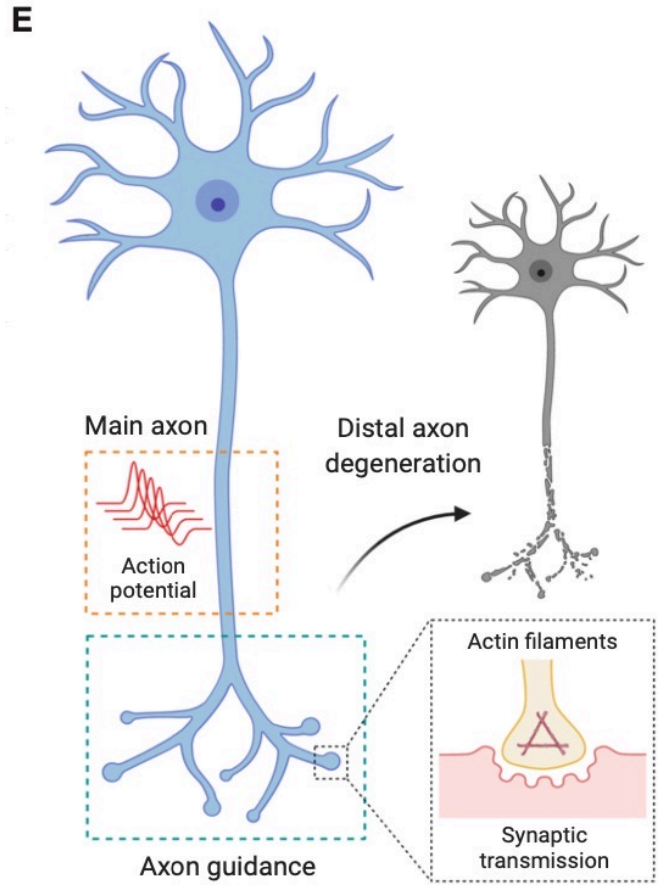
- Overlap with many well-known ALS genes

  C9orf72, *ATXN2*, *SIGMAR1*

- Significantly enriched with known ALS genes
  - Based on a well-curated ALS gene list (*n*=260) *P*=5.2e-3 (OR=2.07)
  - ClinVar *P*=3e-2 (OR=3.06)

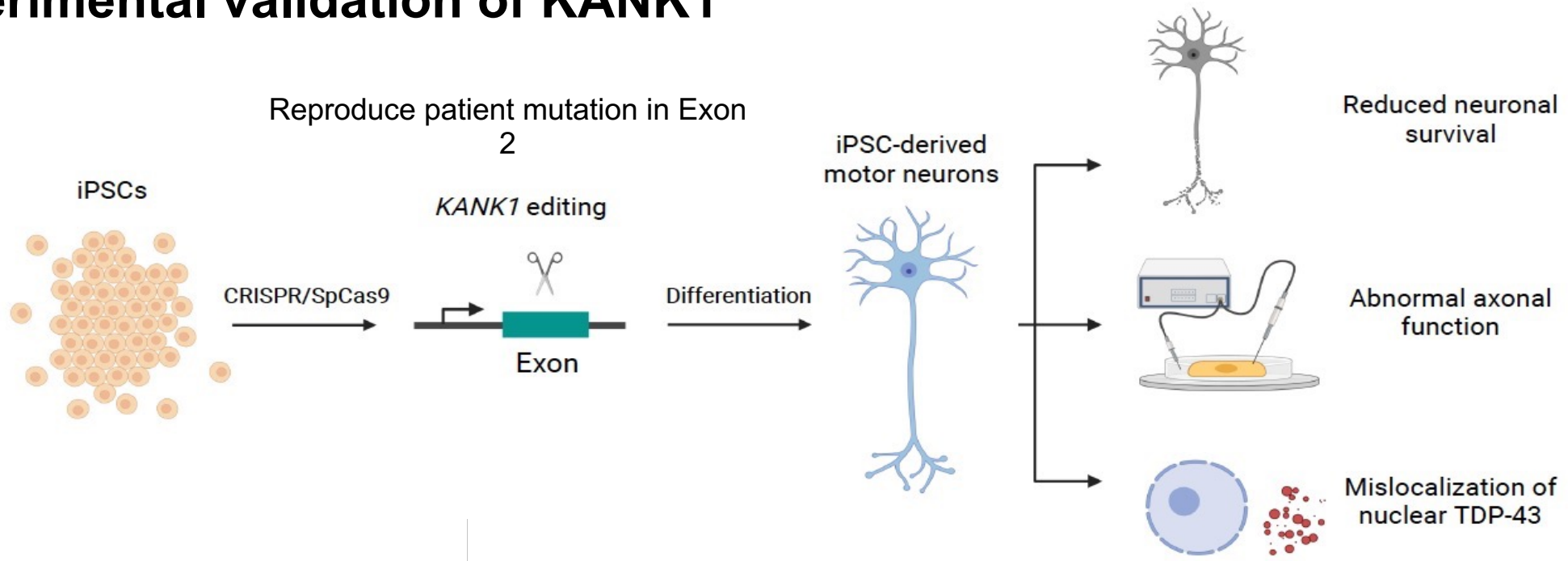


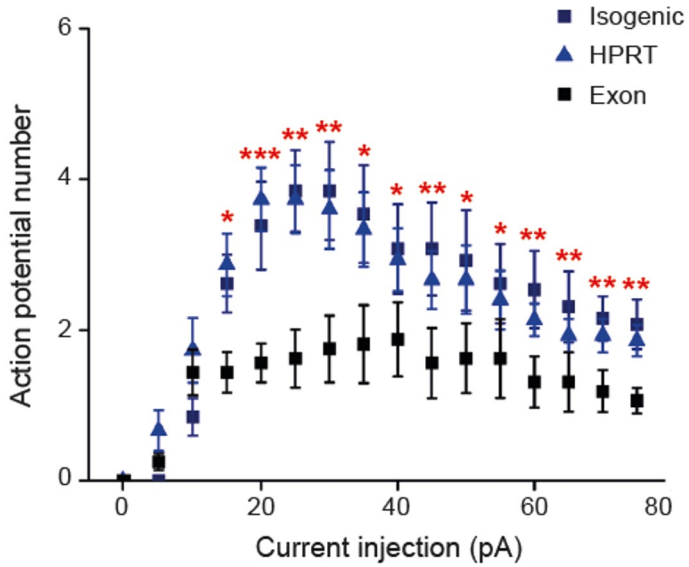
Down-regulated in both ALS brain tissues and iPSC MNs



Function in Distal regions of axons

Experimental validation of KANK1

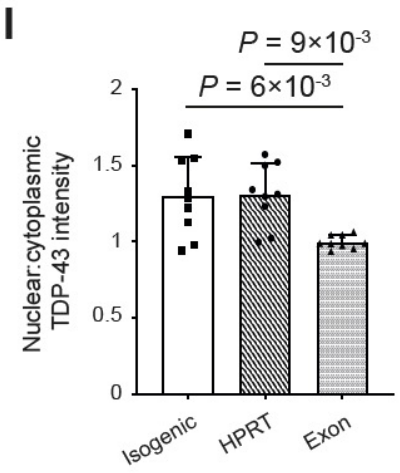


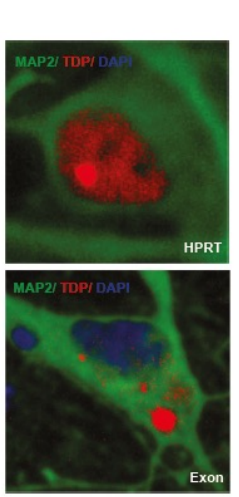


KANK1 ALS mutations cause:

hypoexcitability

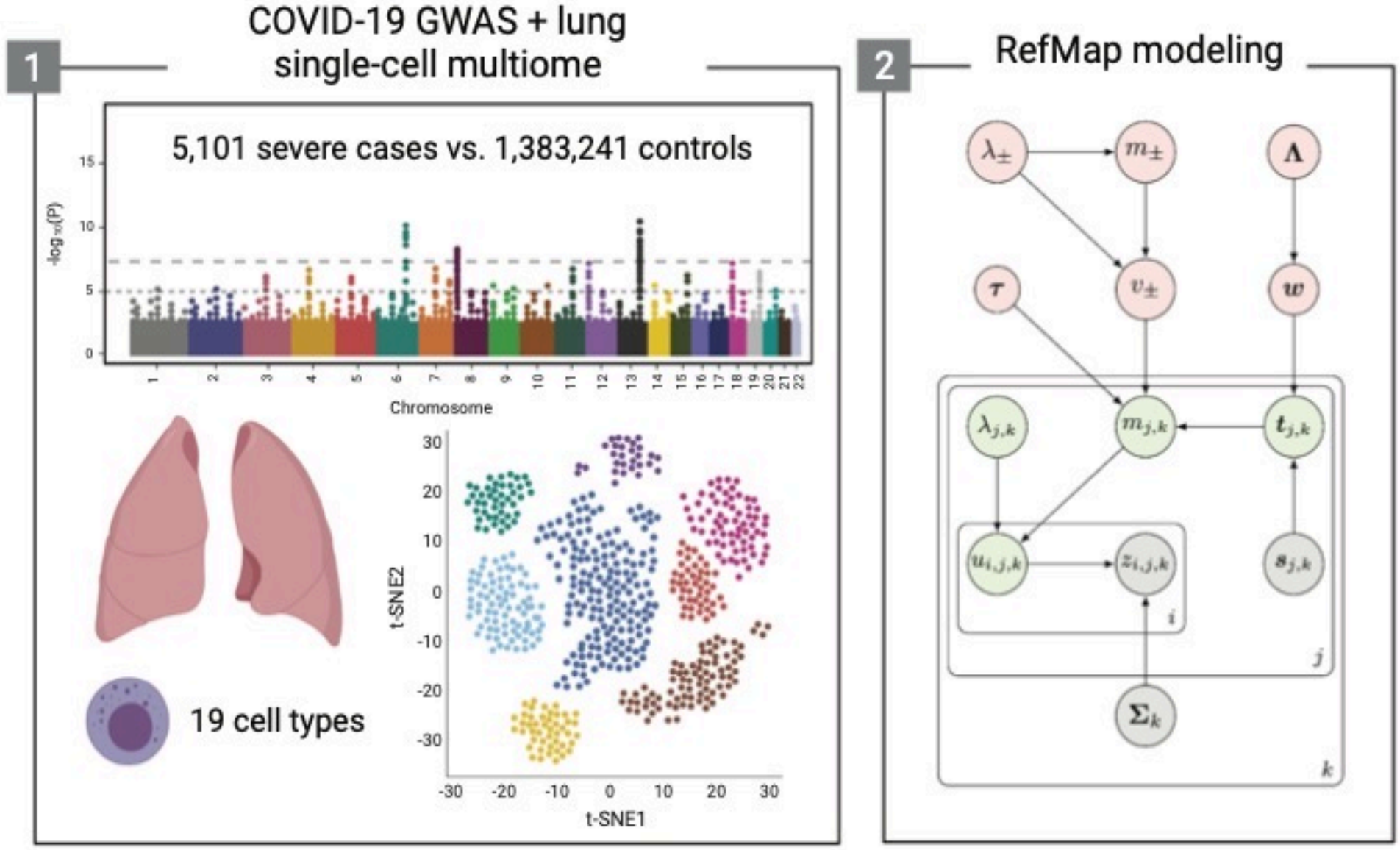
TDP-43 mislocalization





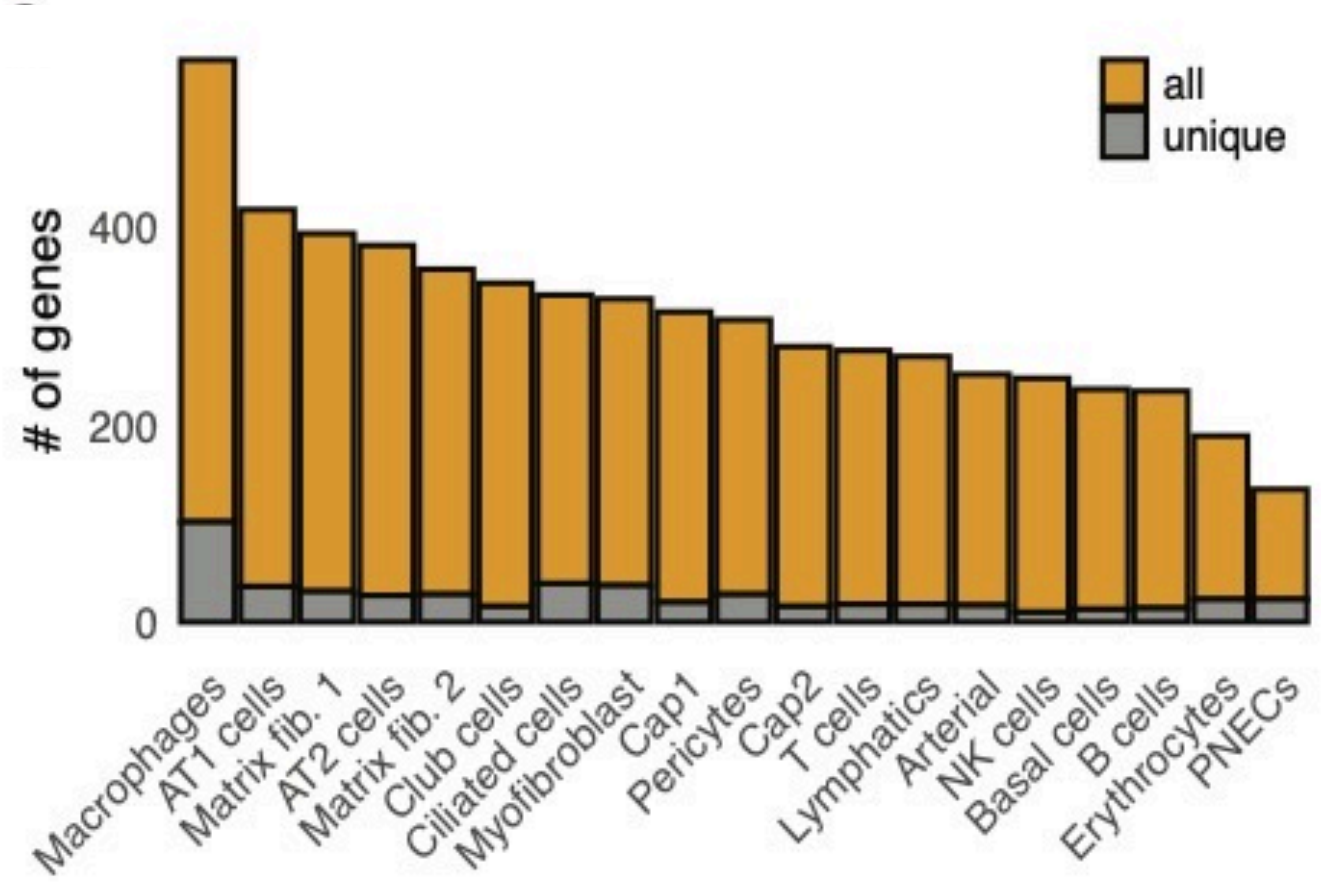
Sai Zhang, Johnathan Cooper-Knock .. *Neuron* 2022

COVID Severity: GWAS + Single-Cell Multiome → Genes

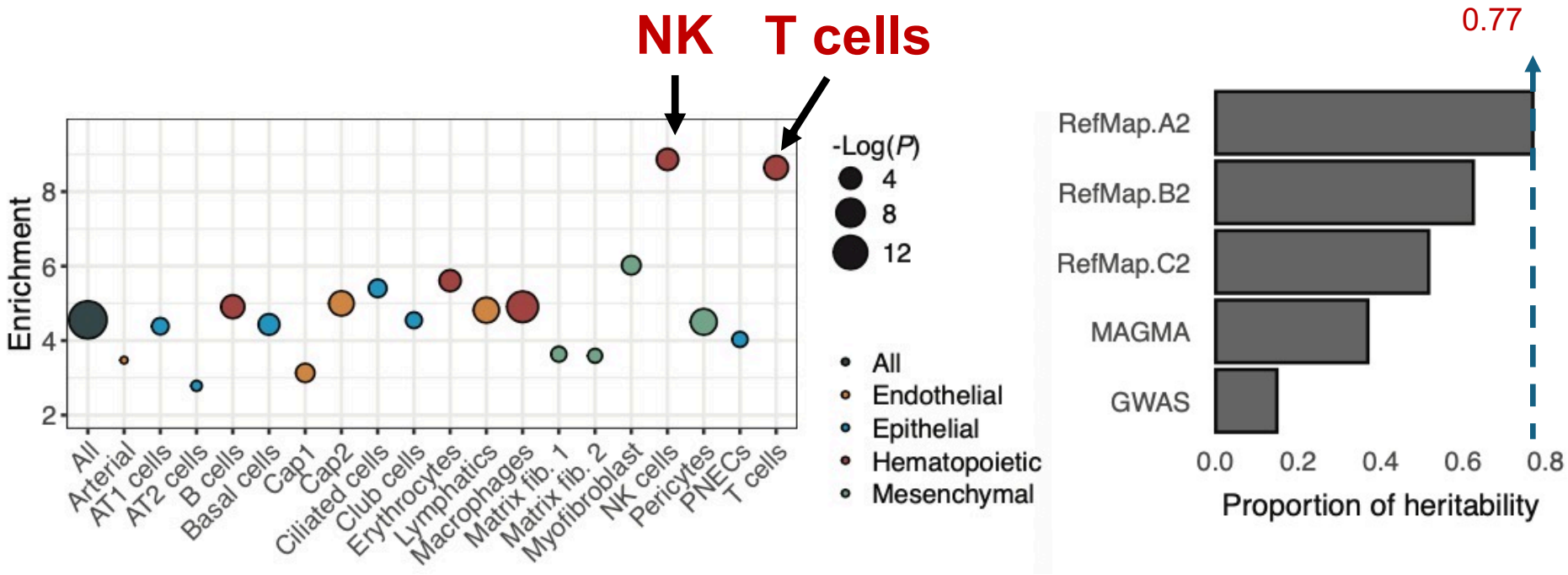


Mapping Target Genes Per Cell Type

6,662 1kb regions →
1,375 genes across 19 cell types

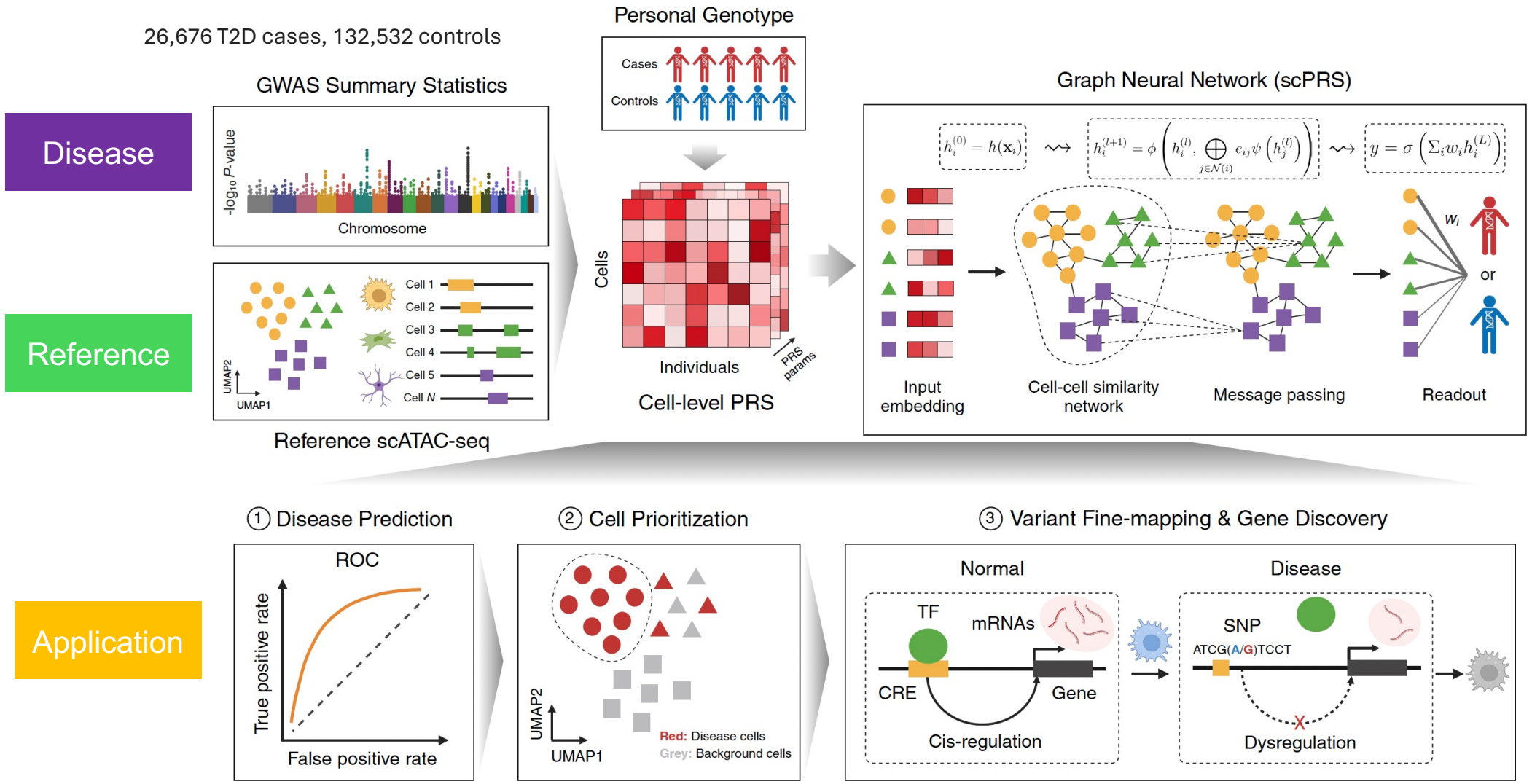


Heritability Analysis Identifies NK Cells and Other Immune Cells



Failed activation of NK cells in the initial stage of SARS-CoV-2 infection is permissive of viral replication, leading to fatal hyperinflammation

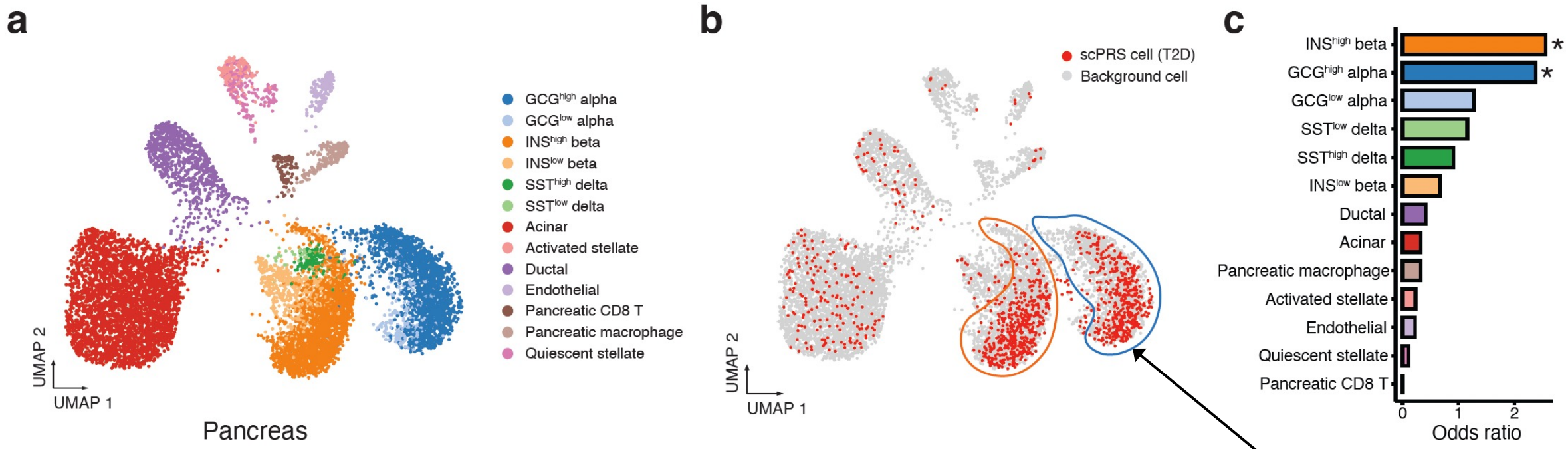
Single-cell PRS (scPRS)



GWAS data from McCarty et al 2017; snATAC data from Chiou et al 2021 Nature

Zhang et al 2025 **Nature Biotech**

scPRS prioritizes T2D-critical cells



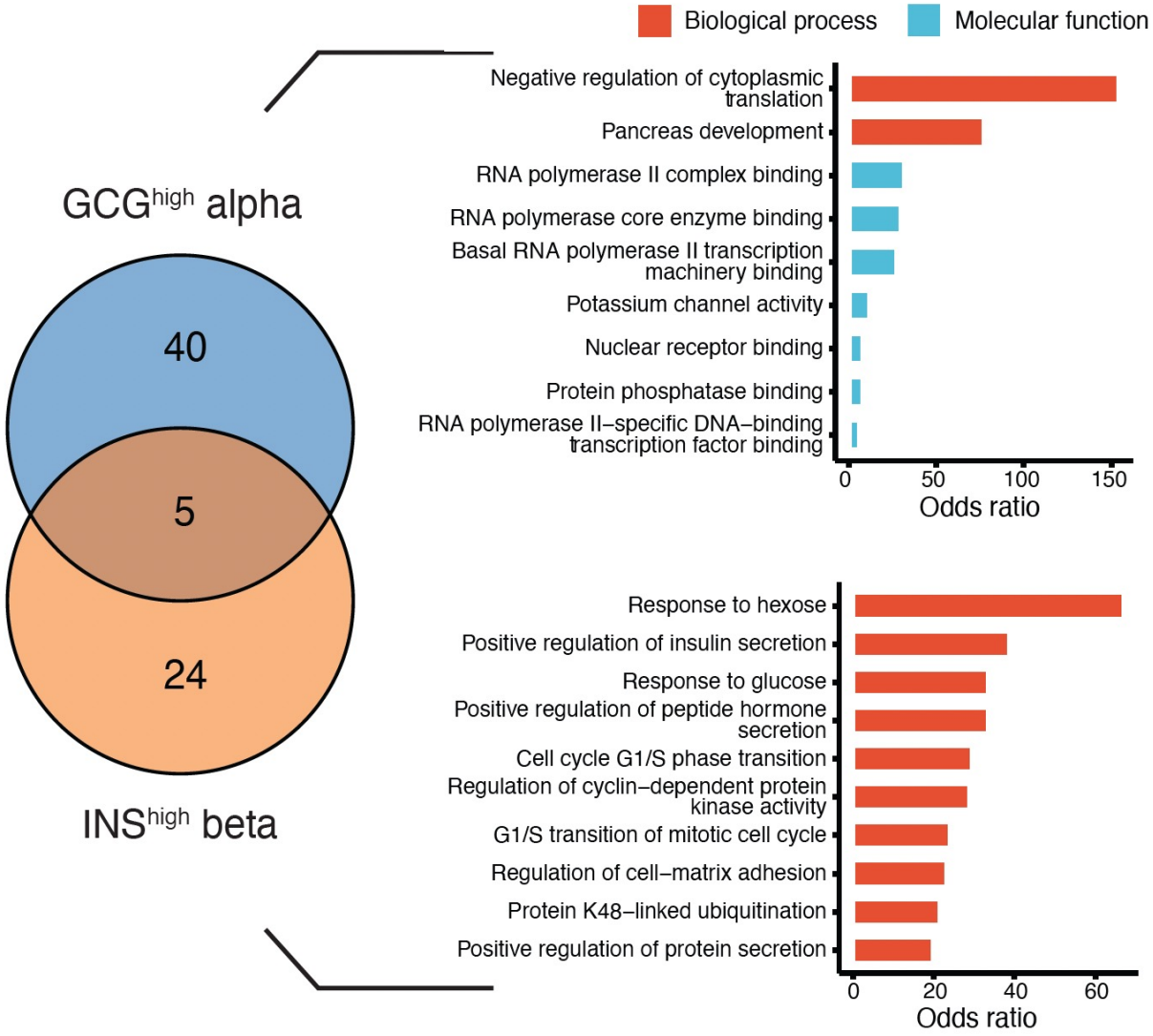
- **GCG^{high} alpha:** counterpart of beta cells responsible for producing glucagon, increasingly linked to T2D
- **INS^{high} beta:** beta cell dysfunction and cell death are key processes in the development of T2D

We for the first time link alpha cell to T2D genetically!

scPRS identifies T2D risk genes and pathways

Beta cell genes, 19 out of 29 (65.5%) are new.

Alpha cell genes, 32 out of 45 (71.1%) are new.



Conclusions

1. Novel AI approaches have led to new genes involved in AAA, ME/CFS, ALS, Long COVID

   →Increase in heretability explained (e.g ALS 6% → 36%)

2. By combine AI approaches with single cell data we can identify relevant cell types Severe COVID: NK & T-cells; Diabetes: alpha cells

3. A focus on cell-specific genetic drivers has enabled us to find genetic associations where traditional approaches are underpowered.

Acknowledgements

Stanford
Jingjing li
Sai Zhang
Cuiping Pan
Fereshteh Jahanbani
Varuna Chander
Martin Kjellberg
Phil Tsao

Durga Thota
Rajan Douglas Maynard
Tristan Chou
Michael Zeineh

Florida
Menghui Liu
Sai Zhang

UK
Johnathan Cooper-Knock

ALS Work:
USC
Justin Ichida

ME/CFS Cornell
Katherine A. Glass
David S. Lu
Faraz Ahmed
Andrew Grimson
Jennifer K. Grenier
Maureen R. Hanson

Carnegie Mellon
Martin Jinye Zhang

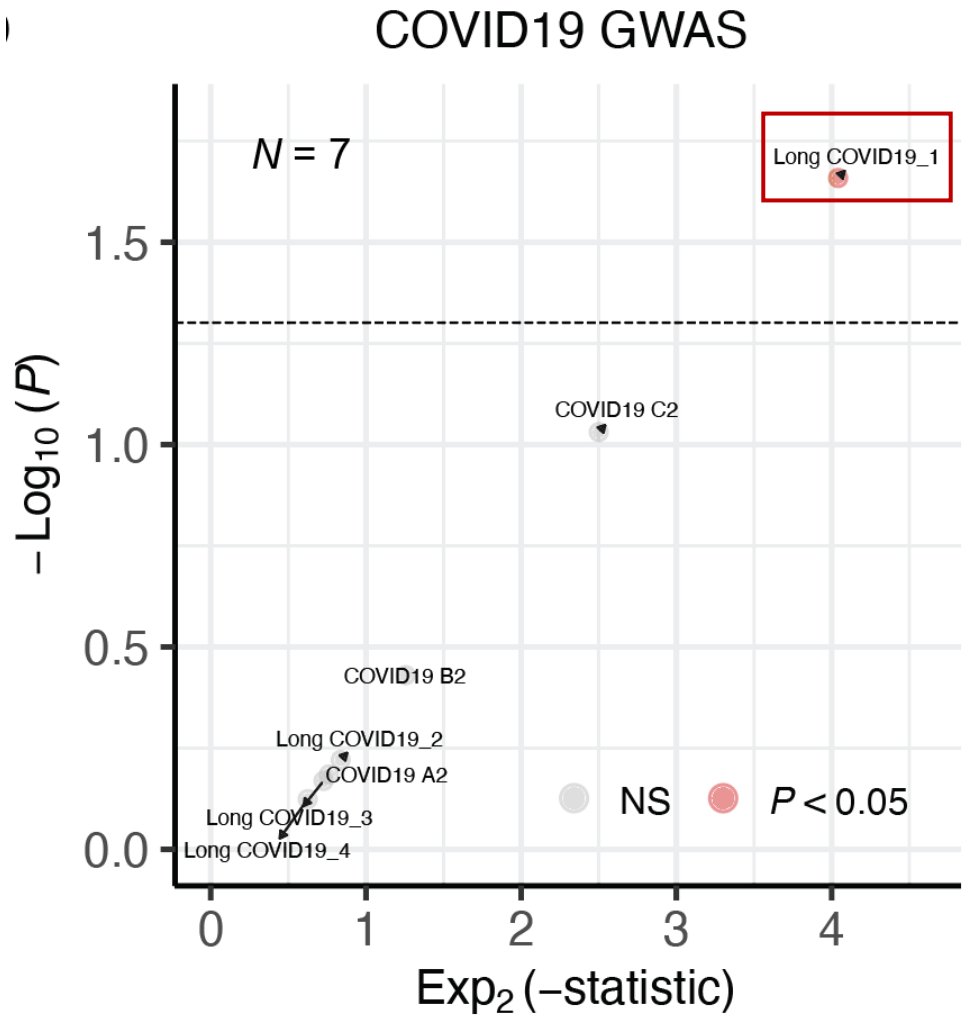
China
Han Li

Support: MPS: NIH, Vinod Khosla, Candace & Bert Forbes

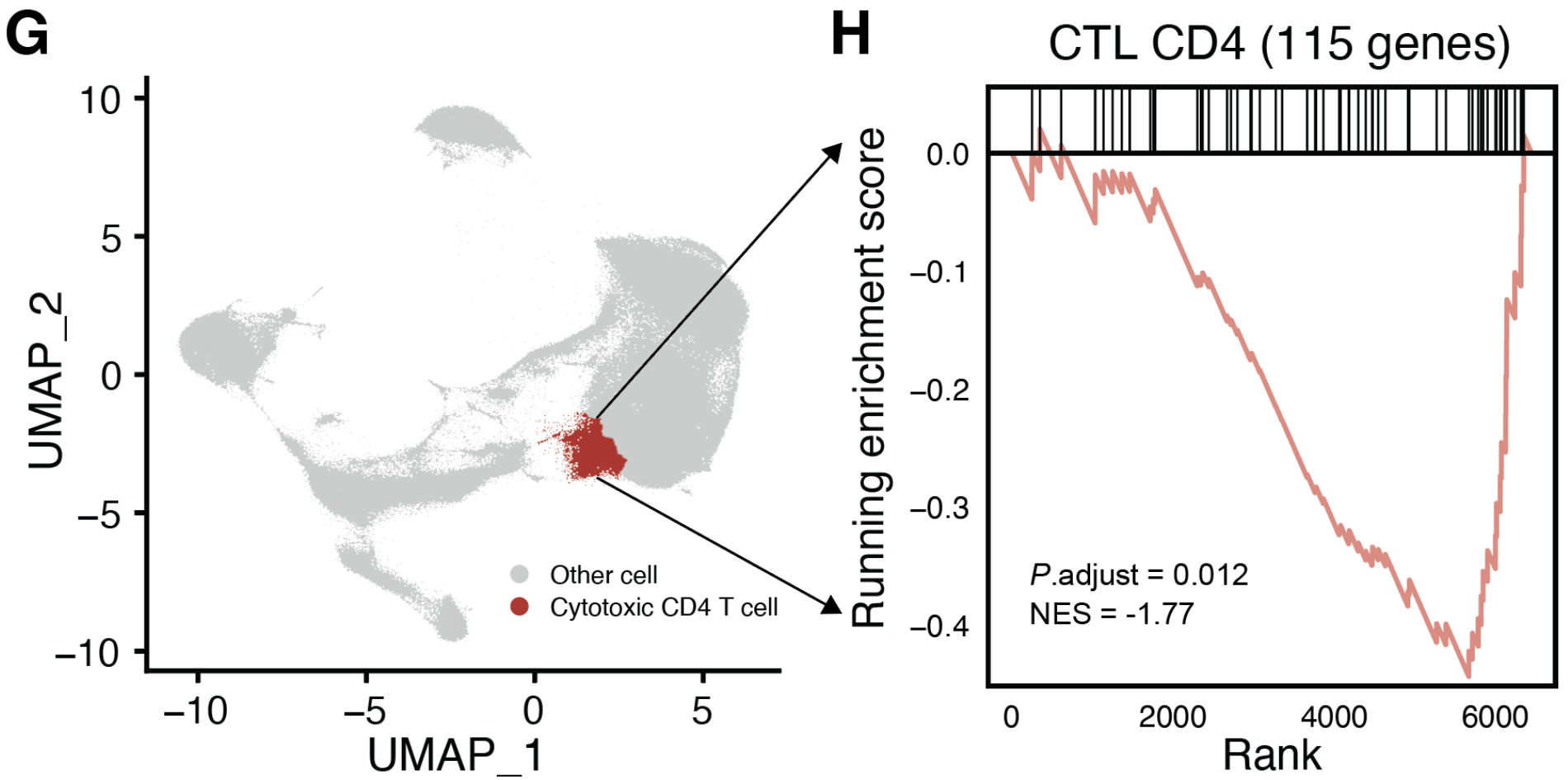
MJZ: Shurl and Kay Curci Foundation.

MH: NIH U54NS105541, U54AI178855

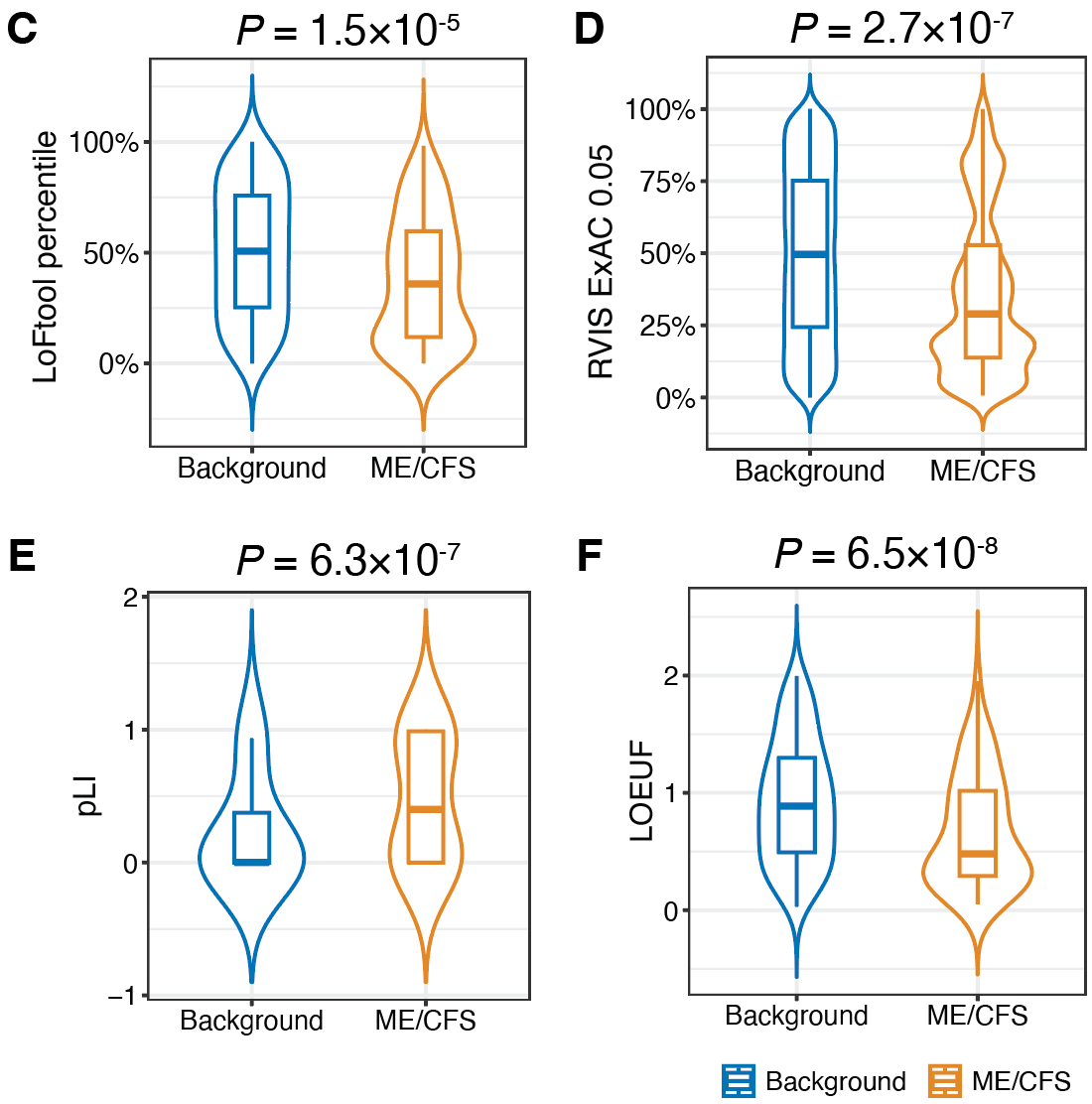
ME/CFS is genetically correlated with long COVID19



ME/CFS genes lower expressed in patient cytotoxic T cells



ME/CFS genes are intolerant to LoF



Identifying Genes Associated with AAA

268 AAA cases

133 AAA controls

50X

Whole genome deep sequencing

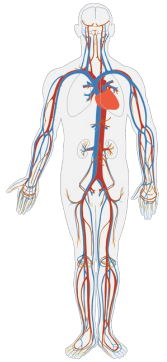
EHR mutation network gene

HEAL: a machine learning framework

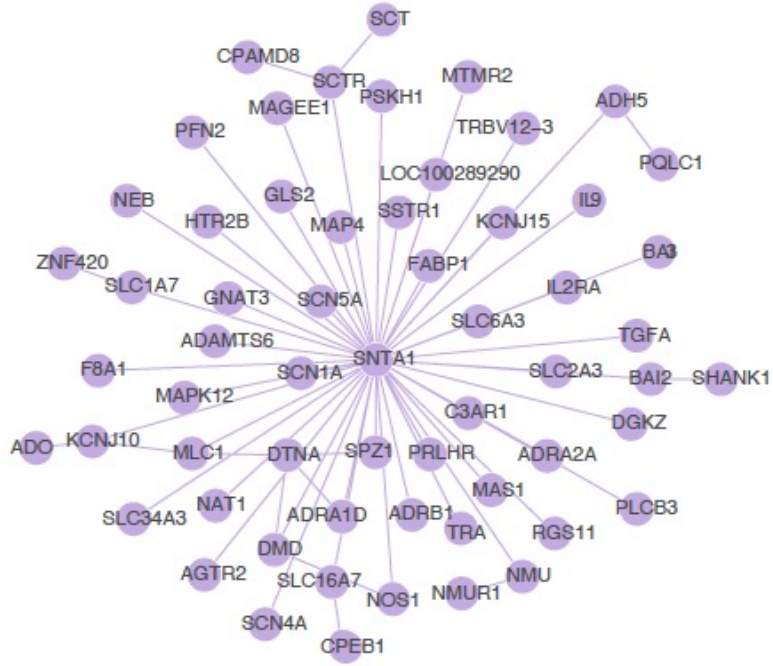
Call pathogenic mutations

HNRNPCL1 HLA-DQB2 CBWD3 PYGM KCNC3

Identify 60 AAA genes

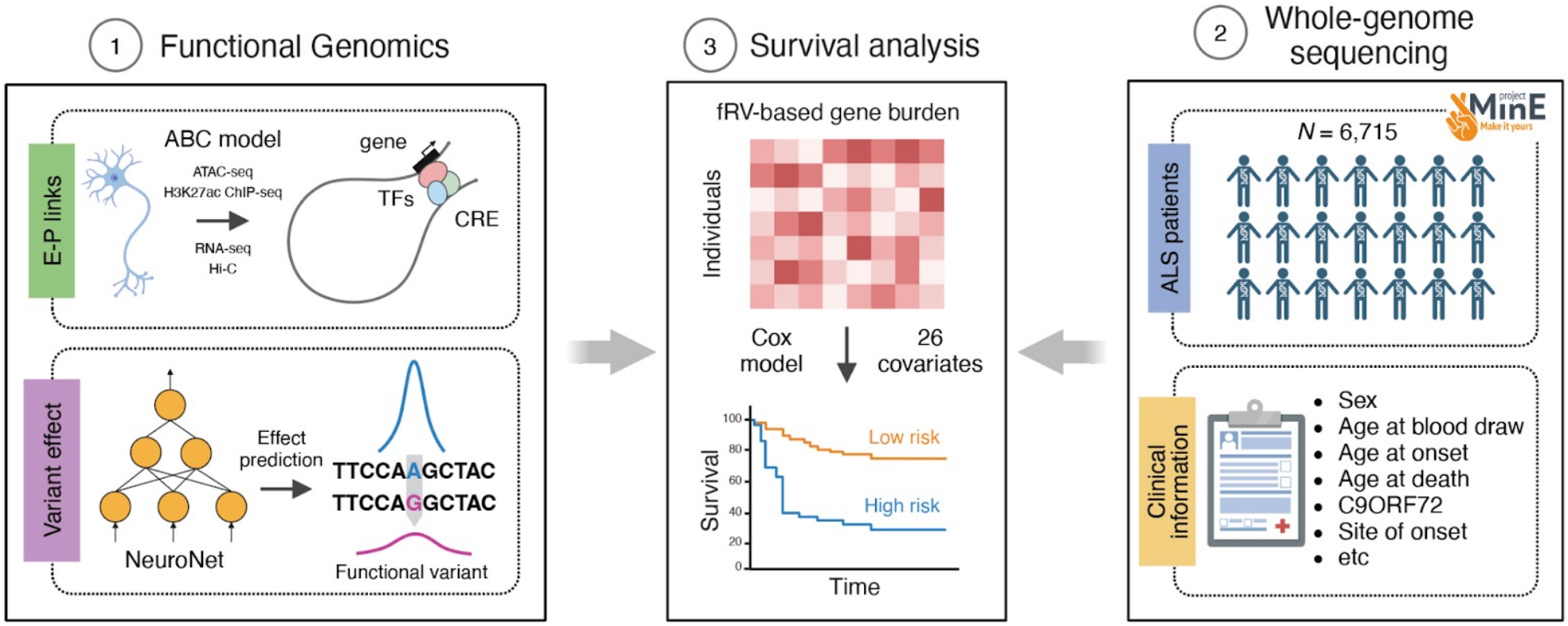


Blood Circulation, Blood Pressure, Cardiomyopathy



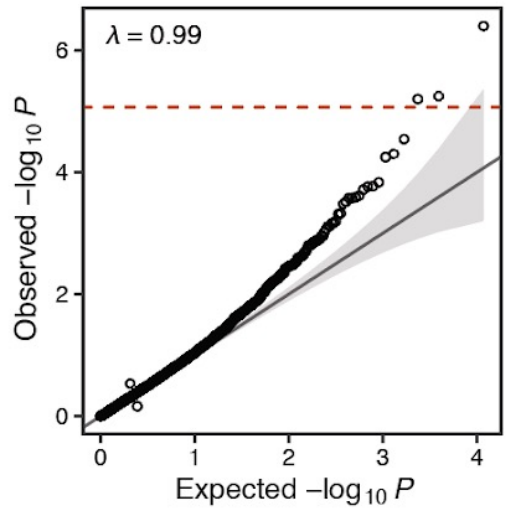
Jingjing LI, Cuiping Pan, Sai Zhang .. Phil Tsao, Cell 2018

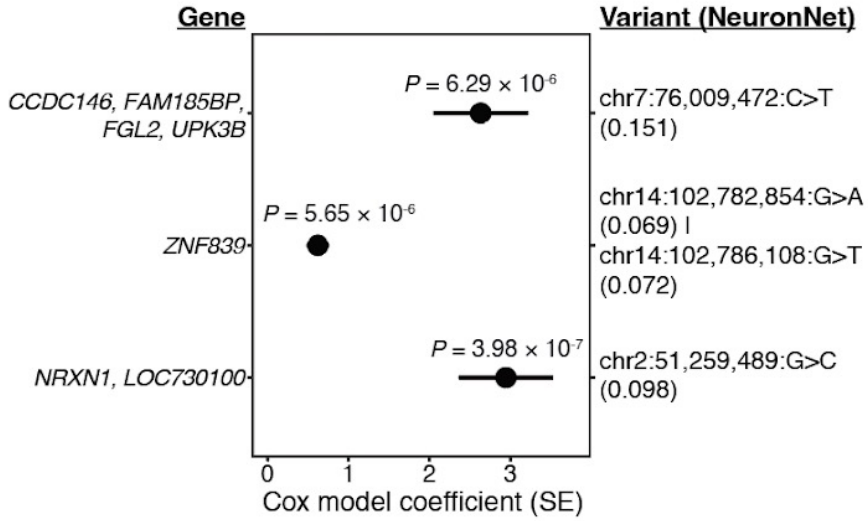
Non-coding mutations which increase motor neuron expression of *CCDC146* shorten ALS survival



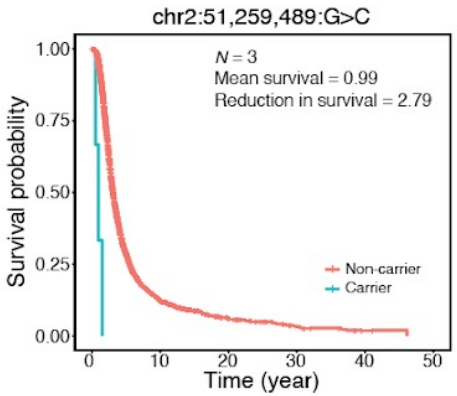
We identified non-coding rare variants which are functional within motor neurons and are linked to ALS survival

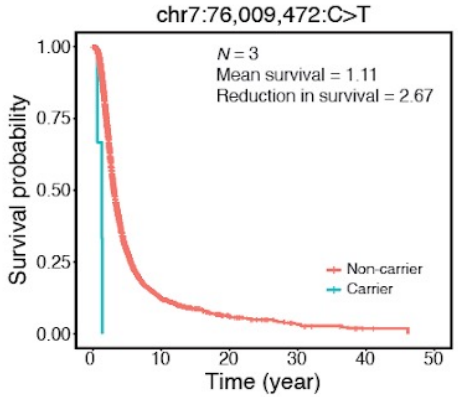
Rare variant burden testing using non-coding variants prioritised by NeuroNet

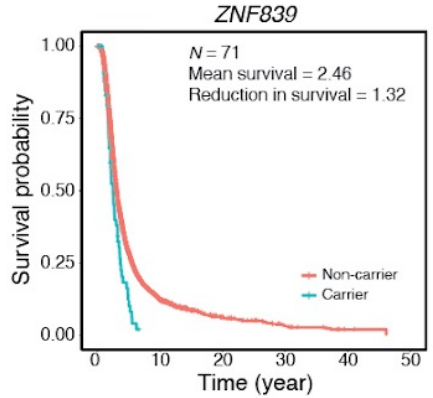




We identified non-coding variants linked to ALS survival but there was some ambiguity regarding target genes which required experimental interrogation.

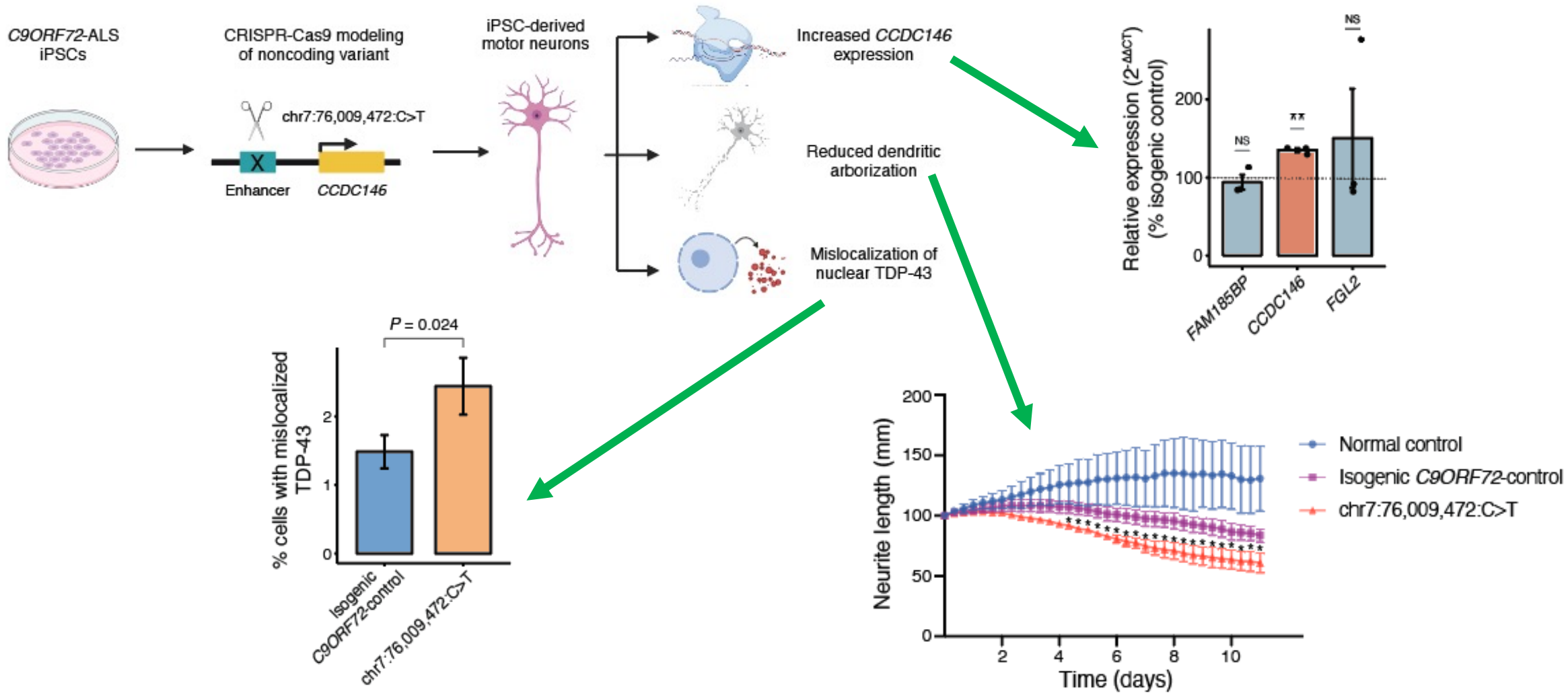




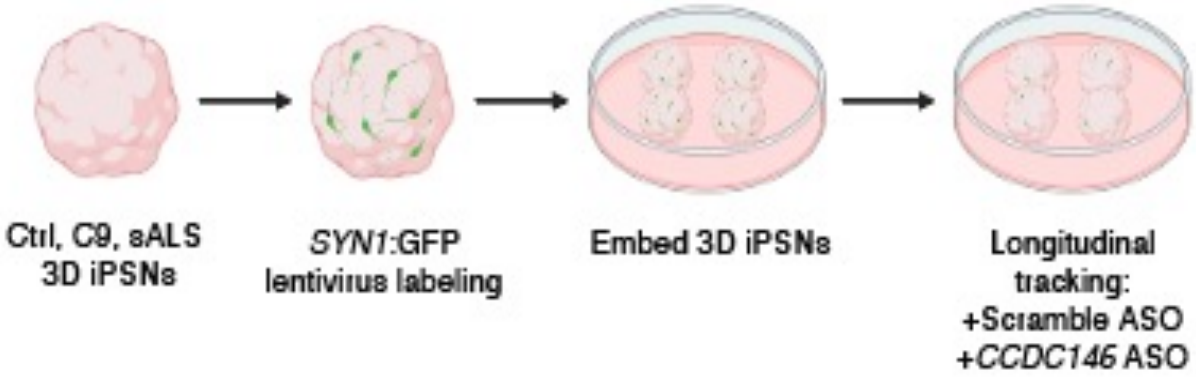


Non-coding variants are associated with >50% reduction in ALS survival.

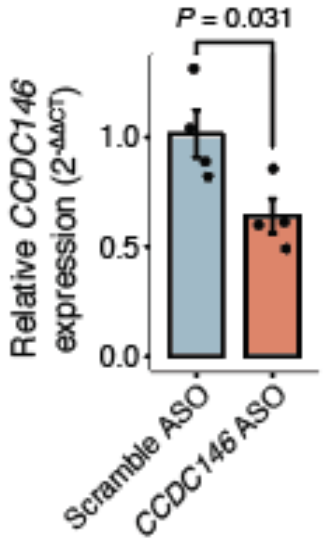
The chr7:76,009,472:C>T variant increases CCDC146 expression in MN and exacerbates ALS-associated phenotypes



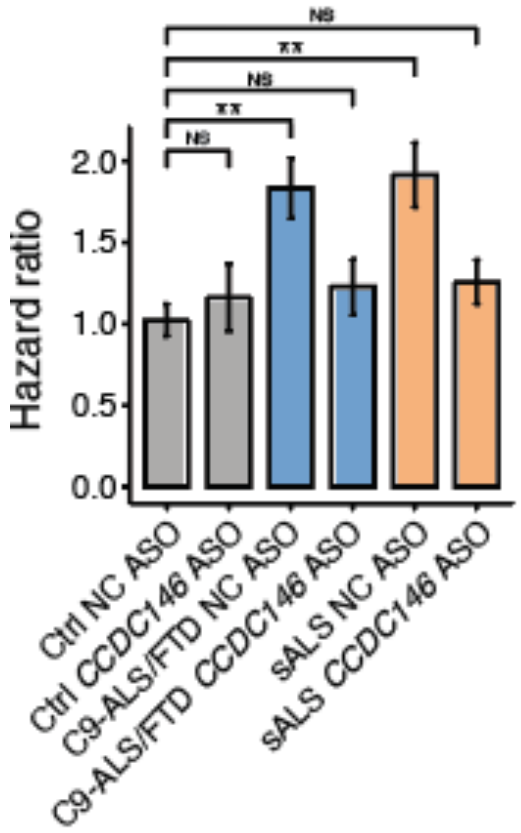
ALS-associated survival defects in iPSC-derived neurons are completely rescued by ASO which reduces *CCDC146* expression



We tested an ASO knockdown of *CCDC146* in a 3D cultures of patient-derived neurons.

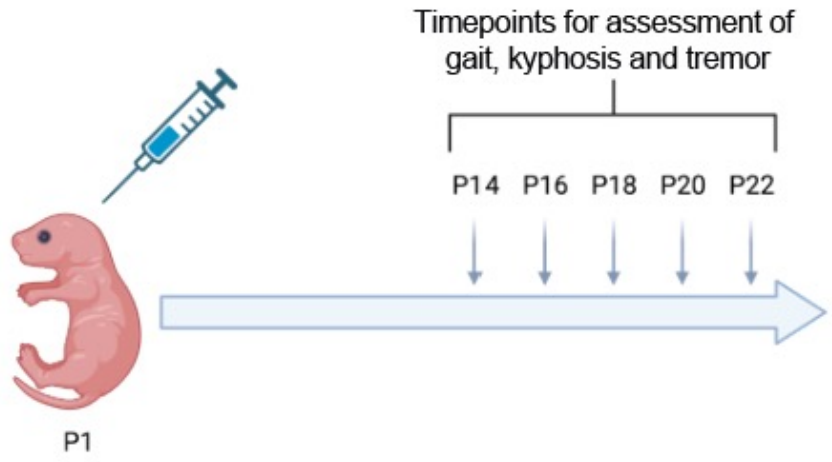


The ASO effectively reduces *CCDC146* expression



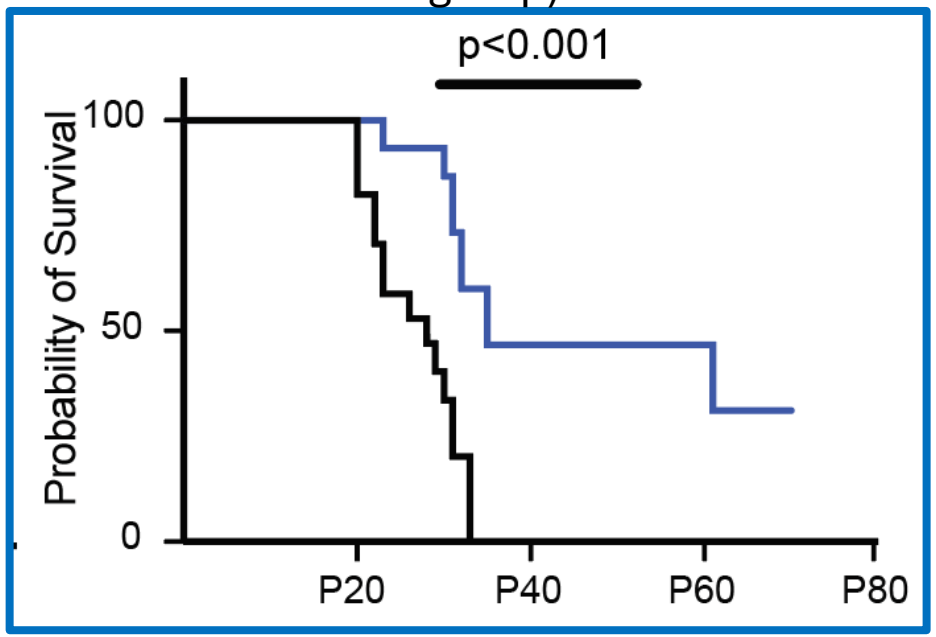
The ASO is not toxic to control neurons (grey) but completely rescues ALS patient neurons from glutaminergic insult. This has now been extended to n=15 different patient lines.

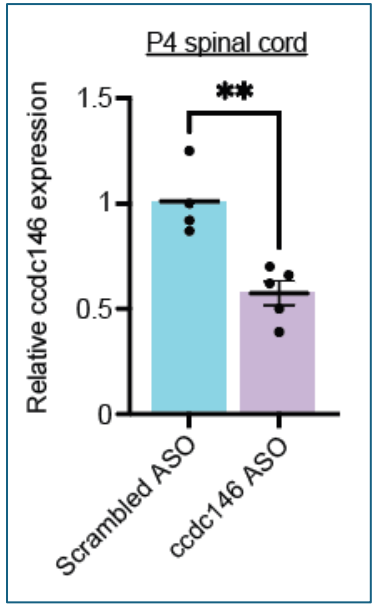
ASO against murine *ccdc146* rescues motor phenotypes and extends survival in an aggressive mouse model of ALS



TAR4/4 model, Wils et al, 2010, *PNAS*

The ASO extends survival in TAR4/4 mice (n=15 each group)





The ASO effectively reduces *ccdc146* expression

Knockdown of CCDC146 is likely to be well tolerated

- Humans and mice without CCDC146 present with non-syndromic male infertility only (Muroňová et al, 2024, *eLife*).
- ASO against CCDC146 was NOT toxic in 3D cultures of neurons derived from neurologically normal individuals, compared to scrambled ASO.
- ASO against murine ccdc146 is NOT toxic in non-transgenic mice.