Biobank Project:
Progress & Plans

Carlos Cruchaga. PhD

Jen Gentsch, BSN, RN
JoAnne Norton, MSN, RN, PMHCNS-BC

Current Biobank

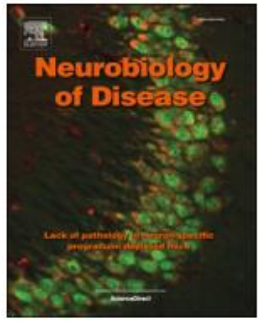
- Obtain baseline blood for DNA
- DNA for more than >3,000 Dystonia Coalition participants.
  - Around 50% are from cervical dystonia (CD)
- DNA has been instrumental to identify new genes for dystonia

Brief history of Dystonia genetics

Contents lists available at ScienceDirect

Neurobiology of Disease

journal homepage: www.elsevier.com/locate/ynbdi



Review

Dystonia genes and their biological pathways

H.A. Jinnah[a,*], Yan V. Sun[b]

[a] Departments of Neurology, Human Genetics and Pediatrics, Emory University School of Medicine, Atlanta, GA, United States of America
[b] Department of Epidemiology, Emory University Rollins School of Public Health, Atlanta, GA, United States of America

Jinnah, et al., 2019

Brief history of Dystonia genetics

Table 1
Genes associated with isolated dystonia.

Gene	Protein	Function	Inheritance	Typical phenotype
ANO3	Anoctamin-3	Related to family of calcium-activated chloride channels	AD (~50% penetrant)	Adult-onset craniocervical dystonia with prominent tremor
CIZ1	Cip1-interacting zinc finger protein 1	DNA binding and replication	AD (reduced penetrance)	Adult-onset cervical dystonia
COL6A3	α3 subunit of type VI collagen	Major component of collagen in most connective tissues	AR	Childhood-onset cervical and upper body
GNAL	G-protein subunit alpha L	Cell surface guanine-nucleotide binding protein involved in signal transduction	AD (~50% penetrance)	Adult-onset focal or segmental involving neck
THAP1	Thanatos-associated domain-containing apoptosis protein 1	DNA binding and transcription	AD (~60% penetrance)	Adolescent-onset craniocervical and upper body
TOR1A	Torsin family 1 member A	Endoplasmic reticulum protein chaperone	AD (~30% penetrance)	Childhood-onset limb or generalized

Abbreviations: AD, autosomal dominant.

Table 2
Selected dystonia genes associated with dopamine signaling.

Gene	Protein	Function	Inheritance	Typical phenotype
ADCY5	Adenylate cyclase 5	Couples dopamine receptors to cAMP messenger systems	AR	Childhood-onset mixed motor dystonia with chorea and dystonia
DDC	Aromatic amino acid decarboxylase	Enzyme involved in dopamine synthesis	AR	Infantile encephalopathy, oculogyric crises, dysautonomia, dystonia
GCH1	GTP-cyclohydrolase 1	Rate-limiting step in tetrahydropterin synthesis	AD (partial penetrance)	DRD
GNAL	G-protein subunit alpha L	Cell surface guanine-nucleotide binding protein involved in dopamine receptor signal transduction	AD (partial penetrance)	Adult-onset focal or segmental involving neck
GNAO1	G-protein subunit alpha 01	Cell surface guanine-nucleotide binding protein involved in dopamine-receptor signal transduction	AR	Childhood-onset mixed motor dystonia with chorea and dystonia
HPRT1	Hypoxanthine-guanine phosphoribosyl transferase	Purine metabolism enzyme that results in marked dopamine deficiency without neurodegeneration	XL	Lesch-Nyhan disease
LRRK2	Leucine-rich repeat kinase 2	Protein kinase associated with dopamine neuron degeneration	AD	Parkinson's disease, often with dystonia
POLG	Polymerase gamma 1	Mitochondrial enzyme that results in loss of dopamine neurons	AR	Markedly varied from infancy to later adulthood
PINK1	PTEN-induced putative kinase 1	Protein kinase associated with dopamine neuron degeneration	AR	Parkinson's disease, often with dystonia
PRKN	Parkin	Ubiquitin-ligase protein associated with dopamine neuron degeneration	AR	Parkinson's disease, often with dystonia
PTPS	6-pyruvoyl-tetrahydropertin synthase	Enzyme involve in tetrahydropertin synthesis	AR	DRD or infantile encephalopathy
SLC18A2	Vesicular monoamine transporter 2	Presynaptic storage of dopamine in vesicles	AR	Infantile-onset parkinsonism, dystonia, dysautonomia
SLC6A3	Dopamine transporter	Presynaptic uptake of dopamine from the synaptic cleft	AR	Infantile to adult onset parkinsonism and dystonia
SNCA	α-synuclein	Presynaptic protein associated with dopamine neuron degeneration	AD	Parkinson's disease, often with dystonia
SPR	Sepiapterin reductase	Enzyme involve in tetrahydropertin synthesis	AR	DRD or infantile encephalopathy
TH	Tyrosine hydroxylase	Enzyme involved in rate-limiting step in dopamine synthesis	AR	DRD or infantile encephalopathy

Abbreviations: AD, autosomal dominant; AR, autosomal recessive; DRD, dopa-responsive dystonia.

Jinnah, et al., 2019

Brief history of Dystonia genetics

Table 3
Genes associated with NBIA disorders.

Gene	Protein	Function	Inheritance	Disorder
ATP13A2	P-type ATPase 13A2	Lysosomal cation pump	AR	Kufor-Rakeb disease
C19orf12	Unknown	Mitochondrial membrane protein	AR	Mitochondrial membrane protein associated neurodegeneration (MPAN)
COASY	CoA synthase	CoA synthesis	AR	COASY protein associated neurodegeneration (CoPAN)
CP	Ceruloplasmin	Heavy metal transport	AR	Aceruloplasminemia
DCAF17	DDB1 and CUL4 associated factor 17	Nucleolar protein	AR	Woodhouse-Sakati disease
FA2H	Fatty acid hydroxylase type 2	Fatty acid metabolism	AR	Fatty acid hydroxylase associated neurodegeneration
FTL	Ferritin light chain	Heavy metal transport	AD	Hereditary neuroferritinopathy
GTPBP2	GTP binding protein	Suspected mRNA metabolism	AR	NA
PANK2	Pantothenate kinase type 2	CoA synthesis	AR	Pantothenate-kinase associated neurodegeneration (PKAN)
PLA2G6	Phospholipase A2	Phospholipid metabolism	AR	PLA2G6 associated neurodegeneration (PLAN)
SCP2	Sterol carrier proteins X and 2	Metabolism of fatty acids	AR	Leukoencephalopathy with dystonia and motor neuropathy
WDR45	β-propeller protein	Autophagosome protein	XL	β-propeller associated neurodegeneration (BPAN)

Abbreviations: AD, autosomal dominant; AR, autosomal recessive, CoA, coenzyme A; XL, X-linked.

DYSTONIA COALITION

Jinnah, et al., 2019

Brief history of Dystonia genetics

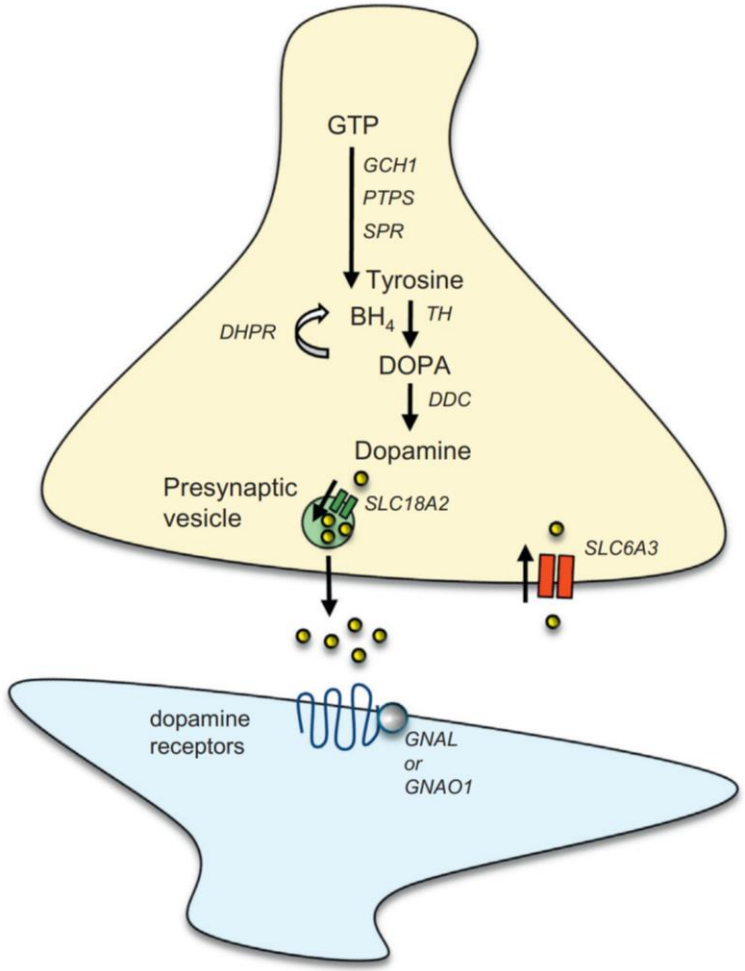
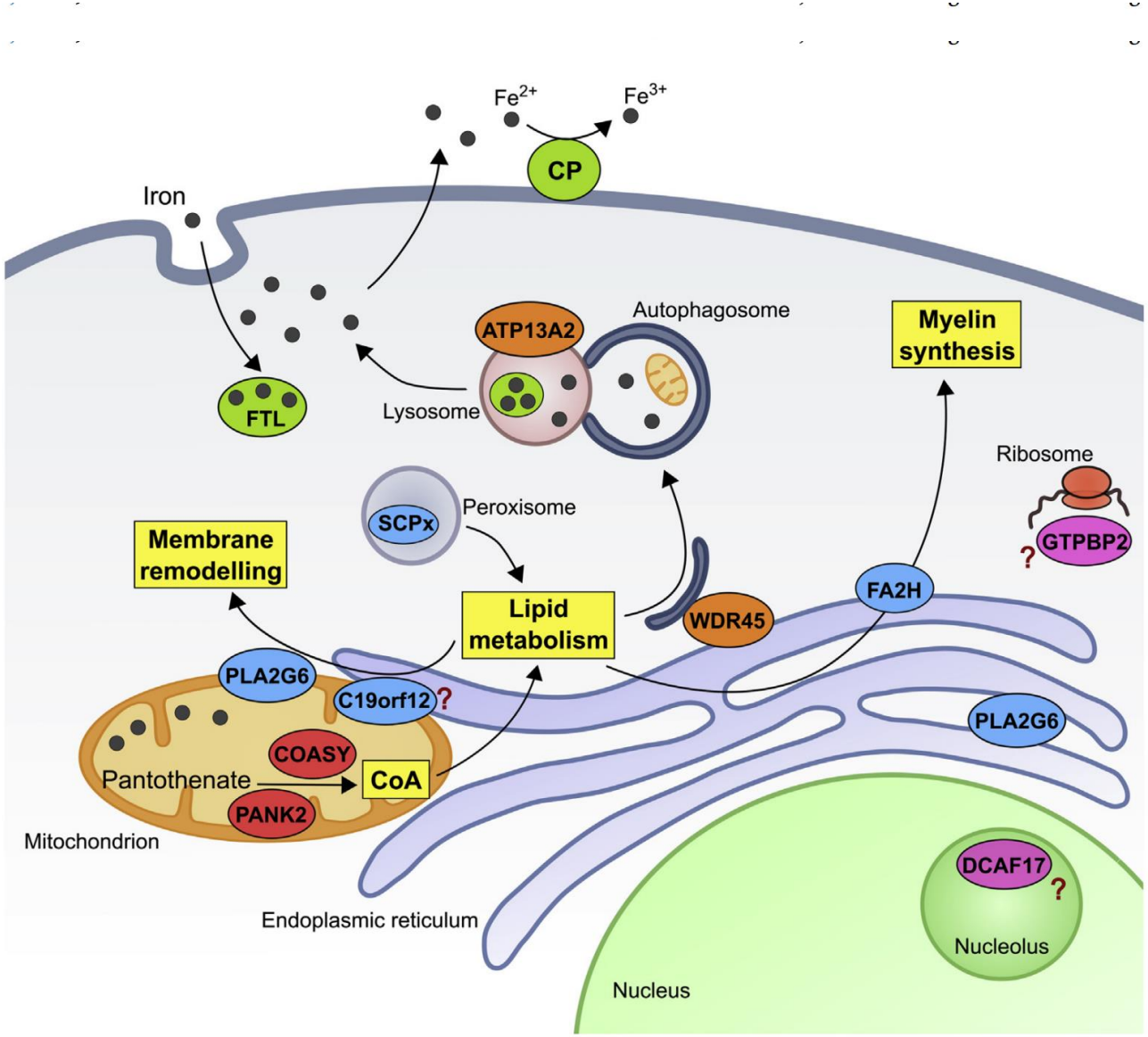
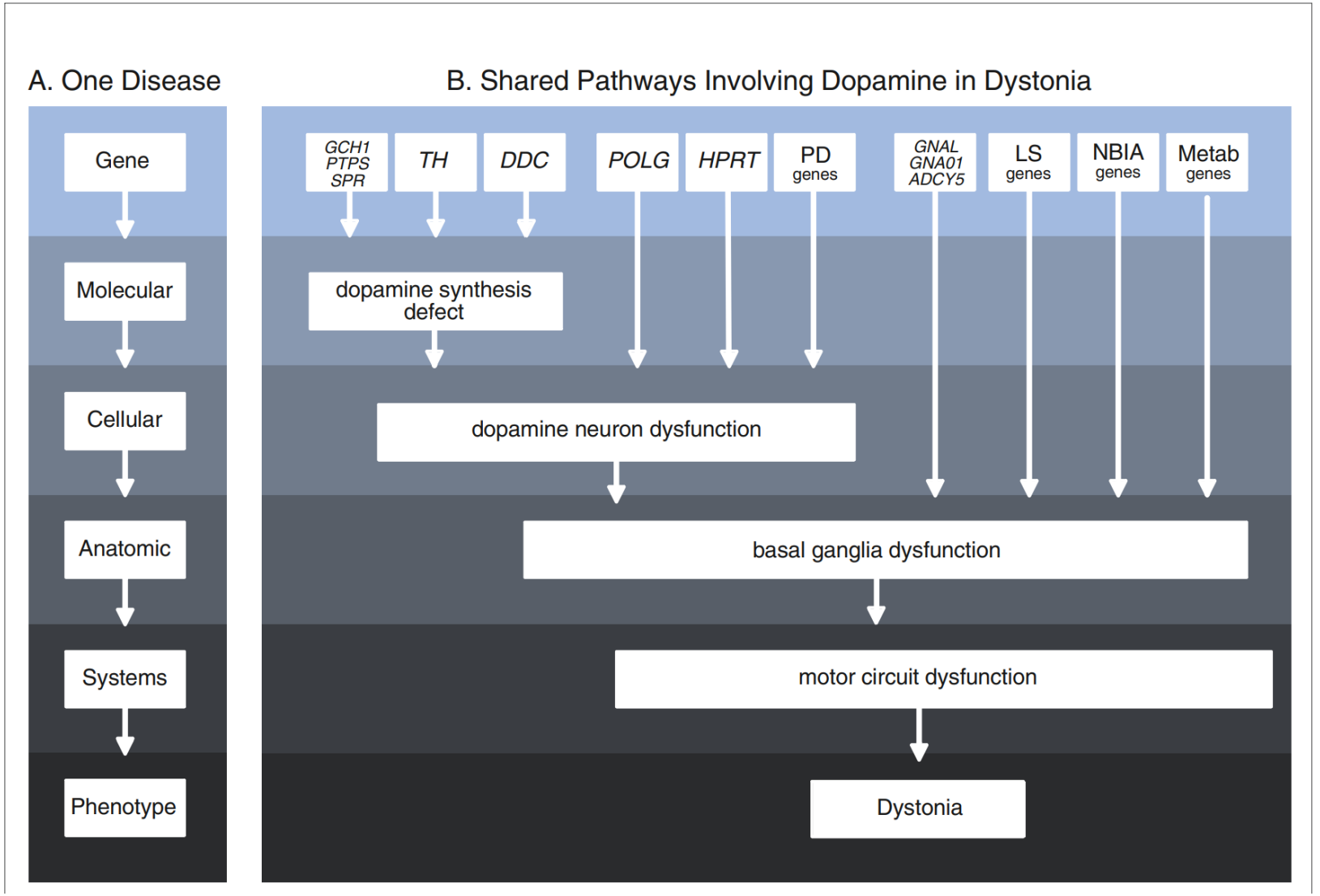


Fig. 1. Dopaminergic signaling in dystonia. The presynaptic dopamine neurons



Jinnah, et al., 2019

Brief history of Dystonia genetics



Jinnah, et al., 2019

Genetic Characterization of DC participants

Mark S. LeDoux, MD, PhD
Satya R. Vemula, PhD
Jianfeng Xiao, MD, PhD
Misty M. Thompson, PhD
Joel S. Perlmutter, MD
Laura J. Wright, MA
H.A. Jinnah, MD, PhD
Ami R. Rosen, MS
Peter Hedera, MD, PhD
Cynthia L. Comella, MD
Anne Weissbach, MD
Johanna Junker, MD
Joseph Jankovic, MD
Richard L. Barbano, MD, PhD
Stephen G. Reich, MD
Ramon L. Rodriguez, MD
Brian D. Berman, MD
Sylvain Chouinard, MD
Lawrence Severt, MD, PhD
Pinky Agarwal, MD
Natividad P. Stover, MD
On behalf of the Dystonia Coalition Investigators, Dystonia Genetic Consortium

Genetic Characterization of DC participants

Mark S. LeDoux, MD, PhD
Satya R. Vemula, PhD
Jianfeng Xiao, MD, PhD
Misty M. Thompson, PhD
Joel S. Perlmutter, MD
Laura J. Wright, MA
H.A. Jinnah, MD, PhD
Ami R. Rosen, MS
Peter Hedera, MD, PhD
Cynthia L. Comella, MD
Anne Weissbach, MD
Johanna Junker, MD
Joseph Jankovic, MD
Richard L. Barbano, MD, PhD
Stephen G. Reich, MD
Ramon L. Rodriguez, MD
Brian D. Berman, MD
Sylvain Chouinard, MD
Lawrence Severt, MD, PhD
Pinky Agarwal, MD
Natividad P. Stover, MD
On behalf of the Dystonia Coalition Investigators, Dystonia Genetic Consortium

Neurology: Genetics

Table 3 Bioinformatic analysis of SVs identified in participants with CD

						Minor allele frequency				In silico pathogenicity/deleteriousness					Splicing (human splice finder 3.0)			
Gene	Participant (s)	cDNA	Protein	ACMG classification	dbSNP	EVS	1KG	ExAC	DC	SIFT	MutationTaster	PolyPhen-2	CADD raw	CADD Phred	Splice site distance	HSF matrices	ESE finder	ESR sequences
GNAL	DYS1579	c.40C>T	p.Q14*	Pathogenic					0.0005	NA	Disease causing	NA	2.72	22.5	142	None	None	ESE site broken
GNAL	n = 5	c.66C>T	p.R22R	Uncertain significance	rs73397885	0.01	0.008	0.00562	0.0025	Tolerated	Polymorphism	NA	2.43	22.1	79	New donor site	ESE site broken	None
GNAL	DYS497	c.139C>A	p.L47I	Likely pathogenic					0.0005	Damaging	Disease causing	Benign	2.55	22.3	6	None	ESE site broken	ESE site broken
GNAL	DYS734	c.214C>G	p.P72A	Likely pathogenic					0.0005	Tolerated	Disease causing	Benign	1.49	15.9	4	None	None	New ESS site
GNAL	n = 135	c.218+19C>A	NA	Likely benign	rs1895689	0.000307	0.348	0.2968	0.078	NA	NA	NA	1.16	14.1	19	None	None	None
GNAL	n = 74	c.932-7T>G	NA	Benign	rs3892113	0.06	0.044	0.0628	0.041	NA	NA	NA	0.05	4.1	7	None	None	None
GNAL	n = 33	c.1014C>T	p.A338A	Uncertain significance	rs41289504	0.01	0.0034	0.008831	0.016	Tolerated	Disease causing	NA	2.01	19.2	15	None	ESE site broken	New ESS site
GNAL	DYS41	c.1018G>A	p.G340S	Likely pathogenic	rs142792291	0.0001538		0.0000249	0.0005	Tolerated	Disease causing	Probably damaging	2.13	20.3	19	New acceptor site	None	New ESS site
GNAL	DYS87	c.1060G>A	p.V354M	Pathogenic					0.0005	Damaging	Disease causing	Probably damaging	2.57	22.4	61	None	ESE site broken	New ESS site
THAP1	DYS1729	c.57C>T	p.P19P	Uncertain significance	rs146087734	0.00023		0.000099	0.0005	Tolerated	Disease causing	NA	1.93	18.6	14	None	None	None
THAP1	DYS1706	c.71+9C>A	NA	Likely pathogenic	rs200209986	0.001999		0.001291	0.0005	NA	NA	NA	0.93	12.6	9	None	None	None
THAP1	DYS88	c.153C>G	p.S51R	Pathogenic					0.0005	Damaging	Disease causing	Probably damaging	5.06	27.1	81	None	None	None
THAP1	DYS1068	c.427A>G	p.M143V	Uncertain significance	rs374512193	0.00007		0.00007413	0.0005	Tolerated	Polymorphism	Benign	0.85	12.1	159	New acceptor site	None	None
TOR1A Exon 5	DYS266	c.823A>G	p.K275E	Uncertain significance	rs148036363	0.0002307		0.0003789	0.0005	Tolerated	Disease causing	Probably damaging	2.27	21.4	74	None	ESE site broken	None
TOR1A Exon 5	DYS1754	c.907_909 delGAG	p.E303del	Pathogenic	rs724159981				0.0005	NA	Disease causing	NA	2.73	22.5	158	None	ESE site broken	New ESS site
TOR1A Exon 5	DYS1565	c.962C>T	p.T321M	Uncertain significance				0.00001647	0.0005	Damaging	Disease causing	Probably damaging	2.06	19.6	213	None	None	New ESS site

cDNA/protein (*GNAL*-NM_001142339.2/NP_001135811.1, *THAP1*- NM_018105.2/NP_060575.1, *TOR1A*-NM_000113.2/NP_000104.1).
Abbreviations: 1KG = 1000 Genomes Project; ACMG = American College of Medical Genetics and Genomics; CADD = Combined Annotation-Dependent Depletion; CD = cervical dystonia; cDNA = complementary DNA; dbSNP = Single Nucleotide Polymorphism database; DC = Dystonia Coalition; ESE = exonic splicing enhancer; ESR = exonic splicing regulator; ESS = exonic splicing silencer; EVS = Exome Variant Server;

Common variants with low effect size

EMORY ROLLINS SCHOOL OF PUBLIC HEALTH

Genome-wide Association Study Identifies Common Genetic Variants Associated with Cervical Dystonia

Yan V. Sun[1,2], Chengchen Li[1], Qin Hui[1], Joel S. Perlmutter[3], Samantha Ruehl[4], Christine Klein[5], Joseph Jankovic[6], Richard L. Barbano[7], Stephen G. Reich[8], J. Douglas Bremner[9,10], Viola Vaccarino[1], Arshed A. Quyyumi[11], H. A. Jinnah[12], on behalf of the Dystonia Coalition Investigators

[1]Department of Epidemiology, Rollins School of Public Health, Emory University, Atlanta, Georgia, USA; [2]Department of Biomedical Informatics, Emory University School of Medicine, Atlanta, Georgia, USA; [3]Department of Neurology and Neurological Surgery, Washington University School of Medicine, St. Louis, MO, USA; [4]Department of Neurological Sciences, Rush Medical College, Chicago, IL, USA; [5]The Institute of Neurogenetics, University of Luebeck, Luebeck, Germany; [6]Department of Neurology, Baylor College of Medicine, Houston, TX, USA; [7]Department of Neurology, University of Rochester Medical center, Rochester, NY, USA; [8]Department of Neurology, University of Maryland School of Medicine, Baltimore, MD, USA; [9]Atlanta VA Medical Center, Decatur, Georgia, USA; [10]Departmnt of Radiology, Emory University School of Medicine, Atlanta, Georgia, USA; [11]Department of Medicine, Division of Cardiology, Emory University School of Medicine, Atlanta, Georgia, USA; [12]Department of Neurology and Human Genetics, Emory University School of Medicine, Atlanta, Georgia, USA

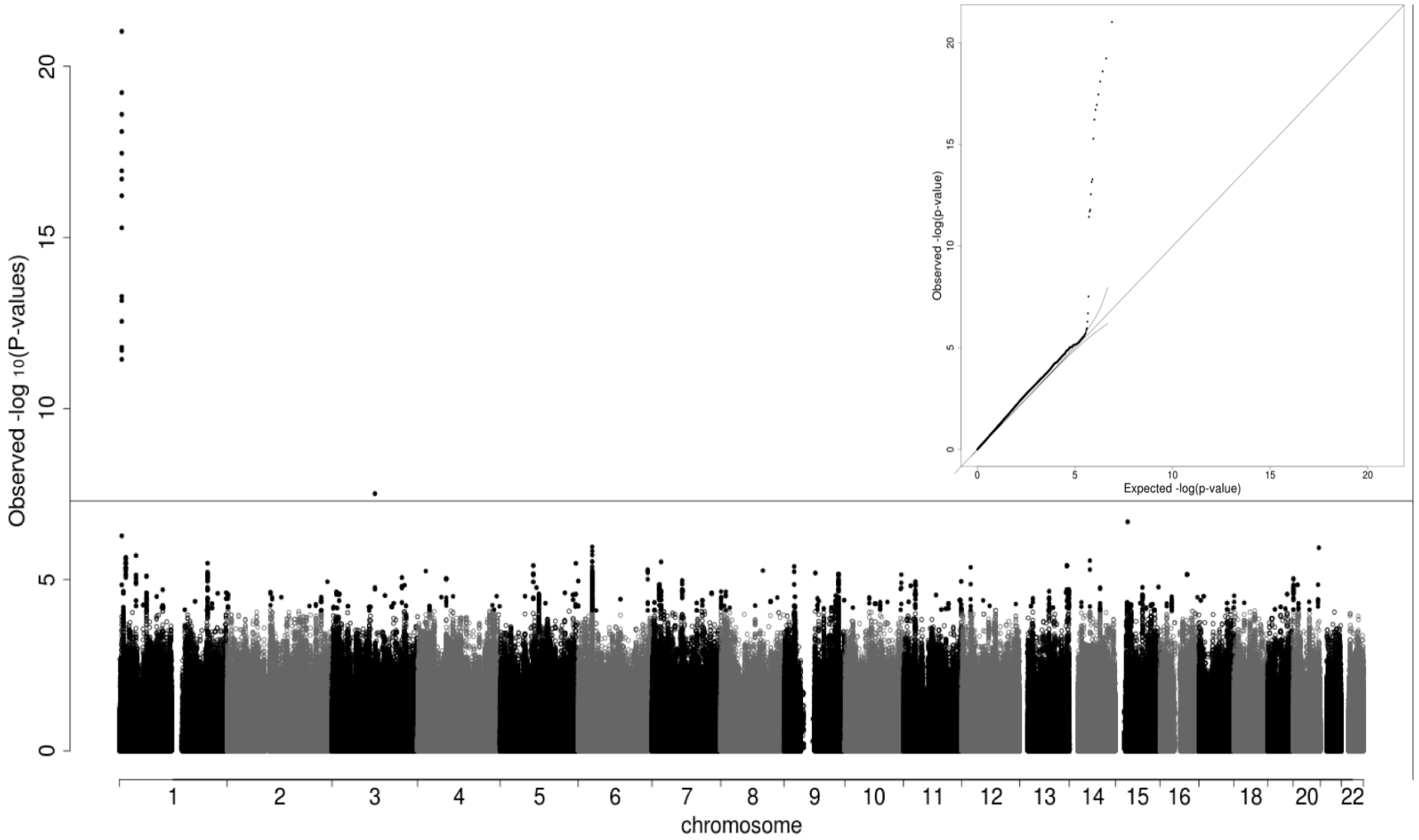
EMORY UNIVERSITY SCHOOL OF MEDICINE

919 cases and 1,491 controls

Other DC samples with GWAS data

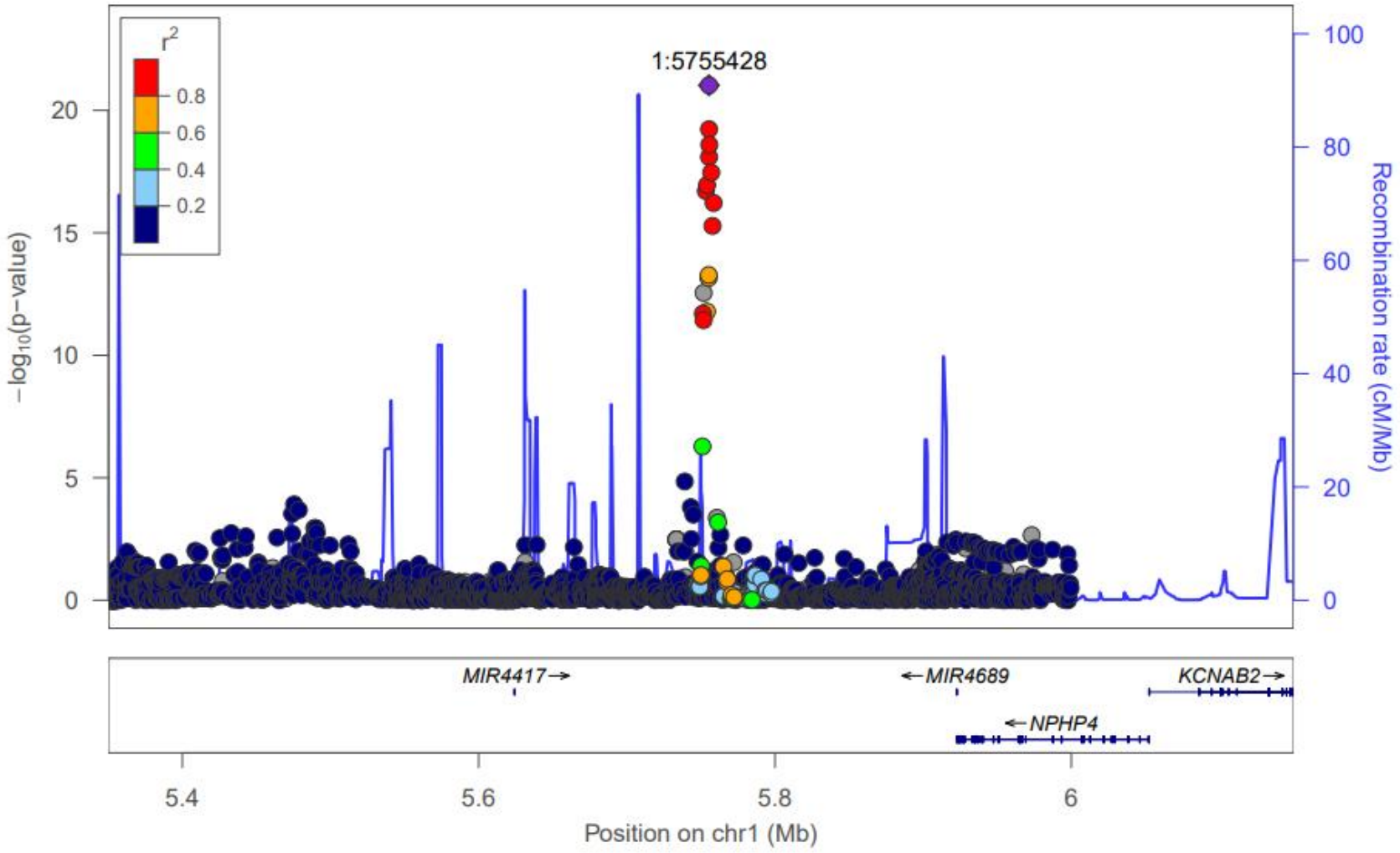
Type	Count
Focal	1446
Segmental	395
Multifocal	114
Hemi	8
Generalized	102

Common variants with low effect size



Sun et al, Unpublished

Common variants with low effect size



Sun et al, Unpublished

Common variants with low effect size

- largest GWAS for any type of dystonia to date
- 16 common genome-wide significant variants (p-value<5×10^{-8}) in 2 distinct loci
- Chromosome 1 is close to *NPHP4* and *KCNAB2*.
  - The *NPHP4* gene encodes nephrocystin 4, which is involved in renal tubular development and function
  - associated with kidney disease, sometimes with neurological dysfunction
- Chromsome 3 signal is close to *KCNAb2*
  - *KCNAB2* gene encodes a beta-2 subunit that modifies the properties of the voltage-gated potassium channel encoded by *KCNA4.*
  - highly expressed in many regions of the nervous system, where it regulates neuronal excitability and neurotransmitter release

Sun et al, Unpublished

Goals for the Biobank

- To extend the current DNA repository by targeting BSP, LD and limb dystonia subjects as well as multiplex families.
    - Sporadic and familial presentations
    - Longitudinal (each visit)

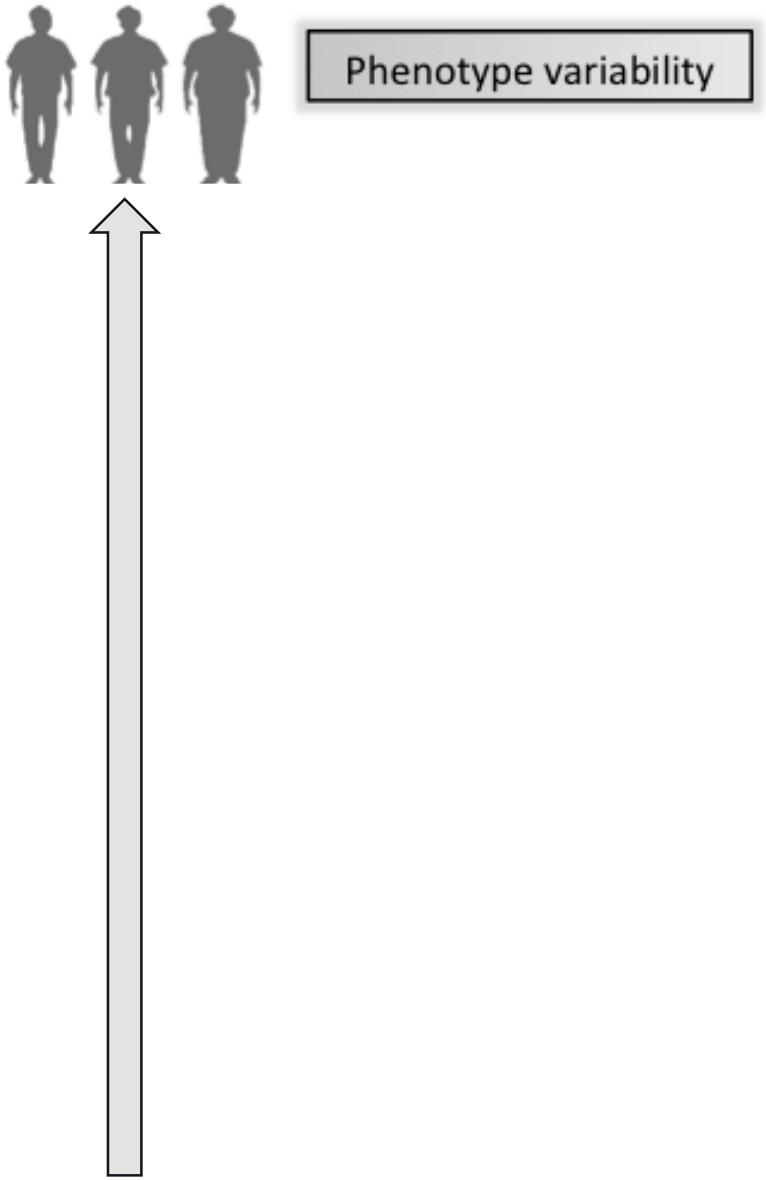
- To develop a centralized repository of other blood-based materials
    - DNA
    - RNA
    - Plasma

- To identify novel genetic and proteomic factors for dystonia risk

- to identify genetic and proteomic factors that influence spread of dystonia

DYSTONIA COALITION

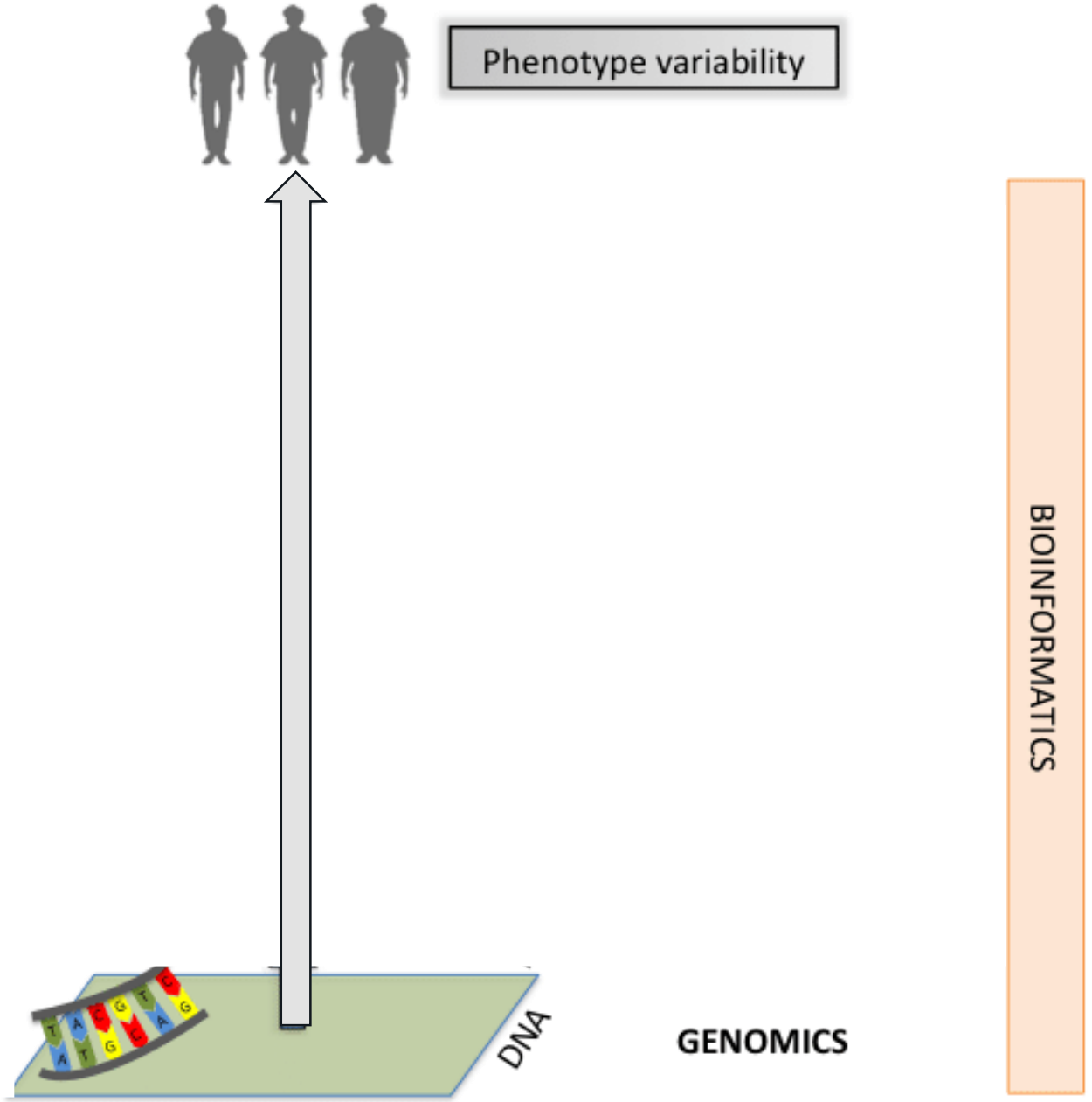
Goal: to Understand the Phenotypic Variability



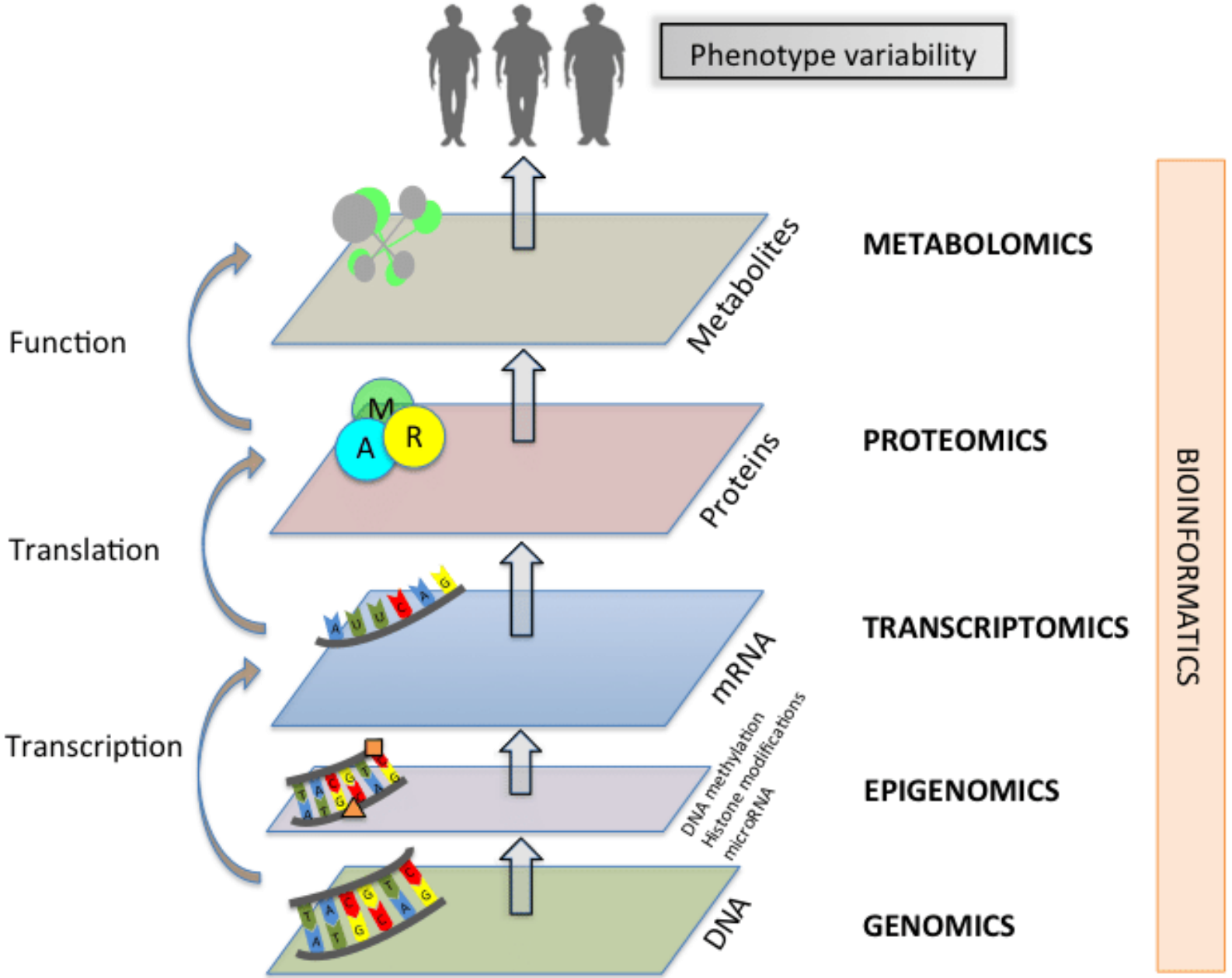
- Identify novel genes and pathways
- Identify novel molecular biomarkers
- Create better prediction models
- Identify novel targets and drugs

Goal: to Understand the Phenotypic Variability

Genetics is just the First Layer

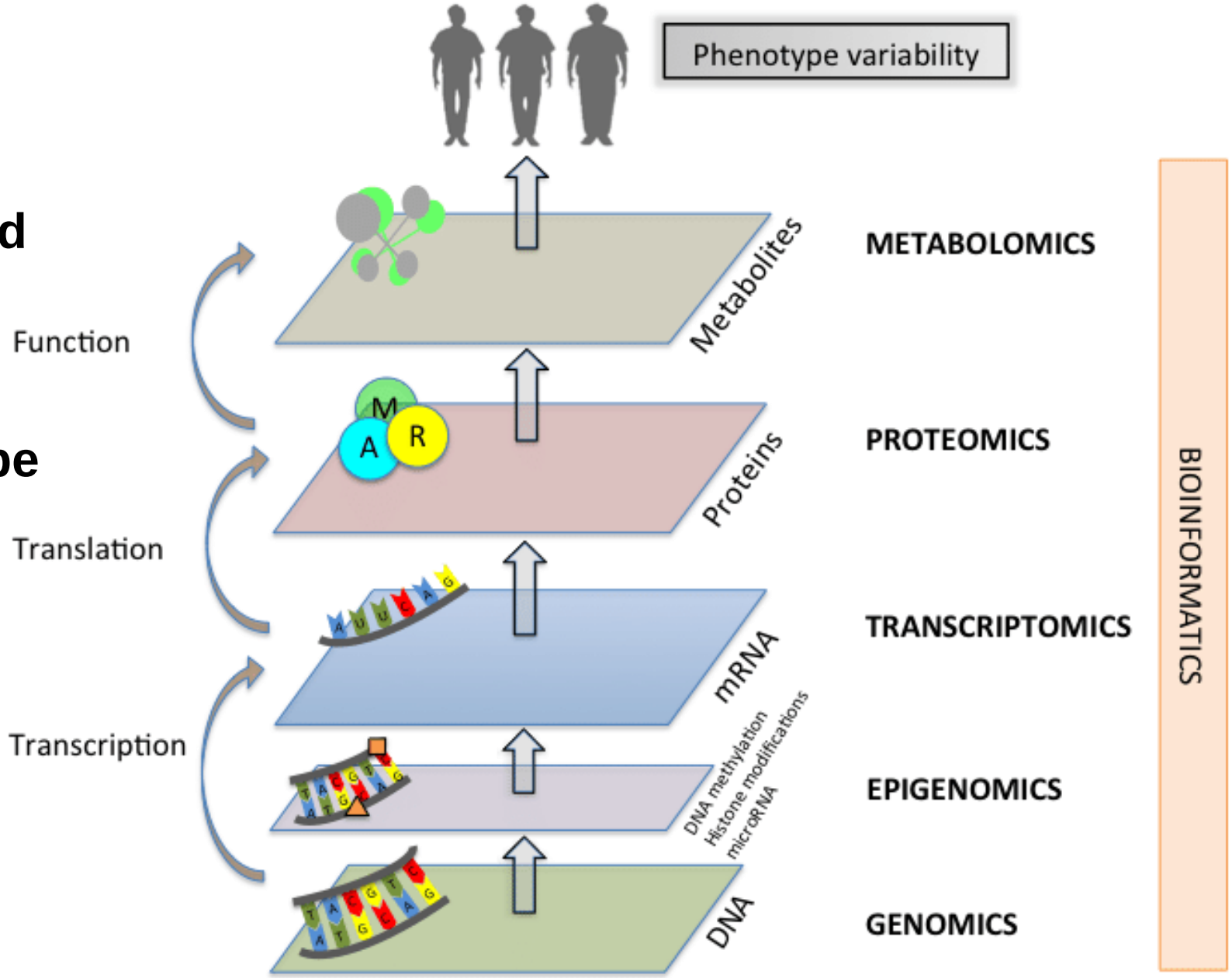


Goal: to Understand the Phenotypic Variability



Goal: to Understand the Phenotypic Variability

- **We need to go beyond just GWAS/WGS**
- **Molecularly Phenotype Clinical cohorts**
- **Generating multiple layers of omic data**

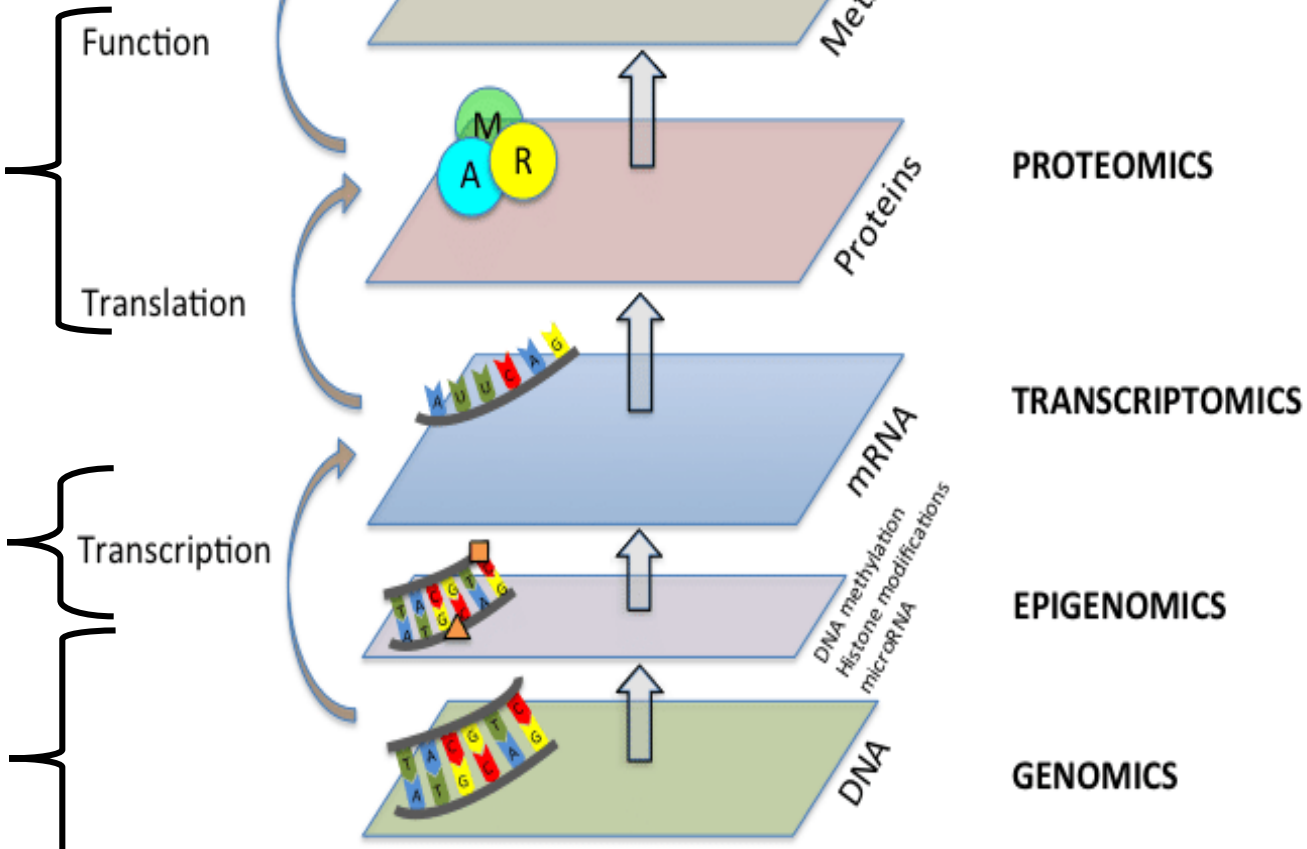


Genetics is just the first layer

Plasma

Blood-RNA

Blood-DNA

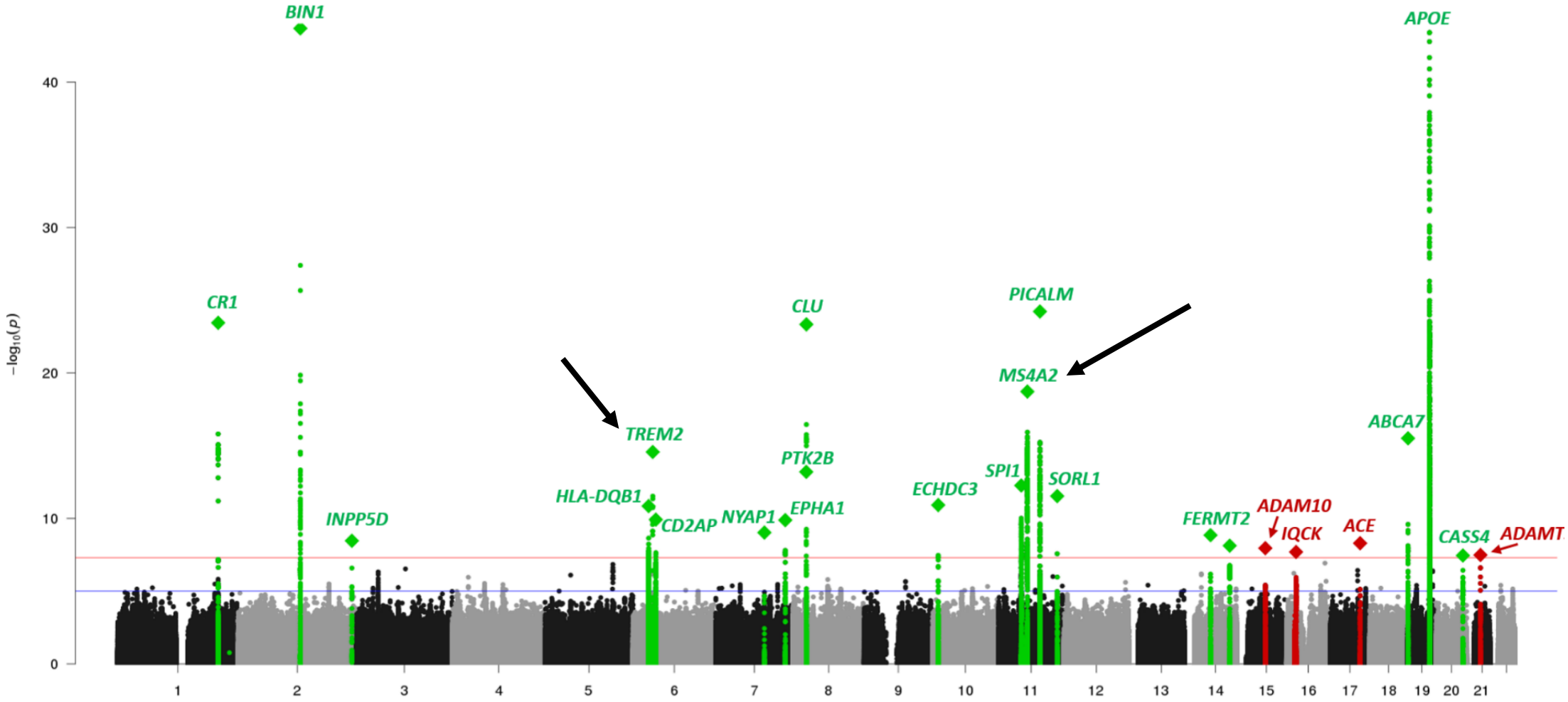


A Metabolomic Study of Cervical Dystonia

- Plasma samples from 100 cases with idiopathic cervical dystonia and 100 controls
- 7,346 metabolic features remained after quality control
- and 289 demonstrated significant differences between cases and controls
- 9 biological processes to be significantly associated at $p<0.05$, 5 pathways were related to carbohydrate metabolism, 3 pathways were related to lipid metabolism.

Liu et al, Unpublished

Meta-analysis of genetic association with diagnosed Alzheimer's disease identifies novel risk loci and implicates Abeta, Tau, immunity and lipid processing

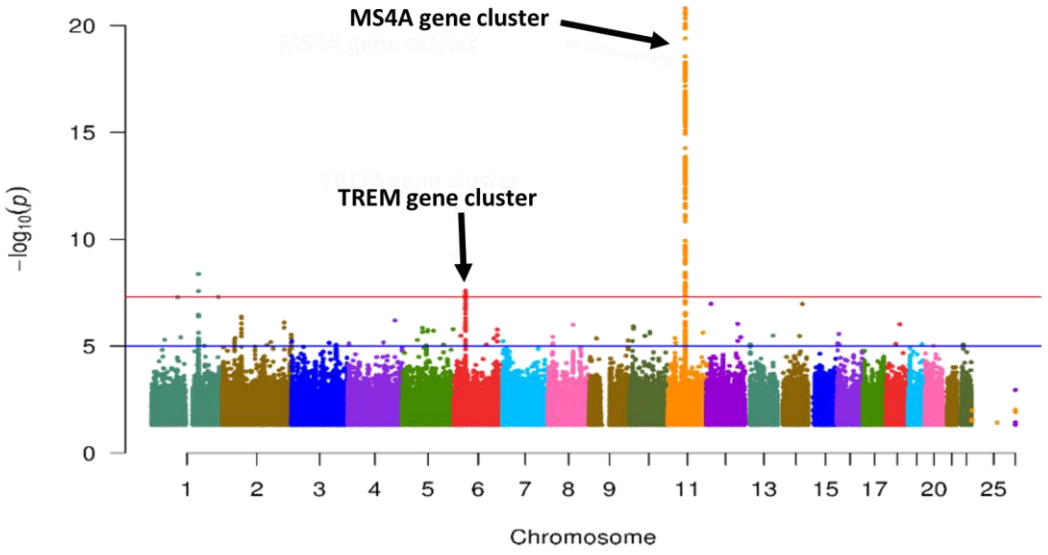


Kunkle et al, 2019, N=89,769

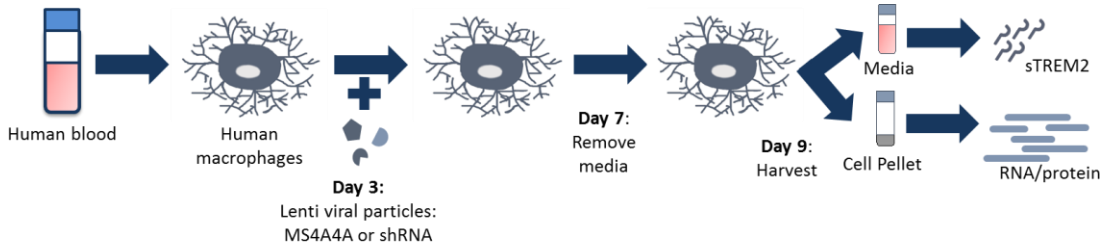
Cruchaga LAB

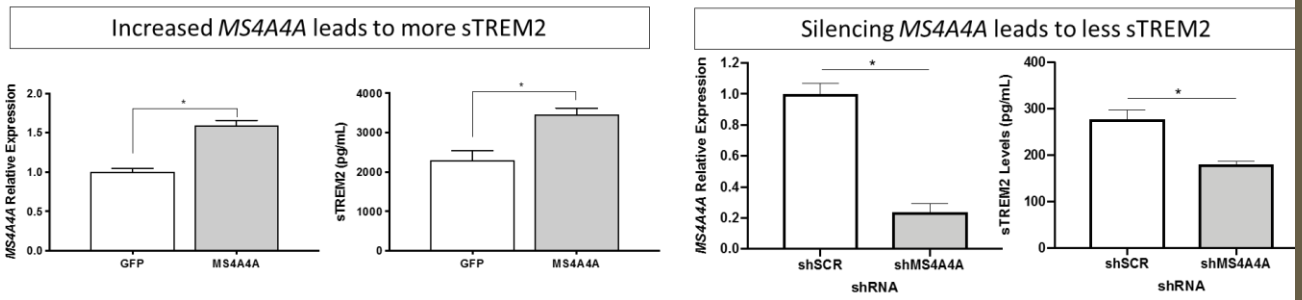
MS4A4A modulates AD risk by regulating *sTREM2* levels

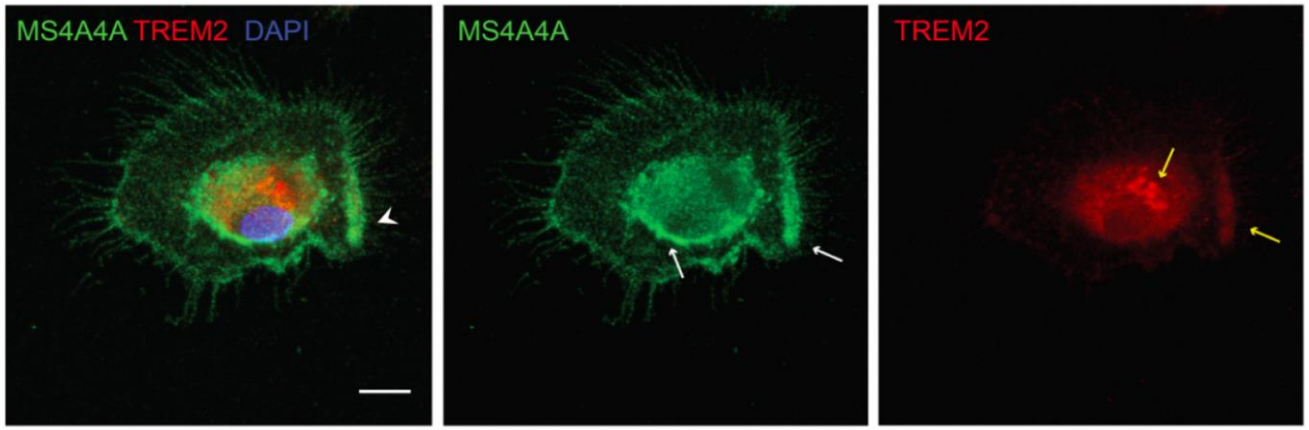
GWAS for CSF sTREM2



MS4A4A affects sTREM2







MS4A4A and TREM2 co-localized

Deming Y, Filipello F, Cignarella F et al., *Science Translational Medicine* In Press

Multi-omics: Biomarker Identification

- **Proteomic data in**
  - 1,300 CSF, 650 plasma, and 450 brain samples with Alzheimer Disease
  - To identify specific proteins that are differentially expressed in AD cases vs controls and create novel prediction models
  - Mendelian Randomization to identify novel biomarkers

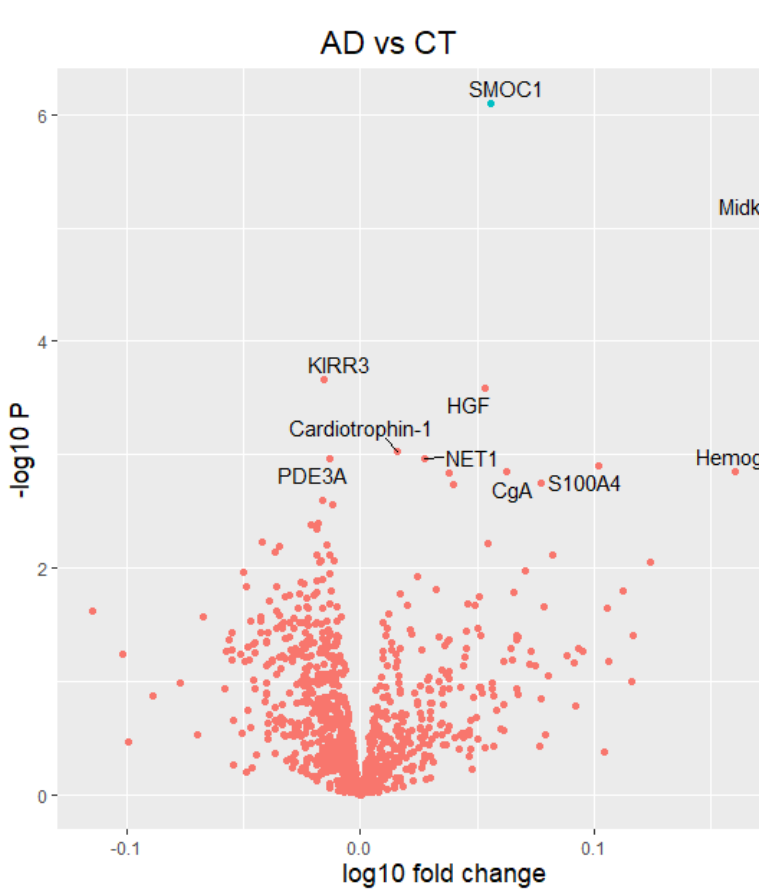
Beyond classic CSF biomarkers

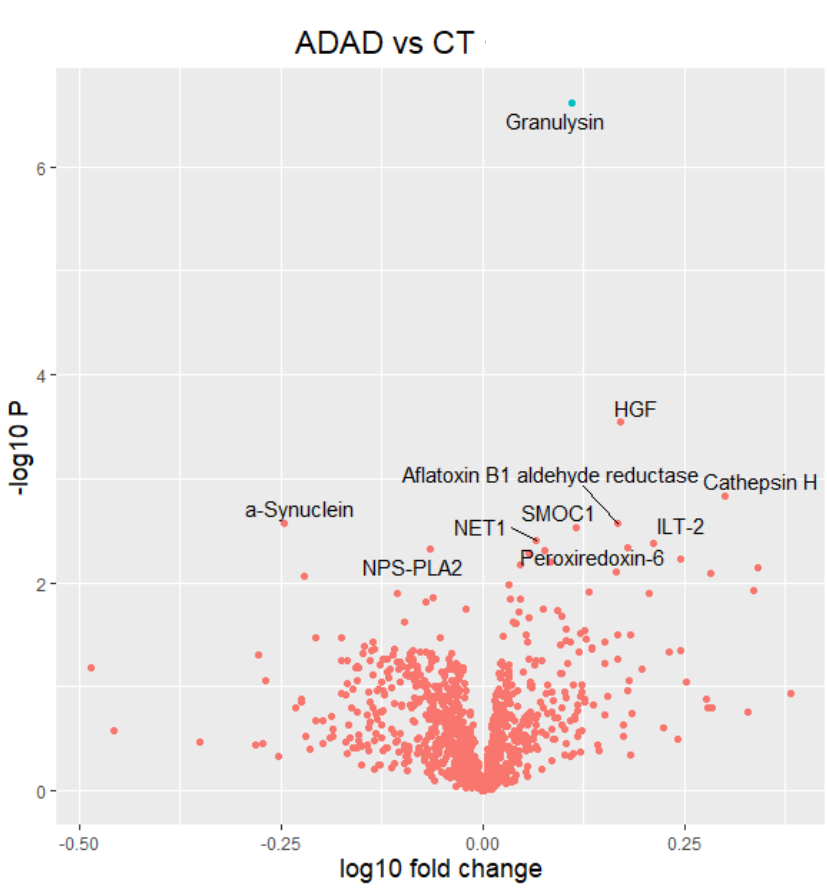
Overview of sample size for multi-tissue proteomics

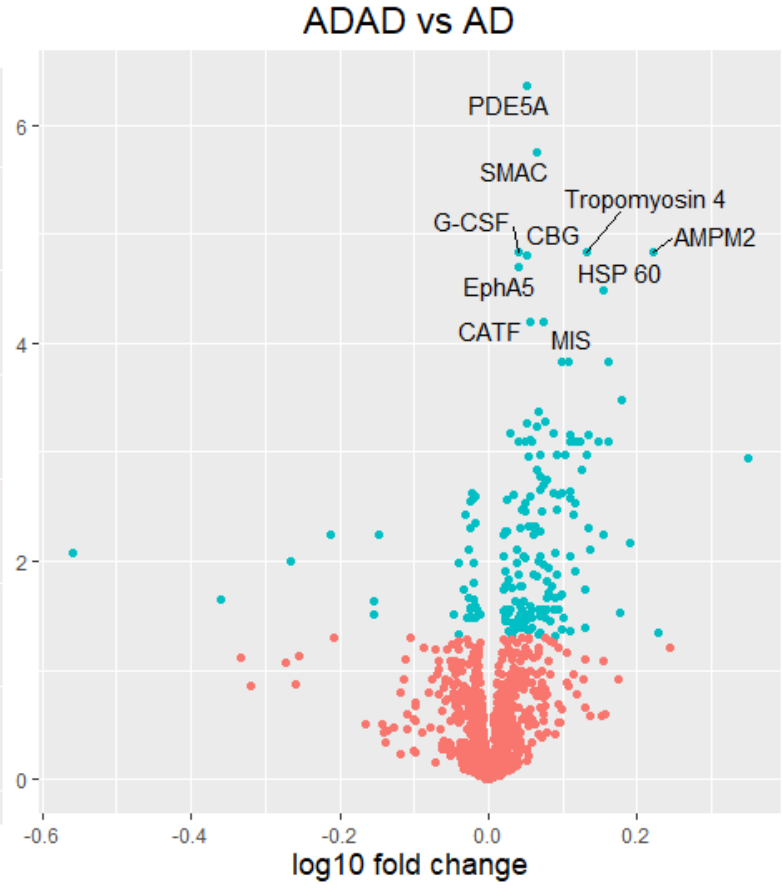
Tissue	Number of Samples	Number of Proteins
CSF	835	713
Plasma	529	931
Brain	380	1079

Differential protein abundance in brain tissue

Analysis	Proteins with FDR P < 0.05
AD vs Control	2
ADAD vs Control	1
ADAD vs AD	182

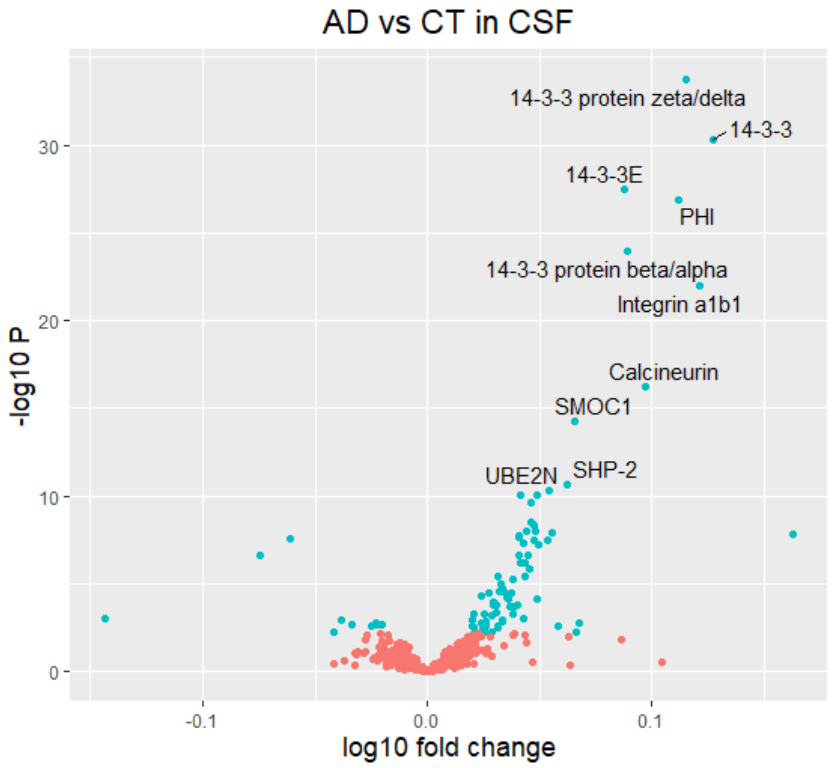


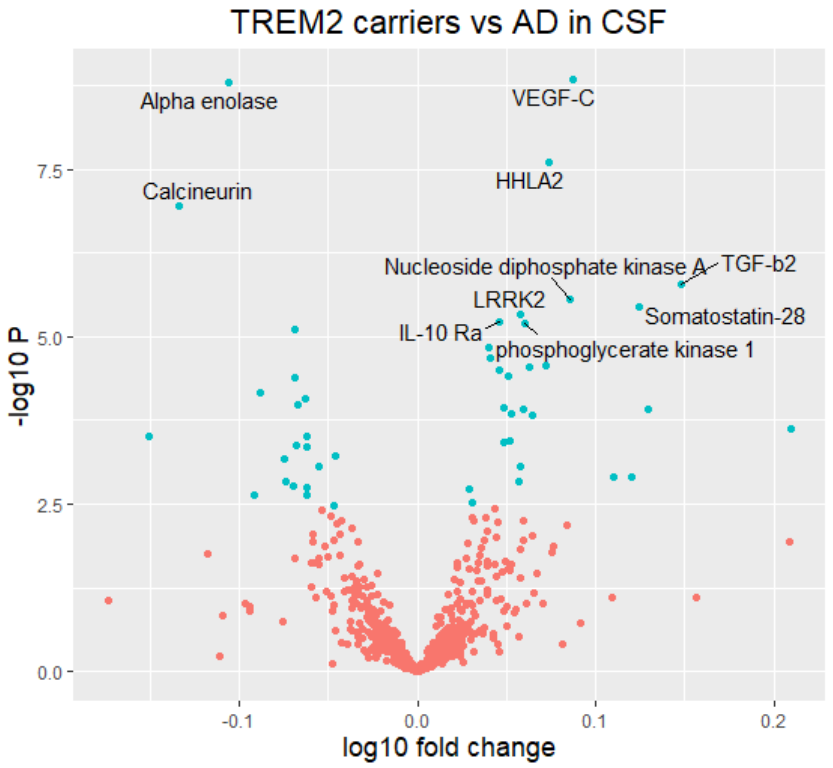


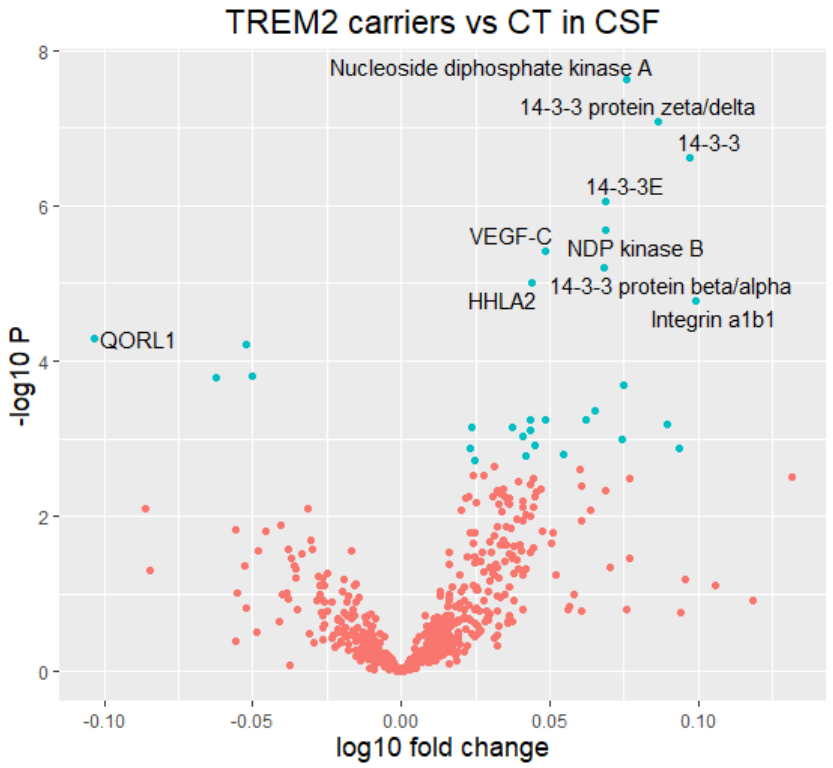


Differential Protein Abundance in CSF

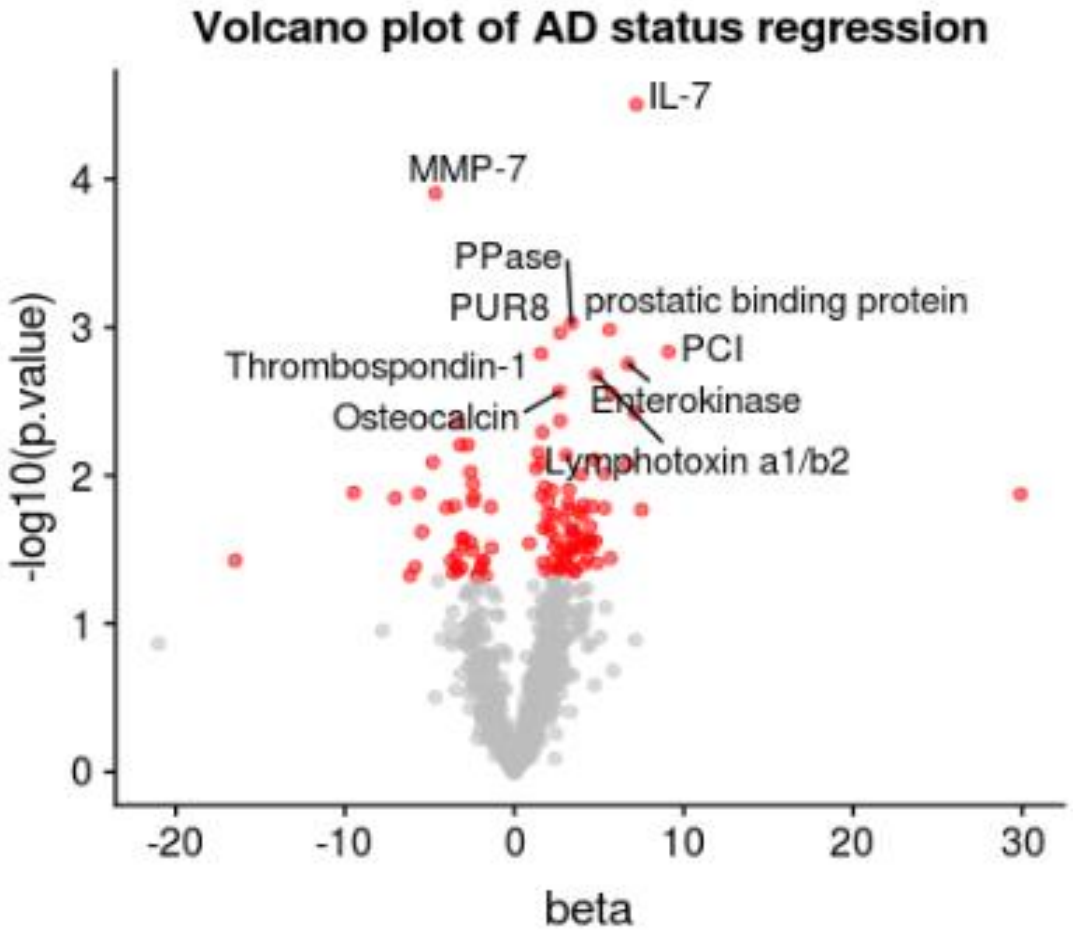
Analysis	Proteins with FDR P < 0.05
AD (N=206) vs Control (N=514)	85
TREM2 (N=47) vs AD (N=175)	48
TREM2 (N=47) vs Control (N=498)	30

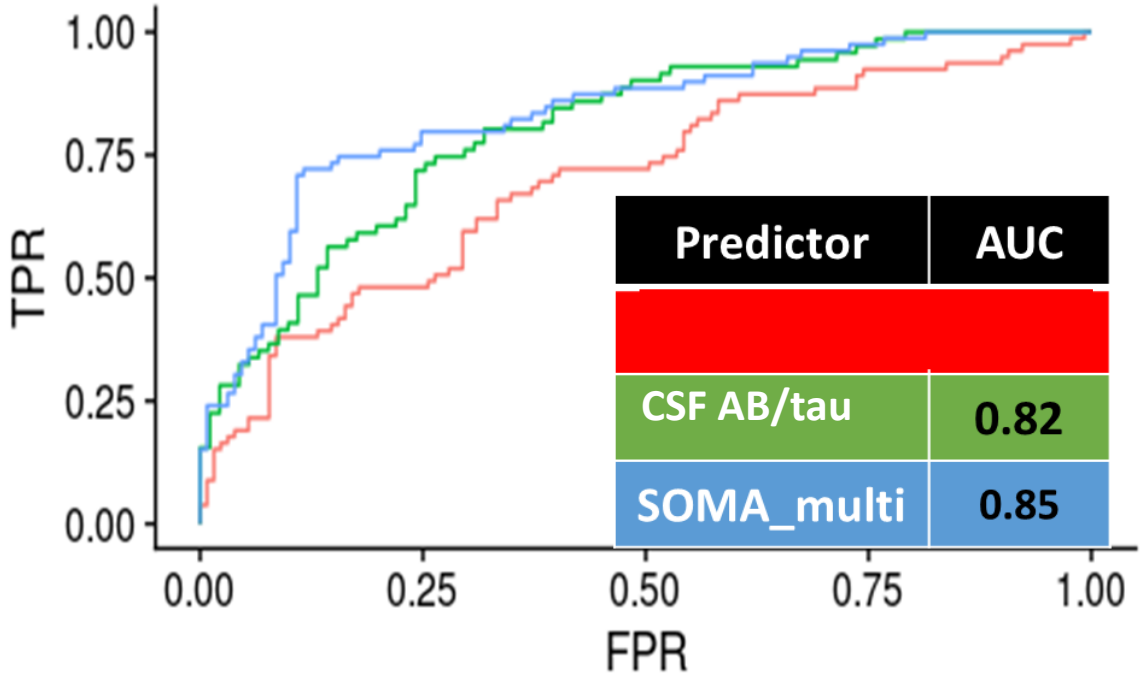






Plasma Biomarker discovery





Summary

- **The goals is to molecularly characterized the DC cohort**
  - Genetic (GWAS, WGS), epigenomics (longitudinal), transcriptomics (longitudinal), proteomic (longitudinal), metabolomic and lipidomics

- **Deep molecular phenotyping of well clinically characterized cohorts will lead to the identification of:**
  - Novel genes and pathways implicated on the diseases
  - A deeper understanding pathologic events
  - Novel molecular phenotypes
  - Novel therapeutic targets

- **The multi-omic data (genetic, epigenetic, transcriptomic, proteomic, metabolomic, between others) will allow to a more personalize prediction of disease risk and treatment**

Q & A

Biobank (BB) Project
Washington University in St. Louis
Project Manager: Carlos Cruchaga. PhD cruchagac@wustl.edu
Study Coordinator: Jen Gentsch j.gentsch@wustl.edu
Study Coordinator: JoAnne Norton nortonj@wustl.edu