Exploring the transcriptome for novel biomarker discovery

AMP-AD Biomarker Pilot Study

Nilüfer Ertekin-Taner, MD, PhD
Professor of Neurology and Neuroscience
Mayo Clinic Florida

Spring ADC Meeting
May 2nd, 2019

AMP-AD Team

Mayo Clinic Florida
Nilüfer Ertekin Taner, MD, PhD
Steven G. Younkin, MD, PhD
Dennis Dickson, MD
Minerva Carrasquillo, PhD
Mariet Allen, PhD
Joseph Reddy, PhD
Xue Wang, PhD
Kim Malphrus
Thuy Nguyen
Sarah Lincoln

University of Florida
Todd Golde, MD, PhD.
Jada Lewis, Ph.D.
Paramita Chakrabarty, PhD
Yona Levites, PhD

Institute for Systems Biology
Nathan Price, PhD
Cory Funk, PhD
Hongdong Li, PhD
Paul Shannon

Banner Sun Health
Tom Beach, MD

New York Genomics Center (WGS)

Mayo Clinic Genomics Core
Bruce Eckloff (RNASeq)
Julie Cunningham, PhD (GWAS)

Mayo Clinic Bioinformatics Core
Asha Nair, PhD

Mayo Clinic Epigenomics Development Laboratory
Tamas Ordog, MD
Jeong-Hong Lee, PhD

Mayo Clinic IT
Andy Cook
Curt Younkin

Mayo Clinic Study of Aging
Ronald Petersen, MD, PhD

AMP-AD consortium

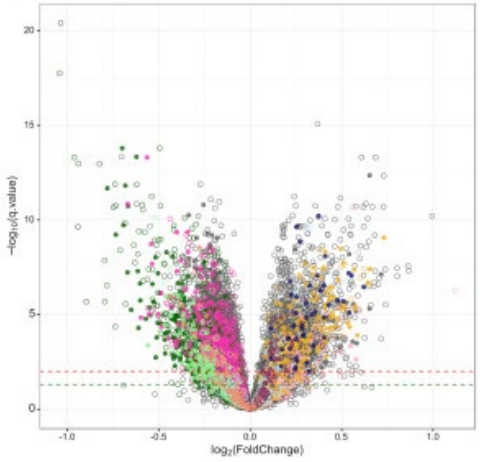
Overarching Goal

A System Approach to Targeting Innate Immunity in AD (U01 AG046193)

To identify and validate targets within innate immune signaling, and other pathways that can provide disease-modifying effects in AD using a multifaceted systems level approach.

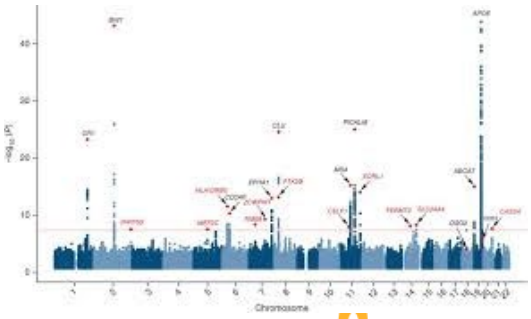
AMP-AD Phase I

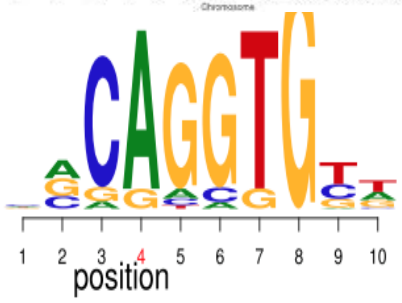
Original Aim 1: To detect transcript alterations in innate immunity genes in mice and humans.



- RNAseq human and mice brains.
- Differential expression.
- Protein/Nanostring validation
- Expression quantitative trait loci (eQTL).

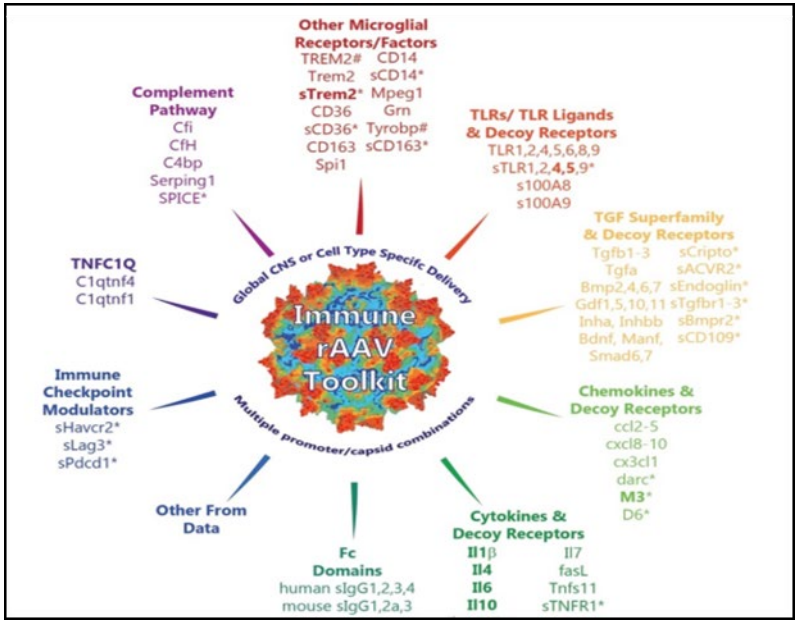
Original Aim 2: To assess AD risk conferred by variants in innate immunity genes from Aim 1.





- Test eQTL for effects on AD risk
- Functionally annotate AD risk variants for effects on gene expression.
- Expression/Transcription factor networks.

Original Aim 3: To manipulate innate immune states in vivo.



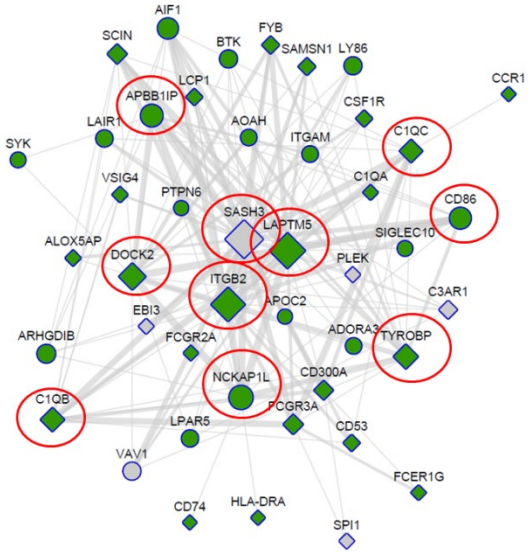
- rAAV based genetic manipulation in mice and cells.
- Evaluate Aß, tau, neurodegeneration in model systems.

Original Aim 4: To determine outcome of gene manipulation in wild type mice.

Behavioral studies in nontransgenic mice.

AMP-AD Phase-I: Outcomes

"Immune" co-expression networks enriched for AD risk genes.



AD candidate genes: *CD33, FERMT3*, HLA-DRB1/6, INPP5D, MS4A, PLD4*, RIN3, TYROBP, TREM2*

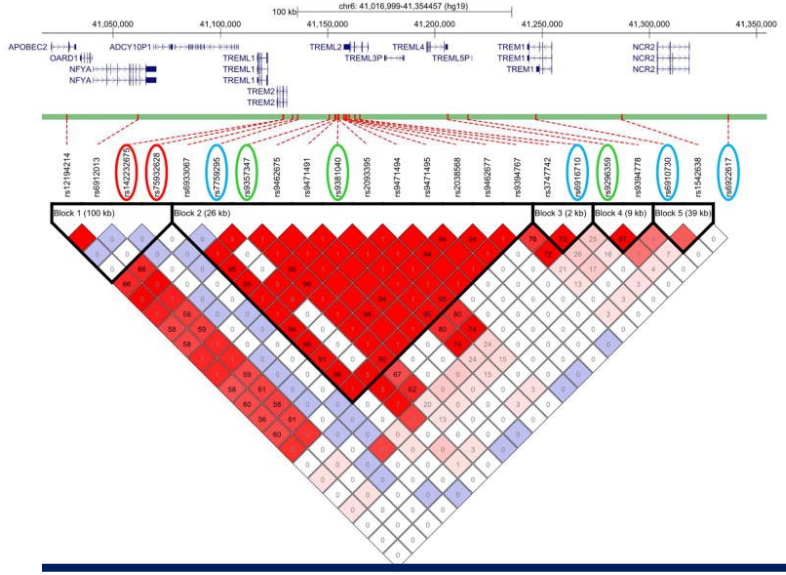
gene_symbol	CRND8: 3 month		CRND8: 6 month		CRND8: 12 month		APPPS1: 12 month	
	Direction	p-value	Direction	p-value	Direction	p-value	Direction	p-value
SASH3	Up-in-Tg	NS	Up-in-Tg	NS	Up-in-Tg	1.59E-02	Up-in-Tg	1.62E-02
LAPTM5	Up-in-Tg	NS	Up-in-Tg	1.34E-06	Up-in-Tg	6.58E-11	Up-in-Tg	7.63E-07
ITGB2	Up-in-Tg	NS	Up-in-Tg	NS	Up-in-Tg	NS	Up-in-Tg	NS
DOCK2	Up-in-Tg	NS	Up-in-Tg	NS	Up-in-Tg	NS	Up-in-Tg	NS
NCKAP1L	Up-in-Tg	NS	Up-in-Tg	1.01E-05	Up-in-Tg	1.28E-07	Up-in-Tg	2.68E-05
TYROBP	Up-in-Tg	**4.28E-04**	Up-in-Tg	**4.62E-06**	Up-in-Tg	**9.42E-15**	Up-in-Tg	**1.16E-09**
CD86	Up-in-Tg	NS	Up-in-Tg	5.29E-03	Up-in-Tg	7.32E-06	Up-in-Tg	2.84E-04
APBB1IP	Up-in-Tg	NS	Up-in-Tg	3.32E-04	Up-in-Tg	2.17E-09	Up-in-Tg	3.21E-04
C1QC	Up-in-Tg	NS	Up-in-Tg	5.02E-04	Up-in-Tg	1.73E-07	Up-in-Tg	6.39E-05
C1QB	Up-in-Tg	NS	Up-in-Tg	5.55E-03	Up-in-Tg	2.47E-07	Up-in-Tg	5.77E-04

- Modulation of innate immunity proteins influences Aß and tau pathophysiology (Chakrabarty et al., Neuron, 2015; Li et al., FASEB, 2015).
- AD risk genes *PLCG2, ABI3 and TREM2* have higher levels in AD brains, Aß models and reside in immune networks (Sims et al., Nature Genetics, 2017; Conway et al., Molec Neurodeg, 2018).
- 20 "immune" targets undergoing testing in model systems

AMP-AD Phase I: Outcomes

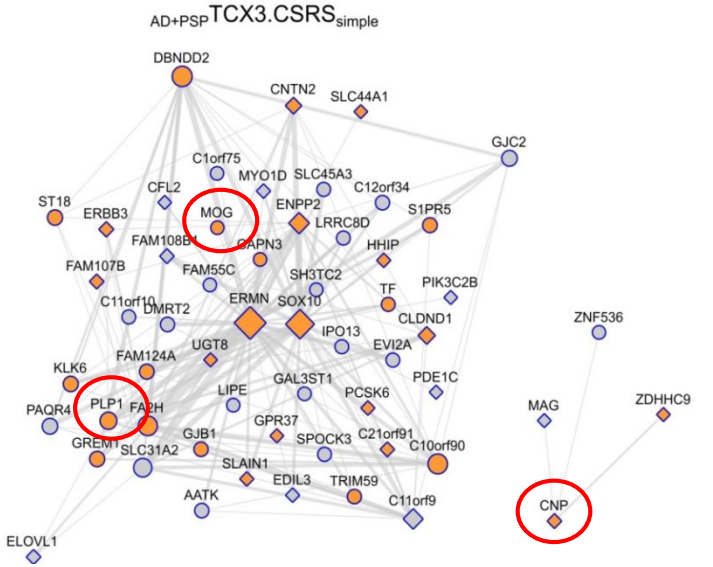
Regulatory variant at *TREM* locus.

An intronic variant at the *TREM* locus is associated with higher brain *TREM2* and *TREML1* levels and resides in a TF binding site (*Carrasquillo et al., Alzheimer's and Dementia, 2016*). **- Omics integration, directionality, mechanism.**



Myelin networks in AD.

Conserved brain myelination networks are altered in AD and PSP (*Allen et al., Alzheimer's and Dementia, 2017*). **- Comparative transcriptomics, novel targets.**



- Disseminated rAAV vectors to the AMP-AD research community
- Many AD candidate risk genes have strong eQTL and/or differential expression in brain (*Allen et al, Neurology Genetics 2015, 2017; Ridge et al., Genome Medicine, 2017; Mukherjee et al., Alzheimer's and Dementia, 2017*).

AMP-AD Biomarker Pilot

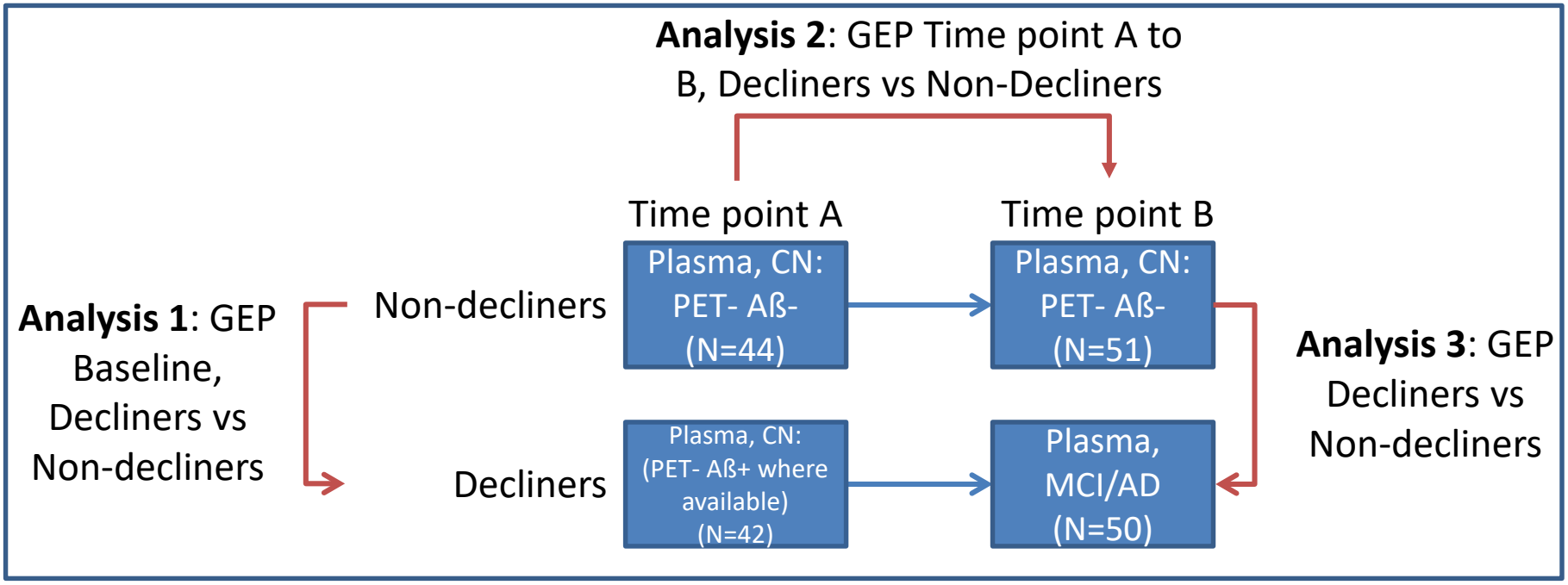
Hypothesis: **Brain transcriptional changes that occur in AD may be reflected in the periphery (plasma or blood).**

- Genetic/epigenetic risk factors ***upstream*** of disease neuropathology may lead to perturbations in gene expression levels/splicing. *E.g. Transcriptional changes as a consequence of regulatory disease variants.*
- Gene expression changes ***downstream*** of disease neuropathology may be detectable in both brain and periphery. *E.g. Transcriptional changes reflecting neuronal loss, microglial activation etc.*

Goal: **Establish and validate a scalable and reliable approach to identify AD biomarkers leveraging brain and plasma transcriptome from well-characterized cohorts.**

Approach

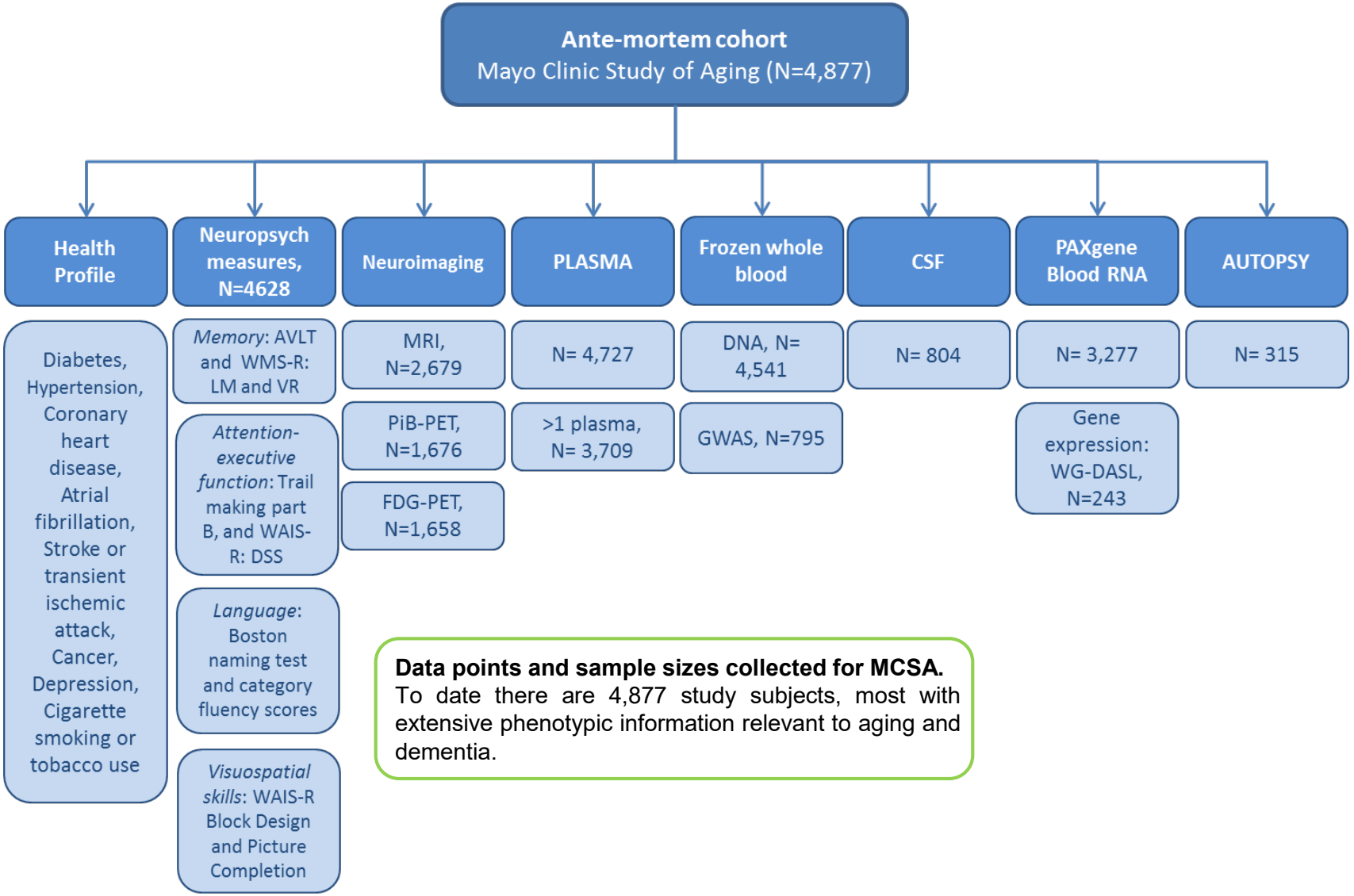
Gene Expression Profiling (GEP) utilizing RNAseq on plasma samples



Time point A (Baseline): All participants are clinically normal (CN).
Time point B: Decliners-Incident MCI/AD vs. Non-decliners-CN

- Analysis 1: **Preclinical biomarker,** differences between decliners and non-decliners prior to clinical conversion to MCI/AD.
- Analysis 2: **Biomarkers of rate of decline,** plasma expression changes over time in decliners vs. non-decliners.
- Analysis 3: **Biomarkers of clinical impairment,** differences between decliners and non-decliners, for comparison with brain GEP results from AMP-AD.

Cohort



Data points and sample sizes collected for MCSA. To date there are 4,877 study subjects, most with extensive phenotypic information relevant to aging and dementia.

Cohort Demographics

Demographics		
Plasma RNAseq Cohort (n = 104)	**Decliners (n = 51)**	**Non-Decliners (n = 53)**
% APOE e4+	39.2%	35.8%
% Female	40.0%	43.4%
Mean Age at Time Point A (SD)	79.6 (6.7)	79.1 (6.3)
Mean Age at Time Point B (SD)	82.1 (6.8)	81.6 (6.5)
Mean Age difference Time Point B-A (SD)	2.5 (0.5)	2.5 (0.6)
Date Range Time Point A	3/2006-6/2015	1/2005-4/2015
Date Range Time Point B	10/2008-12/2017	8/2008-11/2017

- All 104 participants are part of the Mayo Clinic Study of Aging.
- Decliners and non-decliners are matched for ages, sex, dates of collection.
- Plasma for Time Point A was available for 42 decliners and 44 non-decliners.
- Plasma for Time Point B was available for 50 decliners and 51 non-decliners.
- RNA was extracted from a total of 187 unique plasma samples (961 aliquots).

Plasma cell-free mRNA and miRNA expression

MCSA Frozen plasma

Draw years: 2005-2017

N=104 participants,
187 plasma samples,
961 aliquots

Total RNA extraction
(Qiagen miRNeasy
Serum/Plasma Kit)

+ DNase I treatment

+Spike-in control: miR-39

NEXTflex Small RNA-Seq Kit
v3
(Bio Scientific) for **miRNAseq**

SMARTer Pico RNA-Seq
(Clontech) for **mRNAseq**

Human Neuropathology
custom panel (NanoString) for
targeted analysis (760 genes)

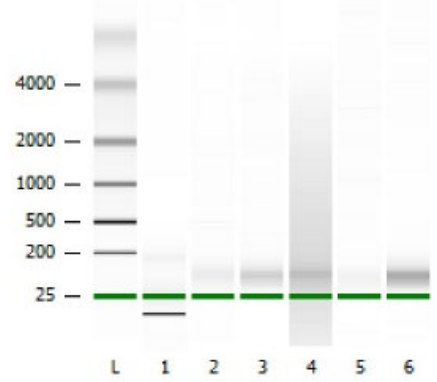
Neuropsychology measures : memory, attention executive function, language, visuospatial skills
Neuroimaging : amyloid PiB-PET, FDG-PET
Clinical history: smoking, obesity, hypertension, type-2 diabetes
Clinical diagnosis : AD, amnestic MCI, normal

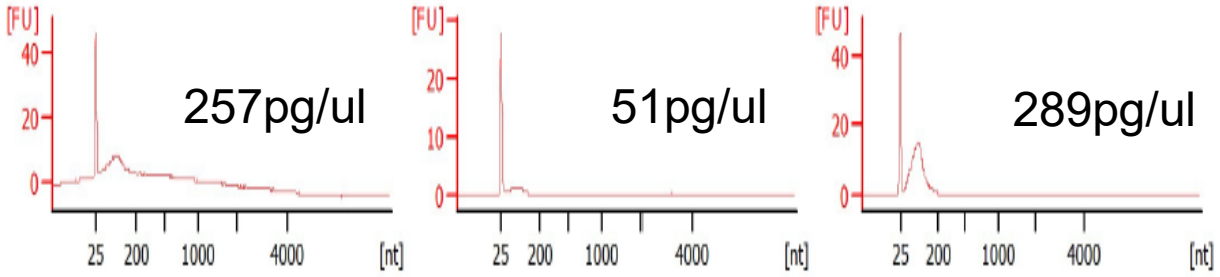
RNA extraction from plasma: Sample QC

✓ Nanodrop Spectroscopic absorbance at 415nm for detection of low level hemolysis (A_{415} readings >0.2classified as being hemolyzed): **961 plasma aliquots extracted/QC'ed.**

A_{415} 0.2 → 0.01

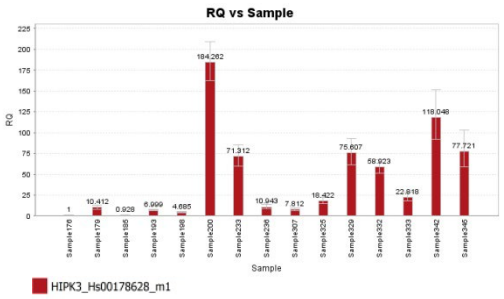
✓ Agilent Bioanalyzer Total RNA 6000 Pico used to determine cfRNA concentration. (Absence of ribosomal bands indicates no cellular RNA contamination).

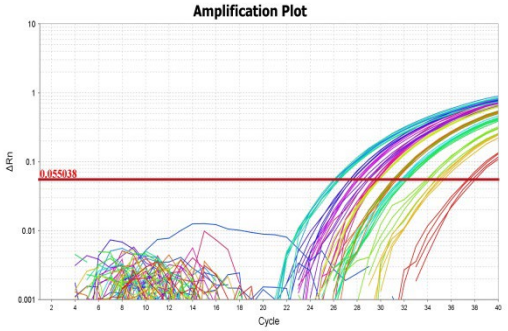




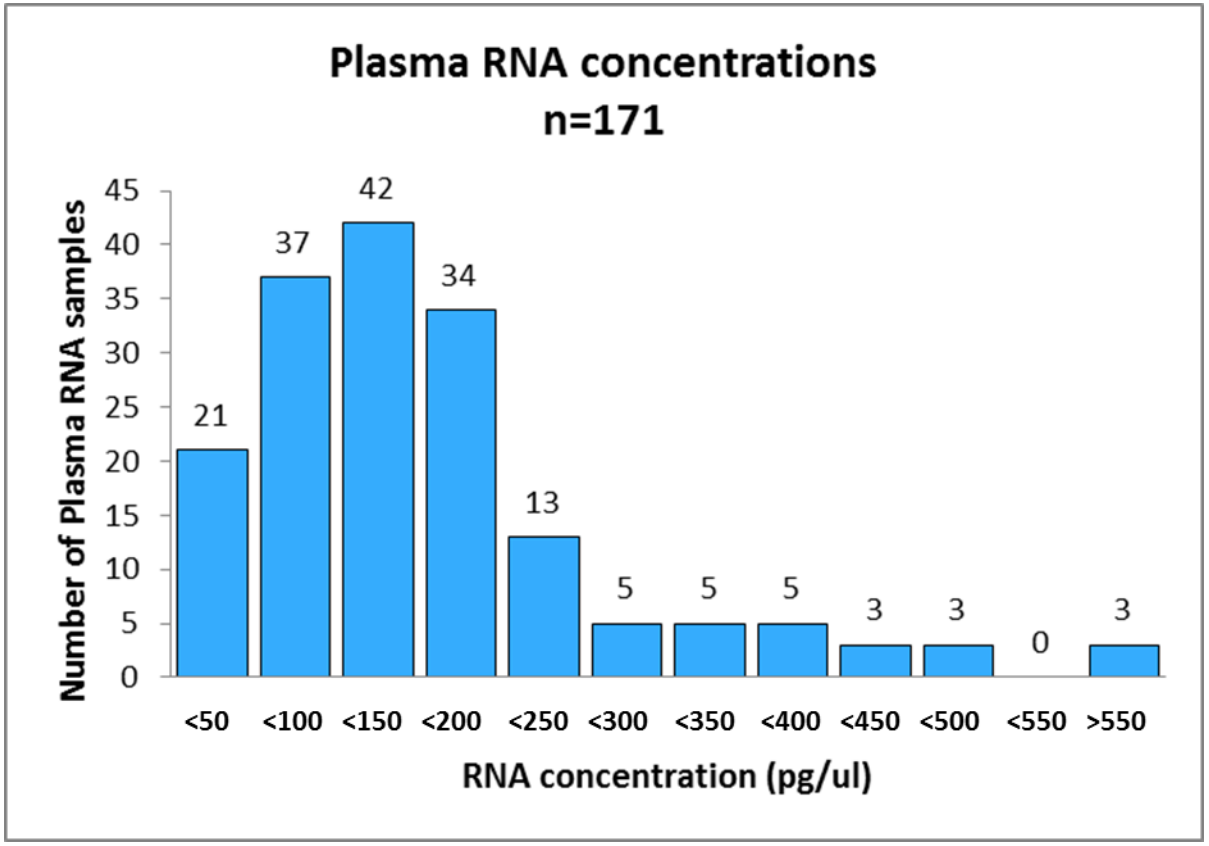
Concentration (pg/ul): mean = 159; min = 14, max = 857

✓ Taqman qPCR for detection of expressed genes HIPK3,FLI1 and Cel _miR-39 Spike in control(Qiagen)



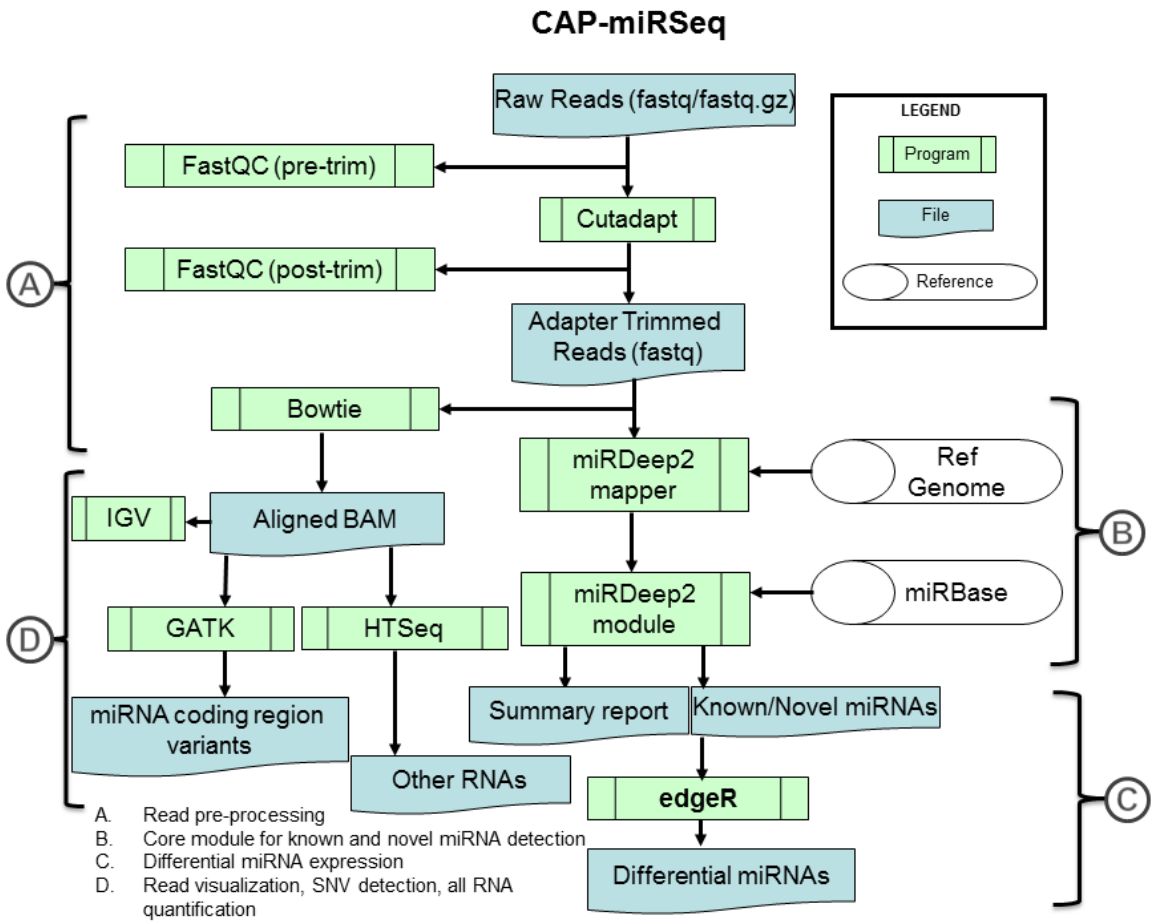


RNA extraction from plasma: Concentrations



- For the 187 unique plasma samples that will undergo RNAseq, RNA was extracted from a total of 961 plasma aliquots in order to obtain sufficient RNA.
- Median RNA concentration = 134 pg/ul, mean±standard deviation = 159±121.9, range 14-857 pg/ul.

Plasma miRNAseq pilot: Bioinformatics pipeline



CAP-miRSeq: a comprehensive analysis pipeline for microRNA sequencing data

Zhifu Sun, Jared Evans, Aditya Bhagwate, Sumit Middha, Matthew Bockol, Huihuang Yan, and Jean-Pierre Kocher

BMC Genomics. 2014; 15(1): 423. PMCID: PMC4070549

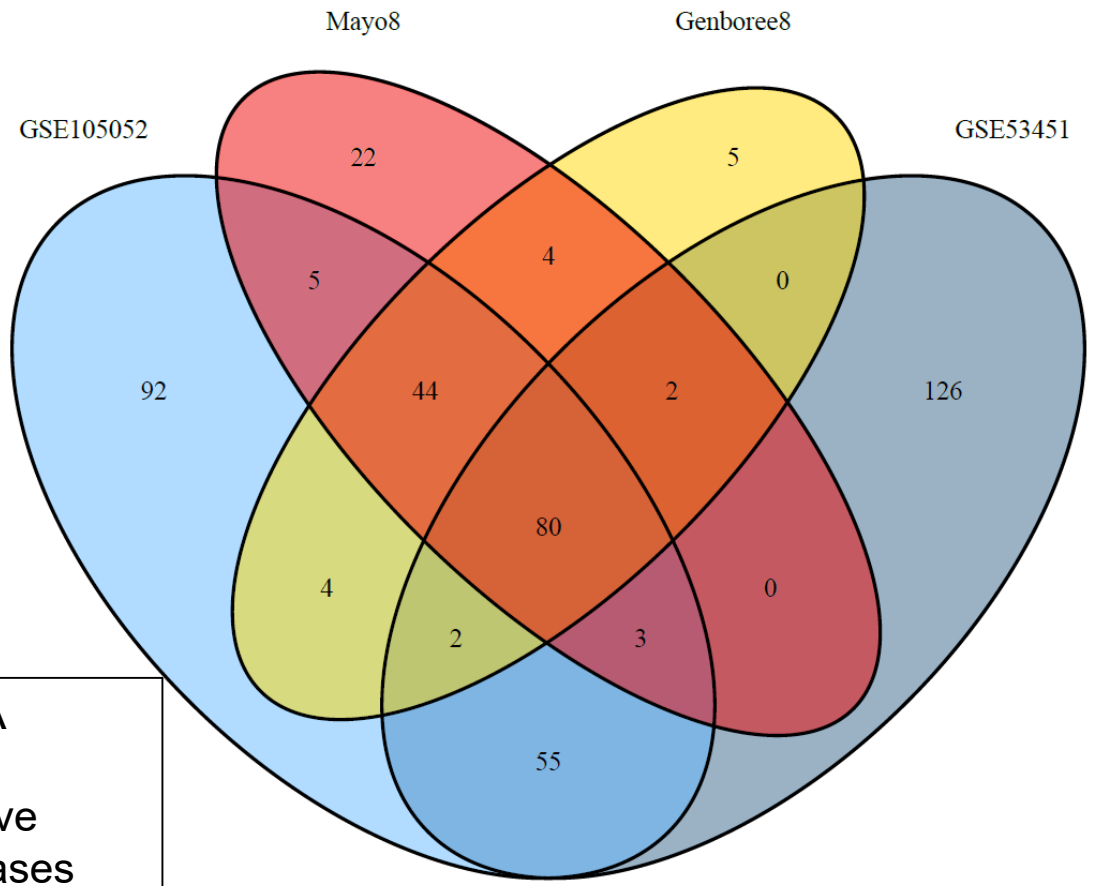
Plasma miRNAseq pilot preliminary results

Dataset	miRNA
GSE105052	285
GSE53451	268
Mayo8	160
Genboree8	141

GSE105052: Plasma miRNA profiles of 25 Friedreich's ataxia patients and 17heathy subjects, by deep sequencing using Illumina HiScan SQ.

GSE53451: Plasma exosomes miRNA sequenced from 7 humans

miRNAs with median raw read count >= 1



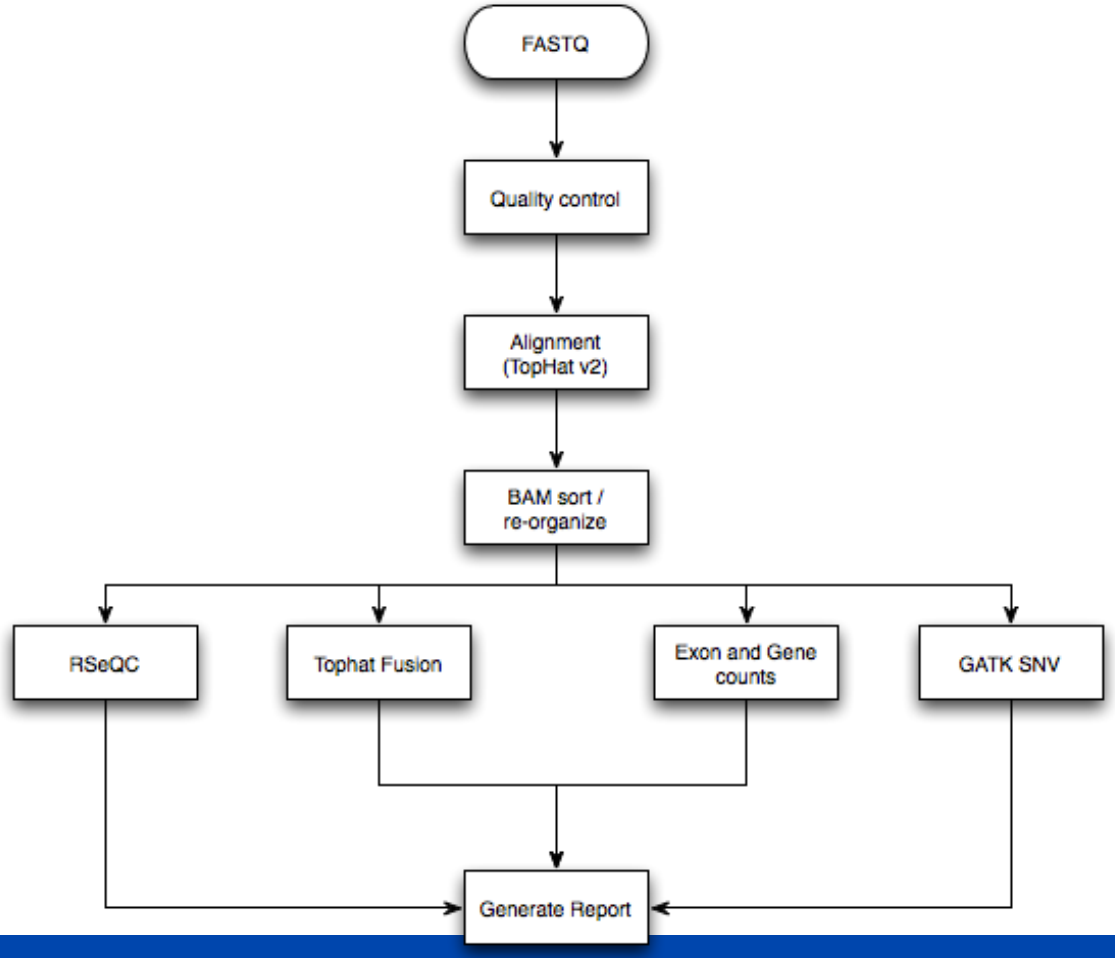
- CAP-miR >Genboree for miRNA detection
- Mayo Clinic plasma samples have ~50% overlap with public databases

Plasma mRNAseq pilot: Bioinformatics pipeline

MAP-RSeq: Mayo Analysis Pipeline for RNA sequencing

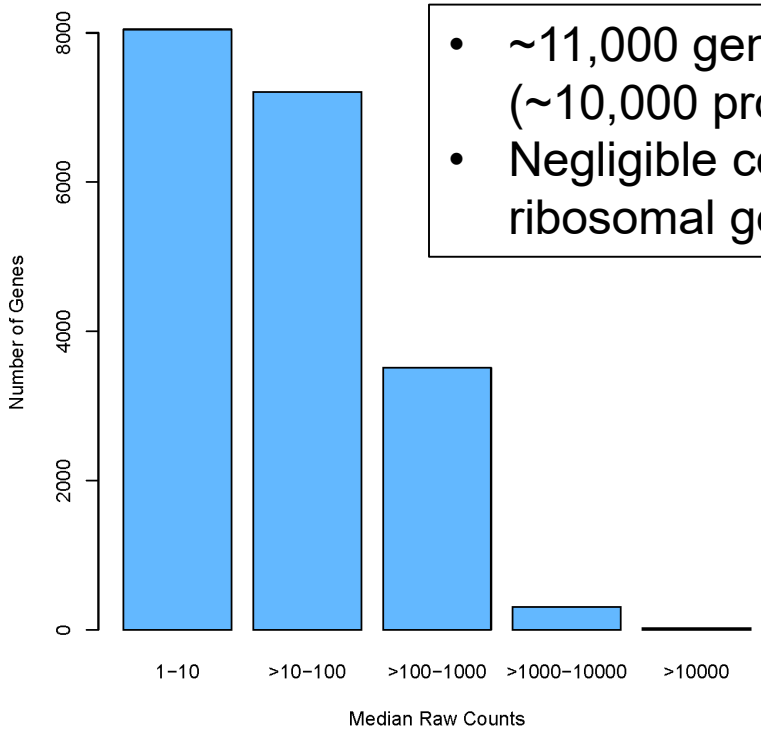
Kalari *et al. BMC Bioinformatics* 2014, **15**:224

BMC Bioinformatics



Plasma mRNAseq pilot: Preliminary results

Sample	Total counts	Counts in Genes	Intergenic	Intronic	Globins	Ribo	UTR
s_65289_RNA_Plasma	10894248	2557724 (23.5)	15%	14%	27 (0%)	136 (0%)	25%
s_66883_RNA_Plasma	8283524	1598128 (19.3)	20%	22%	63 (0%)	82 (0%)	22%
s_66959_RNA_Plasma	16646080	5135575 (30.9)	7%	7%	66 (0%)	104 (0%)	30%



- ~11,000 genes with >10 raw reads (~10,000 protein coding).
- Negligible counts for globin and ribosomal genes.

Plasma mRNA Expression: Targeted Panel

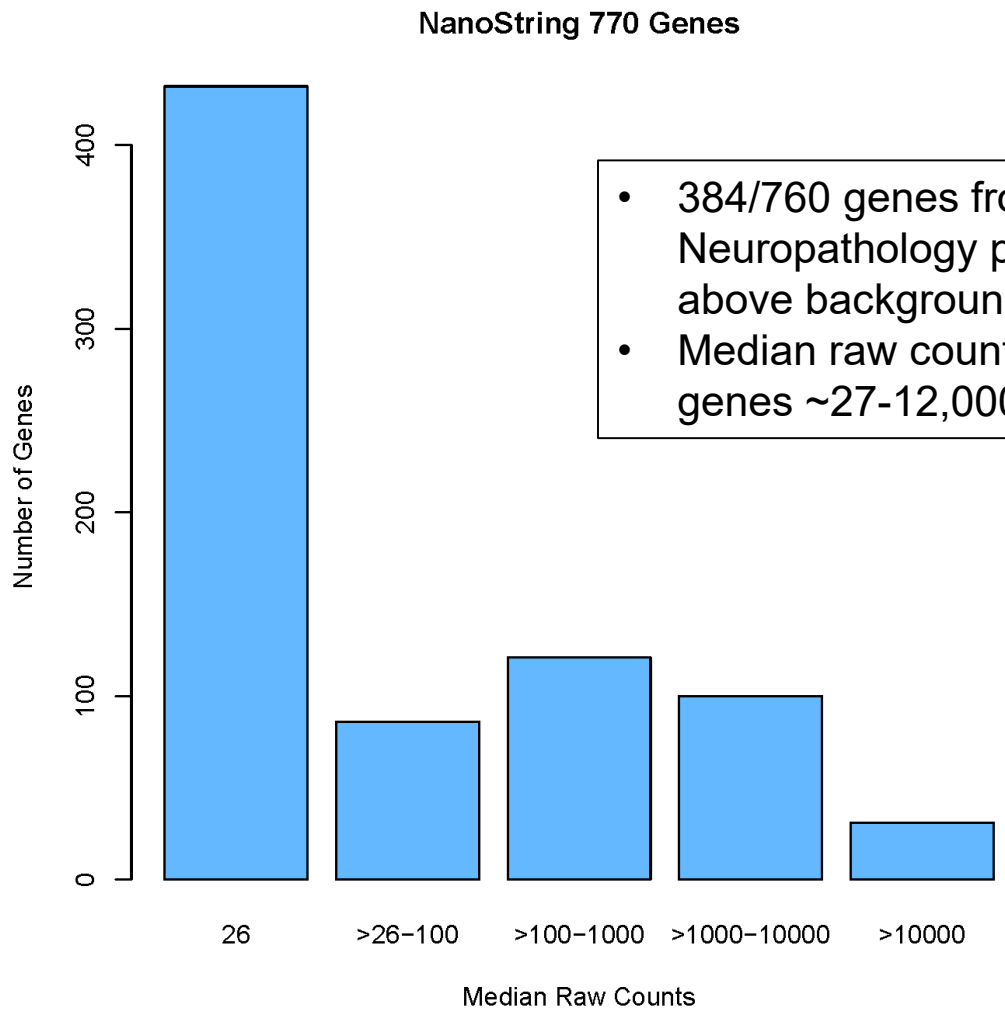
nanoString TECHNOLOGIES

nCounter® Human Neuropathology Panel - Gene and Probe Details

Official Symbol	Accession	Alias / Previous Symbol	Official Full Name
APOE	NM_000041.2	AD2	apolipoprotein E
APP	NM_000484.3	AD1	amyloid beta precursor protein
BDNF	NM_170732.4		brain-derived neurotrophic factor
CD14	NM_000591.2		CD14 molecule
CD33	NM_001177608.1	SIGLEC3, SIGLEC-3, p67, FLJ00391	CD33 molecule
CLU	NM_203339.2	CLI, APOJ, SGP-2, SP-40, TRPM-2, KUB1, CLU1, CLU2	clusterin
CR1	NM_000573.3	CD35, KN	complement component 3b/4b receptor 1 (Knops blood group)
CRH	NM_000756.1	CRF, CRH1	corticotropin releasing hormone
CSF1R	NM_005211.2	FMS, C-FMS, CSFR, CD115	colony stimulating factor 1 receptor
CX3CR1	NM_001337.3	GPR13, CMKBRL1, CMKDR1, V28, CCRL1	chemokine (C-X3-C motif) receptor 1
CXCL16	NM_001100812.1	SR-PSOX, CXCLG16, SRPSOX	chemokine (C-X-C motif) ligand 16
HDAC1	NM_004964.2	RPD3L1, HD1, GON-10	histone deacetylase 1
HDAC2	NM_001527.1	RPD3, YAF1	histone deacetylase 2
HDAC6	NM_001321225.1	KIAA0901, JM21, HD6, FLJ16239, PPP1R90	histone deacetylase 6
HDAC7	NM_001098416.2	HDAC7A, DKFZP586J0917	histone deacetylase 7
IL10	NM_000572.2	CSIF, TGIF, IL10A, IL-10	interleukin 10
IL1B	NM_000576.2	IL1F2, IL-1B, IL1-BETA	interleukin 1 beta
IL1R1	NM_001320984.1	IL1R, IL1RA, D2S1473, CD121A	interleukin 1 receptor, type I
IL6	NM_000600.3	IFNB2, IL-6, BSF2, HGF, HSF	interleukin 6
NINJ2	NM_016533.4		ninjurin 2
NTRK1	NM_001012331.1	TRK, TRKA, MTC	neurotrophic tyrosine kinase, receptor, type 1
SORL1	NM_003105.5	C11orf32, gp250, LR11, LRP9, SorLA, SorLA-1	sortilin-related receptor, L(DLR class) A repeats containing
SP1	NM_003109.1		Sp1 transcription factor
SPI1	NM_003120.1	PU.1, SPI-A, OF, SFPI1, SPI-1	Spi-1 proto-oncogene
STAT1	NM_139266.1	STAT91, ISGF-3	signal transducer and activator of transcription 1
TGFB1	NM_000660.4	TGFB, DPD1, CED, TGFbeta	transforming growth factor beta 1
TLR4	NM_138554.2	hToll, CD284, TLR-4, ARMD10	toll like receptor 4
TNF	NM_000594.2	TNFA, TNFSF2, DIF, TNF-alpha	tumor necrosis factor
TNFRSF1A	NM_001065.2	TNFR1, TNF-R, TNFAR, TNFR60, TNF-R-I, CD120a, TNF-R55	tumor necrosis factor receptor superfamily member 1A
TREM2	NM_018965.3	TREM-2, Trem2a, Trem2b, Trem2c	triggering receptor expressed on myeloid cells 2

Table 1. Subset of the 760 gene transcripts that can be tested with the Nanostring neuropathology panel.

Plasma mRNA Targeted Panel: Results



- 384/760 genes from the NanoString Neuropathology panel were detected above background (>26 counts).
- Median raw counts for detectable genes ~27-12,000.

Plasma mRNA Expression: Mayo Clinic Custom Panel

- NanoString nCounter™ **Mayo Clinic Custom CodeSet Design**:
- 50 genes enriched for those implicated in inflammation and/or vascular disease.
- 49 transcripts already detected in plasma by either RNAseq or the NanoString Human Neuropathology panel pilots.
- Florida Consortium for African-American AD Studies (FCA^{3}DS)* case-control cohort will be evaluated with this panel for 442 plasma RNA samples.
  - Mayo Clinic Florida (N. Ertekin-Taner/N. Graff-Radford).
  - Miami Mount Sinai (M. Greig-Custo/R. Duara)
  - University of Florida (M. Wicklund)
- Differential gene expression and eQTL analyses will be performed.

A2M	CCL2	HDAC1	IL4	PICALM
ABCA7	CD14	HDAC2	ITGAL	SORL1
ACTB	CD33	HDAC6	ITGB2	SP1
AIF1	CLU	HDAC7	LRP1	SPI1
AKAP9	CR1	ICAM1	LY86	STAB1
AOAH	CRH	IL16	LY96	STAT1
APOE	CSF1R	IL18	LYZ	TGFB1
APP	CX3CR1	IL1B	MS4A6A	TLR4
BIN1	CXCL16	IL1RN	NINJ2	TNFRSF1A
BLNK	HCK	IL33	OAZ1	TNFRSF1B

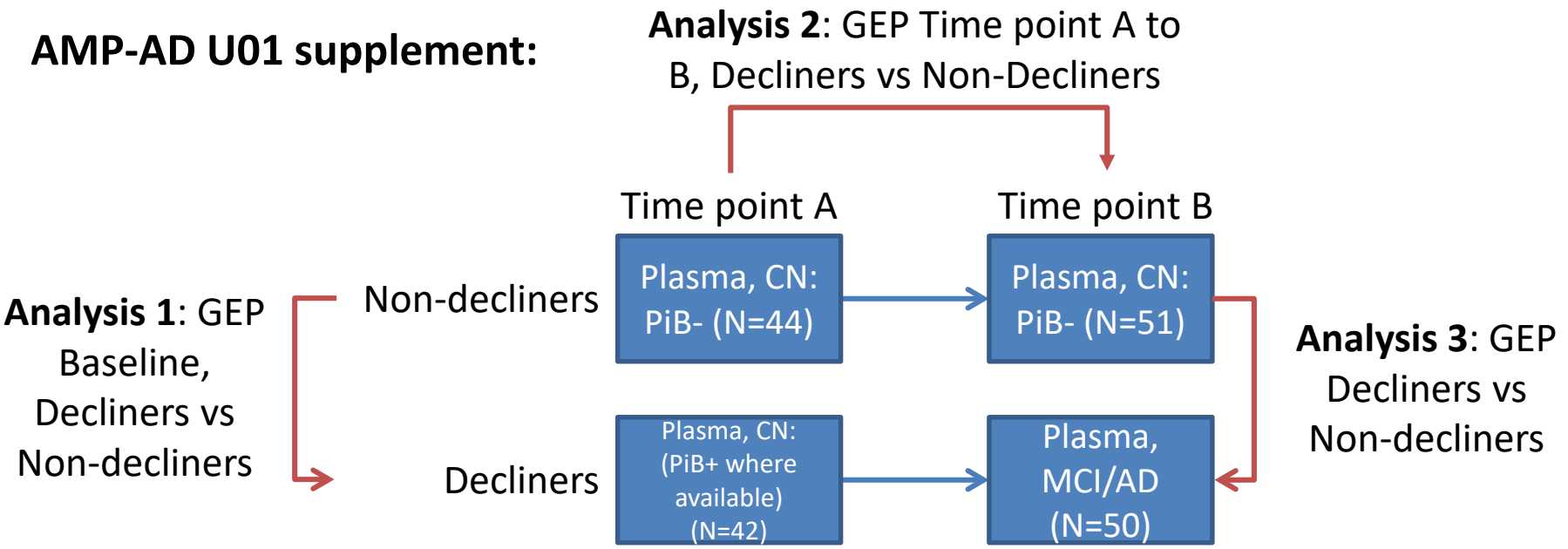
MAYO CLINIC

***Funded by Florida Department of Health Ed and Ethel Moore Alzheimer's Disease Grant**

Summary of Upcoming Biomarker Datasets

- Plasma RNAseq – AMP-AD U01 biomarker pilot
- Plasma miRNAseq – AMP-AD U01 biomarker pilot
- Plasma Nanostring (African-Americans) – FCA^{3}DS
- PAXgene RNAseq – M^{2}OVE AD RF1

AMP-AD U01 supplement:



Expected Outcomes

- Establish ***scalable methodology*** for measurement of cell-free transcriptome in archived plasma samples.
- Discovery of transcripts and transcriptional networks from plasma:
  - That are altered prior to clinical MCI/AD: Candidate ***pre-clinical biomarkers***.
  - That have differential rate of change in decliners vs. non-decliners: Candidate ***biomarkers of rate of decline***.
  - That are altered in incident MCI/AD: Candidate ***biomarkers of clinical impairment***.
- Cross-comparisons with brain transcripts and transcriptional networks that are altered in AD to identify ***CNS transcriptional changes that are reflected in the periphery***.
- Identification of ***eQTL that may drive transcript biomarkers***.

Future Implications

- ***To follow progression*** of specific aspects of the disease pathophysiology as reflected in peripheral transcripts and transcriptional networks (e.g. inflammation, synaptic physiology, lipoproteins).
- Plasma transcriptional biomarkers that have strong correlations with AD pathology may be utilized in the ***stratification and enrichment*** of subjects for clinical trials.
- These biomarkers can be used to follow ***response to therapy*** for those drugs that target specific transcripts/proteins or their networks.


MAYO
CLINIC

Other blood gene expression studies

All study samples come from PAXgene blood from the Mayo Clinic Study of Aging

WG-DASL

GHR Foundation

- N: 44 AD, 56 aMCI, 82 CN-PiB+, 61 CN-PiB-
- 29,377 probes (> 20,000 expressed in > 50% of samples).
- DEG analysis to identify early (CN-PiB+ vs CN-PiB-), disease relevant (AD vs CN-PiB-) changes.
- Future directions: Co-expression network analysis, PiB correlations and eQTL.

NanoString: Transcripts

Mayo Clinic ADRC

- N: WG-DASL cohort (and 8 internal replicates)
- 78 targets + housekeeping and cell type markers.
- AD candidate genes: overall expression and alternatively spliced exons.
- DEG analysis between all groups.
- Future directions: Exon eQTL and PiB correlations.

NanoString: Biomarker

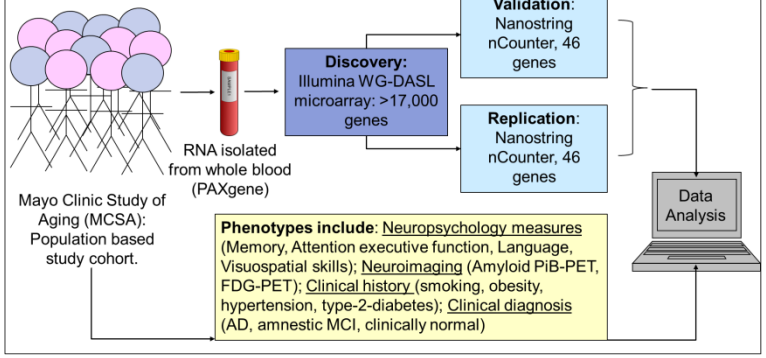
Mayo Clinic CIM

- Validation: WG-DASL cohort
- Replication: 116 CN PiB+, 209 CN PiB- (and 8 internal replicates)
- 47 targets (WG-DASL DEG's) + housekeeping and cell type markers.
- Future Directions: Validation of WG-DASL measures and DEG replication analysis is ongoing.

NanoString Biomarker

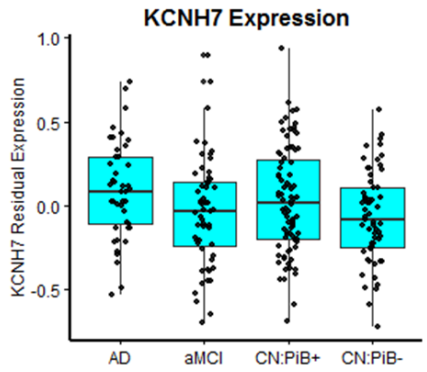
Approach

Hypothesis: Disease relevant gene expression changes can be detected in blood. If true, these may be used as a biomarker to predict disease risk and/or identify subsets of individuals that may benefit from targeted therapies aimed at prevention or treatment.

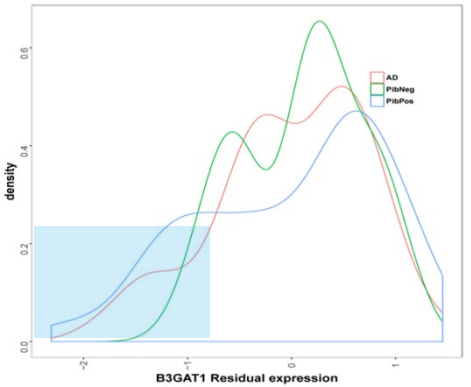


Target Gene Selection

DEG, N=28 genes



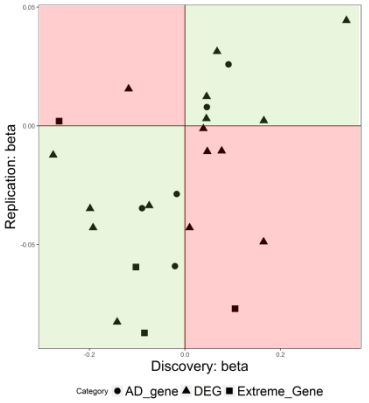
Extreme, N=9 genes



AD Candidate	5	*ABCA7; CLU; APP; PSEN1; PSEN2*
Housekeeping	4	*ACTB; TXNL1; GAPDH; EIF4A2*
Cell Type	4	*CD3G; CD8B1; CD20; NAMPT*

Replication

25/46 genes measured reliably with NanoString; 17/25 consistent direction



Samples

Cohort	Diagnosis	N	Mean Age (SD)	Females (%)	Mean RIN (SD)
Discovery	AD	44	87.39 (5.6)	24 (55%)	9.62 (0.44)
Discovery	aMCI	56	82.82 (5.5)	25 (45%)	9.66 (0.31)
Discovery	CN-PiB+	82	81.41 (5.7)	32 (39%)	9.58 (0.41)
Discovery	CN-PiB-	61	78.48 (4.4)	29 (48%)	9.68 (0.32)
Replication	CN-PiB+	116	74.11 (7.7)	58 (50%)	9.48 (0.49)
Replication	CN-PIB-	209	71.08 (7.5)	100 (48%)	9.35 (0.54)

Results: Validation