MODEL-AD
Bioinformatics and Data Management Core

Greg Carter (Head) The Jackson Laboratory
Ali Mortazavi (Head) University of California, Irvine
Lara Mangravite (Co-Head) Sage Bionetworks
Andrew Saykin (Co-Head) Indiana University
Bioinformatics Overview

<table>
<thead>
<tr>
<th>DATA SOURCES</th>
<th>CORE ACTIVITY</th>
<th>NEW RESOURCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADSP</td>
<td>Variant identification</td>
<td>Coding and noncoding candidates</td>
</tr>
<tr>
<td>ADNI</td>
<td>Variant prioritization</td>
<td>Ranked list of top variants</td>
</tr>
<tr>
<td>ENCODE/RoadMap AMP-AD Allen Brain Atlas</td>
<td>Human-to-mouse variant translation</td>
<td>New genetic models</td>
</tr>
<tr>
<td>MouseExomizer HMDC ATAC-seq</td>
<td>Mouse-to-human phenotype mapping</td>
<td>Validated AD mouse models</td>
</tr>
<tr>
<td>AMP-AD ADNI M²OVE-AD</td>
<td>Data distribution</td>
<td>Expanded AMP-AD Knowledge Portal</td>
</tr>
<tr>
<td>MODEL-AD</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Variant Prioritization

Systematic assessment of LOAD loci

- Significance in multiple studies
- Predicted effect on function
- Human-mouse sequence conservation
- Differential expression in AD
- Noncoding variant effects

➢ Ali Mortazavi, UCI BDMC

Lipid homeostasis/vascular
- APOE, APOC1, MTHFR
- CR1, CD33, TREM2, MYO1C, PSMA1, HMHA1, IL1RAP, TYROBP, STAT3, CSF1R, SPI1, STAT4, PLCG2

Immune
- CLU
- HLA-DRB5
- NCR2, PLXNC1
- RHBD2, PTK2B, PDGFA
- KIF21B, ERC2, PICALM

Mitochondria
- MTHFDL1, TOMM40, SPG7
- CD2AP, INPP5D, MS4A4E, MS4A4A, PVRL2, ANK1, ASGR2, CDH23, GUCY2D, ITM2C, UNX1, MYO10, PCNT, PODXL, PSTPIP1, SLC15A4, SLC16A3, CEACAM1, SNX1

Membrane/ECM

Synaptic Signaling
- BIN1, SLC24A4, KCNN4, BCHE, SLC6A17, HTR4, CLASP2
- CLASP2
- KIF21B, ERC2, PICALM

- WDR81
- FERMT2, SORL1
- ABCA7
- NCR2, PLXNC1
# Variant Summary Metrics

<table>
<thead>
<tr>
<th>Gene</th>
<th>significant association</th>
<th>SNP/gene replication</th>
<th>pathogenic</th>
<th>conserved</th>
<th>differential expression in AD</th>
<th>AD biology</th>
</tr>
</thead>
<tbody>
<tr>
<td>EXO5</td>
<td>✓</td>
<td>✓</td>
<td>✗</td>
<td>✓</td>
<td>↓</td>
<td>?</td>
</tr>
<tr>
<td>CLASP2</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>◼</td>
<td>✓ Reelin Signaling</td>
</tr>
<tr>
<td>MS4A6E</td>
<td>✓</td>
<td>✓</td>
<td>✗</td>
<td>✗</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>SORL1</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>◼</td>
<td>✓ Retromer Trafficking</td>
</tr>
<tr>
<td>PLCG2</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>◼</td>
<td>?</td>
</tr>
<tr>
<td>MAPT</td>
<td>✓</td>
<td>✓</td>
<td>✗</td>
<td>✗</td>
<td>◼</td>
<td>✓ Tau pathology</td>
</tr>
<tr>
<td>MTMR4</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>◼</td>
<td>✓ TGF-beta signaling</td>
</tr>
<tr>
<td>SHC2</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✗</td>
<td>◼</td>
<td>?</td>
</tr>
</tbody>
</table>
Cross-Species Phenotype Alignment

clinical study cohorts

‘Omics

Neuropathology

Brain Imaging

mouse models
### Human Genomics of AD via AMP-AD

<table>
<thead>
<tr>
<th>Study</th>
<th>N</th>
<th>Brain Regions</th>
</tr>
</thead>
<tbody>
<tr>
<td>ROS/MAP</td>
<td>700</td>
<td>dorsolateral prefrontal cortex</td>
</tr>
<tr>
<td>Mt Sinai Brain Bank</td>
<td>300</td>
<td>frontal pole, superiortemporal gyrus, parahippocampal gyrus, inferiorfrontal gyrus</td>
</tr>
<tr>
<td>Mayo Clinic</td>
<td>270</td>
<td>cerebellum, temporal cortex</td>
</tr>
</tbody>
</table>
AMP-AD Gene Modules

Neuronal Modules
- Serotonergic synapse
- Glutamatergic synapse
- Neuronal System

Microglia Modules
- Osteoclast differentiation (TREM2, TYROBP)
- Staphylococcus aureus infection
- Complement and coagulation cascades

Astrocyte Modules
- Sphingolipid metabolism

Mixed Modules
Mouse Model Transcriptomes

![PCA plot of mouse model transcriptomes](image)

- APOEε4/ε4
- APOE-/-
- BIN1-/+ 
- CD2AP-/+ 
- CLU-/- 
- APP/PS1
- ApoE4
- Apoe
- B6
- Bin1
- Cd2ap
- Clu
- App

**PC 1 (13.5%)**

**PC 2 (10%)**
Mouse Gene Modules via WGCNA

ivory module, p:9.66×10^{-6}

orange module, p:4.64×10^{-13}

skyblue3 module, p:4.64×10^{-13}

lightcyan1 module, p:1.85×10^{-5}
Mouse Gene Modules via WGCNA

- **neuroimmune**
  - Antigen processing and presentation
  - Staphylococcus aureus infection
  - Osteoclast differentiation
  - Phagosome
  - Allograft rejection
  - Endocytosis
  - Oxidative phosphorylation
  - Parkinson's disease
  - Fc gamma R-mediated phagocytosis
  - Bacterial invasion of epithelial cells
  - Spliceosome
  - Carbon metabolism
  - Citrate cycle (TCA cycle)
  - Biosynthesis of amino acids
  - RNA transport
  - Nicotine addiction
  - African trypanosomiasis
  - Malaria
  - Pyrimidine metabolism
  - Purine metabolism
  - MAPK signaling pathway
  - Apoptosis
  - FoxO signaling pathway
  - Ribosome biogenesis in eukaryotes
  - Steroid biosynthesis
  - Terpenoid backbone biosynthesis
  - Synthesis and degradation of ketone bodies
  - Fatty acid metabolism

- **neurometabolism**

- **immune**

- **cell death**

- **lipid metabolism**

FDR

- 0.03
- 0.02
- 0.01
Human-Mouse Transcriptome Alignments

- Synapse
- Neurosignaling
- Inflammation
- Microglia
- Astrocyte
- Lipid processing
- Metabolism
nanoString Neuropath Analysis

- 5xFAD, APOE4, and APOE4.TREM2R47H
- Six months of age, whole brain
- Three female replicates
- Compared to 30 AMP-AD modules

- Glycolysis, gluconeogenesis
- Negative regulation of neuron apoptosis

- Inflammation, complement, microglia
- Synaptic, BNDF signaling

- 5xFAD, APOE4, and APOE4.TREM2R47H
- Six months of age, whole brain
- Three female replicates
- Compared to 30 AMP-AD modules

- Glycolysis, gluconeogenesis
- Negative regulation of neuron apoptosis

- Inflammation, complement, microglia
- Synaptic, BNDF signaling
Integrating AMP-AD WGS Data

Imported 1800+ whole genomes from AMP-AD Knowledge Portal

- QC checks for quality scores, sample duplication, etc
- LD pruning, MAF filtering
- PCA for population structure
- Comparison to 1000 Genomes to validate populations
Human-Mouse Neuroimmune Similarity

Ivory mouse module
- Upregulated in ApoE\(^{-/-}\), ApoE4, APP/PS1 mice
- Overlap with human immune modules from AMP-AD \((p = 10^{-29})\)
- Contains \textit{TYROBP}, \textit{TREM2}, \textit{C1QA}, \textit{CSF1R}

\[\text{AMP - AD}\]
- Human Variants
  - SNPs in \textit{TREM2}, \textit{CSF1R}, etc.
- Human Module
  - mRNA of \textit{TYROBP}, \textit{TREM2}, etc.

\[\text{Model - AD}\]
- Mouse Module
  - mRNA of \textit{TYROBP}, \textit{TREM2}, etc.
- Mouse Models
  - ApoE\(^{-/-}\), ApoE4, APP/PS1

\text{ivory module, } p: 9.66 \times 10^{-6}\]
Data Dissemination

Data sharing online

- Mouse genetic information: variant(s), strain background
- Mouse phenotype data: RNA-seq, imaging, etc.
- Preclinical data: standards, protocols, results
- Preclinical results searchable on AlzPED
UCI BDMC Activities

1. Support variant identification and prioritization
   - Focus on non-coding variants
   - Coordinate with IU/Jax/SAGE

2. Reanalyze publicly available data to support variant prioritization in mouse

3. Analyze UCI RNA-seq data produced by center

4. Submit RNA-seq results to Synapse

Focus on #2 and #3 today
Using publicly available chromatin marks in mouse to guide element selection

From Gosselin 2017
Enhancers controlling gene expression can be very far from their gene.

To complicate matters, the Sonic Hedgehog limb enhancer is in the intron of another gene that it does not regulate.
Majority of GWAS SNPs map to open chromatin elements outside of gene coding sequences

(Maurano, 2012)
Distal GWAS SNPs mapping to cognate promoters

(Maurano, 2012)
Topologically associated domains defined by HiC identify interacting regions

From Dixon, 2013
Some of the GWAS hits are on the same TAD – do they interact?
RNA SEQUENCING PIPELINE

TISSUE

RNA

Tissue lyser homogenization

Qiacube RNA extraction

SMART-seq2 construction of Illumina compatible libraries

Wood lab

Data analysis

10-20 million reads per library

Libraries sequenced on Nextseq500

Mortazavi lab
Mouse versus human ages

• 2 month old BL6 mouse would correspond to a teenager

• 8 month old BL6 mouse would correspond to a 35 year old human

• 22 month old BL6 mouse would correspond to a 65 year old human

• RNA-seq data in young mice → early disruption and biomarkers

• RNA-seq data in older mice → better match to to LOAD ?
CSF1R – an AMP-AD target

CSF1R+/- het mice have:

- Impaired memory
- Normal brain size
- Impaired myelination
- Increased microglia

(Chitu, 2016)
Differential Expression Analysis

Frontal Cortex RNA, 8 months old
6 WT vs 6 CSF1R hets

APP and APOE are significantly higher in the CSF1R +/- mice
Pathway analysis flags AD among others

David
P-value 1.9E-3
FDR 7.5E-2
Humanizing Aβ as a platform to introduce GWAS variants

### Alzheimer's Association International Conference (AAIC)

**O1-01** Development of New Models and Analysis Methods: Novel Model Systems to Study Dementia, Sunday, July 22, 2018: 8:00 AM - 9:30 AM, McCormick Place, Room - 184

**O1-01-04** Haβ-KI: A Knock-in Mouse Model for Sporadic Alzheimer's Disease

---

**Humanizing Aβ:**

- HuAβ: DAEFRHDGVEYHVHQKLVFFAEDVGSNKGAIGLMVGGVIA
- moAβ: DAEFGHDSGFEVRHQKLVFFAEDVGSNKGAIGLMVGGVIA

**Secretase Sites:**

- Aβ 1-42
- Aβ 11-42
- β-Secretase +1
- β-Secretase +11
- β-Secretase +17
- α-Secretase
- γ-Secretase +40 +42
Differential expression analysis of hAβKI vs WT by age

2 genotypes x 2 time-points x 2 sexes x 2 replicates = 32 mice

WT22mon_vs_WT2mon

h22mon_vs_h2mon

Log2(Fold Change)

-log10(P-value)

P<=0.05,FDR<=0.1
P<=0.05,FDR>0.1
P>0.05,FDR>0.1

Differential expression analysis of hAβKI vs WT by age
Differential expression analysis of hAβKI vs WT by genotype

Talk by David Baglietto-Vargas
DHCR7 is more highly expressed in hAbKI

- Last enzymatic step in cholesterol synthesis in the adult brain
- What is the cell type overexpressing Dhcr7?
Upcoming UCI BDMC Activities

1. Support variant identification and prioritization
   - Focus on non-coding variants
   - Coordinate with IU/Jax/SAGE

2. Reanalyze publicly available data to support variant prioritization in mouse

3. Analyze UCI RNA-seq data produced by center

4. Submit RNA-seq results to Synapse

5. Analyze single-cell RNA-seq data from aging WT and AD mouse models