Alzheimer’s Disease and Genetics

• Where are we now?
• What can genetics findings be used for?
• What do we expect to achieve?
• What is the final goal?

Is technology driving what we do?
Or
Is technology being used to answer questions?
Alzheimer’s Disease and Genetics

• Where are we now?
• What can genetics findings be used for?
• What do we expect to achieve?
• What is the final goal?

Is technology driving what we do?
Or
Is technology being used to answer questions?
Where are we now?

Rare Variants

<table>
<thead>
<tr>
<th>Early-Onset AD Rare variants</th>
<th>Late-Onset AD Rare variants</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. <strong>APP</strong></td>
<td>1. <em>PSEN2</em> (early and late onset)</td>
</tr>
<tr>
<td>2. <strong>PSEN1</strong></td>
<td>2. <strong>APP</strong> (protective variant)</td>
</tr>
<tr>
<td>3. <strong>PSEN2</strong></td>
<td>3. <strong>TREM2</strong></td>
</tr>
<tr>
<td></td>
<td>4. <strong>UNC5C</strong></td>
</tr>
<tr>
<td></td>
<td>5. <strong>TREML2</strong></td>
</tr>
<tr>
<td></td>
<td>6. <strong>PLXNA4</strong></td>
</tr>
<tr>
<td></td>
<td>7. <strong>AKAP9</strong></td>
</tr>
</tbody>
</table>
Late-Onset AD Common variants

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td><strong>APOE</strong></td>
</tr>
<tr>
<td>2.</td>
<td><strong>SORL1</strong></td>
</tr>
<tr>
<td>3.</td>
<td><strong>CR1</strong></td>
</tr>
<tr>
<td>4.</td>
<td><strong>CLU</strong></td>
</tr>
<tr>
<td>5.</td>
<td><strong>PICALM</strong></td>
</tr>
<tr>
<td>6.</td>
<td><strong>BIN1</strong></td>
</tr>
<tr>
<td>7.</td>
<td><strong>CD2AP</strong></td>
</tr>
<tr>
<td>8.</td>
<td><strong>EPHA1</strong></td>
</tr>
<tr>
<td>9.</td>
<td><strong>MS4A4A</strong></td>
</tr>
<tr>
<td>10.</td>
<td><strong>ABCA7</strong></td>
</tr>
<tr>
<td>11.</td>
<td><strong>HLA-DRB5/HLA-DRB1</strong></td>
</tr>
<tr>
<td>12.</td>
<td><strong>PTK2B</strong></td>
</tr>
<tr>
<td>13.</td>
<td><strong>SLC24A4/RIN3</strong></td>
</tr>
<tr>
<td>14.</td>
<td><strong>MAPT</strong></td>
</tr>
<tr>
<td>15.</td>
<td><strong>CASS4</strong></td>
</tr>
<tr>
<td>16.</td>
<td><strong>INPP5D</strong></td>
</tr>
<tr>
<td>17.</td>
<td><strong>MEF2C</strong></td>
</tr>
<tr>
<td>18.</td>
<td><strong>NME8</strong></td>
</tr>
<tr>
<td>19.</td>
<td><strong>ZCWPW1</strong></td>
</tr>
<tr>
<td>20.</td>
<td><strong>CELF1</strong></td>
</tr>
<tr>
<td>21.</td>
<td><strong>FERMT2</strong></td>
</tr>
<tr>
<td>22.</td>
<td><strong>TREM2L/TREM2</strong></td>
</tr>
<tr>
<td>23.</td>
<td><strong>GLIS3</strong></td>
</tr>
<tr>
<td>24.</td>
<td><strong>ABCG1</strong></td>
</tr>
<tr>
<td>25.</td>
<td><strong>GalNAc</strong></td>
</tr>
<tr>
<td>26.</td>
<td>Intergenic – chr 9</td>
</tr>
<tr>
<td>27.</td>
<td><strong>FRMD4A</strong></td>
</tr>
</tbody>
</table>

---

**APOE**

- $P < 5 \times 10^{-8}$
- $OR = 1.08 - 1.37$, **5.22**
- $MAF = 3.9\% - 49\%$
GWAS signals: 90-95% of causative variants are in non-promoter regulatory elements

Genomic information
- eQTL (GTEX)
- Roadmap
- Encode
- Fantom 5

New Technology
- Capture C methods
- CRISPR
Known AD Genes

1. APP  
2. PSEN1  
3. PSEN2  
4. APOE  
5. SORL1  
6. CR1  
7. ABCA7  
8. MAPT  
9. TREM2  
10. UNC5C  
11. PLXNA4  
12. AKAP

Other GWAS loci

"New wine from old bottles"
Alzheimer’s Disease and Genetics

• Where are we now?
• What can genetics findings be used for?
• What do we expect to achieve?
• What is the final goal?

Is technology driving what we do?
Or
Is technology being used to answer questions?
• Prediction

• Mechanism

• Drug targets
• Prediction

• Mechanism

• Drug targets
<table>
<thead>
<tr>
<th>Gene</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>APP</td>
<td>Autosomal dominant</td>
</tr>
<tr>
<td>PSEN1</td>
<td>Highly penetrant</td>
</tr>
<tr>
<td>PSEN2</td>
<td>Early-onset</td>
</tr>
</tbody>
</table>

**Rare! (< 0.1%?)**

Highly predictive
Applies to a very small number of cases

Used to design prevention trials
### APOE

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Odds ratio (95% Confidence interval)</th>
<th>Control frequency (percent)</th>
<th>Case frequency (percent)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ε3/ε3</td>
<td>1.0 (referent)</td>
<td>60.9</td>
<td>36.4</td>
</tr>
<tr>
<td>ε2/ε2</td>
<td>0.6 (0.2 – 2.0)</td>
<td>0.8</td>
<td>0.2</td>
</tr>
<tr>
<td>ε2/ε3</td>
<td>0.6 (0.5 – 0.8)</td>
<td>12.7</td>
<td>4.8</td>
</tr>
<tr>
<td>ε2/ε4</td>
<td>2.6 (1.6 – 4.0)</td>
<td>2.6</td>
<td>2.6</td>
</tr>
<tr>
<td>ε3/ε4</td>
<td>3.2 (2.8 – 3.8)</td>
<td>21.3</td>
<td>41.1</td>
</tr>
<tr>
<td>ε4/ε4</td>
<td>14.9 (10.8 – 20.6)</td>
<td>1.8</td>
<td>14.8</td>
</tr>
<tr>
<td>ε4/ε4</td>
<td>35.07 (23.8 – 51.8)</td>
<td>onset age 60 – 69 years</td>
<td></td>
</tr>
</tbody>
</table>

### APOE

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Percent in Controls</th>
<th>Life Time Risk – Age 85 years (male/female)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ε3/ε3</td>
<td>60.9%</td>
<td>7% – 12%</td>
</tr>
<tr>
<td>ε3/ε4</td>
<td>21.3%</td>
<td>22% - 35%</td>
</tr>
<tr>
<td>ε4/ε4</td>
<td>1.8%</td>
<td>52% – 68%</td>
</tr>
</tbody>
</table>

Life-time risk – risk to develop AD between birth and a specific age (85 year in table above)

Genin et. al. Molec. Psychiatry 16, 903 (2011)
Other rare variant genes

- **TREM2**

<table>
<thead>
<tr>
<th>Variant</th>
<th>control carrier frequency</th>
<th>Odds ratio</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>R47H</td>
<td>0.4%</td>
<td>2.29</td>
<td>4.31 x 10^{-12}</td>
</tr>
<tr>
<td>R62H</td>
<td>1.26%</td>
<td>1.67</td>
<td>5.64 x 10^{-12}</td>
</tr>
</tbody>
</table>

Two additional loci:
- OR = 1.58 (risk allele frequency in controls = 1.28%)
- OR = 2.29 (risk allele frequency in controls = 0.4%)

African Americans: 
- **ABCA7**
  - OR = **1.79 (CI, 1.47 – 2.12); risk allele controls = 7%**

Caucasians: 
- **ABCA7**
  - OR = **1.11 (CI, 1.11 – 1.19) risk allele controls = 19%**

Reitz et al. JAMA 309, 1483 (2013)
<table>
<thead>
<tr>
<th>Chr.</th>
<th>Closest gene(^2)</th>
<th>MAF (SE)(^3)</th>
<th>OR (95% CI)</th>
<th>Meta P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>CR1</td>
<td>0.197 (0.012)</td>
<td>1.18 (1.14-1.22)</td>
<td>5.7x10(^{-24})</td>
</tr>
<tr>
<td>2</td>
<td>BIN1</td>
<td>0.409 (0.017)</td>
<td>1.22 (1.18-1.25)</td>
<td>6.9x10(^{-44})</td>
</tr>
<tr>
<td>6</td>
<td>CD2AP</td>
<td>0.266 (0.010)</td>
<td>1.10 (1.07-1.13)</td>
<td>5.2x10(^{-11})</td>
</tr>
<tr>
<td>7</td>
<td>EPHA1</td>
<td>0.338 (0.010)</td>
<td>0.90 (0.88-0.93)</td>
<td>1.1x10(^{-13})</td>
</tr>
<tr>
<td>8</td>
<td>CLU</td>
<td>0.379 (0.010)</td>
<td>0.86 (0.84-0.89)</td>
<td>2.8x10(^{-25})</td>
</tr>
<tr>
<td>11</td>
<td>MS4A6A</td>
<td>0.403 (0.012)</td>
<td>0.90 (0.87-0.92)</td>
<td>6.1x10(^{-16})</td>
</tr>
<tr>
<td>11</td>
<td>PICALM</td>
<td>0.358 (0.008)</td>
<td>0.87 (0.85-0.89)</td>
<td>9.3x10(^{-26})</td>
</tr>
<tr>
<td>19</td>
<td>ABCA7</td>
<td>0.190 (0.012)</td>
<td>1.15 (1.11-1.19)</td>
<td>1.1x10(^{-15})</td>
</tr>
<tr>
<td>6</td>
<td>HLA-DRB5/HLA-DRB1</td>
<td>0.276 (0.012)</td>
<td>1.11 (1.08-1.15)</td>
<td>2.9x10(^{-12})</td>
</tr>
<tr>
<td>8</td>
<td>PTK2B</td>
<td>0.366 (0.012)</td>
<td>1.10 (1.08-1.13)</td>
<td>7.4x10(^{-14})</td>
</tr>
<tr>
<td>11</td>
<td>SORL1</td>
<td>0.039 (0.004)</td>
<td>0.77 (0.72-0.82)</td>
<td>9.7x10(^{-15})</td>
</tr>
<tr>
<td>14</td>
<td>SLC24A4/RIN3</td>
<td>0.217 (0.009)</td>
<td>0.91 (0.88-0.94)</td>
<td>5.5x10(^{-9})</td>
</tr>
<tr>
<td>2</td>
<td>INPP5D</td>
<td>0.488 (0.018)</td>
<td>1.08 (1.05-1.11)</td>
<td>3.2x10(^{-8})</td>
</tr>
<tr>
<td>5</td>
<td>MEF2C</td>
<td>0.408 (0.010)</td>
<td>0.93 (0.90-0.95)</td>
<td>3.2x10(^{-8})</td>
</tr>
<tr>
<td>7</td>
<td>NME8</td>
<td>0.373 (0.012)</td>
<td>0.93 (0.90-0.95)</td>
<td>4.8x10(^{-9})</td>
</tr>
<tr>
<td>7</td>
<td>ZCWPW1</td>
<td>0.287 (0.016)</td>
<td>0.91 (0.89-0.94)</td>
<td>5.6x10(^{-10})</td>
</tr>
<tr>
<td>11</td>
<td>CELF1</td>
<td>0.316 (0.022)</td>
<td>1.08 (1.05-1.11)</td>
<td>1.1x10(^{-8})</td>
</tr>
<tr>
<td>14</td>
<td>FERMT2</td>
<td>0.092 (0.009)</td>
<td>1.14 (1.09-1.19)</td>
<td>7.9x10(^{-9})</td>
</tr>
<tr>
<td>20</td>
<td>CASS4</td>
<td>0.083 (0.006)</td>
<td>0.88 (0.84-0.92)</td>
<td>2.5x10(^{-8})</td>
</tr>
<tr>
<td>Chr.</td>
<td>Closest gene²</td>
<td>MAF (SE)³</td>
<td>Overall</td>
<td>OR (95% CI)</td>
</tr>
<tr>
<td>------</td>
<td>---------------</td>
<td>-----------</td>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td>1</td>
<td>CR1</td>
<td>0.197 (0.012)</td>
<td></td>
<td>1.18 (1.14-1.22)</td>
</tr>
<tr>
<td>2</td>
<td>BIN1</td>
<td>0.409 (0.017)</td>
<td></td>
<td>1.22 (1.18-1.25)</td>
</tr>
<tr>
<td>6</td>
<td>CD2AP</td>
<td>0.266 (0.010)</td>
<td></td>
<td>1.10 (1.07-1.13)</td>
</tr>
<tr>
<td>7</td>
<td>EPHA1</td>
<td>0.338 (0.010)</td>
<td></td>
<td>0.90 (0.88-0.93)</td>
</tr>
<tr>
<td>8</td>
<td>CLU</td>
<td>0.379 (0.010)</td>
<td></td>
<td>0.86 (0.84-0.89)</td>
</tr>
<tr>
<td>11</td>
<td>MS4A6A</td>
<td>0.403 (0.012)</td>
<td></td>
<td>0.90 (0.87-0.92)</td>
</tr>
<tr>
<td>11</td>
<td>PICALM</td>
<td>0.358 (0.008)</td>
<td></td>
<td>0.87 (0.85-0.89)</td>
</tr>
<tr>
<td>19</td>
<td>ABCA7</td>
<td>0.190 (0.012)</td>
<td></td>
<td>1.15 (1.11-1.19)</td>
</tr>
<tr>
<td>6</td>
<td>HLA-DRB5/HLA-DRB1</td>
<td>0.276 (0.012)</td>
<td></td>
<td>1.11 (1.08-1.15)</td>
</tr>
<tr>
<td>8</td>
<td>PTK2B</td>
<td>0.366 (0.012)</td>
<td></td>
<td>1.10 (1.08-1.13)</td>
</tr>
</tbody>
</table>

**OR or (1/OR) range = 1.08 – 1.22**

**Prediction?**
200 risk genes
Risk allele is common (0.1 – 0.9)
GRR = 1.04
Disease prevalence is 10%

Prediction

• Some rare variants are highly predictive
• Some rare variants will be modestly predictive
• *APOE* is a major contributor to risk
• Common variants have a limited contribution to risk assessment

• Will find more common variants
• May find more rare-variants – modest predictive value
• Unlikely that a common large effect gene (*e.g.* *APOE*) will be found
• Prediction
• Mechanism
• Drug targets

Expected outcomes:
• Find genes that identify specific mechanisms
• Multiple genes in the same pathway
• Effect direction
  o High risk allele loss or gain of function
  o High-risk allele increase of decrease expression
Mechanisms/pathways

Aβ metabolism

*APP, PSEN1, PSEN2*

Innate immune system – microglial cells

*TREM2, CR1, MS4 region, two new genes*

Cholesterol metabolism (?)

*APOE, ABCA7*

Intracellular vesicle trafficking

*SORL1, ABCA7*

Synaptic dysfunction/membrane function

*PICALM, BIN1, EPHA1*
Can Genetic discoveries be used for drug target identification?

Alzheimer’s Disease
Aβ antibodies
Presenilin inhibitors
BACE1 inhibitors
APOE
Coronary artery disease

- PCSK9: Proprotein convertase subtilisin/kexin type 9

What about small-effect genes and drug development?
**Biological, clinical and population relevance of 95 loci for blood lipids**

A list of authors and their affiliations appears at the end of the paper.

Plasma concentrations of total cholesterol, low-density lipoprotein cholesterol, high-density lipoprotein cholesterol and triglycerides are among the most important risk factors for coronary artery disease (CAD) and are targets for therapeutic intervention. We screened the genome for common variants associated with plasma lipids in >100,000 individuals of European ancestry. Here we report 95 significantly associated loci ($P < 5 \times 10^{-8}$), with 59 showing genome-wide significant association with lip traits for the first time. The newly reported associations include single nucleotide polymorphisms (SNPs) near known lipid regulators (for example, CYP7A1, NPC1L1 and SCARB1) as well as in scores of loci not previously implicated in lipoprotein metabolism. The 95 loci contribute not only to normal variation in lip traits but also to extreme lipid phenotypes and have impact on lip traits in three non-European populations (East Asians, South Asians and African Americans). Our results identify several novel loci associated with plasma lipids that are also associated with CAD. Finally, we validated three of the novel genes—GALNT2, PP1R3B and TTC39B—with experiments in mouse models. Taken together, our findings provide the foundation to develop a broader biological understanding of lipoprotein metabolism and to identify new therapeutic opportunities for the prevention of CAD.

<table>
<thead>
<tr>
<th>Gene*</th>
<th>P-value</th>
<th>Trait</th>
<th>Effect size (mg/dl)</th>
<th>drug</th>
</tr>
</thead>
<tbody>
<tr>
<td>HMG co-A reductase</td>
<td>$9 \times 10^{-47}$</td>
<td>TC</td>
<td>+2.84</td>
<td>statins</td>
</tr>
<tr>
<td><em>NPC1L1</em></td>
<td>$3 \times 10^{-11}$</td>
<td>TC</td>
<td>+2.01</td>
<td>ezetimibe</td>
</tr>
<tr>
<td>PCSK9</td>
<td>$2 \times 10^{-28}$</td>
<td>LDL</td>
<td>+2.01</td>
<td>alirocumbab</td>
</tr>
<tr>
<td><em>APOE</em></td>
<td>$9 \times 10^{-147}$</td>
<td>LDL</td>
<td>+7.14</td>
<td>none</td>
</tr>
</tbody>
</table>
We estimate that selecting genetically supported targets could double the success rate in clinical development.
Drug targets

• Genetic studies can lead to successful drug development

• Effect size:
  — Does not predict druggability
  — Small effect genes are potential drug targets.
Tesla Syndrome

It’s new
It’s shiny
It must be better!

I want one!
Tesla Syndrome

Is technology driving what we do?

Or

Are we using better technology to answer important questions?
Next Generation DNA sequencing
  • whole exome
  • whole genome

RNASeq

Histone acetylation/methylation

Methylation

DNase hypersensitivity

XYZomics
Whole exome sequencing

- DNA sequence for all (most) exons
- $\sim$600/sample
- Coding/splice junction mutations
- 5’ and 3’ UTR
- Some small RNAs

Whole genome sequencing

- DNA sequence for all (most) $3 \times 10^9$ nucleotides
- $1,250 – 1,350/sample$
- All coding, non-coding, intergenic mutations
- Can get all (most) structural variants
**Pros** - Whole exome sequencing
  - cheaper
  - Less costly to process data
    ($40/subject versus $140/subject)
  - Easy (sort of) to interpret

**Cons**
  - Miss 98% of genetic variability
  - Limited resolution for structural variants

**Pros** - Whole genome sequencing
  - Get all (most) mutations
  - **Potential for all structural variants**

**Cons**
  - More expensive to produce/process/store
  - More difficult to interpret ALL the data
Genomic resources

- Encode
- RoadMap
- GTex
- Fantom5
- Other databases

Enhance interpretation of intergenic and intronic genetic variation
Structural Variants (SVs) Introduction

Type
- Insertions
- Deletions
- Inversions
- Translocations
- Copy number variation (CNV)

Size
- Inversely related to frequency
- 1bp to very large
• SVs account for more of our genetic variability than single nucleotide variants (SNPs)
• Currently miss SVs in the 1 bp to ~5,000 bp range
Structural Variants (SVs)

- SVs - Alzheimer’s disease/other neurologic disorders
  - *APP* duplication
  - *SNCA* duplication
  - *PSEN1* indel
  - *PMP22* deletion/duplication
  - *MAPT* inversion/CNVs
  - Loss-of-function deletion – *ABCA7*

Whole genome sequencing will allow us to see SVs of all sizes – not previously genotyped
Imputation

1. Combined data from different genotyping platforms

2. Test variants not directly genotyped: rare-variants

Reference panel

- 30,000 Genomes

Rare-Variant GWAS

Rare variant detected by sequencing

Rare variant carrier imputed
Concordance: HRC vs. WGS

<table>
<thead>
<tr>
<th>MAF Ranges</th>
<th>“Best Guess” % Concordance</th>
<th>“Stringent” % Concordance</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.2-1%</td>
<td>99.731</td>
<td>99.956</td>
</tr>
<tr>
<td>1-3%</td>
<td>99.522</td>
<td>99.924</td>
</tr>
<tr>
<td>3-5%</td>
<td>99.373</td>
<td>99.876</td>
</tr>
<tr>
<td>5-10%</td>
<td>99.389</td>
<td>99.847</td>
</tr>
<tr>
<td>10-15%</td>
<td>99.239</td>
<td>99.646</td>
</tr>
<tr>
<td>15-20%</td>
<td>99.068</td>
<td>99.424</td>
</tr>
</tbody>
</table>

Genotypes Used: 90,251,702 (HRC) vs. 76,642,843 (WGS)

- Hard-call genotypes from imputation
  - “Best Guess” – call goes to any genotype with prob>0.5
  - “Stringent” – call goes to genotypes only with prob>0.9
- Comparing HRC imputation of ADNI GWAS vs. ADNI WGS genotype calls in 213 ADNI samples
- Only looking at alternate genotype concordance (R/A; A/A)
Imputation:
- Uses existing GWAS genotypes
- Impute to an allele frequency of 0.1%

- Most SVs
- Disease specific mutations
- Ethnic groups not in reference panel

“New wine from old bottles”

Analyze rare variants not directly genotyped

Rare-Variant GWAS
Whole genome sequencing:

• detect all SVs

• detect rare variants not in reference panels
  (e.g. Alzheimer’s disease – specific variants)

• variants in different ethnic groups

Use both whole genome sequencing and imputation
Future:

• Use imputed genotypes for replication studies
• Generate reference panels from:
  o Different ethnic groups
  o Disease populations
  o Sequence data processed for SVs
Goal: Completely resolve the genetics of AD – all genetic variation that alters risk

The longer the list of valid risk/preventative genes;

• The better the chance of finding a druggable target
• The more we will understand about disease mechanism
Alzheimer’s Disease Genetics Consortium

University of Pennsylvania
Li-San Wang
Adam Naj
Laura Cantwell
Beth Dombroski
Otto Valladeras
Sherry Beecher

Mass. General
Deborah Blacker

Standford
Tom Montine

Mt. Sinai school of Medicine
Alison Goate

University of Washington
Bud Kukull
Duane Beekly
Debbie Tsuang

Indiana University
Tatiana Foroud
Kelly Michelle Faber

University of Miami
Peggy Pericak-Vance
Gary Beecher
Eden Martin
Kara Hamilton
Brian Kunkle

Boston University
Lindsay Farrer
Kathryn Lunetta

Case Western
Jonathan Haines
Will Bush

Columbia
Richard Mayeux
Christiane Reitz
Badri Vardarajan
Jennifer Manly

NIA
Marilyn Miller

NIA/NIH, Alzheimer’s Association
The End