Genetics of AAA Disease

Philip S Tsao, PhD

Stanford Cardiovascular Institute
Stanford University School of Medicine
VA Palo Alto Health Care System
AAA - Risk Factors

- Dilation of the infrarenal aorta (>3.0cm)
- AAAs are common, lethal
  - 2-6% men; 1-2% women >60y/o
- Associated with atherosclerosis but with distinct differences
  - Transmural inflammation
  - SMC apoptosis
  - Impaired extracellular remodeling
  - Progressive luminal expansion
- Family Hx; Smoking; HTN; DM(-)
- No effective therapy for early disease
- Intervention is guided by imaging (ultrasound, CT scans, etc)
**Epidemiology Studies: AAA has a Strong Genetic Component**

- Family history studies estimated a 70% genetic component of AAA
- Twins study (265 twins) suggested a 70% heritability of AAA
- 1st degree relatives have 2x risk
Background: Single Nucleotide Polymorphism (SNP; or Variant SNV)

All Humans are at least 99.5% similar in their genetic code.

Most common of form genetic variation occurs when a single nucleotide differs between individuals.

Single Nucleotide Polymorphism (SNP)

On average ~ 1/300 base pairs

Conclusion:

Uncorrected multiple testing and flexible study design (particularly testing many inheritance models and subgroups, and failure to check for Hardy-Weinberg equilibrium) contributed to apparently false associations being reported. Heterogeneity, possibly due to the case mix, geographical, temporal, and environmental variation between different studies, was evident. Polymorphisms in nine genes had strong or moderate support on the basis of the literature at this time.
Associating DNA Variation with AAA

**GWAS: common variants**
- All variants aggregated by gene (gene burden test)
- GWAS; pathway analysis
  - DAB2IP
  - LRP1
  - LDLR
  - SORT1
  - IL6R
  - CDKN2BAS1/ANRIL
  - SMYD2
  - LINC00540
  - near PCIF1/MMP9/ZNF335/ERG

**Less-frequent Variants**
- MAF<=5% in 1KG, CG, NHLBI ESP
- Association test: c-alpha
  1. Input VCF
  2. Remove blacklisted genes
  3. Select less-frequent variants
  4. General functional relevance
  5. Association test (case vs ctrl)
  6. AAA relevance w/ Ingenuity curation

305 variants, 38 genes

**Rare Variants**
- Not present in 1KG calls
- HEAL-mutational burden case/ctrl; PPI network

GWAS: common variants

Less-frequent Variants

Rare Variants
Decode the Genome for AAA: 480 Deeply Sequenced Human Genomes

Patient Recruitment
- Stanford Hospital
- VAPAHCS
- Kaiser Northern CA

DNA Isolation from Blood

WGS
- HiSeq 2000: ~50x
- Omni2.5M array

Variant Calling/QC
- BWA/GATK (cloud)
- sample-level QC
- variant-level QC

After stringent QC: 401 samples, 23.7M SNVs, 7.9M INDELs
Characteristics of the Study Population

Control: <= 2.5cm

Grey area: 2.5 – 3.5 cm

Case: >= 3.5cm

<table>
<thead>
<tr>
<th></th>
<th>Control</th>
<th>Grey area</th>
<th>Case</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total (N=401)</td>
<td>103</td>
<td>99</td>
<td>199</td>
</tr>
<tr>
<td>female (N=)</td>
<td>11</td>
<td>13</td>
<td>12</td>
</tr>
<tr>
<td>male (N=)</td>
<td>92</td>
<td>86</td>
<td>187</td>
</tr>
</tbody>
</table>
Associating DNA Variation with AAA

GWAS: common variants
- all variants aggregated by gene (gene burden test)
- GWAS; pathway analysis
  - DAB2IP
  - LRP1
  - LDLR
  - SORT1
  - IL6R
  - CDKN2BAS1/ANRIL
  - SMYD2
  - LINC00540
  - near PCIF1/MMP9/ZNF335
  - ERG

Less-frequent Variants
- MAF<=5% in 1KG, CG, NHLBI ESP
- association test: c-alpha
  1. Input VCF
  2. remove blacklisted genes
  3. select less-frequent variants
  4. general functional relevance
  5. Association test (case vs ctrl)
  6. AAA relevance w/ Ingenuity curation

305 variants, 38 genes

Rare Variants
- not present in 1KG calls
- HEAL-mutational burden case/ctrl; PPI network

GWAS:
- common variants
- all variants aggregated by gene (gene burden test)
- GWAS; pathway analysis

DAB2IP
LRP1
LDLR
SORT1
IL6R
CDKN2BAS1/ANRIL
SMYD2
LINC00540
near PCIF1/MMP9/ZNF335
ERG

60 genes → 40 PPI modules
Less Frequent Variants: Case vs Ctrl

**Input variants**
- 20M
- 17.3M
- 10.9M
- 145.8K
- 4.4K

**Removed variants on blacklisted genes**
- Removed variants w/ AF >5% in 1KG, CG, NHLBI ESP

**General Functional Relevance**
- Association test (p<0.04, genes, pathways)

**AAA relevance** *(305 variants, 38 genes)*

- Related to Abdominal Aortic Aneurysm in the curated Ingenuity databases

**Experimental evidence**
- Pathogenic; Likely pathogenic;
- Uncertain significance

**Gain of function of a gene**
- Established in literature; gene fusion; microRNA binding site
- Inferred activating mutation by Ingenuity; predicted gain-of-function by BSIFT

**Loss of function of a gene**
- Frameshift, in-frame, stop codon change, missense, likely splice site loss up to 2 bp into intron;
- Deleterious to a microRNA; promoter loss; enhancer
Less Frequent Variants Associated with AAA

305 variants, 38 genes

- Heterogeneity: 50% variants occur in only one genome
- Functionality: 80% of the variants are on promoters and exons
- Pathogenicity: 40 variants (13%) are likely pathogenic or pathogenic
305 variants, 38 genes

Processes Enriched in AAA Patients

- Immunology (257)
- Physiology (257)
- Developmental biology (254)
- Apoptosis (247)
- Immune cells (244)
- Leukocytes (244)
- Necrosis (237)
- Molecular biology (234)
- Cell migration and motility (231)
- Cell growth (227)
- Cell proliferation (224)
- Cellular metabolism (224)
- Development of embryo (219)
- Development of organism (219)
- Developmental process of embryo... (219)
- Cell signaling (262)
- Proliferation of cells (269)
- Cell death and survival (269)
- Cell death (269)
- Cell growth and proliferation (271)
- Cell biology (305)
- Organismal death and survival (196)
- Developmental process of tumor... (201)
- Transcriptional regulation (201)
- Organogenesis (202)
- Development of organ (202)
- Regulation of immune response (211)
- Tissue development (219)
- Developmental process of tissue... (219)
- Developmental process of organ... (219)
Pathways Enriched in AAA Patients

(305 variants, 38 genes. p-values: 1.2E-6 - 2.3E-2)
# Diseases Enriched by the 38 Genes

<table>
<thead>
<tr>
<th>Name</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neurological Disease</td>
<td>4.96E-12</td>
</tr>
<tr>
<td>myocardial infarction</td>
<td>6.79E-11</td>
</tr>
<tr>
<td>Reproductive System Disease</td>
<td>2.07E-10</td>
</tr>
<tr>
<td>acute coronary syndrome</td>
<td>3.52E-10</td>
</tr>
<tr>
<td>Skeletal and Muscular Disorders</td>
<td>4.01E-10</td>
</tr>
<tr>
<td>Inflammatory Disease</td>
<td>1.78E-09</td>
</tr>
<tr>
<td>disease of central nervous system</td>
<td>2.65E-09</td>
</tr>
<tr>
<td>Connective Tissue Disorders</td>
<td>2.65E-09</td>
</tr>
<tr>
<td>Fibrosis</td>
<td>2.81E-09</td>
</tr>
<tr>
<td>encephalopathy</td>
<td>2.96E-09</td>
</tr>
<tr>
<td>Psychological Disorders</td>
<td>4.27E-09</td>
</tr>
<tr>
<td>Infarction</td>
<td>2.05E-08</td>
</tr>
</tbody>
</table>

**myocardial infarction genes**

<table>
<thead>
<tr>
<th>Gene</th>
<th>Cases</th>
<th>Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGTR1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CD40LG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CD44</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CETP</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CSF1R</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ESR1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>F2R</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Expression of the 38 Genes in Two Types of AAA Animal models

<table>
<thead>
<tr>
<th>WGS</th>
<th>Elast</th>
<th>ANGII</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Down regulated: 5 genes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GNA11</td>
<td>0.25</td>
<td>0.35</td>
</tr>
<tr>
<td>PPARA</td>
<td>0.11</td>
<td>0.45</td>
</tr>
<tr>
<td>IFNA21</td>
<td>0.5</td>
<td>0.59</td>
</tr>
<tr>
<td>FKBP4</td>
<td>0.5</td>
<td>0.61</td>
</tr>
<tr>
<td>FGFR1</td>
<td>0.49</td>
<td>1.42</td>
</tr>
<tr>
<td>Up regulated: 6 genes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TNFRSF4</td>
<td>3.1</td>
<td>1.85</td>
</tr>
<tr>
<td>PLK2</td>
<td>2.67</td>
<td>2.92</td>
</tr>
<tr>
<td>CD44</td>
<td>7.2</td>
<td>2.98</td>
</tr>
<tr>
<td>IKBKE</td>
<td>4.3</td>
<td>3.01</td>
</tr>
<tr>
<td>SPARC</td>
<td>2.72</td>
<td>3.65</td>
</tr>
<tr>
<td>CSF1R</td>
<td>6.12</td>
<td>4.47</td>
</tr>
</tbody>
</table>

• 21/38 genes (~60%) differentially regulated in ANGII mouse model
• 14/38 genes (~40%) differentially regulated in ELAST mouse model
• Shared pathways/cascades (e.g.):
  - Cell cycle
  - Cell growth and proliferation
  - Apoptosis
  - Adhesion
  - NF-kB cascade
Associating DNA Variation with AAA

GWAS: common variants
- all variants aggregated by gene (gene burden test)
- GWAS; pathway analysis
  - DAB2IP
  - LRP1
  - LDLR
  - SORT1
  - IL6R
  - CDKN2BAS1/ANRIL
  - SMYD2
  - LINC00540
  - near PCIF1/MMP9/ZNF335
  - ERG

Less-frequent Variants
- <=5% in 1KG, CG, NHLBI ESP
- association test: c-alpha
  1. Input VCF
  2. remove blacklisted genes
  3. select less-frequent variants
  4. general functional relevance
  5. Association test (case vs ctrl)
  6. AAA relevance w/ Ingenuity curation

- 305 variants, 38 genes

Rare Variants
- not present in 1KG calls
- HEAL-mutational burden case/ctrl; PPI network

GWAS:
- common variants
  - all variants aggregated by gene (gene burden test)
  - GWAS; pathway analysis

- DAB2IP
- LRP1
- LDLR
- SORT1
- IL6R
- CDKN2BAS1/ANRIL
- SMYD2
- LINC00540
- near PCIF1/MMP9/ZNF335
- ERG
Rare Variants by Gene Burden: 60 Genes Segregated Case from Ctrl

60 genes

Enriched for immune-related functions:
- interferon-gamma-mediated signaling (FDR=0.07)
- MHC class II receptor activity (FDR=1.6e-4)
- T cell co-stimulation (FDR=0.07)

Consistent with transcriptional change in AAA tissue samples (transcriptome data from: Biros, Gabel et al. 2015)
40 PPI Modules Enriched in AAA Genomes
40 PPI Modules Enriched in AAA Genomes
### Reported AAA-genes vs Modules

#### Table 4: Assessment of evidence of association.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Polymorphism</th>
<th>Amount of evidence</th>
<th>Replication</th>
<th>Protection from bias</th>
<th>Overall assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>M_693: HLA-A</td>
<td>rs7633818</td>
<td>A</td>
<td>C</td>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td>M_694: HLA-B</td>
<td>rs10757278</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>Strong</td>
</tr>
<tr>
<td>M_75: MMP12</td>
<td>rs4643994</td>
<td>A</td>
<td>A</td>
<td>C</td>
<td>Weak</td>
</tr>
<tr>
<td>M_438: NOS1</td>
<td>rs5186</td>
<td>A</td>
<td>B</td>
<td>B</td>
<td>Moderate</td>
</tr>
<tr>
<td>M_75: PLA2G4C</td>
<td>rs964184</td>
<td>A</td>
<td>C</td>
<td>C</td>
<td>Weak</td>
</tr>
<tr>
<td>M_75: PLA2G4D</td>
<td>rs934901</td>
<td>A</td>
<td>C</td>
<td>C</td>
<td>Weak</td>
</tr>
<tr>
<td>M_433: PLA2G16</td>
<td>rs1367117</td>
<td>A</td>
<td>C</td>
<td>C</td>
<td>Weak</td>
</tr>
<tr>
<td>M_672: PLA2G2A</td>
<td>rs333</td>
<td>B</td>
<td>C</td>
<td>C</td>
<td>Weak</td>
</tr>
<tr>
<td>M_75: PLA2G4C</td>
<td>rs3991244</td>
<td>A</td>
<td>C</td>
<td>C</td>
<td>Weak</td>
</tr>
<tr>
<td>M244: PHACTR2</td>
<td>rs7025426</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>Strong</td>
</tr>
<tr>
<td>ELN</td>
<td>rs2071307</td>
<td>B</td>
<td>C</td>
<td>C</td>
<td>Weak</td>
</tr>
</tbody>
</table>

**Legend:**
- A: Association
- B: Benign
- C: Confirmed
- R: Replication
- S: Strong
- M: Moderate
- W: Weak
Genetic Data adds to Clinical Data

![Receiver Operating Characteristic (ROC) curve](image)

- True positive rate (Sensitivity)
- False positive rate (1-Specificity)

- Genome: 0.690
- EHR: 0.775
- Genome+EHR: 0.803
Crosstalk of the Two Analyses

- **Two Levels of Overlap (gene + regulator):** 5 genes (13%)

<table>
<thead>
<tr>
<th>Gene</th>
<th>Regulator</th>
<th>Module</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPARC</td>
<td>MIR29A</td>
<td>M_244:SPARC</td>
</tr>
<tr>
<td>IL3</td>
<td>GATA2</td>
<td>M_414:GATA2</td>
</tr>
<tr>
<td>RNASE4</td>
<td>GATA2</td>
<td>M_414:GATA2</td>
</tr>
<tr>
<td>CD40LG</td>
<td>GATA2</td>
<td>M_414:GATA2</td>
</tr>
<tr>
<td>MAPK1</td>
<td>MIR217</td>
<td>M_244:MIR606HG, MIR217, MIR16-1</td>
</tr>
</tbody>
</table>

- **One Level of Overlap (gene + regulator):** 13 genes (34%)

<table>
<thead>
<tr>
<th>Gene</th>
<th>Regulator</th>
<th>Module</th>
</tr>
</thead>
<tbody>
<tr>
<td>ESR1</td>
<td>FOX1</td>
<td>M_695:FOX1</td>
</tr>
<tr>
<td>MAP3K4</td>
<td>GATA2</td>
<td>M_414:GATA2</td>
</tr>
<tr>
<td>IK3B</td>
<td>ZNF354C</td>
<td>M_244:ZNF56</td>
</tr>
<tr>
<td>CD44</td>
<td>ELF1</td>
<td>M_310:ELF3</td>
</tr>
<tr>
<td>CREM</td>
<td>ELF1</td>
<td>M_310:ELF3</td>
</tr>
<tr>
<td>CFH</td>
<td>-</td>
<td>M_75:CFHR2</td>
</tr>
<tr>
<td>IFN21</td>
<td>-</td>
<td>M_75:IFNA16</td>
</tr>
<tr>
<td>FKBP4</td>
<td>-</td>
<td>M_402:FKBP1</td>
</tr>
<tr>
<td>TNFRSF7F</td>
<td>-</td>
<td>M_244:TNFRSF1</td>
</tr>
<tr>
<td>LRR23</td>
<td>-</td>
<td>M_144:LRR375</td>
</tr>
<tr>
<td>F2R</td>
<td>-</td>
<td>M_417:F2RL3</td>
</tr>
<tr>
<td>WDR20</td>
<td>-</td>
<td>M_75:WDR88</td>
</tr>
<tr>
<td>CSF11R</td>
<td>-</td>
<td>M_633:CSF2</td>
</tr>
</tbody>
</table>

**Recurring modules**

- **M_577:** Blood vessel development, aneurysm
- **M_75**
- **M_414**
- **M_244**
- **M_144:** Microbe-based processes
Summary

- **400+ WGS** for identifying the genetic architecture for AAA
- **38 genes** identified by analyzing less-frequent, functional variants combined with curated databases
- **60 genes and 40 interaction modules** identified by analyzing rare variants and PPI network
- Significant **overlap and complementarity** in these two analyses
- More **global analysis** incorporating all variants, including structural variants and copy number variations, are ongoing.
ACKNOWLEDGEMENTS

● Collaborators
Stanford University:
   Michael Snyder
   Wing H Wong
Ingenuity:
   Nathan Pearson
Google Genomics:
   Nicole Deflaux
   Jonathan Bingham

● Analysis Team
   Cuiping Pan
   Jingjing Li
   Sai Zhang
   Xin Ma
   Joshua Spin

● Sample prep
   Alicia Deng

● IT Support
   Somalee Datta
   Keith Bettinger

● Funding:
   National Institute of Health
   US Department of VA