Genome-Wide Association Study of Drug Addiction* in Caucasians and African Americans

*and some other interesting traits

Rick Sherva, PhD
Assistant professor
Biomedical Genetics
Genetics of Cocaine Dependence in Two Populations

- RC2-DA028909-01
- Continuation of long running study on genetics of addiction
- Earlier phases used linkage/candidate gene approaches
- Despite finding linkage peaks, few genes or variants were identified for CD or OD
- Recently obtained genome-wide SNP data
GWAS: Opportunities and Challenges

• GWAS data allows genetic QC
  – Familial relationships (IBD sharing)
  – Gender (X-chromosome heterozygosity)
  – Race (principal components analysis)

• There were issues...
Familial relationships, etc.

- Drug addicts can use $100
  - Will say someone is their relative who isn’t
  - Will enroll at multiple recruitment sites
  - Will say a different person is their relative at site B
  - Will hurry through a long, boring questionnaire

- Also, scientists and lab techs make errors
Very high retention rate

- 509 pairs with IBD issues
- 132 unresolved IBD issues
- 26 unresolved gender mismatches
- 960 rescued
Ancestry
Primary phenotypes

• SSADDA
  – Modified version of the SSAGA
  – DSM-IV drug dependence (cocaine, opioids, alcohol, tobacco, stimulants, marijuana)
  – Psychiatric diseases (depression, mania, PTSD, anxiety disorder, phobias)
  – Electronic interviewer form with built in consistency checks/skip patterns
Other traits

• OD subtype scores
  – PCA–based data reduction, high heritability
  – Type 4: Heavy use
  – Type 5: Early-onset, co-morbid, heavy user group

• FTND

• NEO (extraversion, agreeableness, openness, neuroticism, conscientiousness)

• Sexual behavior (lifetime partner count)

• Other (handedness, eye color)
The Numbers

• Discovery set
  – 3318 AA (442 families, 1592 women)
  – 2379 EA (331 families, 982 women)
  – Mean age ~39, ~55% male

• Two replication sets
  – SAGE (1311 AA and 2752 EA)
  – Internal (761 AA and 1679 EA)
AA

• Cocaine dependence
  – 2482 cases, 186 exposed controls
• Opioid dependence
  – 691 cases, 640 exposed controls
• Alcohol dependence
  – 1739 cases, 1101 exposed controls
• Tobacco dependence
  – 1909 cases, 1101 exposed controls
• 498 “clean” controls
• Cocaine dependence
  – 1806 cases, 293 exposed controls
• Opioid dependence
  – 1435 cases, 299 exposed controls
• Alcohol dependence
  – 1421 cases, 475 exposed controls
• Tobacco dependence
  – 1838 cases, 449 exposed controls
• 63 “clean” controls
High rates of comorbid dependence

• 1300 with two dependencies
• 1280 with three
• 460 dependent on alcohol, tobacco, cocaine and opiates
• Nicotine and cocaine most common combination \((r = .36)\)
• Only alcohol and opiates uncorrelated
Special considerations for addiction statistical genetics

• Comorbidity (co-linearity vs. independence)

• Exposure (environment)

• Free will
“I’m starting to really like the smell of cocaine.”
Model selection

• DSM-IV symptom count
  – SNP = OD + CD + AD + TD + Age + Sex + PCs
  – Maximize information content
  – Assure findings are substance-specific

• Case-control with exposed controls
  – AFF = SNP + Age + Sex + PCs
  – Limit bias due to misclassifying genetically susceptible but non-exposed individuals

• Linear and logistic regression models solved using GEE
Nicotine: a small mystery
...but see a signal in AAs
Alcohol: a positive control

ADH4, ADH1B

ADH1B, ADH1C
Cocaine symptom count

[Graph showing the distribution of symptom counts across different chromosomes.]

Chromosome
Cocaine case-control
Opioid symptom count

KCNC1
Opioid case control

PITPNM3

Chromosome
Phase I replication: SAGE

• The Study of Addiction: Genetics and Environment
• AAs and EAs
  – COGA (Alcohol)
  – FSCD (Cocaine)
  – COGEND (Nicotine)
• Has related and unrelated participants (mostly unrelated)
• Mostly mid-western
## Cocaine signals

<table>
<thead>
<tr>
<th>Model</th>
<th>Gene</th>
<th>Population</th>
<th>Top-ranked p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case-Control</td>
<td>POTEH</td>
<td>all</td>
<td>3.09E-08</td>
</tr>
<tr>
<td>Case-Control</td>
<td>ATP2B2*</td>
<td>AA</td>
<td>1.22E-07</td>
</tr>
<tr>
<td>Case-Control</td>
<td>ZSWIM5</td>
<td>all</td>
<td>3.36E-07</td>
</tr>
<tr>
<td>Case-Control</td>
<td>CNTNAP2*</td>
<td>AA</td>
<td>3.51E-07</td>
</tr>
<tr>
<td>Case-Control</td>
<td>C1QL2</td>
<td>EA</td>
<td>4.08E-07</td>
</tr>
<tr>
<td>Symptom Count</td>
<td>GP6</td>
<td>all</td>
<td>6.95E-07</td>
</tr>
<tr>
<td>Symptom Count</td>
<td>CPE*</td>
<td>all</td>
<td>1.30E-06</td>
</tr>
</tbody>
</table>
## Opioid signals

<table>
<thead>
<tr>
<th>Model</th>
<th>Gene</th>
<th>Population</th>
<th>Top-ranked value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Symptom Count</td>
<td>KCNG2*</td>
<td>AA</td>
<td>8.69E-10</td>
</tr>
<tr>
<td>Symptom Count</td>
<td>APBB2</td>
<td>AA</td>
<td>1.82E-09</td>
</tr>
<tr>
<td>Symptom Count</td>
<td>KCNC1*</td>
<td>AA</td>
<td>7.47E-09</td>
</tr>
<tr>
<td>Symptom Count</td>
<td>PARVA</td>
<td>AA</td>
<td>1.75E-08</td>
</tr>
<tr>
<td>Case-Control</td>
<td>DPP6*</td>
<td>AA</td>
<td>2.33E-07</td>
</tr>
<tr>
<td>Case-Control</td>
<td>ANO10</td>
<td>AA</td>
<td>1.53E-06</td>
</tr>
</tbody>
</table>
Phase II Replication: Internal

• 2000+ previously un-genotyped individuals

• Unrelated, EA-heavy

• Panel of 1500 SNPs selected across ~15 traits on a custom chip

• Had identical phenotype information as the discovery
SNPs that replicated
Pathway Analysis

• Identify potentially important risk variants with sub genome wide significance

• Uses a network of interacting genes as the basis for a statistical test

• Relies on a database of known gene x gene interactions, biological processes, regulatory networks, etc.
Ingenuity Pathway Analysis

• Calculate number of independent tests within every gene

• Identify genes with a corrected SNP P-value < 10E-3 (EA N= 221, AA N =254)

• Identify networks and pathways with an overrepresentation of these genes
CD network in AA

Linked to receptors for NMDA, dopamine, and GABA
CD network in EA

Hippocampal morphology

M-opioid receptor
Dopamine, ions and addiction

Dopamine burst firing requires glutamatergic input, NMDA receptor activation, opening of high-threshold calcium currents, and activation of calcium-activated potassium currents to terminate the burst. Changes in synaptic inputs, calcium currents, or calcium-activated potassium currents could alter burst patterns of firing in dopamine neurons.

Remaining Work

• Still waiting for replication genotyping on several of the top SNPs

• Next gen sequencing in DISC1
  – Different variants in fams and unrelated?

• Gene x gene interaction tests on Ca/K genes

• Integrating personality SNPs
Conclusions

• No robust, strong effect “Sheen gene” variants for OD or CD

• The inherited susceptibility to OD and CD is more likely in pathways related to learning and memory, synaptic vesicle biochemistry and ion homeostasis in neurons
Now for something completely different...

Lifetime Sexual Partners

N: 5073
Mean: 30.18
## i35a signals

<table>
<thead>
<tr>
<th>Type</th>
<th>Gene</th>
<th>P All</th>
<th>Dir</th>
<th>AA P</th>
<th>EA P</th>
<th>Male P</th>
<th>Female P</th>
</tr>
</thead>
<tbody>
<tr>
<td>intron</td>
<td>PADI4</td>
<td>1.14E-03</td>
<td>--------+</td>
<td>8.08E-02</td>
<td>4.08E-03</td>
<td>4.13E-07</td>
<td>5.52E-01</td>
</tr>
<tr>
<td>intergenic</td>
<td>KLF3/TLR10</td>
<td>1.47E-04</td>
<td>--------</td>
<td>4.12E-05</td>
<td>2.18E-01</td>
<td>9.49E-01</td>
<td>5.50E-08</td>
</tr>
<tr>
<td>30K 5'</td>
<td>NDUFS6</td>
<td>5.37E-06</td>
<td>++++</td>
<td>2.28E-05</td>
<td>3.09E-02</td>
<td>4.09E-04</td>
<td>3.91E-03</td>
</tr>
<tr>
<td>intron</td>
<td>SCIN</td>
<td>7.94E-05</td>
<td>+++++++</td>
<td>2.92E-01</td>
<td>4.21E-06</td>
<td>1.01E-02</td>
<td>2.52E-03</td>
</tr>
<tr>
<td>8K 3'</td>
<td>LOC100130301</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>9.92E-08</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>9.47E-09</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>intron</td>
<td>CRTAC1</td>
<td>NA</td>
<td>NA</td>
<td>2.11E-06</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>NA</td>
<td>NA</td>
<td>2.36E-08</td>
<td>++++?++</td>
<td>3.05E-08</td>
<td>NA</td>
<td>NA</td>
<td>1.85E-05</td>
</tr>
<tr>
<td>intron</td>
<td>DNAJB13</td>
<td>5.93E-05</td>
<td>---</td>
<td>7.36E-05</td>
<td>1.47E-01</td>
<td>9.86E-01</td>
<td>1.88E-08</td>
</tr>
<tr>
<td>22K 5'</td>
<td>CCND2</td>
<td>NA</td>
<td>NA</td>
<td>3.29E-11</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>1K 5'</td>
<td>CCDC62</td>
<td>1.34E-04</td>
<td>++------</td>
<td>5.81E-02</td>
<td>4.17E-04</td>
<td>4.81E-07</td>
<td>7.99E-01</td>
</tr>
<tr>
<td>intron</td>
<td>CCDC62</td>
<td>1.37E-04</td>
<td>++++++++</td>
<td>4.57E-02</td>
<td>6.14E-04</td>
<td>9.34E-06</td>
<td>3.82E-01</td>
</tr>
<tr>
<td>intron</td>
<td>LTBP2</td>
<td>3.17E-09</td>
<td>--------</td>
<td>3.751E-07</td>
<td>1.07E-03</td>
<td>1.41E-06</td>
<td>4.27E-04</td>
</tr>
<tr>
<td>NA</td>
<td>NA</td>
<td>1.14E-03</td>
<td>++++++</td>
<td>2.66E-01</td>
<td>4.26E-04</td>
<td>4.56E-01</td>
<td>3.86E-08</td>
</tr>
</tbody>
</table>
Collaborators

Lindsay Farrer
Ryan Koesterer
John Farrell
Scott Melville
Alexan Mardigian

Thanks to the Genome Science Institute for the IPA license