Network Medicine Approaches to COPD

Edwin K. Silverman, M.D., Ph.D.

Channing Division of Network Medicine
Pulmonary and Critical Care Medicine
Brigham and Women’s Hospital
Boston, Massachusetts
Edwin K. Silverman: Conflicts of Interest

1) Personal financial relationships with commercial interests relevant to medicine, within past 3 years:
   • Consulting: GlaxoSmithKline, Merck
   • Lecture Fees (Honoraria): Merck, Novartis

2) Personal financial support from a non-commercial source relevant to medicine, within past 3 years:

   No relationships to disclose

3) Personal relationships with tobacco industry entities:

   No relationships to disclose
Waves of Discovery in Complex Disease Genetics

- Linkage Analysis
- Candidate Gene Association
- Genome-wide Association
- Exome/Genome Sequencing
- Systems Genetics/Networks
What Is a Network?

A collection of points (nodes) that are joined in pairs by lines (edges). A graphical approach to visualize and analyze relationships between variables of interest.

(Adapted from M. Newman, Networks: An Introduction, 2010)
What Is Network Medicine?

The study of cellular, disease, and social networks which aims to quantify the complex interlinked factors contributing to individual diseases.

(Adapted from Barabasi, NEJM 2007; 357:404)

Key components of Network Medicine:
--Holistic rather than reductionist approach
--Construction of molecular disease networks
--Non-linear responses of complex systems
--Emergent properties from entire network
--Investigates responses of networks to various types of perturbation
--Employs systems biology methods
High Throughput Assessment of Multiple Biological Processes

From Network Medicine: Complex Systems in Human Disease and Therapeutics, edited by Loscalzo/Barabasi/Silverman
Types of Networks Utilized in Network Medicine

**Correlation Network**
- **Nodes:** Omics Data for a Gene
- **Edges:** Correlation between Omics Data

**Gene Regulatory Network**
- **Nodes:** Transcription Factors (Circles) and Genes (Squares)
- **Edges:** Gene Regulatory Relationship

**Protein-Protein Interaction Network**
- **Nodes:** Protein
- **Edges:** Physical Interactions
Approaches to Complex Diseases in Channing Division of Network Medicine

- Defining Molecular Pathways
- Identifying Optimal Disease Phenotypes
- Integrating Multiple -omics Data Types

Building Disease Networks

Developing and Validating New Disease Classifications

Developing New Treatments and Preventions
Possible Functional Impacts of Genetic Variants on the Protein-Protein Interaction Network

“Normal”

Loss/Gain of interaction

Alter interaction strength

None

Loss of many interactions
Approach for Genome-wide Association Studies in COPD (Hardin, J COPDF 2014)

1) Subject enrollment → 2) Genotyping → 3) Quality Control

4) Association analysis → 5) Meta-analysis

International COPD Genetics Consortium
COPD GWAS (Hobbs/Cho, Nat Genet 2017)

• Included discovery in 15,256 COPD cases and 47,936 controls from 26 studies with genotyping of select top results ($P < 5 \times 10^{-6}$) in 9,498 COPD cases and 9,748 controls from UK-BiLEVE.
## Replication of COPD GWAS Regions

<table>
<thead>
<tr>
<th>Locus</th>
<th>First Author (Year)</th>
<th>SNP</th>
<th>Odds Ratio</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>HHIP</td>
<td>Van Durme (2010)</td>
<td>rs13118928</td>
<td>0.80</td>
<td>2 x 10^{-4}</td>
</tr>
<tr>
<td>HHIP</td>
<td>Zhou (2012)</td>
<td>rs13118928</td>
<td>0.68</td>
<td>2 x 10^{-3}</td>
</tr>
<tr>
<td>HHIP</td>
<td>Wain (2015)</td>
<td>rs1032296</td>
<td>0.88</td>
<td>7 x 10^{-7}</td>
</tr>
<tr>
<td>FAM13A</td>
<td>Young (2011)</td>
<td>rs7671167</td>
<td>0.79</td>
<td>0.01</td>
</tr>
<tr>
<td>FAM13A</td>
<td>Xie (2015)</td>
<td>rs7671167</td>
<td>0.53 (CC genotype)</td>
<td>8 x 10^{-8} (CC)</td>
</tr>
<tr>
<td>CHRNA3/5</td>
<td>Wilk (2012)</td>
<td>rs1051730</td>
<td>1.17</td>
<td>3 x 10^{-7}</td>
</tr>
<tr>
<td>CHRNA3/5</td>
<td>Hardin (2012)</td>
<td>rs8034191</td>
<td>1.89</td>
<td>7 x 10^{-7}</td>
</tr>
<tr>
<td>IREB2</td>
<td>Chappell (2011)</td>
<td>rs2568494</td>
<td>1.30</td>
<td>5 x 10^{-4}</td>
</tr>
<tr>
<td>MMP1/MMP12</td>
<td>Hunninghake (2012)</td>
<td>rs2276109</td>
<td>N/A</td>
<td>4 x 10^{-5}</td>
</tr>
<tr>
<td></td>
<td>Arja (2014)</td>
<td>rs2276109</td>
<td>0.48</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>Jackson (2016)</td>
<td>rs17368582</td>
<td>0.71</td>
<td>5 x 10^{-6}</td>
</tr>
</tbody>
</table>
GWAS: Strengths and Weaknesses

- **Strengths**
  - Multiple genome-wide significant results found in many complex diseases
  - GWAS associations have often been replicated by multiple studies
  - Genotyping and Analysis approaches are well-established

- **Weaknesses**
  - Functional variants identified in a small minority of loci
  - Odds ratios for identified GWAS loci are low
  - GWAS loci (at least in isolation) are not very useful for prediction
  - Much of the estimated heritability remains unexplained
### Functional Genetics of COPD Lung Tissue Population (Morrow/Hersh)

<table>
<thead>
<tr>
<th></th>
<th>COPD cases N=111</th>
<th>Control smokers N=40</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>63.3 (± 6.6)</td>
<td>65.7 (± 9.0)</td>
<td>n.s.</td>
</tr>
<tr>
<td>Female sex</td>
<td>59 (53.2%)</td>
<td>25 (62.5%)</td>
<td>n.s.</td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>African American</td>
<td>18 (16.2%)</td>
<td>5 (12.5%)</td>
<td>n.s.</td>
</tr>
<tr>
<td>White</td>
<td>90 (81.1%)</td>
<td>34 (85.0%)</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>3 (2.7%)</td>
<td>1 (2.5%)</td>
<td></td>
</tr>
<tr>
<td>Pack-years of smoking</td>
<td>61.3 (± 26.3)</td>
<td>33.6 (± 21.0)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>FEV₁ % predicted</td>
<td>26.5 (± 9.4)</td>
<td>98.7 (± 12.5)</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>FEV₁/FVC</td>
<td>0.32 (± 0.10)</td>
<td>0.79 (± 0.05)</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>
No Differential Expression of COPD GWAS Genes in Lung Tissue *(Morrow/Hersh)*

<table>
<thead>
<tr>
<th>Rank</th>
<th>Gene</th>
<th>P Value</th>
<th>Adjusted P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1680</td>
<td>MMP12</td>
<td>0.02</td>
<td>0.30</td>
</tr>
<tr>
<td>3277</td>
<td>HHIP</td>
<td>0.04</td>
<td>0.42</td>
</tr>
<tr>
<td>6372</td>
<td>AGER</td>
<td>0.11</td>
<td>0.59</td>
</tr>
<tr>
<td>6537</td>
<td>IREB2</td>
<td>0.12</td>
<td>0.60</td>
</tr>
<tr>
<td>8216</td>
<td>DLC1</td>
<td>0.16</td>
<td>0.66</td>
</tr>
<tr>
<td>8688</td>
<td>CHRNA5</td>
<td>0.18</td>
<td>0.67</td>
</tr>
<tr>
<td>13798</td>
<td>FAM13A</td>
<td>0.34</td>
<td>0.80</td>
</tr>
<tr>
<td>19429</td>
<td>CHRNA3</td>
<td>0.53</td>
<td>0.89</td>
</tr>
<tr>
<td>22640</td>
<td>RIN3</td>
<td>0.64</td>
<td>0.93</td>
</tr>
<tr>
<td>29796</td>
<td>TGFB2</td>
<td>0.89</td>
<td>0.98</td>
</tr>
</tbody>
</table>

Note: Top differentially expressed gene was HMGB1, a known interactor with AGER
Are Interactors with COPD GWAS Genes More Likely To Be Differentially Expressed in Lung Tissue? (Morrow/Hersh)

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Number of genes or proteins</th>
<th>Enrichment p-value for FEV(_1) Genes</th>
<th>Enrichment p-value for COPD Genes</th>
</tr>
</thead>
<tbody>
<tr>
<td>IREB2 RNA immunoprecipitation seq</td>
<td>4008</td>
<td>3.6 x 10(^{-11})</td>
<td>2.4 x 10(^{-5})</td>
</tr>
<tr>
<td>IREB2 trans-eQTLs at p&lt;0.05</td>
<td>1612</td>
<td>1.4 x 10(^{-10})</td>
<td>0.0002</td>
</tr>
<tr>
<td>HHIP Tandem affinity purification</td>
<td>216</td>
<td>1.8 x 10(^{-6})</td>
<td>0.06</td>
</tr>
<tr>
<td>HHIP trans-eQTLs at p&lt;0.05</td>
<td>1560</td>
<td>2.4 x 10(^{-5})</td>
<td>0.09</td>
</tr>
<tr>
<td>Hhip(^{+/−}) mouse 6 mo. smoking experiment, genotype x smoke interaction</td>
<td>492</td>
<td>2.7 x 10(^{-6})</td>
<td>0.0001</td>
</tr>
<tr>
<td>FAM13A trans-eQTL at p&lt;0.05</td>
<td>1753</td>
<td>&lt;1.0 x 10(^{-12})</td>
<td>2.7 x 10(^{-6})</td>
</tr>
<tr>
<td>FAM13A Tandem affinity purification</td>
<td>97</td>
<td>0.3</td>
<td>1</td>
</tr>
</tbody>
</table>
Weighted Gene Co-Expression Network Analysis in COPD vs. Control Lung Tissue
(Morrow/Hersh)

• Methods:
  – WGCNA method analyzes correlation in gene expression; retains scale-free
degree distribution of coexpression network
  – WGCNA produces a set of modules (labeled by color), each containing a set of
unique genes
Cyan Module from WGCNA Analysis of Lung Tissue Population (Morrow/Hersh)

Note: GO Analysis highlighted B cell proliferation and signaling
Disease Modules, Disease Neighborhoods, and the Interactome
Dense module searching approach for GWAS in protein-protein interaction networks

Searches for subnetworks within the protein-protein interaction network containing multiple genes with low GWAS p-values

Single most significant SNP represents the gene, but later modified to enable gene-based p-values

Uses a seed gene-based approach to build a disease subnetwork

Determines if there is an excess of low p-values in a subnetwork using permutation testing

Uses all of the GWAS data, not just the top SNPs
Beyond GWAS in COPD: Probing the Landscape between Gene-Set Associations, Genome-Wide Associations and Protein-Protein Interaction Networks (McDonald, Hum Hered 2014; 78: 131)

- dmGWAS with COPDGene as Discovery and GenKOLS as Replication Population
- Found Ubiquitin C, which interacts with 7,232 proteins in PPI network
- Gene set enrichment of 10 shared genes in the consensus modules +/- UBC implicated IL7 signaling
Mapping of COPD Seed Genes in the Interactome
(A. Sharma)

COPD seed genes (n=11)

Mendelian Syndromes: SERPINA1, ELN, FBLN5

COPD GWAS: HHIP, FAM13A, IREB2, CHRNA3, CHRNA5, RIN3, TGFB2, MMP12

Network proximity calculations based on Random Walk–DADA method
COPD Disease Network Module within the Protein-Protein Interaction Network (A. Sharma)
COPD Disease Network Module: Expanded with FAM13A Interactors (A. Sharma)
COPD Disease Network Module: Differential Expression of Module Members (A. Sharma)
Network-based stratification of tumor mutations

Matan Hofree¹, John P Shen², Hannah Carter², Andrew Gross³ & Trey Ideker¹–³
Applying Network-Based Stratification to COPD (Chang/Castaldi, Genomics 2016)

• Rationale
  – Gene expression array data is noisy
  – External information about groups of genes expected to act together could guide solutions
  – NBS merges data-driven pattern finding with protein interaction network data to arrive at more “biologically” informed subtypes

• Methods:
  – Peripheral blood gene expression microarrays
  – Primary analysis in ECLIPSE (n=229) with replication in COPDGene (n=135) COPD cases and smoking controls
  – Feature selection – 1,812 genes with expression levels associated with $\text{FEV}_1$, $\text{FEV}_1/\text{FVC}$, or CT emphysema
## NBS Clinical Replication in COPDGene
*(Chang/Castaldi, Genomics 2016)*

### ECLIPSE

<table>
<thead>
<tr>
<th></th>
<th>C1</th>
<th>C2</th>
<th>C3</th>
<th>C4</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>29</td>
<td>89</td>
<td>93</td>
<td>18</td>
</tr>
<tr>
<td>FEV&lt;sub&gt;1&lt;/sub&gt;</td>
<td>49</td>
<td>67</td>
<td>82</td>
<td>87</td>
</tr>
<tr>
<td>FEV&lt;sub&gt;1&lt;/sub&gt;/FVC</td>
<td>0.43</td>
<td>0.53</td>
<td>0.63</td>
<td>0.69</td>
</tr>
<tr>
<td>Emphysema</td>
<td>20</td>
<td>13</td>
<td>6</td>
<td>6</td>
</tr>
</tbody>
</table>

### COPDGene

<table>
<thead>
<tr>
<th></th>
<th>C1</th>
<th>C2</th>
<th>C3</th>
<th>C4</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>19</td>
<td>50</td>
<td>61</td>
<td>5</td>
</tr>
<tr>
<td>FEV&lt;sub&gt;1&lt;/sub&gt;</td>
<td>49</td>
<td>67</td>
<td>72</td>
<td>87</td>
</tr>
<tr>
<td>FEV&lt;sub&gt;1&lt;/sub&gt;/FVC</td>
<td>0.48</td>
<td>0.59</td>
<td>0.63</td>
<td>0.70</td>
</tr>
<tr>
<td>Emphysema</td>
<td>10</td>
<td>5</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

*Note: Clustering without network constraint was much less reproducible*
Evolution of Complex Disease Genetic Studies
(Silverman/Loscalzo, Discovery Med 2012; 14: 143)

First Generation Genetic Studies
- Genetic Variants
- Disease
- Low Power To Detect Associations

Second Generation Genetic Studies
- Genetic Variants
- Single -Omics Data Type
- ??
- Disease

Third Generation Genetic Studies
- Genetic Variants
- Proteomics
- Transcriptomics
- Metabolomics
- Disease Subtypes
Approaches for Genetic Network Analysis

- Assess Genetic Variants, -omics, and Disease Phenotypes
- Genetic Association of SNPs to -omics Data
- Identify Network of SNPs and Related Analytes in Single -omics Data Type
- Interpret Association Results Using Interactome Network
- Functional Studies to Build Molecular Networks (e.g., RNAi)
- Build Genetic Networks from Association Results (e.g., dmGWAS)
- Build Epistasis Networks (e.g., ViSEN)
- Analyze SNPs, Disease Phenotypes, and Multiple -omics Data Types to Build Disease Networks
- Genetic Association of SNPs to Disease
- Genetic Association of SNPs to -omics Data

Interpret Association Results Using Interactome Network
COPD as a Model of Complex Disease

- Gene Discovery
  - GWAS
  - WGS
  - Epigenetics

- COPD Genetic Architecture
  - Systems/Networks
  - Integrative Omics
  - Machine Learning
  - Natural History

- Phenotype
  - Subtyping
  - Imaging

- Function
  - Cells
  - Mice
  - Functional Variants

- Networks
Collaborators

- **Transcontinental COPD Genetics Study (Korea):** Young Soo Shim, JJ Yim, Woo Jin Kim, DK Kim, Sei Won Lee, Myung Goo Lee
- **International COPD Genetics Consortium (Korea):** JJ Yim, Woo Jin Kim, DK Kim, Yeon Mok Oh, Mi Kyeong Lee, Sang Do Lee
- **ECLIPSE Genetics Study:** Michael Cho, DK Kim, Wayne Anderson, Sreekumar Pillai, Xiangyang Kong, David Lomas, ECLIPSE Steering/Scientific Committees
- **Norway Case-Control Study:** Per Bakke, Amund Gulsvik, Sreekumar Pillai, Craig Hersh, Dawn DeMeo, Michael Cho
- **Functional Genetics of COPD:** Xiaobo Zhou, Augustine Choi, Suzanne Cloonan, Dawn DeMeo, Craig Hersh, Jarrett Morrow, Jeanine D’Armiento, John Quackenbush, Kimberly Glass, John Platig, Amitabh Sharma, Yang-Yu Liu, Caroline Owen, Mark Perrella, Bart Celli, Miguel Divo, Zhiqiang Jiang, Taotao Lao, Raphael Bueno, Gerard Criner
- **COPDGene:** James Crapo, Barry Make, John Hokanson, Doug Everett, Terri Beaty, Michael Cho, Peter Castaldi, David Lynch, George Washko, Raul San Jose Estepar, James Ross, Merry-Lynn McDonald, Craig Hersh, Dawn DeMeo, Emily Wan, Brian Hobbs, Robert Busch, Lystra Hayden, Adel El-Boueiz, Megan Hardin, Jin Hwa Lee, Sung Ho Won, and 21 Clinical Centers
- **Funding:** NIH R01 HL089856 and R01 HL089897 (COPDGene), R01 HL111759 (COPD Networks), P01 HL105339 (COPD Functional Genetics PPG), P01 HL114501 (IPF/COPD PPG), R01 HL086936 (D’Armiento), and GlaxoSmithKline (ECLIPSE)