Gene-expression Profiles as Signatures (of Response) to Environmental Exposures

Smoking and the airway “field of injury” as a novel paradigm for the exposome

Avrum Spira, M.D., M.Sc.
Section of Computational Biomedicine
The Pulmonary Center
Boston University Medical Center
Disclosure

- Founder and consultant to Allegro Diagnostics Inc.
The Airway “Field of Injury” Hypothesis

Smoking (and other inhaled exposures) alters epithelial cell gene expression throughout the respiratory tract.

Variability in epithelial cell genomic response to and damage from smoking linked to tobacco-associated lung disease.
The bronchial airway transcriptome in smoking and lung cancer

Smoking impacts airway gene expression
- PNAS 2004; NAR 2005;

Subset of changes are irreversible upon cessation and can serve as biomarker of past exposure
- Genome Biology 2007

Airway gene expression can serve as an early diagnostic biomarker for lung cancer
- Nature Medicine 2007; Cancer Prevention Research 2008

RNA- U133A Affymetrix array (~22,500 genes)
The reversible and irreversible impact of smoking on bronchial airway gene-expression

Biomarker of past exposure
Airway gene-expression as a diagnostic biomarker for lung cancer

Validation study of 80 gene biomarker on large independent multicenter cohort by Allegro Diagnostics Inc.

- 800 subjects recruited at 14 centers in US, Canada and Europe
- IDE filed with FDA

Cytology obtained has low sensitivity especially for early stage disease. This results in clinical dilemma as to who should go for surgical resection.

80 gene biomarker that can distinguish smokers with and without lung cancer
- Sensitivity 80%; specificity 84% in two independent cohorts
- 90% sensitive for stage 1 disease
- 95% sensitivity and 95% NPV when combined with cytology obtained at bronchoscopy
- Independent of clinical and radiographic predictors of disease

Nature Medicine 2007; CAPR 2008
Extending the field to the upper airway
Mouth Scraper

FIG. 1

Biotechniques 2004
Technique for Obtaining RNA from Nasal Mucosal Brushings

- 1% Lidocaine topically to inferior turbinate
- Cytosoft® brushings from “interior” of inferior turbinate
- Immerse in RNA later
The nose-mouth-airway study as part of NIEHS Genes and Environment Health Initiative

Collect bronchial, buccal, and nasal epithelial cells from current (n=13) and never (n=14) smokers

Large (>200nts) and small RNAs isolated from same samples

Large RNA run on all-exon arrays (1.4 million exons)

Small RNAs run on miRNA microarrays (476 microRNAs)

Zhang et al. Physiological Genomics 2010
Genes that change in common across respiratory tract in smokers

Gene expression = smoking status + tissue type + smoking*tissue + patient(random effect)
The nose-bronch relationship
Upper airway biomarkers developed as part of the GEI

- Nasal and buccal gene-expression as biomarker of past exposure

- Nasal biomarker of second-hand smoke exposure

- Buccal biomarker of cumulative exposure (pack-yrs) to tobacco smoke

- Nasal gene-expression signature post-smoking cessation
Why develop a biomarker of secondhand smoke exposure

• 20-30% increased risk of heart disease and lung cancer

• Significant morbidity in children exposed
  - lung and ear infections
  - asthma attacks (number and severity)
Moving nasal biomarkers of secondhand exposure to children

- Columbia Center for Children’s Environmental Health (CCCEH) cohort

- Disease Investigation Through Specialized Clinically-Oriented Ventures in Environmental Research (DISCOVER) cohort
Developing a gene-expression signature of smoking cessation

- Kinetics of airway gene-expression changes post-cessation within an individual
- Baseline signature to compare with PREP
Leverage existing clinical trial sponsored by AllegroDx Inc as part of EDRN.

AllegroDx Inc is validating diagnostic gene-expression signature for lung cancer

-IDE filed with FDA

600 current and former smokers undergoing bronchoscopy for suspect lung cancer at 14 medical center in US, Canada, and Europe

miRNA/mRNA from nose isolated at CLIA facility and shipped to BU

mRNA arrays and miRNA-seq mRNA signature validation
Extending the airway “field of injury” to microRNA

- Small, non-coding RNAs ~22-25 nts
- Bind to 3’-UTR of mRNA
- Downregulates gene through several mechanisms
- Implicated in development, cell stress response, cancer.

A MORE ROBUST BIOMARKER ??
- relatively resistant to degradation
- microRNA changes upstream to many mRNA

Impact of smoking on bronchial airway microRNA expression

MicroRNA regulate the gene-expression response to smoking

Schembri et al PNAS. 2009
# Microarray vs. RNA-seq

## Microarrays
- Limited by prior knowledge on what is expressed
- Analog output with limited dynamic range
- Cross hybridization
- Cost $$
- Throughput moderate
- Computation ++

## RNA-seq
- Pure discovery**
- Digital output with large dynamic range
- Alt splicing
- SNP in exons
- Cost $$$$$$4
- Throughput low
- Computation ++++4
Deep sequencing the airway transcriptome

**Large RNA** sequenced on Illumina Solexa system
- 36 base pair reads
- 30 million reads per sample

**Small RNA (15-40 bp)** sequenced on ABI SOLID platform
- 30 base pair sequence
- 200 million reads per run (4 samples multiplexed in single run)

4 pools: Never smoker, current smoker, smoker with lung cancer, smoker with benign lung disease
mRNA-seq identifies novel smoking- and cancer-related gene expression changes in the airway

Beane et al. 2011
Summary

• Airway epithelial gene expression reflects host response to and damage from tobacco smoke
  – Heterogeneity in genomic response may serve as early diagnostic biomarker for lung cancer (and COPD)

• This “field of injury” extends to the nose and mouth epithelium, although the nose appears to be a better surrogate to the bronchus

• Upper airway gene expression can serve as biomarker of tobacco smoke exposure and potentially a non-invasive lung cancer marker
  – Ever exposure
  – Second hand exposure
  – Post-smoking cessation

• RNA-seq may serve as the most comprehensive discovery tool for transcriptomic biomarkers of exposure
## Acknowledgements

<table>
<thead>
<tr>
<th>Boston University</th>
<th>Univ of Utah: Andrea Bild</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marc Lenburg</td>
<td>NCI: Eva Szabo, Nat Rothman, Qing Lan</td>
</tr>
<tr>
<td>Jerome Brody</td>
<td>Vanderbilt University: Pierre Massion</td>
</tr>
<tr>
<td>Joshua Campbell</td>
<td>Hopkins: Pat Breysse, Greg Diette</td>
</tr>
<tr>
<td>Gang Liu</td>
<td>Columbia University: Steve Chillrud, Rachel Miller</td>
</tr>
<tr>
<td>Sherry Zhang</td>
<td>UCLA: Steve Dubinett, Brigitte Gomperts</td>
</tr>
<tr>
<td>Ji Zhang</td>
<td>LRRI: Steve Belinsky</td>
</tr>
<tr>
<td>Joe Guerrein</td>
<td>NIEHS: Dan Shaughnessy, David Balshaw</td>
</tr>
<tr>
<td>Adam Gower</td>
<td></td>
</tr>
<tr>
<td>Christina Anderlind</td>
<td></td>
</tr>
<tr>
<td>Catalina Perdomo</td>
<td></td>
</tr>
</tbody>
</table>