How Does Lung Cancer Develop in Never-Smokers?
Looking for Answers in Genomic Sequencing

Alice Berger, Ph.D.
Assistant Professor
Innovators Network Endowed Chair
Fred Hutchinson Cancer Research Center
Seattle, WA

I have no relevant financial relationships in the last 24 months to disclose.
Population: 30,995
Lung cancer is the leading cause of cancer deaths in men and women

<table>
<thead>
<tr>
<th>Estimated Deaths</th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lung &amp; bronchus</td>
<td>72,500</td>
<td>63,220</td>
</tr>
<tr>
<td>Prostate</td>
<td>33,330</td>
<td>42,170</td>
</tr>
<tr>
<td>Colon &amp; rectum</td>
<td>28,630</td>
<td>24,570</td>
</tr>
<tr>
<td>Pancreas</td>
<td>24,640</td>
<td>22,410</td>
</tr>
<tr>
<td>Liver &amp; intrahepatic bile duct</td>
<td>20,020</td>
<td>13,940</td>
</tr>
<tr>
<td>Leukemia</td>
<td>13,420</td>
<td>12,590</td>
</tr>
<tr>
<td>Esophagus</td>
<td>13,100</td>
<td>10,140</td>
</tr>
<tr>
<td>Urinary bladder</td>
<td>13,050</td>
<td>9,680</td>
</tr>
<tr>
<td>Non-Hodgkin lymphoma</td>
<td>11,460</td>
<td>8,480</td>
</tr>
<tr>
<td>Brain &amp; other nervous system</td>
<td>10,190</td>
<td>7,830</td>
</tr>
<tr>
<td>All sites</td>
<td>321,160</td>
<td>285,360</td>
</tr>
</tbody>
</table>

135,700 deaths per year! Much greater than the population of Morgantown, WV!

American Cancer Society Facts & Figures
Risk factors for cancer can be intrinsic (genetic) or extrinsic (environmental) factors.

**Intrinsic:**
- e.g. *BRCA1* mutations

**Extrinsic:**
- e.g. smoking exposure
1986

American Lung Association

2019

GO2 foundation for lung cancer

FACT FRIDAY

Anyone with lungs can get lung cancer.
A growing proportion of lung cancer is in "never-smokers"

Pelosof et al., JNCI 2017
Lung cancer in never-smokers is the 7th leading cause of cancer deaths

Sun et al., Nat Rev Cancer 2007
Women account for the majority of lung cancer cases in never-smokers

Kenfield SA et al., 2008
The mutational landscape of lung cancer is affected by smoking status, gender

Heavy smokers:

Light/Non-smokers:

Men:

Women:

TCGA, Nature 2014
Motivating questions

• Are there new mutations in lung cancers from never-smokers that could be targeted with cancer therapies?

• Can the mutational profile of the cancers themselves, tell us something about the intrinsic or extrinsic genetic risk factors?
  • Define mutational signatures
  • Demonstrate how smoking alters mutational signatures
  • Illustrate how intrinsic risk factors may also play a role

Challenge: The majority of lung cancers sequenced to date are from smokers.
Motivating questions

- Are there new mutations in lung cancers from never-smokers that could be targeted with cancer therapies?

- Can the mutational profile of the cancers themselves, tell us something about the intrinsic or extrinsic genetic risk factors?
  - Define mutational signatures
  - Demonstrate how smoking alters mutational signatures

Challenge: The majority of lung cancers sequenced to date are from smokers.
The Women’s Health Initiative (WHI) lung cancer cohort

Increased cohort size

Post-menopausal women
Understudied lung cancer population

Extensive metadata
History of second-hand smoke exposure, occupational exposure and demographics
Sub-aim 1: Identification of known and novel somatic mutations: The first goal is to identify somatic mutations occurring in this cohort, including single nucleotide variants (SNV), insertion-deletion (Indel), structural variant (SV) and copy number variant (CNV). This would be the first step in establishing the framework of the genomic landscape of this cohort and will identify potential therapeutic targets specific to this cohort.

Samples, histopathology and sequencing: Lung cancer samples from the WHI cohort were reviewed for histological subtype (Figure 2), smoking history and DNA availability. 76 lung cancer samples were identified with confirmed adenocarcinoma histology and sufficient tumor DNA available for sequencing. The final cohort consists of 67 samples light/never smokers and a control group of 9 patients with a history of heavy smoking.

DNA extraction has been completed for all 76 tumor samples and their matched normal/germline samples (peripheral blood). Sequencing of all tumor-normal pairs has also been completed using a customized exome sequencing approach. The bait set utilized for sequencing covers the entire coding region of the human genome (using the Agilent Exome v5) but also regions known to be translocated in cancer. This bait set has been previously validated in the targeted "OncoPanel" sequencing assay performed at the Dana-Farber Cancer Institute Center for Cancer Genome Discovery. Samples were sequenced using the Illumina HiSeq2500 platform with paired-end reads. Sequence alignment has been completed for all samples and quality control metrics evaluated in-depth. The sequencing data passes all established metrics and is of high quality. Normal samples were sequenced to a median fold coverage of 79x and tumor samples to depth of 94x (Figure 3). In collaboration with the Translational Genomic Repository we have established a RedCAP data repository to securely and accurately curate all metadata associated with sequencing data of each sample.

SNV and Indel Analysis: To identify somatic SNV's and Indels consensus mutational calling approach has been developed using two mutational calling algorithms, MuTect2 and Strelka. The benchmarking of this strategy has been completed and high confidence mutation set has been obtained, which includes only those calls that pass both callers. This high confidence mutation data set was interrogated for mutations known to occur recurrently in lung adenocarcinoma. Several known mutations were identified in this cohort including mutations in EGFR, TP53 and KRAS (Figure 4). Next to determine novel mutations the calls will be assessed through MutSig2CV, to determine significantly mutated genes. MutSig2CV, is an algorithm that assigns gene-level significance scores based on mutation recurrence in each gene and adjusted for a gene-specific background.
“Germline” vs. “somatic” mutations

GERMLINE

- Variant present in gamete
- Variation present in every cell in the body, including those of the germline
- Variants can be passed on to offspring

SOMATIC

- Variant absent in gamete
- Variation can occur during fetal development or any time after birth in any cell of the body, except those of the germline
- Variants are not passed on to offspring

https://www.genomicseducation.hee.nhs.uk/cancer-genomics/
Increased mutational burden in lung cancers from smokers

Never-smokers have lower mutational burden.

Never-smokers have lower mutational burden.
Identifying somatic mutations in tumors

ACTGGCATCGATCGAGAGACCTAGCGTGTATAGCCGGTAGGAATGGC…

*
...like finding typos in 200 volumes of the phone book
What is a “mutational signature”? e.g. in one gene
What is a “mutational signature”? 

Many genes...

...seemingly unrelated
What is a “mutational signature”?

Same type of mutation!

Evidence of the cause

Second-hand smoke?
Aging?
Radon?
Something else?
Smokers and never-smokers have different mutational signatures
Conclusion

• Anyone with lungs can get lung cancer

• Differing mutational signatures in smokers and never-smokers point to different underlying etiology

• Future direction: etiology of lung cancer in younger vs. older patients may also differ
  • Lack of chromosomal translocations in post-menopausal women
Acknowledgements

Berger Lab:
Shriya Kamlapurkar
April Lo
Maria McSharry
Sitapriya Moorthi, Ph.D.
Phoebe Parrish
Amanda Riley

Collaborators:
Garnet Anderson, Ph.D.
Amy Paguirigan, Ph.D.
Gavin Ha, Ph.D.