What is the NCI Center for Cancer Genomics (CCG)?
TCGA: The Pipeline for Comprehensive Characterization of the Tumor Genome

Tissue Sample

Pathology QC

DNA & RNA Isolation, QC

Sequencing

Expression, CNA & LOH, Epigenetics

Data Storage at DCC & CGHub

Integrative Analysis

Comprehensive Characterization of a Cancer Genome

GDAC
TCGA Tumor Project Progress

* Only accepting AA cases/500 target reached
Whither the NCI Center for Cancer Genomics (CCG)?
Open Questions in Cancer Genomics

- What is the full extent of driver mutations and genetic pathways in cancer?
- What is the contribution of intratumor genetic heterogeneity to progression and treatment response?
- What is the genetic basis of metastasis?
The 10K Concept
Targeted Therapy of Lung Adenocarcinoma From Cancer Genomics

Lung adenocarcinoma with \textit{EGFR} deletion mutant in exon 19

Before treatment

Erlotinib treatment (2 months)
Identifying Novel Genomic Targets in Lung adenocarcinoma

1984 - 2003

No known genotype

KRAS

2004

No known genotype

EGFR

KRAS

2009

No known genotype

EGFR

KRAS

BRAF

ALK PIK3CA

HER2

2012

No known genotype

EGFR

KRAS

BRAF

ALK PIK3CA

HER2

ROS1

MET

RET

Matt Meyerson
Significantly mutated genes in 230 lung adenocarcinomas

Mostly known genes

Mixture of novel significant genes and false positives

Juliann Chmielecki, Mara Rosenberg, Matt Meyerson
High lung cancer mutation rates pose a major problem in identifying significantly mutated genes

- Genes near statistical threshold may be true positives (oncogenes or tumor suppressors), or false positives

- Known recurrently mutated genes (e.g. $ERBB2$, $CTNNB1$) aren't detected as significant regardless of method used

- In the end, a much larger sample size will be required to elucidate "all" causative mutations in lung adenocarcinoma
10K Goals

• **Oncogenes and Tumor Suppressors**
  Define comprehensive set of driver genes with $\geq 1\%$ frequency in a particular cancer subtype

• **Genetic Pathways**
  Identify epistatic or cooperative relationships between cancer genes that are altered in $\geq 1\%$ cases

• **Interactions**
  Investigate relationship of somatic alterations to germline variations & exposures (e.g. tobacco)

• **Clinical Implications**
  Correlate genetics to clinical outcomes (e.g. local growth vs. $1^\circ / 2^\circ$ metastasis) and treatment response
The Problem: High Background Mutation Rate in Cancer

Mike Lawrence and Gaddy Getz
Lung Adenocarcinoma has Extensive Genetic Damage

Mike Lawrence and Gaddy Getz
Large Numbers of Tumors Needed to Discover Less Common Oncogenes and Tumor Suppressors

Background mutation rate = 10 mutations/Mb

Power

% of cases with mutation in a given gene

# of cases
- 500
- 1000
- 2000
- 5000
- 10000
- 20000

Gaddy Getz
Large Numbers of Tumors Needed to Discover Less Common Oncogenes and Tumor Suppressors

Background mutation rate = 10 mutations/Mb

<table>
<thead>
<tr>
<th># of cases</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>500</td>
<td></td>
</tr>
<tr>
<td>1000</td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td></td>
</tr>
<tr>
<td>5000</td>
<td></td>
</tr>
<tr>
<td>10000</td>
<td></td>
</tr>
<tr>
<td>20000</td>
<td></td>
</tr>
</tbody>
</table>

% of cases with mutation in a given gene

Gaddy Getz
Large Numbers of Tumors Needed to Discover Less Common Oncogenes and Tumor Suppressors

Background mutation rate = 10 mutations/Mb

- Power
- % of cases with mutation in a given gene

# of cases:
- 500
- 1000
- 2000
- 5000
- 10000
- 20000

Gaddy Getz
Mutual Exclusion of Genetic Aberrations Defines Genetic Pathways in Cancer

TCGA, Nature 2012 487:330
Co-occurrence of Genetic Aberrations Defines Genetic Pathways in Cancer

Ngo et al. Nature 2006 441:106
Large Case Numbers Needed to Assign Less Common Cancer Genes to Genetic Pathways

Co-occurrence of mutations in gene A and B (2X enrichment)

Mutual exclusion of mutations in gene A and B (0.5x enrichment)

# cases for p<0.05 with 80% power

Gene B mutation (% of cases)

Gene B mutation (% of cases)
How to find 10K tumor biopsies?
Genomic Analysis of FFPE Biopsies: A Game Changer

SNV Mutation Discovery

Frozen (n=634)

FFPE (n=759)

128 (20%)

506

253 (33%)

TCGA Pilot

BCM
10K Tumor Biopsies

Sample criteria

- FFPE or frozen biopsy samples large enough for whole exome and RNA-seq analysis (i.e. not FNAs)
- Clinical annotation and treatment response necessary
- Matched normal tissue in most (maybe not all) cases
- Consent for genomic analysis
- Likely focus on common cancers (lung, colon, breast, prostate etc.)
Building a 10K Study

10K Study
CCG

Study 1: TCGA

Study 2: Completed Clinical Trials
- NCI-sponsored
- Institutional

Study 3: Prospective Clinical Trials
(e.g. Alchemist)

Study 4: Epi cohorts
- Completed case-control
- Prospective (e.g. PLCO)

Focused Investigation by Study
10K Integration across Studies
TKI-sensitizing EGFR mutations:
10% in Western population
Up to 50% in Asian population
Enriched in:
• females
• non-smokers
• younger patients
Multiple tests in clinical use
No FDA-approved clinical assay

ALK Rearrangement
5-7% in Western population
FDA approved companion diagnostic:
Vysis Break Apart FISH probe
ALChEMIST
Tissue Flow

Consent & Register: A151216 Screening & Follow-up Protocol (n=~7000)

Pre-op Cohort
- SOP-driven FF/FFPE
- After resection, buffy coat

Post-op Cohort
- Assess FFPE
- buffy coat

CLIA-approved LAB
- EGFR mutation test (sequencing)
- ALK rearrangement (FISH)

TCGA pipeline
- Genomic sequencing
- Transcriptome
- Methylation

E4512: Erlotinib
A081105: Crizotinib
Other Adjuvant Studies

Barbara Conley, Jeff Abrams
ALChEMIST
Beyond Treatment Endpoints

- Molecular profiling studies on large cohort (~ 7000 pts)

- Ability to re-profile at relapse in about 50% of cases ("natural genomic history")

- Opportunity to collect epidemiologic info spanning tobacco, diet, alcohol and work exposures
Questions?