Cancer proteogenomics in evolution: Assessing targets, therapy and resistance

Michael A Gillette, MD, PhD
Broad Institute of MIT and Harvard

Cancer and Evolution Symposium
14 October, 2020
Proteogenomics is an important tool in the campaign to analyze cancer from an evolutionary perspective

• An evolutionary perspective on cancer requires
  • Awareness of disease taxonomy
  • Knowledge of the environment
    • Tumor / microenvironment interactions
    • Special interest in the immune landscape
  • Detailed understanding of the biological repertoire of cancer
    • Sensitivity to individual differences is essential for personalized / precision medicine
  • Assessment of the effect of specific selective pressures
    • Direct analyses under therapeutic perturbations can give critical insights into response and resistance

• Like evolution in general, the unit of evolution can be variously defined
  • Cancer subtype, individual tumor, subclone, cell
  • Proteomic techniques and instrumentation are rapidly evolving; however, single cell global proteomics and especially PTM analyses are currently aspirational

• Full definitions of cancer taxonomy, tumor microenvironment and biological repertoire are fostered by comprehensive molecular characterization

• Determining impact of selective pressures requires approaches that can be serially applied, ideally in a clinical setting, as tumors evolve under perturbation
The Cancer Genome Atlas (TCGA) illuminated the cancer genome... but coverage of the proteome was sparse

The Cancer Genome Atlas

11,000 cancers ~ 33 cancer types
RPPA: 181 Abs; ~ 130 proteins / phosphosites

Ding et al, Cell, 2018

“Having a complete picture of every genomic change associated with each tumor can help us make personalized treatment decisions.”

www.foundationmedicine.com

Ellis, Gillette, Carr et al, Cancer Discovery, 2013
The Cancer Genome Atlas (TCGA) illuminated the cancer genome... but coverage of the proteome was sparse

“Acknowledgments

Akbani et al., A Pan-cancer Proteomic Perspective on The Cancer Genome Atlas Nat Comm 2014
Nusinow et al., Quantitative Proteomics of the Cancer Cell Line Encyclopedia Cell 2020

www.foundationmedicine.com
Cancer proteogenomics supports integrated multi-omic analyses for more complete characterization of tumors and adjacent normal tissues.

Use genomic, transcriptomic, and proteomic platforms simultaneously to gain a comprehensive understanding of human cancer in order to improve cancer diagnosis and treatment.

“Having a complete picture of every molecular change associated with each tumor can help us make personalized treatment decisions.”

Many processes downstream of the genome can affect the tumor phenotype.

Zhang et al., Clinical Potential of Mass spectrometry-based proteogenomics
Nat Rev Clin Oncol, 2019
CLINICAL PROTEOGENOMICS TUMOR ANALYSIS CONSORTIUM

Goals

- Accelerate understanding of cancer biology
- Proteogenomically characterize tumors
- Produce public resources (data, assays, images, reagents) for hypothesis-driven science
- Support clinically relevant research projects

Achieved through

TUMOR CHARACTERIZATION

Proteome Characterization Centers
Proteogenomic Data Analysis Centers

TRANSLATIONAL RESEARCH

Proteogenomic Translational Research Centers
- pre-clinical and clinical trial samples
- Mechanisms, response, resistance, toxicity

Integrated research consortium that applies standardized comprehensive proteomics and genomics workflows, strict biospecimen collection protocols (optimized for genomics and proteomics) – ensuring rigor & reproducibility
Lung cancer is the leading cause for cancer-associated death in the US and worldwide. Among women, breast cancer leads incidence and is the second leading cause of death.

### Estimated New Cases

<table>
<thead>
<tr>
<th></th>
<th>Males</th>
<th>Females</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prostate</td>
<td>191,930</td>
<td>21%</td>
</tr>
<tr>
<td>Lung &amp; bronchus</td>
<td>116,300</td>
<td>13%</td>
</tr>
<tr>
<td>Colon &amp; rectum</td>
<td>78,300</td>
<td>9%</td>
</tr>
<tr>
<td>Urinary bladder</td>
<td>62,100</td>
<td>7%</td>
</tr>
<tr>
<td>Melanoma of the skin</td>
<td>60,190</td>
<td>7%</td>
</tr>
<tr>
<td>Kidney &amp; renal pelvis</td>
<td>45,520</td>
<td>5%</td>
</tr>
<tr>
<td>Non-Hodgkin lymphoma</td>
<td>42,380</td>
<td>5%</td>
</tr>
<tr>
<td>Oral cavity &amp; pharynx</td>
<td>38,380</td>
<td>4%</td>
</tr>
<tr>
<td>Leukemia</td>
<td>35,470</td>
<td>4%</td>
</tr>
<tr>
<td>Pancreas</td>
<td>30,400</td>
<td>3%</td>
</tr>
<tr>
<td>All Sites</td>
<td>893,660</td>
<td>100%</td>
</tr>
</tbody>
</table>

### Estimated Deaths

<table>
<thead>
<tr>
<th></th>
<th>Males</th>
<th>Females</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lung &amp; bronchus</td>
<td>72,500</td>
<td>23%</td>
</tr>
<tr>
<td>Prostate</td>
<td>33,330</td>
<td>10%</td>
</tr>
<tr>
<td>Colon &amp; rectum</td>
<td>28,630</td>
<td>9%</td>
</tr>
<tr>
<td>Pancreas</td>
<td>24,640</td>
<td>8%</td>
</tr>
<tr>
<td>Liver &amp; intrahepatic bile duct</td>
<td>20,020</td>
<td>6%</td>
</tr>
<tr>
<td>Leukemia</td>
<td>13,420</td>
<td>4%</td>
</tr>
<tr>
<td>Esophagus</td>
<td>13,100</td>
<td>4%</td>
</tr>
<tr>
<td>Urinary bladder</td>
<td>13,060</td>
<td>4%</td>
</tr>
<tr>
<td>Non-Hodgkin lymphoma</td>
<td>11,460</td>
<td>4%</td>
</tr>
<tr>
<td>Brain &amp; other nervous system</td>
<td>10,190</td>
<td>3%</td>
</tr>
<tr>
<td>All Sites</td>
<td>321,160</td>
<td>100%</td>
</tr>
</tbody>
</table>

**Bray CA Cancer Journal for Clinicians 2018**
**Siegel CA Cancer Journal for Clinicians, 2020**
LUAD Discovery samples represent diverse country of origin, smoking status and stage. Genomics and proteomics profiles nearly complete for 110 LUADs & 101 NATs.*

* NAT= Normal Adjacent Tissue

Gillette, et al., Cell 2020
Global characterization of kinase fusions identified novel fusions, allowed assessment of likely functionality, nominated biomarkers and exposed fusion-driven biology.

Gillette, et al., Cell 2020
120 genes had methylation associated with alterations of mRNA, protein and phosphosite expression, suggesting their possible functional significance.

Proteins regulate antigen-specific T-cell activation.
Proteogenomics exposes KEAP1 / NFE2L2 (NRF2) biology and a putative novel regulatory mechanism. Mutation association analysis highlights important outliers seen only in the phosphosite data.
Kinase outlier analyses nominate candidate therapeutic targets

The immune landscape of LUAD shows regulated “cold” and “hot” tumor clusters

STK11 mutant tumors are especially “cold” and associated with neutrophil degranulation

Gillette, et al., Cell 2020
Data provide a resource for global and subtype-specific LUAD biomarker development

Widely expressed Cancer-testis (CT) antigens are prime candidates as both biomarkers and immunogenic targets

Gillette, et al., Cell 2020
Non-negative matrix (NMF)-based multi-omics clustering defined 4 breast cancer clusters. Luminal clustering was discordant with PAM50 assignments. PAM50 Luminal A tumors assigned to the NMF Luminal B-enriched cluster had intermediate prognosis.
Proteogenomic analysis of ERBB2 positive tumors shows “pseudo-ERBB2” samples with ERBB2 amplification but not protein expression. Some of these may have alternative 17q drivers.

DP1: Satpathy et al. Nature Communications 2020

Krug, et al., Cell 2020 (in press)
PAM50 HER2E, PG ERBB2-negative samples had phosphosite evidence of other ERBB and MAPK signaling.

Proteogenomic analysis of the I-TME suggested broader applicability of immunotherapy in breast cancer.

Krug, et al., Cell 2020 (in press)
APOBEC-mediated mutagenesis correlates with an active I-TME in luminal breast cancer. I-TME markers negatively correlated with NER, BER and MMR in luminal samples only. Phosphoproteomic data were consistent with suppressed DNA damage checkpoint activity in luminals.

- Grams of wet weight tissue obtained from a surgical resection for analysis
- Samples of bulk tumors are cryopreserved to obtain uniform sample for DNA, RNA and protein processing
- Typically 1-2 cores per patient (10-20 mg wet weight tissue/core)
- Substantially lower DNA, RNA, protein yield
- Cryopulverization not feasible
- Often embedded in wax (OCT)
Biopsy Trifecta Extraction (BioTExt) allows suite of full suite of genomic and proteomic analyses from a single needle-core biopsy

- Multiplex proteomic analysis
  - 10 different patient cores/plex
- Reserved sample for targeted proteomics
  - MRM, PRM
- Kinome Profiling

Microscaled proteomics (MiProt)
- Genomics
  - WGS
  - WXS
  - RNA-Seq
  - miRNA-Seq

Satpathy et al., Nat Comm 2020
Proteome depth from cores is similar to bulk
Number of quantified phosphosites is reduced
Biology is preserved

From 25 ug peptides/core
• >10,000 proteins
• >20,000 phosphosites

Satpathy et al., Nat Comm 2020
Microscaling technologies have been successfully applied to needle biopsy samples from clinical trials

14 patients:
- 9 pCR
- 5 non-pCR

35 samples:
- 20 Pre-treatment
- 15 On-treatment

Analysis:
- Total observed data points: 27,217
- Average observed per sample: 27,217

- Gene expression (RNA-seq): 23,549
  - Average observed: 19,492
- Somatic mutation (WXS): 369
  - Average observed: 27
- Proteins (TMT11 Proteomics): 11,657
  - Average observed: 16,333
- Phosphopeptides (TMT11 Proteomics): 23,261
  - Average observed: 17,401

Patient ID
- Therapy Response
- Her2 locus:
- CNA
- RNA
- Protein

Response:
- pCR
- non-pCR

PAM50 subtype:
- Her2-enriched
- Basal
- LumA
- LumB
- NA

Therapy:
- Trastuzumab
- Pertuzumab

Copy number amplification
- −0.5 to 3.5

Z-score (RNA, protein, and phosphoprotein)
- −3 to 3

Satpathy et al., Nat Comm 2020

Phosphoproteomics suggests mechanisms of resistance and therapeutic alternatives. Initial model-based verification data are encouraging.
Summary

- Proteogenomics provides a powerful, reproducible and complementary approach to characterizing cancer biology, exploring mechanisms of resistance and identifying potential therapeutic vulnerabilities.

- Proteogenomics should be part of the armamentarium in programs designed to analyze cancer from an evolutionary perspective, helping illuminate:
  - **Disease taxonomy**
    - Revised hormone receptor positive breast cancer subtype assignments
  - **Knowledge of the environment**
    - Immune landscape gives insight into biology and therapeutic options in lung and breast cancer
    - STK11 tumors may be vulnerable to therapies targeting neutrophil degranulation proteins
    - Subsets of luminal tumors nominated for immune therapy
  - **Detailed biological repertoire of cancer and individual tumors**
    - Proteomic and PTM associations with driver mutations, fusion events and promoter methylation
    - Sample-level characterization of vulnerabilities from phosphosite and kinase outlier analyses
    - Subtype- and sample-specific metabolic profiling leveraging acetylproteomics
    - Improved definition of clinically important marker status (ERBB2, Rb) with therapeutic implications
  - **Effects of specific selective pressures**
    - Model systems may improve understanding of mechanisms of resistance
    - Direct analyses of human tumors on treatment can give critical insights into response and resistance
Acknowledgments

PROTEOMICS
- Karsten Krug
- Shankha Satpathy
- Lauren Tang
- Melanie MacMullan
- Shayan Avanesian
- Ramani Kothadi
- Karl Clauser
- Philipp Mertins
- D. R. Mani
- Steve Carr

GENOME ANALYSIS
- David Heiman
- Mike Noble
- Chet Birger
- Yifat Geffen
- Yosef Maruvka
- Gad Getz

- Alla Karpova
- Song Cao
- Steven Foltz
- Matthew Wyczalkowski
- Li Ding

- Lili Blumenberg
- Emily Kawaler
- Kelly Ruggles
- David Fenyo

- Mohan Dhanasekaran
- Marcin Cieslik
- Alexey Nesvizhkii
- Gil Omenn
- Arul Chinnaiyan

- Francesca Petralia
- Xiaoyu Song
- Jiayi Ji
- Boris Reva
- Azra Krek
- Pei Wang

- Eric Jaehnig
- Meenakshi Anurag
- George Miles
- Jonathan Lei
- Suhas Vasaikar
- Chen Huang
- Bing Zhang
- Thomas (“Trey”) Westbrook
- Matthew Ellis