Proteomic signatures of non-small cell lung cancer

David P. Carbone, MD PhD
Vanderbilt Cancer Center
Nashville, TN USA
U.S. Cancer Mortality, 2007

- Lung: 210,000 new cases, 170,000 deaths
- Colon: 100,000 new cases, 50,000 deaths
- Breast: 230,000 new cases, 55,000 deaths
- Prostate: 210,000 new cases, 25,000 deaths

ACS 2007
Improving outcomes in NSCLC with better molecular signatures

• NO proven screening strategy

• NO proven biomarkers for
  – Early detection - more curable
  – Diagnosis - avoid futile thoracotomies or missed cures
  – Prediction of response to therapy - individualized therapy

• We have to move beyond light microscopy in clinical decision-making toward use of a “molecular signature”.
Proteomics

• Analyzes patterns of protein expression
• Potential advantages:
  – Most nucleic acid sequences have their effect via translation into proteins
  – Protein expression is often not tightly associated with RNA expression
  – Ability to detect post-translational modifications - proteolytic, phosphorylation, lipidation, ubiquitination, etc.
Challenges

• Orders of magnitude more complex than the genome.
• Huge range of expression levels.
• Highly unstable.
• No amplification technologies.
Points to consider

- Protein patterns **should** be better biomarkers than RNA/DNA
- Increasing awareness of problems
  - Overfitting, recursive testing of “test sets”, mixing of training and testing sets
  - Systematic bias in analysis or sample collection
  - Mixing of prognostic and predictive endpoints
Early detection - lung cancer diagnosis from peripheral blood

Both tumor-derived proteins and host-reactive proteins potentially important
Blood diagnostics tested in lung cancer

- CYFRA21-1, NSE, CA15-3, CA19-9 and CA125
- VEGF
- Circulating DNA
- IL-6, TNF-alpha, and leptin
- CRP, SAA
- Bcl-2, MIF
- Beta defensins, ADAM8, MMP9
- Many others…
Sensitivity at 95% Specificity

<table>
<thead>
<tr>
<th>Marker</th>
<th>Sensitivity</th>
</tr>
</thead>
<tbody>
<tr>
<td>CEA</td>
<td>26 - 33%</td>
</tr>
<tr>
<td>SCC</td>
<td>39 - 41%</td>
</tr>
<tr>
<td>CYFRA 21.1</td>
<td>36 - 81%</td>
</tr>
</tbody>
</table>
Can combinations of known markers improve the accuracy of the test?

Antibody microarrays,
Luminex analysis
Antibody microarrays

• 84 antibodies, 80 antigens
• 24 lung cancers, 24 healthy and 32 pts with COPD
• CRP (13.3 fold), SAA (2.0 fold), AAT (1.4 fold) and MUC1 (1.4 fold)
Antibody array: evaluating blood for a panel of candidate biomarkers

Antibody array ROC results


CRP
SAA
AAT
MUC1
Unbiased discovery
Discriminant features by MALDI
NCI Clinical Proteomic Technologies Initiative for Cancer

- 5-year initiative
- Build a foundation of technologies, data, reagents and standards, analysis systems, and infrastructure
- Emphasis on standardization of technology platforms
- Accelerate translation of discovery research and clinical applications

http://proteomics.cancer.gov
Approach to serum biomarkers by MS

- Gene expression profiles
- Cancer biology

normal → preneoplasia → tumor

“proteotypes”

candidate markers

measurable in blood?

yes → develop test, evaluate in clinical trial

no
Heatmap view of identified proteins

Total 3273 proteins

- All pools: 753 proteins
- Cancer pools: 1204 proteins
- SCC pool: 891 proteins
- Adeno Ca. pools: 283 proteins
<table>
<thead>
<tr>
<th>Protein Name</th>
<th>SUM-N</th>
<th>SUM-T</th>
</tr>
</thead>
<tbody>
<tr>
<td>Major vault protein</td>
<td>0</td>
<td>48</td>
</tr>
<tr>
<td>Protein DJ-1</td>
<td>0</td>
<td>48</td>
</tr>
<tr>
<td>Splice Isoform Sap-mu-0 of Proacti</td>
<td>0</td>
<td>43</td>
</tr>
<tr>
<td>Hepatoma-derived growth factor</td>
<td>0</td>
<td>42</td>
</tr>
<tr>
<td>29 kDa protein</td>
<td>0</td>
<td>39</td>
</tr>
<tr>
<td>14-3-3 protein theta</td>
<td>0</td>
<td>38</td>
</tr>
<tr>
<td>14-3-3 protein beta/alpha</td>
<td>0</td>
<td>38</td>
</tr>
<tr>
<td>CDNA FLJ45525 fis, clone BRTHA2</td>
<td>0</td>
<td>36</td>
</tr>
<tr>
<td>Galectin-3 binding protein precursor</td>
<td>0</td>
<td>34</td>
</tr>
<tr>
<td>Cargo selection protein TIP47</td>
<td>0</td>
<td>33</td>
</tr>
<tr>
<td>P63 protein</td>
<td>0</td>
<td>32</td>
</tr>
<tr>
<td>60S acidic ribosomal protein P2</td>
<td>0</td>
<td>31</td>
</tr>
<tr>
<td>19 kDa protein</td>
<td>0</td>
<td>31</td>
</tr>
<tr>
<td>Cathepsin B precursor</td>
<td>0</td>
<td>31</td>
</tr>
<tr>
<td>poly(rC)-binding protein 2 isoform</td>
<td>0</td>
<td>29</td>
</tr>
<tr>
<td>14-3-3 protein sigma</td>
<td>0</td>
<td>29</td>
</tr>
<tr>
<td>Splice Isoform A of Chloride intracellular channel protein</td>
<td>0</td>
<td>29</td>
</tr>
<tr>
<td>14-3-3 protein epsilon</td>
<td>0</td>
<td>28</td>
</tr>
<tr>
<td>Chloride intracellular channel protein</td>
<td>0</td>
<td>28</td>
</tr>
<tr>
<td>Glucosidase II beta subunit precursor</td>
<td>0</td>
<td>28</td>
</tr>
<tr>
<td>pyruvate kinase 3 isoform 2</td>
<td>0</td>
<td>28</td>
</tr>
<tr>
<td>Lactotransferrin precursor</td>
<td>0</td>
<td>28</td>
</tr>
</tbody>
</table>
# Candidate marker identification

<table>
<thead>
<tr>
<th>Name</th>
<th>IPI</th>
<th>Control</th>
<th>SCC</th>
<th>Adeno Ca.</th>
</tr>
</thead>
<tbody>
<tr>
<td>CD26</td>
<td>IPI00018953</td>
<td>0</td>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td>CYFRA 21-1</td>
<td>IPI00479145</td>
<td>0</td>
<td>372</td>
<td>330</td>
</tr>
<tr>
<td>NSE</td>
<td>IPI00216171</td>
<td>0</td>
<td>35</td>
<td>53</td>
</tr>
<tr>
<td>KL-6/MUC1</td>
<td>IPI00013955</td>
<td>0</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>Napsin A</td>
<td>IPI00014055</td>
<td>0</td>
<td>0</td>
<td>34</td>
</tr>
<tr>
<td>Plunc</td>
<td>IPI00009856</td>
<td>0</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>SCC</td>
<td>IPI00022204</td>
<td>0</td>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td>CEA</td>
<td>IPI00654795</td>
<td>0</td>
<td>8</td>
<td>13</td>
</tr>
</tbody>
</table>

**Pooled sample shotgun**
Peptide quantitation in serum

![Graph showing peptide quantitation in serum with pg on the x-axis and concentration on the y-axis.]
Shotgun Conclusions

• We identified >5,000 proteins from Normal, Adeno Ca. and SCC
• For tumor marker discovery, shotgun proteomics has promising power to find molecules expressed and post-translationally modified in tumor tissues at the protein level.
• We are in the process of evaluating these markers as preneoplastic markers in tissue and diagnostic markers in peripheral blood.
Diagnosis of preneoplastic lesions

Which will progress to cancer and which will not?

Biomarkers needed to assess efficacy in chemoprevention trials
Spectra obtained from different regions of lung squamous cell carcinoma

Normal Epithelium

Dysplasia

Invasive Tumor

Spectra obtained from different regions of lung squamous cell carcinoma
Tumors

Preinvasive lesions

Normal bronchus
Predictive signatures
Figure 2  IC\textsubscript{50} values for each drug show a range of \textit{in vitro} sensitivity across the different cell lines. For most drugs, tumor cell lines fall into distinct sensitivity and resistant phenotypes.
Prediction of response
Discriminant signals

Figure 3

Vinorelbine: m/z 9154

Paclitaxel: m/z 9154
Disease specific survival:
Stage>1 with adjuvant carbo/taxol

Months

0 1 2 3 4 5 6 7 8 9

1.2
1.0
0.8
0.6
0.4
0.2
0.0

res
sens
Disease specific survival:
Stage>1 without adjuvant carbo/taxol

0 0.2 0.4 0.6 0.8 1
0 1 2 3 4 5 6 7 8 9

res
sens

Months
Disease specific survival:
Stage 1 without adjuvant carbo/taxol

Months
EGFr TKI

Mutations do not identify all patients with clinical benefit (especially in the west...
BR.21: overall survival

HR = 0.70 (0.58–0.85)
Stratified log-rank p < 0.001


Percentage

At risk
Erlotinib 488 255 145 23 4 0
Placebo 243 107 50 9 0 0

Time (months)
• Erlotinib v. Docetaxel

• No survival difference by:
  – EGFR mutations
    • Patients with EGFR mutant tumors did better in both arms
  – Ras mutations
    • Patients with ras mutant tumors did worse in both arms
  – FISH, Smoking status - no difference
Predicting response from pre-treatment sera
Training cohort
Italian A (70)
Japanese (69)

Algorithm
(8 peaks -> good vs poor outcome)

Test cohort
Italian B (67)

Control cohorts: 1 - 3

Test cohort
ECOG-Erlotinib (96)

Overview

Algorithm Generation

Assessment of the prediction

Training cohort
Italian A (70)
Japanese (69)

Test cohort
Italian B (67)

Control cohorts: 1 - 3

Test cohort
ECOG-Erlotinib (96)

TTP
OS

good / poor

good / poor

good / poor

good / poor

TTP
OS

good / poor

good / poor

good / poor
Concordance of prediction results between Vanderbilt and Colorado

97.1 %

<table>
<thead>
<tr>
<th></th>
<th>UCCC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Good</td>
</tr>
<tr>
<td>VU</td>
<td>Good</td>
</tr>
<tr>
<td></td>
<td>Poor</td>
</tr>
<tr>
<td></td>
<td>Undefined</td>
</tr>
</tbody>
</table>
Test cohort (Italian B)

- Good (31 deaths, 37 cases, median 119)
- Poor (21 deaths, 24 cases, median 62)

$p = 0.004$

- Good (29 deaths, 34 cases, median 192.5)
- Poor (23 deaths, 24 cases, median 73)

$p = 0.008$
Prediction using serum or plasma (n=73)
Control test cohorts - chemo, no TKIs

NSCLC patients from Italy (n = 32)

NSCLC patients from Tennessee (n = 61, stage IIIB or IV)

---

**NSCLC patients from Italy (n = 32)**
- Good (19 deaths, 20 cases, median 163)
- Poor (11 deaths, 12 cases, median 141)
  - p = 0.42

**NSCLC patients from Tennessee (n = 61, stage IIIB or IV)**
- Good (2 deaths, 41 cases, median 729)
- Poor (14 deaths, 20 cases, median 312)
  - p = 0.54
Test cohort 3: NSCLC s/p Surgery only (Poland)

- Good (22 deaths, 44 cases, median 1430)
- Poor (11 deaths, 21 cases, median 1233)

$p=0.786$
Benefit in subsets

Smokers

- Good
- Bad

Survival [days]

Males

- Good
- Bad

Survival [days]

SCC

- Good
- Bad

Survival [days]

$\text{p} < 0.0001$

$\text{p} = 0.0002$

$\text{p} = 0.0003$
ECOG validation cohort: Advanced NSCLC, first line erlotinib

- Good (36 deaths, 69 cases, median 306)
- Poor (22 deaths, 27 cases, median 107)

$p=0.0007$
Lung SPECs Trial of erlotinib

210
Untreated
Advanced
NSCLC

Mandatory
biopsy

EGFr TKI
until
progression

Carboplatin/
Taxol
(Avastin)

Biomarkers:
- EGFr mutations
- ras/other mutations
- Proteomics
- FISH
- IHC
- Affy arrays
Potential Trial Schema

Randomize

ARM A

ARM B

Randomize

Standard of care: Chemo

Serum MALDI analysis

Good

Erlotinib

Bad

Chemo

Endpoint: survival arm A vs. arm B
ECOG 1507: Study Schema

IIIB/IV NSCLC
No Prior Tx
PS 0/1
N=306 patients

MALDI

EGFR Favorable Profile
EGFR Unfavorable Profile

RAND

Std of Care:
Carbo/Pac, Carbo/Doc or Carbo/Gem
Bevacizumab will be given to eligible patients

* Investigator will be blinded to MALDI status

Erlotinib

Std of Care*

* Investigator will be blinded to MALDI status
Conclusions

• Protein signatures may aid in early detection, prognostication and prediction of lung cancer.

• MALDI MS analysis of tumor FNAs may be able to predict OS and TTP after treatment with chemotherapy.

• MALDI-TOF MS of pre-treatment peripheral blood may assist in the selection of NSCLC patients who will show improved survival after treatment with EGFR TKIs.

• These are being prospectively tested in the clinic.
Acknowledgements - serum MALDI/gefitinib response

- Vanderbilt-Ingram Cancer Center
  - David Carbone
  - Fumiko Taguchi
  - Richard Caprioli
  - Pierre Massion

- University of Colorado Cancer Center
  - Ben Solomon
  - Fred Hirsch
  - Paul Bunn Jr.
  - Mark Duncan
  - Steve Hunsucker
  - Rafal Dziadziuszek

- BioDesix
  - Heinrich Roder
  - Julia Grigorieva
  - Maxin Tyspin

- Hospital San Raffaele
  - Vanesa Gregorc
  - Anna Spreafico
  - Fugazza Clara
  - Maria Grazia Viganò

- Policlinico Monteluce
  - Vienna Ludovini

- Kanazawa
  - Kazuo Kasahara

- JFCR
  - Makoto Nishio