What do we learn from Pan-Cancer Subtyping?

TCGA Symposium
May 12, 2014

Pan-Can Integrated Subtypes AWG

Leads: Josh Stuart, UCSC
Chris Benz, Buck
Chuck Perou, UNC
Pan-Cancer: Integrative analysis across tumor types

12 Tumor Types

Omics Characterizations

- Mutation
- Copy Number
- Gene Expression
- DNA Methylation
- MicroRNA
- RPPA
- Clinical Data

Findings: nature.com/tcga

Data:

- Synapse - Sage Bionetworks
### Pan-Cancer-12 Dataset

#### 12 Tumor Types
- Leukemia (LAML)
- Lung adenocarcinoma (LUAD)
- Lung squamous (LUSC)
- Kidney (KIRC)
- Bladder (BLCA)
- Endometrial (UCEC)
- Glioblastoma (GBM)
- Head and neck (HNSC)
- Breast (BRCA)
- Ovarian (OV)
- Colon (COAD)
- Rectum (READ)

#### Thematic Pathways

#### Omics Characterizations
- Mutation
- Copy number
- Gene expression
- DNA methylation
- MicroRNA
- RPPA
- Clinical data

#### Defined by Pan-Cancer AWG

#### ~3500 Samples

<table>
<thead>
<tr>
<th>Tumor Type</th>
<th># Samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>AML</td>
<td>173</td>
</tr>
<tr>
<td>Bladder</td>
<td>122</td>
</tr>
<tr>
<td>Breast</td>
<td>845</td>
</tr>
<tr>
<td>Colon</td>
<td>190</td>
</tr>
<tr>
<td>Endometrial</td>
<td>370</td>
</tr>
<tr>
<td>GBM</td>
<td>168</td>
</tr>
<tr>
<td>Head &amp; Neck</td>
<td>303</td>
</tr>
<tr>
<td>Kidney Clear Cell</td>
<td>480</td>
</tr>
<tr>
<td>Lung Adeno</td>
<td>355</td>
</tr>
<tr>
<td>Lung Squamous</td>
<td>259</td>
</tr>
<tr>
<td>Ovarian Serous</td>
<td>265</td>
</tr>
<tr>
<td>Rectum</td>
<td>72</td>
</tr>
</tbody>
</table>

The Cancer Genome Atlas
Need a Map to navigate so much info

Inspiration: Google Maps

• Fixed, learnable coordinate system
• Natural human intuition
• Overlay stores, places, reviews, photos, video...
• And... Its not a heatmap!
Each sample has its own Address

- Each sample = hexagon (the address)
- Hexagons good packing properties
- Colors display attributes: outcomes, mutations, etc
Connecting Signatures to Addresses

- **Spring Layout**: Low energy configuration of competing springs
- **Snap to grid**: Associate one point per hexagon.
- **Samples w/ similar signatures** → **same zip code**.

**Layout Engine**: DrL. Sandia National Labs

Adam Novak, Sahil Chopra, Robert Baertsch
Identify similar samples in the same zip code

• Map address reflects molecular similarity

• “Zip codes” carry information
  – like: Russian Hill, Berkeley, Silicon Valley, Bethesda, ...

Adam Novak, Sahil Chopra, Robert Baertsch
Are disease-specific AWG subtypes recapped in TumorMap?

- BRCA subtypes resolve clearly on mRNA and PARADIGM maps.
Are disease-specific AWG subtypes recapped in TumorMap?

- Good agreement overall.

Newton, Baertsch, UCSC
6 Data Platforms

- mRNA
  - (Hoadley, UNC)

- microRNA
  - (Hoadley, UNC)

- Protein
  - (Akbani, MDACC)

- DNA Copy Number
  - (Cherniack, Broad)

- DNA Methylation
  - (Shen, USC)

- Exome-Mutations
  - (not used)
    - (Uzunangelov, UCSC)
Single Platform Subtypes

• 6 platforms, each produced 8-19 different clusters.
  – DNA Methylation had the most.

• All subtypes show a strong correlation with tissue of origin.
Single Platform Subtypes Recap on TumorMap

mRNA Map

Colors reflect the subtypes obtained using mRNA platform.

- Newton, Baertsch, UCSC
Single Platform Subtypes Recap on TumorMap

Colors reflect the subtypes obtained using each different platform.

- Good agreement w/ most. miRNA still needs work…

Newton, Baertsch, UCSC
Single Platform Subtypes Correlated with Tissue of Origin

- Exome mutation clusters show least amount of tissue correlation, but still appreciable (~70%, Kandoth et al. *Nature* 2013).
Single platform maps are tissue driven

- Each layout driven by a different data platform.

Newton, Baertsch, UCSC
Cluster of Cluster Assignments - COCA

- Question: How do get one cluster solution from many?
- Answer: Democracy!
  - But like the electoral college system:
    Every Platform Gets a Vote for Each of its Clusters
Cluster Of Cluster Assignments (COCA subtypes)

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
<th>I</th>
<th>J</th>
<th>K</th>
<th>L</th>
<th>M</th>
<th>N</th>
<th>O</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>CN.1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>CN.2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>CN.3</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>CN.4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>CN.5</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>Methy.1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>Methy.2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>Methy.3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>Methy.4</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>11</td>
<td>Methy.5</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>12</td>
<td>miRNA.1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>13</td>
<td>miRNA.2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>14</td>
<td>miRNA.3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>15</td>
<td>miRNA.4</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>16</td>
<td>miRNA.5</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>17</td>
<td>miRNA.6</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>18</td>
<td>miRNA.7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>19</td>
<td>PAM50.Basal</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>20</td>
<td>PAM50.HER2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>21</td>
<td>PAM50.LumA</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>22</td>
<td>PAM50.LumB</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>23</td>
<td>PAM50.Normal</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>24</td>
<td>RPPA.Basal</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>25</td>
<td>RPPA,Her2</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>26</td>
<td>RPPA,X</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>27</td>
<td>RPPA,LumA</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>28</td>
<td>RPPA.LumB</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>29</td>
<td>RPPA,Reactive11</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>30</td>
<td>RPPA,Reactive1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Katherine Hoadley, UNC

The Cancer Genome Atlas
Consensus Clustering defines number of groups

At $K=13$, we have 11 main Cluster of Cluster Assignment (COCA) subtypes
11 main subtypes found (plus 2 minor)
96% Agreement with Subtypes that have no mutations classes
11 main COCA Subtypes

- 11 main subtypes found (plus 2 minor)
- ~90% of samples cluster with their tissue
- PARADIGM TumorMap corresponds well to COCA
12 Tissue of Origin Sites Translate into 11 COCA Subtypes

- Lung
  - Adeno
  - Squam
- Head & Neck
- Bladder
- Breast
- Kidney
- Endometrial
- Rectum
- Colon
- Ovary
- GBM
- AML

1. LUAD-enriched
2. Squamous-like
3. Breast Luminal (includes all HER2+)
4. Breast Basal-like
5. Kidney
6. Endo
7. Bladder
8. Rectum & Colon
9. Ovary
10. GBM
11. AML

131/139 Basal-like are in this COCA group

Chuck Perou, UNC
## Mutations according to COCA subtypes

<table>
<thead>
<tr>
<th>Gene</th>
<th>1- Adeno-enriched</th>
<th>2- Squamous</th>
<th>3- BRCA Luminal</th>
<th>4- BRCA Basal</th>
<th>5-CRC</th>
<th>6-UCEC</th>
<th>7-COAD</th>
<th>8-Bladder</th>
<th>9-OV</th>
<th>10-GBM</th>
<th>13-AML</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>TP53</td>
<td>52%</td>
<td>72%</td>
<td>24%</td>
<td>80%</td>
<td>3%</td>
<td>51%</td>
<td>18%</td>
<td>17%</td>
<td>1%</td>
<td>9%</td>
<td>0%</td>
<td>41%</td>
</tr>
<tr>
<td>PIK3CA</td>
<td>7%</td>
<td>19%</td>
<td>40%</td>
<td>4%</td>
<td>3%</td>
<td>51%</td>
<td>18%</td>
<td>17%</td>
<td>1%</td>
<td>9%</td>
<td>0%</td>
<td>20%</td>
</tr>
<tr>
<td>PTEN</td>
<td>3%</td>
<td>4%</td>
<td>4%</td>
<td>3%</td>
<td>4%</td>
<td>51%</td>
<td>18%</td>
<td>17%</td>
<td>1%</td>
<td>9%</td>
<td>0%</td>
<td>10%</td>
</tr>
<tr>
<td>APC</td>
<td>6%</td>
<td>4%</td>
<td>0%</td>
<td>2%</td>
<td>2%</td>
<td>5%</td>
<td>32%</td>
<td>5%</td>
<td>2%</td>
<td>1%</td>
<td>0%</td>
<td>6%</td>
</tr>
<tr>
<td>MLL3</td>
<td>18%</td>
<td>11%</td>
<td>7%</td>
<td>5%</td>
<td>4%</td>
<td>5%</td>
<td>25%</td>
<td>2%</td>
<td>4%</td>
<td>1%</td>
<td>8%</td>
<td>8%</td>
</tr>
<tr>
<td>VHL</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>52%</td>
<td>1%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>7%</td>
<td>6%</td>
</tr>
<tr>
<td>KRAS</td>
<td>24%</td>
<td>0%</td>
<td>1%</td>
<td>0%</td>
<td>0%</td>
<td>20%</td>
<td>46%</td>
<td>2%</td>
<td>1%</td>
<td>1%</td>
<td>4%</td>
<td>7%</td>
</tr>
<tr>
<td>MLL2</td>
<td>10%</td>
<td>20%</td>
<td>2%</td>
<td>0%</td>
<td>0%</td>
<td>20%</td>
<td>46%</td>
<td>2%</td>
<td>1%</td>
<td>1%</td>
<td>4%</td>
<td>7%</td>
</tr>
<tr>
<td>ARID1A</td>
<td>8%</td>
<td>5%</td>
<td>2%</td>
<td>0%</td>
<td>1%</td>
<td>3%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
<td>5%</td>
<td>6%</td>
</tr>
<tr>
<td>PBRM1</td>
<td>2%</td>
<td>3%</td>
<td>0%</td>
<td>2%</td>
<td>32%</td>
<td>2%</td>
<td>0%</td>
<td>5%</td>
<td>0%</td>
<td>1%</td>
<td>0%</td>
<td>5%</td>
</tr>
<tr>
<td>NAV3</td>
<td>20%</td>
<td>11%</td>
<td>1%</td>
<td>1%</td>
<td>2%</td>
<td>1%</td>
<td>5%</td>
<td>2%</td>
<td>5%</td>
<td>2%</td>
<td>1%</td>
<td>5%</td>
</tr>
<tr>
<td>PIK3R1</td>
<td>2%</td>
<td>2%</td>
<td>3%</td>
<td>1%</td>
<td>0%</td>
<td>31%</td>
<td>2%</td>
<td>0%</td>
<td>0%</td>
<td>1%</td>
<td>0%</td>
<td>5%</td>
</tr>
<tr>
<td>NF1</td>
<td>12%</td>
<td>5%</td>
<td>2%</td>
<td>3%</td>
<td>2%</td>
<td>4%</td>
<td>1%</td>
<td>11%</td>
<td>3%</td>
<td>8%</td>
<td>1%</td>
<td>5%</td>
</tr>
<tr>
<td>SETD2</td>
<td>7%</td>
<td>3%</td>
<td>1%</td>
<td>1%</td>
<td>12%</td>
<td>3%</td>
<td>3%</td>
<td>8%</td>
<td>2%</td>
<td>2%</td>
<td>1%</td>
<td>5%</td>
</tr>
<tr>
<td>ATM</td>
<td>7%</td>
<td>4%</td>
<td>2%</td>
<td>2%</td>
<td>3%</td>
<td>2%</td>
<td>6%</td>
<td>8%</td>
<td>1%</td>
<td>2%</td>
<td>0%</td>
<td>4%</td>
</tr>
<tr>
<td>EGFR</td>
<td>11%</td>
<td>4%</td>
<td>1%</td>
<td>0%</td>
<td>2%</td>
<td>1%</td>
<td>2%</td>
<td>0%</td>
<td>1%</td>
<td>25%</td>
<td>1%</td>
<td>4%</td>
</tr>
<tr>
<td>FBXW7</td>
<td>1%</td>
<td>6%</td>
<td>0%</td>
<td>2%</td>
<td>0%</td>
<td>12%</td>
<td>12%</td>
<td>6%</td>
<td>1%</td>
<td>0%</td>
<td>0%</td>
<td>1%</td>
</tr>
</tbody>
</table>

Only 3 Genes > 10% frequency

Chromatin remodelers, as a class, account for many
DNA Copy # according to COCA subtypes
Interconnected mutated networks reveal subtype and tissue preferential

- HotNet2 mutated subnetworks spanning all tumor types.
Do Integrated Subtypes Provide New Prognostic Information?

Improvement with integrated subtypes over clinical and tissue

Tissue $\rightarrow$ Overall Survival

COCA $\rightarrow$ Overall Survival

Improvement w/ integrated subtypes over clinical and tissue
12 Tissue of Origin Sites Translate into 11 COCA Subtypes

131/139 Basal-like are in this COCA group

Chuck Perou, UNC
BLCA samples diverge into 3 integrated subtypes
BLCA divergence in Pan-Can-12

- BLCA diverge into bladder-enriched, squamous, and LUAD-enriched islands
Integrated subtyping of BLCA distinguishes patient outcomes

• COCA clusters distinguish different survival classes for BLCA
Expression determinants of BLCA divergence

- Squamous-like BLCA show significant genomic differences

Squamous-like

BLCA-enriched

3p Loss in Squamous

BLCA-squamous mutated chromatin remodelers MLLs KDM5A, EP300
Expression determinants of BLCA divergence

- Higher HER2 and Rab25 in non-squamous BLCA – consistent w/ BLCA AWG
- Markers of EMT expressed in squamous BLCA cases
Gene Programs – functionally coupled genes coregulated across PanCan-12

- Identified 22 sets of functionally-related genes coregulated in PanCan-12.
- Gene Programs can recapitulate the integrated subtypes.

The image illustrates the identification process of gene programs through the steps:

1. Applying a bimodal filter (BI > 1.1 in at least one cancer) to 6,898 signatures/modules, evaluating on 3602 samples covering 12 cancers, resulting in 3300 signatures/modules + 4 attractor metagenes.
2. Applying the WGCNA algorithm.
3. Resulting in 22 non-redundant gene programs (GP).
Gene Programs: Surrogates of Integrated Subtypes

- 90% classification accuracy (LDA)

Denise Wolf, UCSF
Gene Programs: Surrogates of Integrated Subtypes

- 90% classification accuracy (LDA)

Denise Wolf, UCSF
Viewing Gene Programs on the TumorMap

ER Signaling “Weather Map”

BRCA Luminals Show High ER signaling
Gene Programs: Surrogates of Integrated Subtypes

- 90% classification accuracy (LDA)

Denise Wolf, UCSF
Viewing Gene Programs on the TumorMap

HIF1A “Weather Map”

GP12_Hypoxia/Glycolosis program

KIRC w/ high hypoxia due to VHL mutations
Gene Programs: Surrogates of Integrated Subtypes

- 90% classification accuracy (LDA)

Denise Wolf, UCSF
Viewing Gene Programs on the TumorMap

Tissue View

MYC Amplification Targets

Basal Signaling

Merge

Subtypes w/ specific gene programs patterns

Denise Wolf, UCSF
Gene Program markers of BLCA divergence

• Squamous BLCA cases show higher GP17

Global Overview of GP17

Restricted to BLCA

“Squamous island”

“Bladder Island”
Oncogenic Tp63 forms are more active in Squamous vs BRCA/Basal (or OV) TP53 mutants

- Most targets in network higher activity in Squamous (more “squares in diagram”)

Christina Yau, Buck
Oncogenic Tp63 forms are more active in Squamous vs BRCA/Basal (or OV) TP53 mutants.
Published as a Resource

Datasets and subtypes provide pivot for further analyses

All datasets hosted on a Synapse project page
- Links to all relevant PanCan-12 datasets
To appear in *Cell*

Stacking cancers and genomic platforms

Hoadley and Zhu

- Vote for your favorite cover image
- Pancakes, Part II – Original Pancake House – 7703 Woodmont Ave.

Zhong Chen
Summary

- Analysis of 12 tumor types w/ 6 platforms display tissue-of-origin as dominant
- Integrated analysis reveals 11 major groups, with some tumor types merging together (HNSCC, Lung Squamous, some Bladder) and others separating (breast luminal vs. Basal-like)
- 1:10 re-classified cases based on the map.
  - Rate similar to EGFR mutations in NSCLC cancers
  - Convergences and Divergences of tissues
- Intriguing subtype-specific differences in TP53 pathway activity between OV, BRCA-Basals, and the Squamous tumors
- Classification adds prognostic information independent of tissue and stage.
  - E.g. COCA clusters define clear prognostic groups for BLCA
- Clearly more investigation will be beneficial; especially those that subtract away tissue-of-origin signals (see Verhaak paper)
Acknowledgments

UCSC Cancer Genomics
• Adam Ewing
• Chris Wilks
• Sofie Salama
• Steve Benz

UCSC Genome Browser Staff
• Mark Diekens
• Melissa Cline
• Jorge Garcia
• Erich Weiler

Buck Institute for Aging
• Christina Yau

Collaborators
• Li Ding, WashU
• Matthew Eills, WashU
• Elaine Mardis, WashU
• Rick Wilson, WashU
• Cyriac Kandoth, WashU

• Joe Gray, OHSU
• Laura Heiser, OHSU
• Nuria Lopez-Bigas, UPF
• Abel Gonzalez, UPF
• Adam Margolin, Sage
• Larsson Omberg, Sage

ORACLE  PCF  LINCS
NSF  Hitachi
Supplemental
• LumB BRCA (and HER2’s) have high MYC but low basal signaling.
• 3 distinct DNA methylation subtypes revealed:
  – One all BRCA-basals, one all luminals and HER2, and one mixed
BRCA DNA Methylation Subtypes

- Created a contrast between BRCA methylation subtypes {1,2} vs {3}.
- SCNA in {1,2}: deletions in CNTN5 (11q22.1) RB1 (13q14.2), ITM2B (13q14.3)
- SCNA in {3}: chr3-12484849-12485147 amplification
- Slightly more TP53 in BRCA-Methylation subtype 3 (BM3) (P<0.04).
- More mutations in ORF KIAA0947 in BM1 and BM2.
BRCA Methylation Subtypes (indep of transcriptional subtypes)

- Correcting for Subtype.
- Restrict only to the luminals. Enriched in BRCA-Meth-2 vs 3: RB1 deletions (P < 0.0076)
- Restrict to basals. Only found one amplicon on 6p21 (POLR1C, POLH, KLK2, CUL9, ...)
- POLH could be *very* interesting as it copies past thymidine dimers and causes hypermutation rates. Do the subset of patients have higher mutation rates?
BLCA Subtypes

- Bladder cancers are split into two subtypes on the map.
- One set tightly associated with squamous cancers of HNSC and LUSC
- The other with the rest of the BLCA
The non-squamous BLCA tumors are characterized by mutations in several genes including ERBB2, BAP1, STAG2, PDGFRA, and the ORF KIAA0947.
The DNA methylation pancan subtypes correspond better to the TumorMap miRNA clusters than the miRNA pancan subtypes!
Mutation Clusters: To Be Continued

- Clustering by mutations gives more tissue-orthogonal clustering.
- But mutation clustering is difficult.
- This solution used prior pathway knowledge to help unify evidence.
BRCA transcriptional subtypes also reflected in the somatic copy number data.

- BRCA basals similar to OV.
  - And to a lesser extent, the Squamous group (HNSC/LUSC)
- BRCA luminal subtypes similar to UCEC and LUAD
Interesting minor BRCA subtype

- A luminal area distinct from the major luminal BRCA area.
- What distinguishes this subtype?
Minor BRCA subtype

- Absence of amplifications
  - chr8p11 amplifications absent in the subtype
    - Genes: LETM2 (8p11.23), WHSC1L1 and POLB (8p11.2)
  - chr11q14 amps absent
    - Genes: ALG8 (11q14.1), KCTD14 (11q14.1)
- Absence of mutations
  - GATA3, MLL3, MAP2K4, PTEN, NCOR1, SYNE1, DMD, PIK3R1, NF1, SPEN, BRCA2, CTCF, TBX3,