PARADIGM, Tumor Map

Yulia Newton
TGCT Update
July 2, 2015
Agenda

• PARADIGM

• Tumor Map
PARADIGM Pipeline Overview

mRNA RNASeq
155 samples, 20503 genes

CNV Gistic
155 samples, 24204 genes

19514 common HUGO gene names
mRNA Data Is More Informative Than CNV In Pathway Space

Correlation between the results of PARADIGM based on mRNA only and PARADIGM based on mRNA+CNV:
# Note About Histology Calls

**Histology diagnosis:**

<table>
<thead>
<tr>
<th>Histology diagnosis</th>
<th>Histology %:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Seminoma; NOS</td>
<td></td>
</tr>
<tr>
<td>Non-Seminoma; Embryonal Carcinoma; Non-Seminoma; Yolk Sac Tumor</td>
<td>95</td>
</tr>
<tr>
<td>Non-Seminoma; Teratoma (Mature); Non-Seminoma; Yolk Sac Tumor</td>
<td>95</td>
</tr>
<tr>
<td>Non-Seminoma; Embryonal Carcinoma</td>
<td></td>
</tr>
<tr>
<td>Non-Seminoma; Embryonal Carcinoma</td>
<td></td>
</tr>
<tr>
<td>Non-Seminoma; Embryonal Carcinoma</td>
<td></td>
</tr>
<tr>
<td>Non-Seminoma; Embryonal Carcinoma</td>
<td></td>
</tr>
<tr>
<td>Non-Seminoma; Yolk Sac Tumor</td>
<td></td>
</tr>
<tr>
<td>Non-Seminoma; Embryonal Carcinoma; Non-Seminoma; Yolk Sac Tumor</td>
<td>80</td>
</tr>
<tr>
<td>Non-Seminoma; Embryonal Carcinoma; Non-Seminoma; Teratoma</td>
<td>35</td>
</tr>
<tr>
<td>Non-Seminoma; Teratoma (Mature); Seminoma; NOS; Non-Seminoma; Embryonal Carcinoma</td>
<td>40</td>
</tr>
<tr>
<td>Non-Seminoma; Teratoma (Mature); Non-Seminoma; Yolk Sac Tumor</td>
<td>95</td>
</tr>
<tr>
<td>Non-Seminoma; Embryonal Carcinoma; Non-Seminoma; Teratoma (Mature); Non-Seminoma</td>
<td>40</td>
</tr>
<tr>
<td>Non-Seminoma; Embryonal Carcinoma</td>
<td></td>
</tr>
</tbody>
</table>
Clustering solution k=5

- k = 5, silhouette = 0.83

Based on top 1000 varying IPLs
Match Primaries

TCGA-2G-AAFG

TCGA-2G-AAFG-01

TCGA-2G-AAFG-05

TCGA-2G-AAGI

TCGA-2G-AAGI-01

TCGA-2G-AAGI-05

TCGA-2G-AAGY

TCGA-2G-AAGY-01

TCGA-2G-AAGY-05

TCGA-2G-AAHP

TCGA-2G-AAHP-01

TCGA-2G-AAHP-05

TCGA-2G-AAKG

TCGA-2G-AAKG-01

TCGA-2G-AAKG-05

TCGA-2G-AAKO

TCGA-2G-AAKO-01

TCGA-2G-AAKO-05
Interestingly, dichotomies for pathway analysis include:

- Clust2 vs. all
- Clust3 vs. all
- Clust4+Clust5 vs. all
- Clust4 vs. Clust5
- Clust2 vs. Clust3
Cluster 2 vs. All
Cluster 4 vs. Cluster 5
Tumor Map Method

• Manuscript in progress
  – Based on published panncan12
• Part of the manuscript for the joint gliomas paper (GBM+LGG)
• Projecting similarity space
• Analyzing data in the space of a single data type (e.g. mRNA, methylation, etc.) or combined data types (e.g. mRNA +methylation, mut+CNV)
Tumor Map Method

Samples

Spring-embedded graph layout algorithm

Springs oscillate with the energy inversely proportional to the edge weight

Top N similarities

Project layout onto a 2-D grid

Projection of the sample similarity space into a 2-D grid
Tumor Map Analysis of TGCT mRNA

Expression filter:

<table>
<thead>
<tr>
<th>Samples</th>
<th>Gene X</th>
<th>N(_{\text{unexpressed}}) (=0)</th>
<th>Is &gt; 90%</th>
<th>Yes</th>
<th>No</th>
<th>Filter</th>
<th>Keep</th>
</tr>
</thead>
</table>

E.g. ACCN5 expressed at some level in 11 samples, unexpressed at all in 146 samples
mRNA Space Recapitulates Histology

Colored by histology:

Colored by non-seminoma type:

Colored by seminoma/not:
miRNA Space Correlates With mRNA Space

Colored by histology:

Colored by seminoma/not:

Colored by non-seminoma type:

Those miRNA’s that had expression levels <=50 in more than 50% of cohort were filtered out. 199 features left after the filter.
Some Teratomas Track With Seminomas In CNV Space

Colored by histology:

Colored by seminoma/not:
Non-seminoma
Seminoma

Colored by non-seminoma type:
- Yolk Sac
- Teratoma
- Embryonal
- Mostly Non–Seminoma; Yolk Sac
- Mostly Non–Seminoma; Teratoma
- Mostly Non–Seminoma; Embryonal
- Mostly Seminoma
- Non–Seminoma; Teratoma
- Non–Seminoma; Embryonal
- Unknown
Seminomas and Non-seminomas Are Two Distinct Groups In RPPA Space
Two Possible Seminoma Subtypes In Methylation Space

Colored by histology:

Colored by non-seminoma type:

Colored by seminoma/not:

Those probes that had methylation levels \(\leq 0.05\) in more than 90% of cohort were filtered out. 264143 probes left after the filter.
Multiple Primaries In mRNA Space
Combined Tumor Maps

Samples → mRNA → Methylation → CNV → Mutations → Transformation into a common space → BST score → Weighted average → Tumor Map pipeline
Transformation Into a Common Space (Bivariate Standardization Transformation)

Large-Scale Mapping and Validation of Escherichia coli Transcriptional Regulation from a Compendium of Expression Profiles

Jeremiah J. Faith¹, Boris Hayete², Joshua T. Thaden²,³, Ilaria Mogno²,⁴, Jamey Wierzbowski²,⁵, Guillaume Cottarel²,⁵, Simon Kasif¹,², James J. Collins¹,², Timothy S. Gardner¹,²,*
Combined tumor maps

Colored by histology:

Colored by non-seminoma type:

Colored by seminoma/not:
Next Steps

• PARADIGM
  – Further analysis and refining of the pathways
  – Survival analysis of PARADIGM subtypes
  – Analysis with additional pathway databases (e.g. STRING)
  – Additional dichotomies

• Tumor Map
  – Mutation data
  – Investigate map islands/subtypes
  – TGCT in the context of pancan12 dataset

• Need
  – Survival data