Supporting Subtype Characterization through Integrative Visualization of Cancer Genomics Data Sets

Alexander Lex / Graz University of Technology
Peter J Park / Harvard Medical School
Nils Gehlenborg / Harvard Medical School & Broad Institute
Acknowledgements
Marc Streit / Johannes-Kepler University Linz
Christian Partl / Graz University of Technology
Dieter Schmalstieg / Graz University of Technology
Michael Noble / Broad Institute
Ian Watson / Dana-Farber Cancer Institute
Steven Quayle / Dana-Farber Cancer Institute
Lynda Chin & Gad Getz with the TCGA GDAC Team / Broad Institute

Funding
National Cancer Institute / U24 CA143867
Austrian Science Fund / Caleydoplex project (P22902)
Austrian Research Promotion Agency / inGeneious project (385567)
Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications.

Molecular subclasses of high-grade gliomas predict prognosis, delineate a pattern of disease progression, and resemble stages in neurogenesis.

Gene expression profiling identifies clinically relevant subtypes of prostate cancer.

Integrated Genomic Analysis Identifies Clinically Relevant Subtypes of Glioblastoma Characterized by Abnormalities in PDGFRA, IDH1, EGFR, and NF1.
Integrated Genomic Analysis Identifies Clinically Relevant Subtypes of Glioblastoma Characterized by Abnormalities in *PDGFRA, IDH1, EGFR*, and *NF1*

<table>
<thead>
<tr>
<th>Gene</th>
<th>Proneural</th>
<th>Neural</th>
<th>Classical</th>
<th>Mesenchymal</th>
</tr>
</thead>
<tbody>
<tr>
<td>TP53</td>
<td>mut</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IDH1</td>
<td>mut</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PDGFRA</td>
<td>ge</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EGFR</td>
<td>ge</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NF1</td>
<td>ge</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CDKN2A</td>
<td>ge</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- **Gene Expression**
- **Mutation**
- **TP53 LOH**
- **EGFRvIII**

*Copy Number*
- high level amplification
- low level amplification
- normal copy number
- hemizygous deletion
- homozygous deletion
mRNA

#1  amp

#2  normal

#3  normal

#4  del

Copy Number

Mutation

mut

normal
Fig. 1. VisBricks in action: Four different groups of dimensions with different numbers of clusters per group. The gray arch connects the overviews of the groups. The arches show how the data is distributed over the clusters in that group, thus summarizing the specifics of a dimension group. The clusters themselves are shown in stacked VisBricks above and below the arch depending on whether their average data values are higher or lower than the overall average for the group. Colored ribbons indicate how data items are distributed to clusters of multiple dimension groups.
mRNA

Data
Gene-level TCGA Glioblastoma Multiforme (July 2011)
mRNA

Analyses

GenePattern Consensus NMF Clustering & GISTIC 2
mRNA
mRNA
mRNA
mRNA
mRNA + Copy Number
mRNA + Copy Number
mRNA + Copy Number
mRNA + Copy Number
mRNA + Copy Number
mRNA + Copy Number + miR + methylation
mRNA + Copy Number + miR + methylation
mRNA
mRNA
mRNA
mRNA
mRNA + pathways
mRNA + pathways
mRNA + pathways
mRNA + pathways
mRNA + pathways
Thank you!

nils@hms.harvard.edu

http://gdac.broadinstitute.org | Firehose Analysis Pipeline
http://www.caleydo.org | Caleydo Visualization Framework