

Online tools to analyze TCGA data

Rehan Akbani, Ph.D.

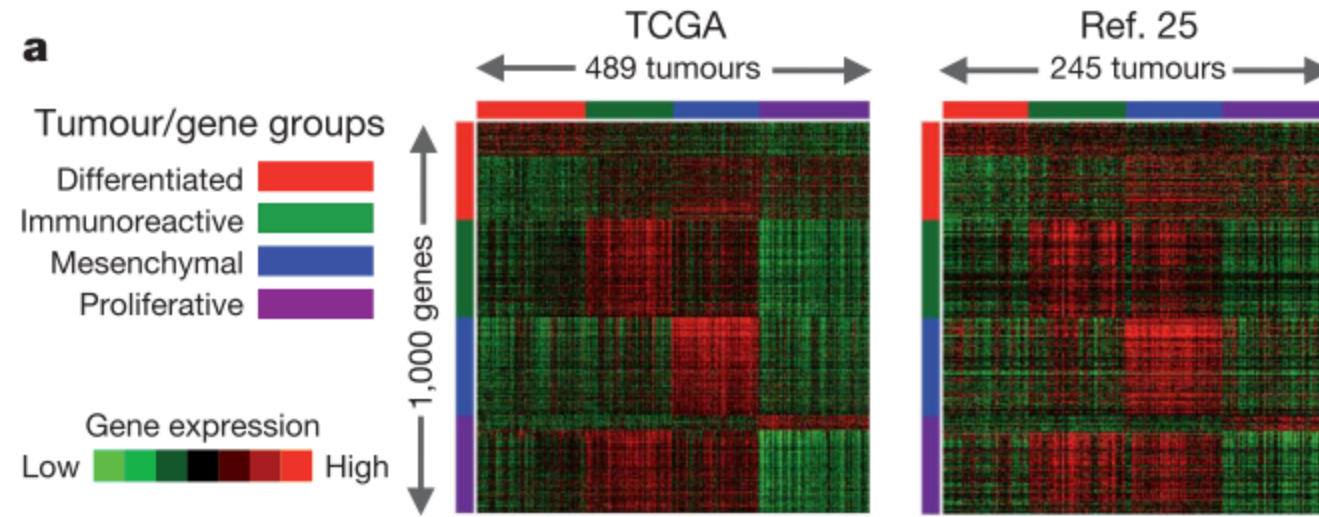
UT MD Anderson Cancer Center

Some Common Analyses Types

- Subtype discovery
 - mRNA, miRNA, DNA methylation, Protein, Copy number
- Outcome prediction
- Pathway alterations
- Copy number profiles
- Cross-tumor comparisons
- Many others
- E.g. TCGA Ovarian Cancer paper

“Integrated genomic analyses of ovarian carcinoma”

Subtype discovery



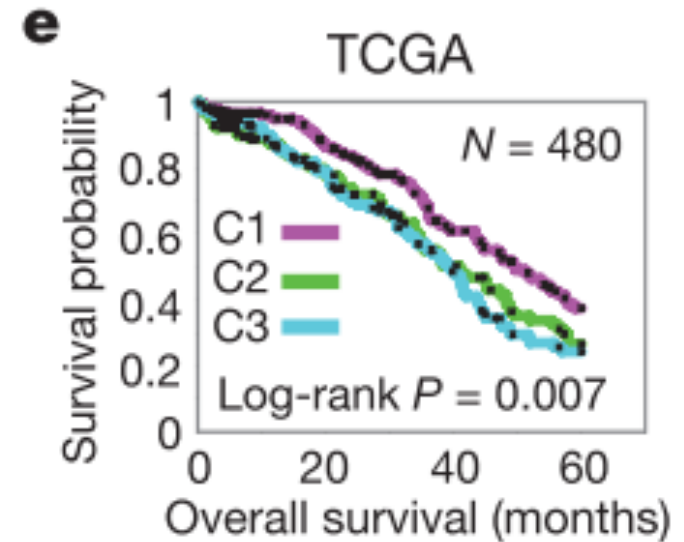
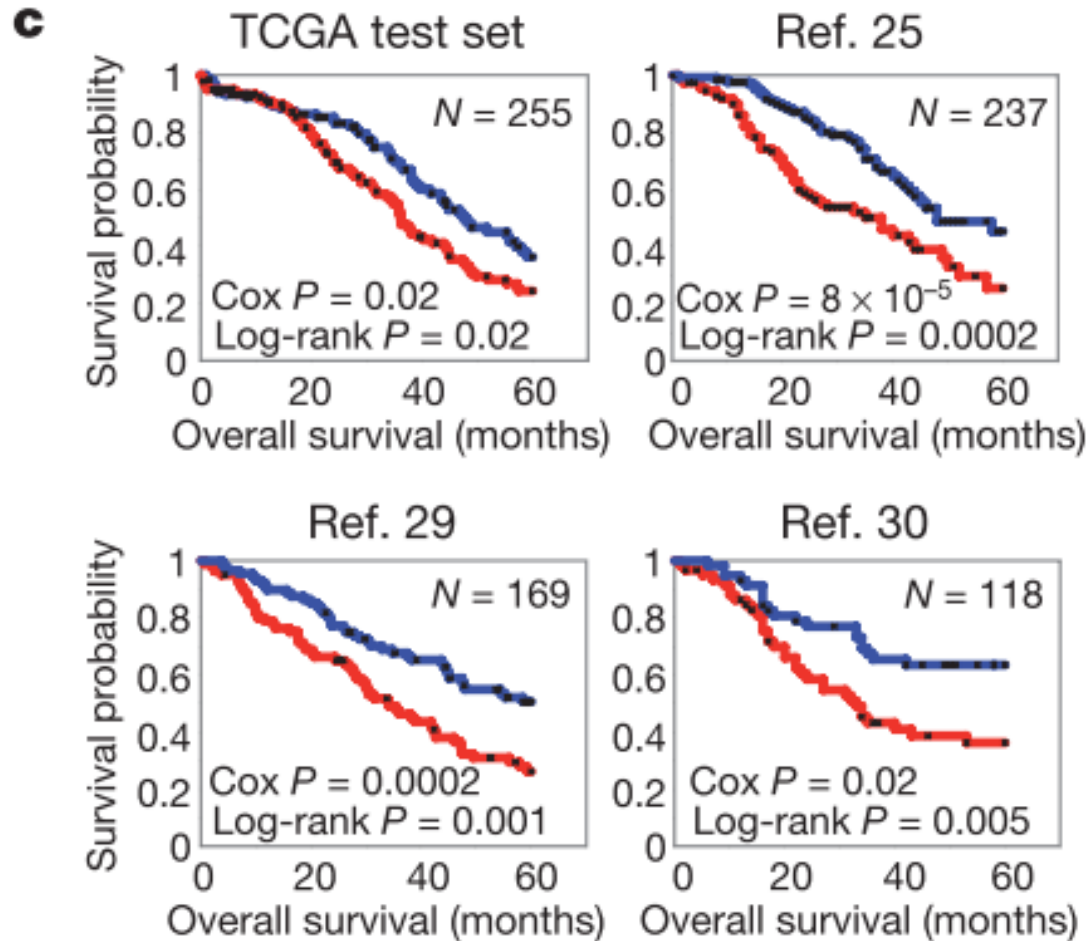
1) Four subtypes found based on mRNA expression

d

		Gene cluster			
		D	I	M	P
miRNA cluster	C1	55	48	15	89
	C2	40	21	51	29
	C3	39	37	43	20

2) Three miRNA based subtypes found

Outcome prediction

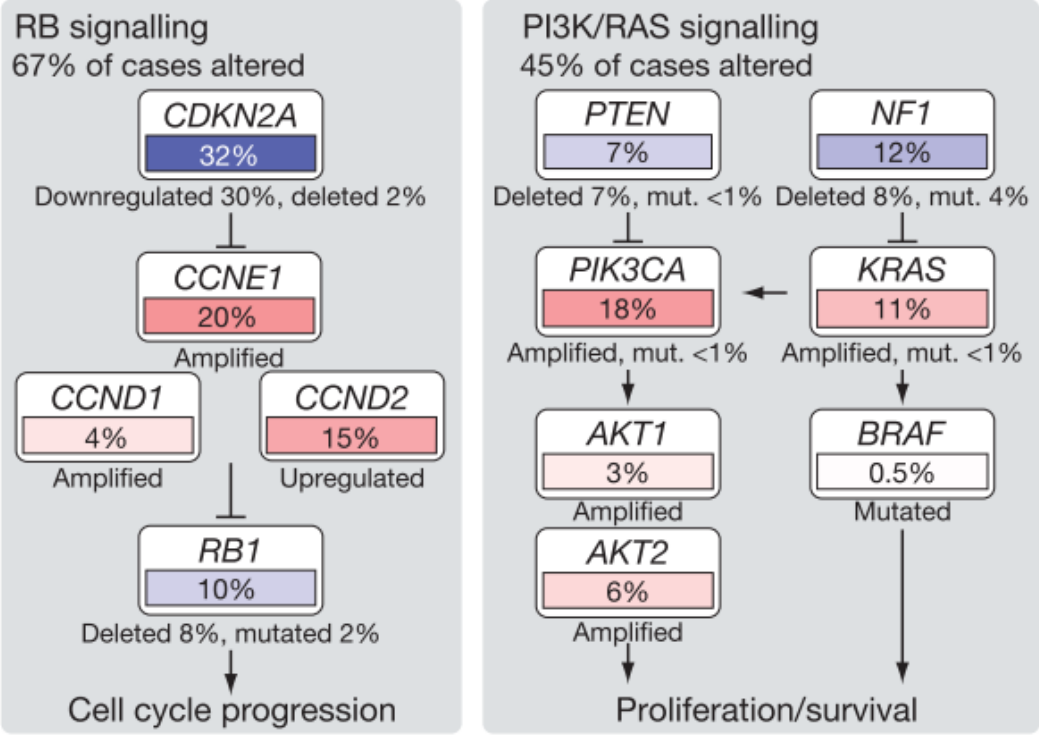


2) Outcome differences in miRNA subtypes

1) Outcome differences in TCGA + 3 other data sets using mRNA expression

Pathway alterations

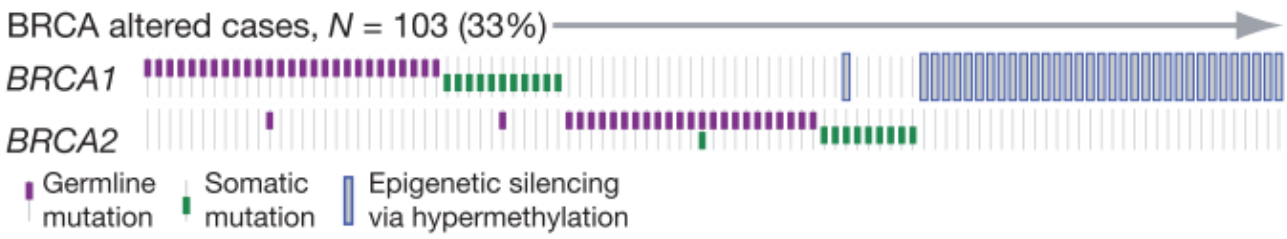
a RB and PI3K/RAS signalling



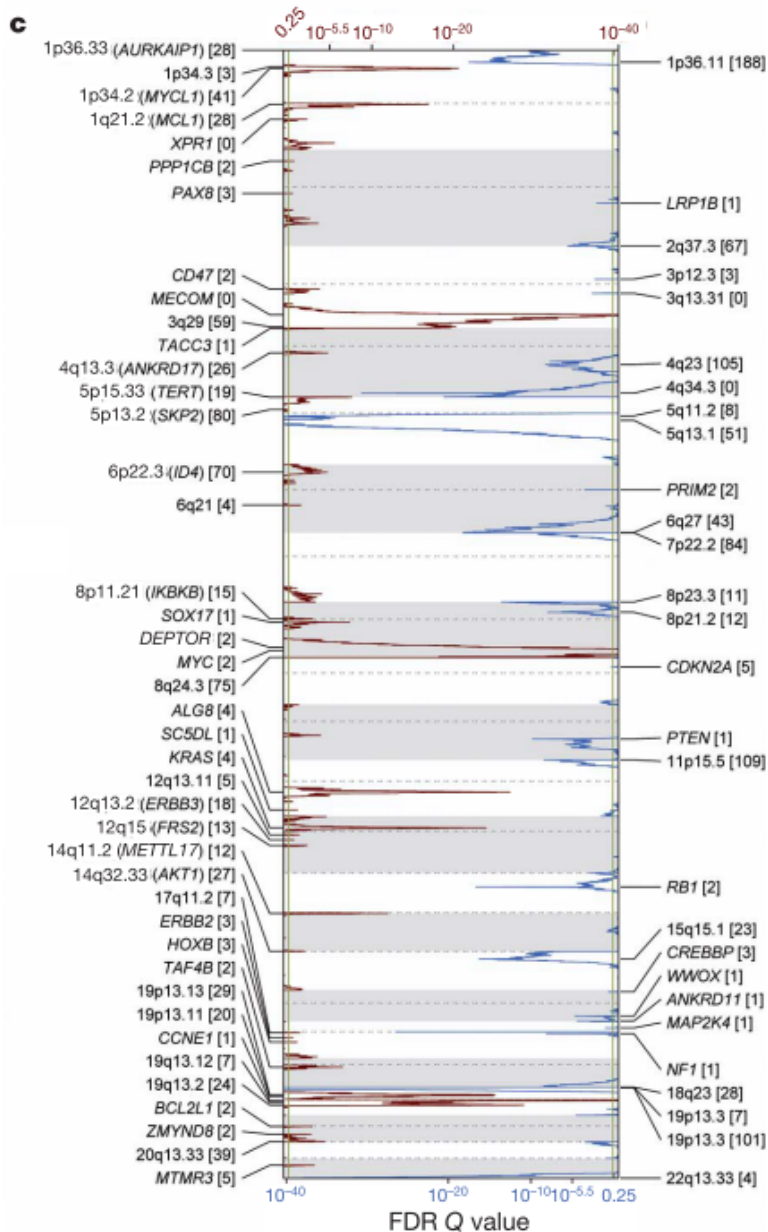
1) Alterations in the RB and PI3K/RAS signaling pathways

2) BRCA1/2 alterations in the HR pathway

c HR alterations

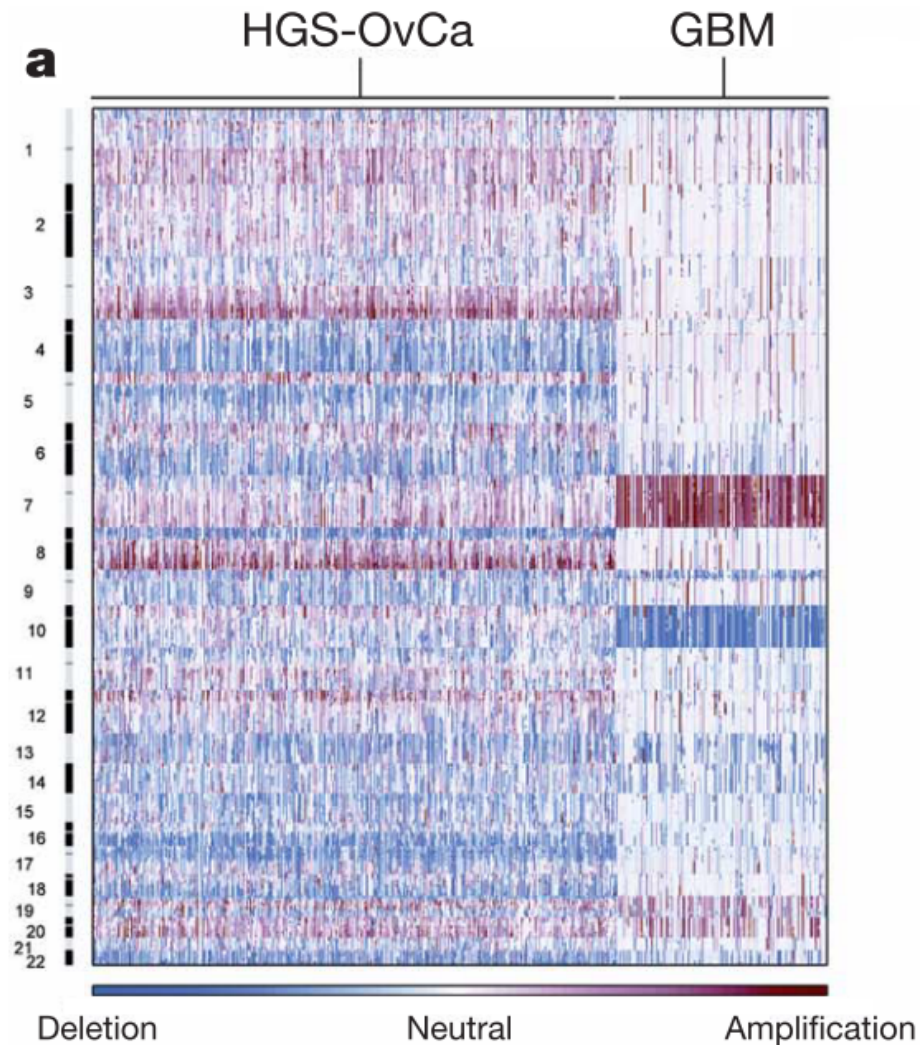


Copy number profiles



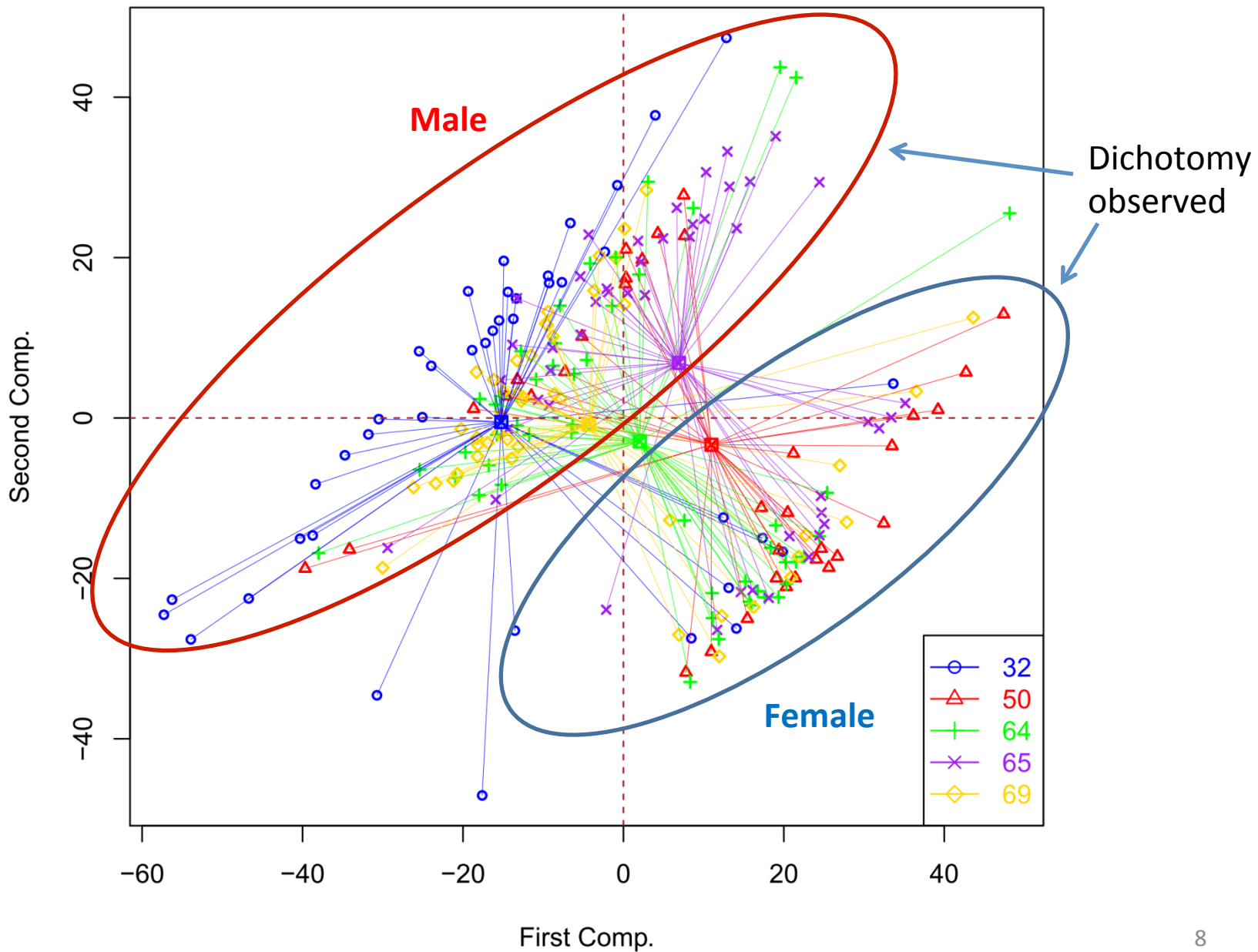
1) Significantly amplified (red) and deleted (blue) chromosome arms

Cross-tumor comparisons

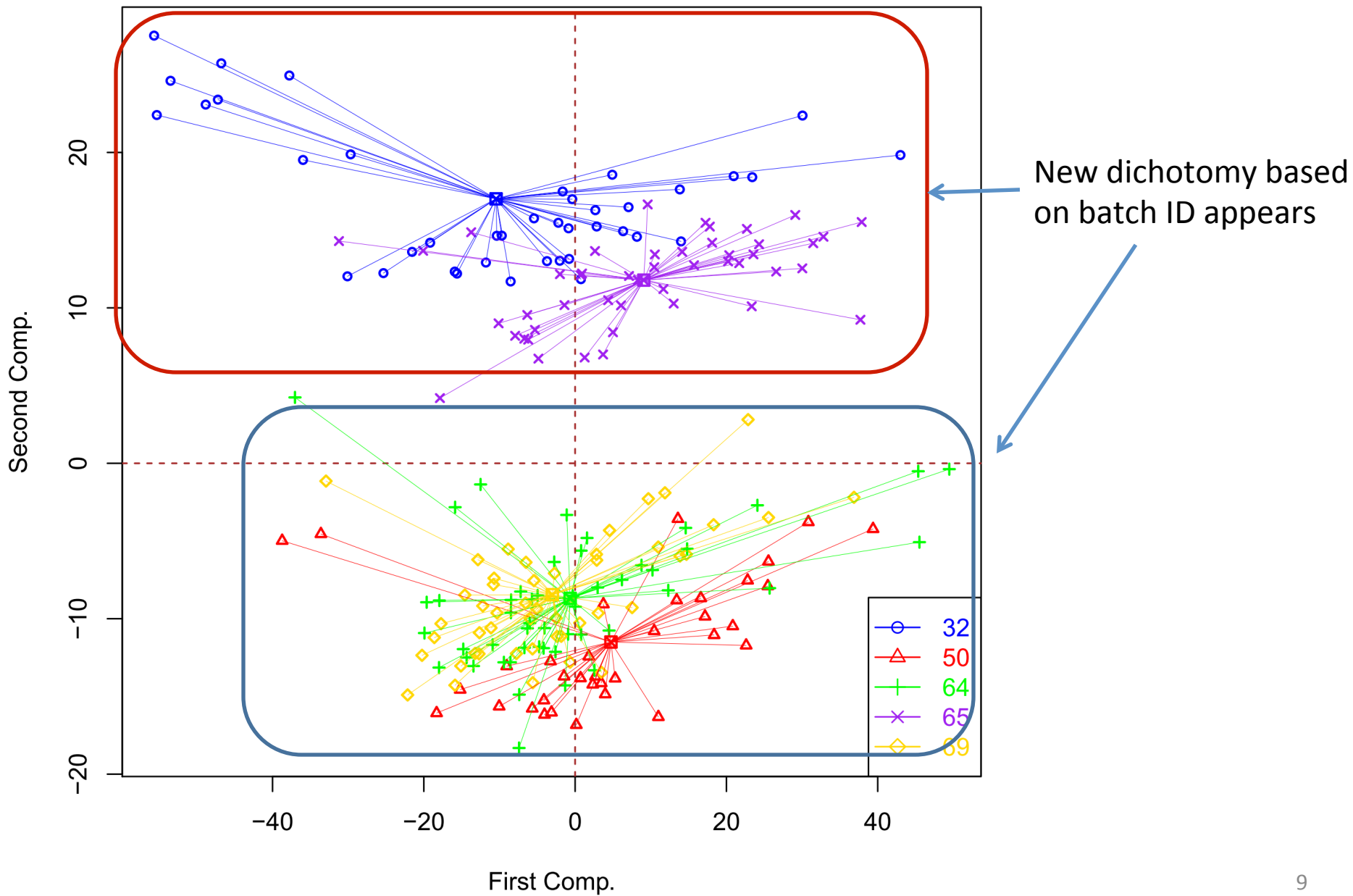


1) Ovarian cancer vs. GBM copy number profiles

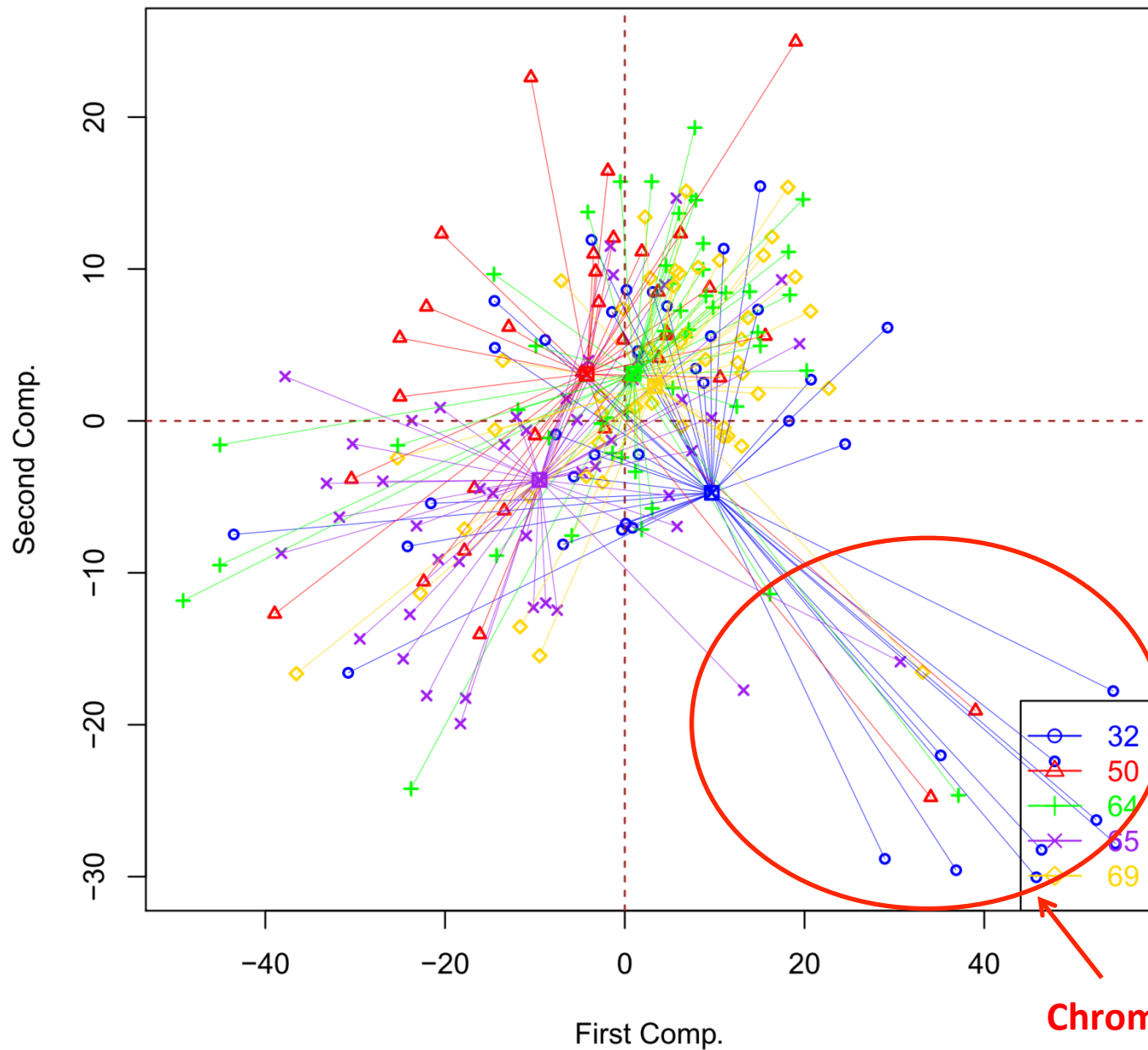
Kidney cancer (KIRC) DNA methylation data (27k)



After removing sex chromosomes



After removing bad probes



Demo and Examples

Introduction

- Go to <http://cBioPortal.org/> by Memorial Sloan Kettering Cancer Center (MSKCC)

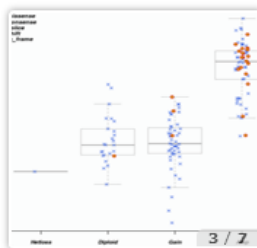


- HOME
- TUTORIALS
- NEWS
- FAQ
- DATA SETS
- ABOUT
- WEB API
- R/MATLAB
- NETWORKS
-
-
-

The cBio Cancer Genomics Portal provides **visualization, analysis** and **download** of large-scale **cancer genomics** data sets.

The portal is developed and maintained by the [Computational Biology Center](#) at [Memorial Sloan-Kettering Cancer Center](#).

Cancer Discovery. May 2012 2; 401. [[Abstract](#)].



Query **Download Data**

Select Cancer Study: All Cancer Studies

Select Data Type Priority: Mutation and CNA Only Mutation Only CNA

Enter Gene Set: Advanced: Onco Query Language (OQL)

Enter HUGO Gene Symbols or Gene Aliases

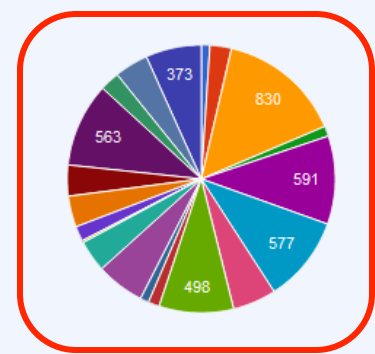
Or Select from Example Gene Sets:

User-defined List

Submit

Data Sets

The Portal contains data for **5547 tumor samples from 20 cancer studies**. [[Details](#)].



Example Queries

- RAS/RAF alterations in colorectal cancer
- BRCA1 and BRCA2 mutations in ovarian cancer
- Protein changes in PTEN-altered ovarian cancer samples
- TP53 and MDM2/4 alterations in GBM
- PTEN mutations in GBM in text format

What's New

The Portal paper has just been published: Ethan Cerami et al. **The cBio Cancer Genomics Portal: An Open Platform for Exploring Multidimensional Cancer Genomics Data**. *Cancer Discovery*. May 2012

Workshop Exercise

Problems:

1. In what percentage of patients with GBM is the RB pathway altered?
2. Which gene in the RB pathway is most frequently altered?
3. Which is the most common type of alteration; amplification, deletion or mutation?
4. Based on the data, is CDKN2A likely to be a tumor suppressor or oncogene? How about CDK4?

Steps

1. In item 1 select “Glioblastoma (TCGA)”

1 Select Cancer Type:

2. In item 2 select the following:

2 Select Genomic Profiles:

- Mutations [?](#)
- Putative copy-number alterations (GBM Pathways) [?](#)
- Putative copy-number alterations (RAE) [?](#)
- mRNA Expression z-Scores [?](#)

3. In item 3 select “All Complete Tumors (seq, mRNA, CNA)”

3 Select Patient/Case Set:

Steps

4. In item 4 select “Glioblastoma: RB Pathway (7 genes)” from the drop-down list below:

4 Enter Gene Set:

Enter gene symbols below:

[Advanced: Onco Query Language \(OQL\)](#)

```
CDKN2A CDKN2B CDKN2C CDK4 CDK6 CCND2 RB1
```

Or Select from Example Gene Sets:

Glioblastoma: RB Pathway (7 genes) ▾

- The list of 7 genes will automatically be filled in
5. Hit the “Submit” button

Result

Summary Network Plots Survival Mutations RPPA Event Map IGV Download Bookmark

OncoPrint [\(What are OncoPrints?\)](#)

Get OncoPrint:

Customize OncoPrint

Case Set: All Complete Tumors: All tumor samples that have mRNA, CNA and sequencing data (122 samples)

Altered in 87 (71%) of cases.

Move the mouse pointer over the OncoPrint below for more details about cases and alterations.

Total All 122 cases -->
altered



Amplification Homozygous Deletion Mutation

Copy number alterations are putative.

Exercise 1

Problem 1:

1. In what percentage of patients with GBM is the RB pathway altered?
2. Which gene in the RB pathway is most frequently altered?
3. Which is the most common type of alteration; amplification, deletion or mutation?
4. Based on the data, is CDKN2A likely to be a tumor suppressor or oncogene? How about CDK4?

Result

Summary Network Plots Survival Mutations RPPA Event Map IGV Download Bookmark

OncoPrint [\(What are OncoPrints?\)](#)

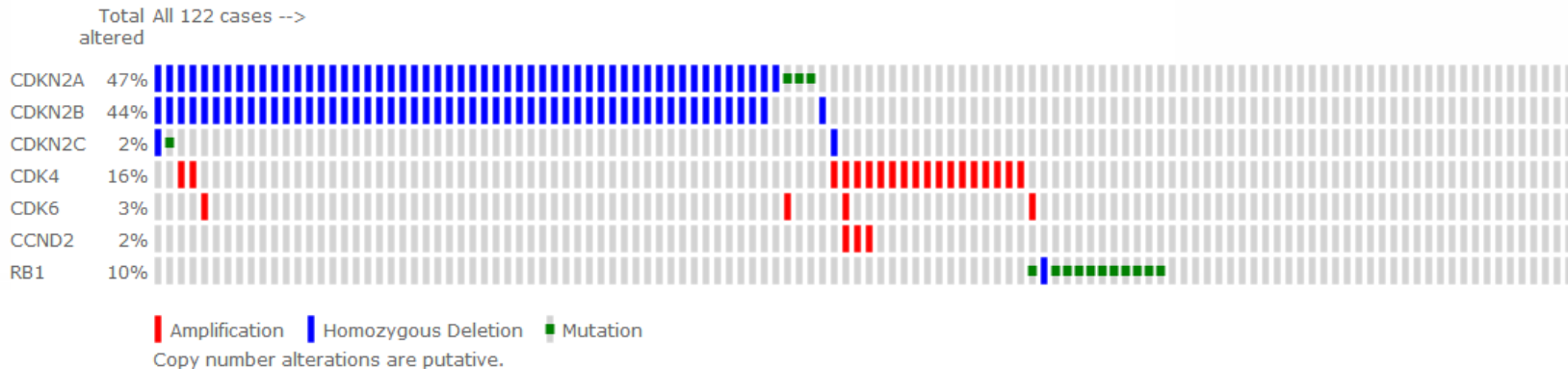
Get OncoPrint:

Customize OncoPrint

Case Set: All Complete Tumors: All tumor samples that have mRNA, CNA and sequencing data (122 samples)

Altered in 87 (71%) of cases.

Move the mouse pointer over the OncoPrint below for more details about cases and alterations.

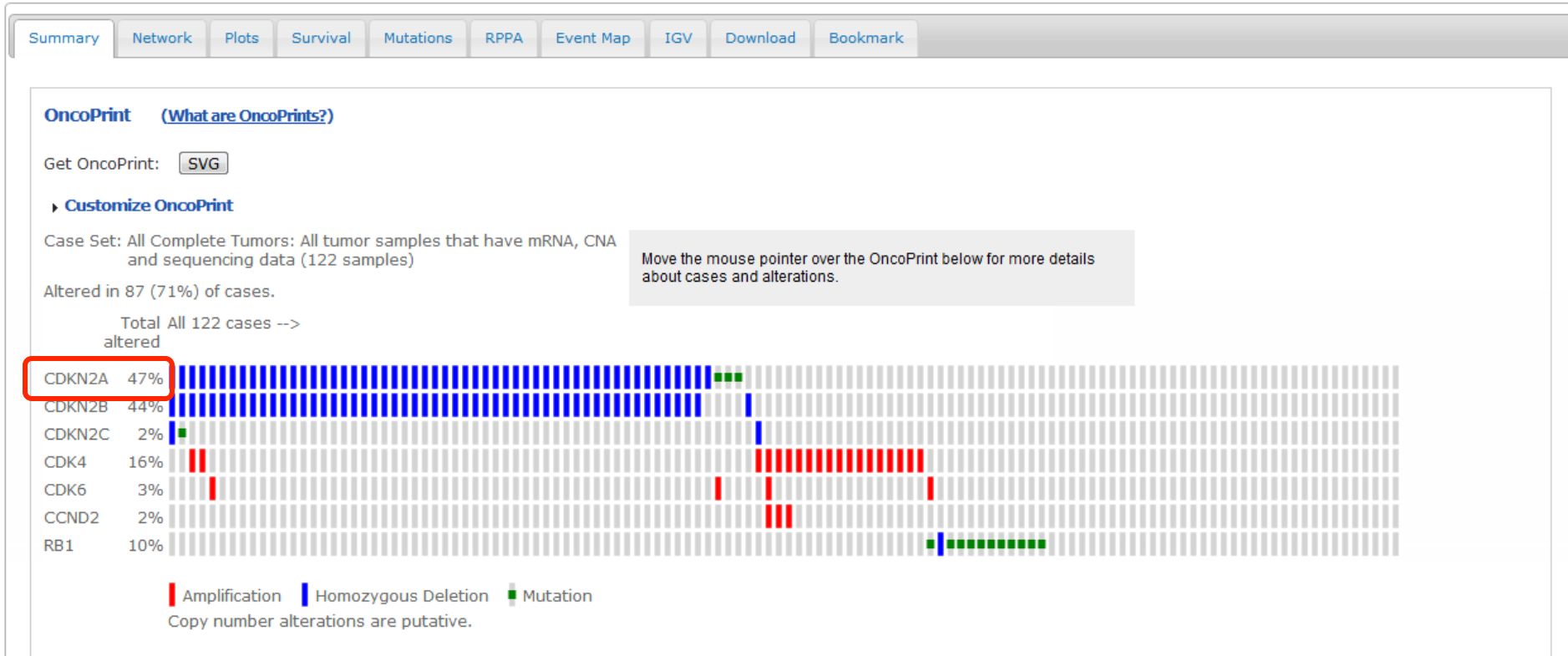


Exercise 1

Problem 1:

1. In what percentage of patients with GBM is the RB pathway altered?
2. Which gene in the RB pathway is most frequently altered?
3. Which is the most common type of alteration; amplification, deletion or mutation?
4. Based on the data, is CDKN2A likely to be a tumor suppressor or oncogene? How about CDK4?

Result



Exercise 1

Problem 1:

1. In what percentage of patients with GBM is the RB pathway altered?
2. Which gene in the RB pathway is most frequently altered?
3. Which is the most common type of alteration; amplification, deletion or mutation?
4. Based on the data, is CDKN2A likely to be a tumor suppressor or oncogene? How about CDK4?

Result

Summary Network Plots Survival Mutations RPPA Event Map IGV Download Bookmark

OncoPrint [\(What are OncoPrints?\)](#)

Get OncoPrint:

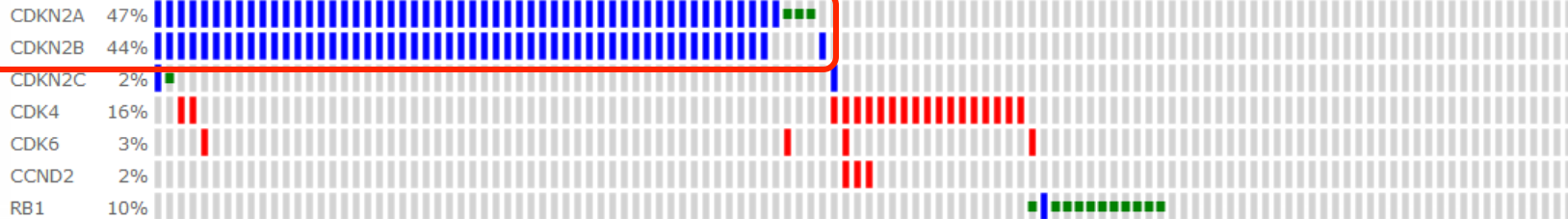
Customize OncoPrint

Case Set: All Complete Tumors: All tumor samples that have mRNA, CNA and sequencing data (122 samples)

Altered in 87 (71%) of cases.

Move the mouse pointer over the OncoPrint below for more details about cases and alterations.

Total All 122 cases -->
altered



Amplification Homozygous Deletion Mutation

Copy number alterations are putative.

Exercise 1

Problem 1:

1. In what percentage of patients with GBM is the RB pathway altered?
2. Which gene in the RB pathway is most frequently altered?
3. Which is the most common type of alteration; amplification, deletion or mutation?
4. Based on the data, is CDKN2A likely to be a tumor suppressor or oncogene? How about CDK4?

Result

Summary Network Plots Survival Mutations RPPA Event Map IGV Download Bookmark

OncoPrint [\(What are OncoPrints?\)](#)

Get OncoPrint:

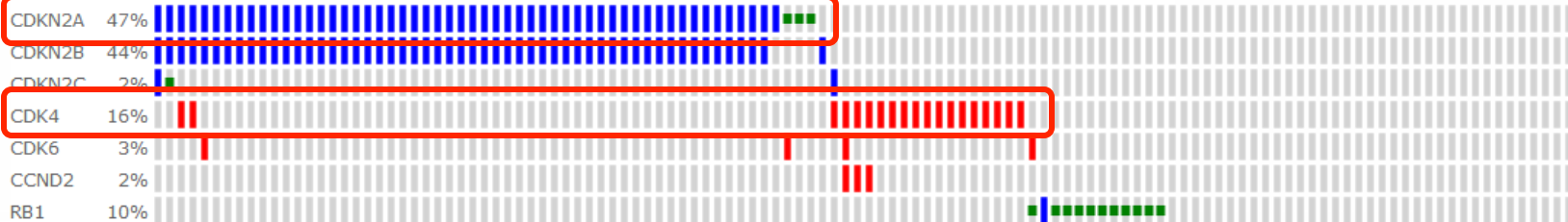
► Customize OncoPrint

Case Set: All Complete Tumors: All tumor samples that have mRNA, CNA and sequencing data (122 samples)

Altered in 87 (71%) of cases.

Total All 122 cases -->
altered

Move the mouse pointer over the OncoPrint below for more details about cases and alterations.



Amplification | Homozygous Deletion | Mutation
Copy number alterations are putative.

Exercise

- Is there a statistically significant difference in overall and progression free survival between RB pathway altered vs. non-altered patients?



- HOME
- TUTORIALS
- NEWS
- FAQ
- DATA SETS
- ABOUT
- WEB API
- R/MATLAB
- NETWORKS
-
-
-

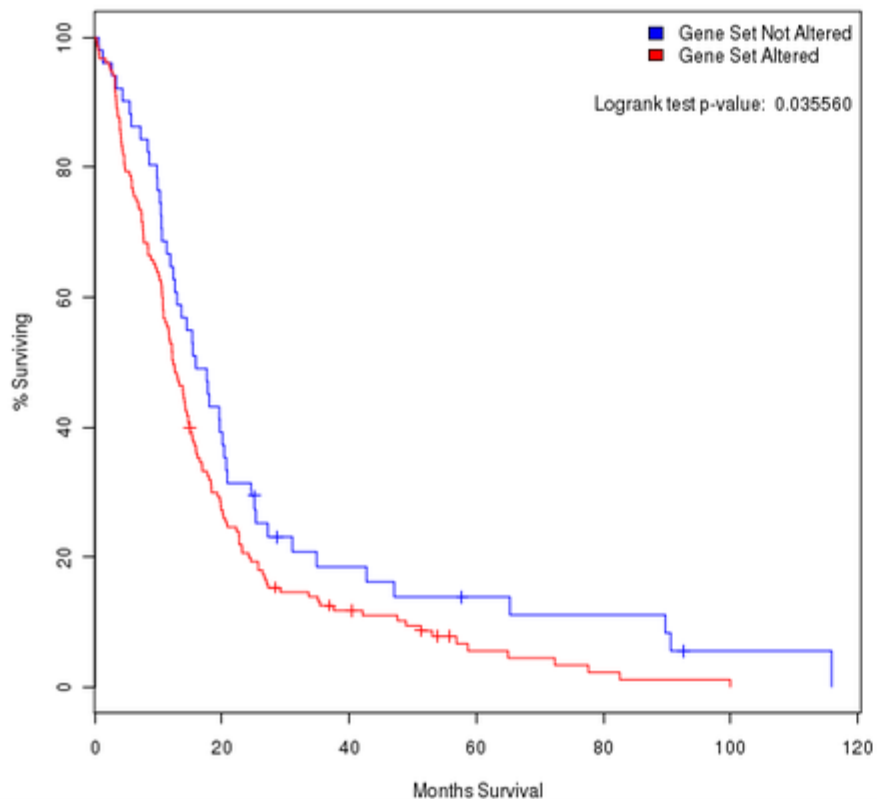
Gene Set / Pathway is altered in 75.2% of all cases.

Glioblastoma (TCGA, Nature 2008)/All Complete Tumors: (206)/Glioblastoma: RB Pathway (7 genes)/7 genes

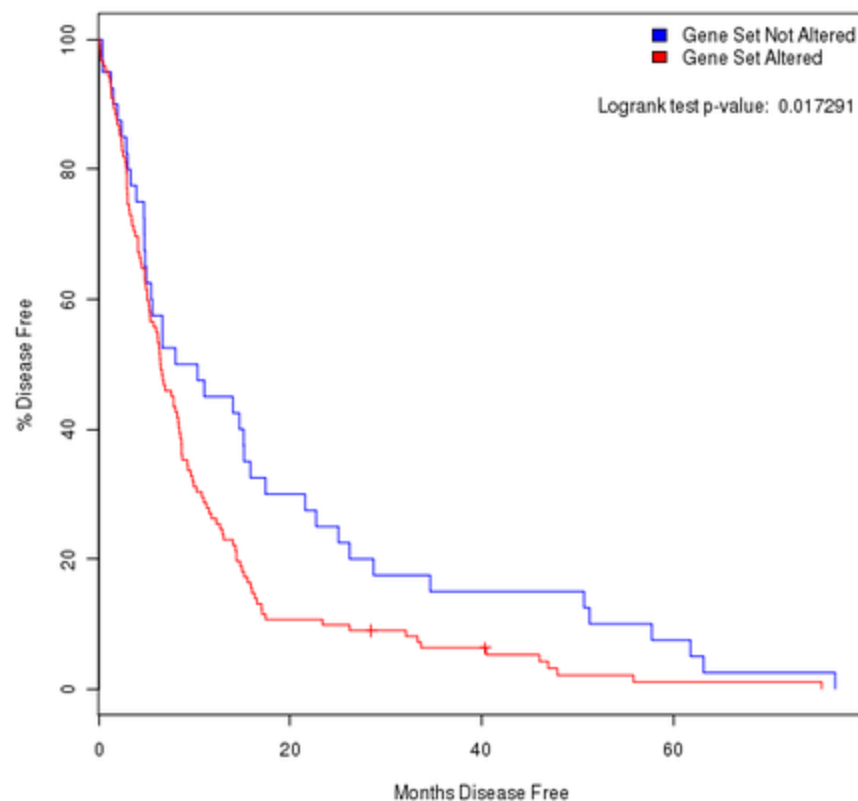
[Modify Query](#)

- Summary
- Network
- Plots
- Survival
- Mutual Exclusivity
- Mutations
- Event Map
- Download
- Bookmark

Overall Survival, Logrank Test P-Value: 0.03556 [\[PDF\]](#)



Disease Free Survival, Logrank Test P-Value: 0.017291 [\[PDF\]](#)



Exercise

- Which genes are altered mutually exclusively?



- HOME
- TUTORIALS
- NEWS
- FAQ
- DATA SETS
- ABOUT
- WEB API
- R/MATLAB
- NETWORKS
-
-
-

Gene Set / Pathway is altered in 75.2% of all cases.

Glioblastoma (TCGA, Nature 2008)/All Complete Tumors: (206)/Glioblastoma: RB Pathway (7 genes)/7 genes

► Modify Query

- Summary
- Network
- Plots
- Survival
- Mutual Exclusivity
- Mutations
- Event Map
- Download
- Bookmark

Gene	CDKN2B	CDKN2A	CDK4	RB1	CDKN2C	CCND2	CDK6
CDKN2B	---	0	0	0.021944	0.058545	0.241304	0.572941
CDKN2A		---	0.000006	0.023168	0.110337	0.348115	0.463442
CDK4			---	0.094052	0.686582	0.009202	0.052594
RB1				---	0.540223	0.262674	0.796155
CDKN2C					---	0.852462	0.887434
CCND2						---	0.057403
CDK6							---

p-values <0.05, as derived via Fisher's Exact test are outlined in red.
p-values are *not* adjusted for FDR.

Legend

strong tendency toward mutual exclusivity

some tendency toward mutual exclusivity

Exercise

- Find out which gene is most commonly mutated in breast cancer
- <https://confluence.broadinstitute.org/display/GDAC/Home>
- Find the distribution of correlations between DNA methylation and mRNA in BRCA