Name&ID: 28/04/2017

Worksheet for cBioPortal

Welcome to cBioPortal, dear friends!

Please, don't be scared of the longevity of the worksheet. Page 4 and 5 belong to *OPTIONAL* R-part.I've just explained every step in detail.Please,read carefully.Thanks!

QUERYING SINGLE CANCER STUDY

When you google "cBioPortal", you will see website of this portal on top of the results page. You need to click to reach the portal.

Once you click, you will see:
"Query" and "*Download"* options.
We'll start with the **Query**

In Query, you'll see options such as :

- *Select Cancer Study*
- *Select Genomic Profiles*
- *Select Patient/Case Set*
- *Enter Gene Set*

1st Please *Select Cancer Study* as **Breast Invasive Carcinoma,TCGA Cell,2015)**

2nd Please *Select Genomic Profiles* of

- ❏ *Mutations*
- ❏ *Putative Copy Number Alterations from GISTIC*
- ❏ *mRNA Expression Data*
  mRNA Expression (RNA seq RPKM)
  *enter a z-score of threshold +/- **1.5***
- ❏ *Protein Expression Z-score (default)*

3rd Please *Select Patient/Case Set* as **All Complete Tumors**

4rd Please *Enter Gene Set* by using Select From Recurrently Mutated Genes (MutSig)
Select all

5th Click to **Submit**

A New Page will be opened. In this new page, you will see titles such as

- *OncoPrint*
- *Mutual Exclusivity*
- *Plots*
- *Mutations*
- *Co-expression*
- *Enrichments*
- *Survival*
- *Network*
- *CN Segments*
- *Download*
- *Bookmark*

6th In *OncoPrint* menu,please write 3 highest mutated genes (based on the percentage values)

7th In *Mutual Exclusivity* menu, please write the number of significant mutations of gene pairs for mutual exclusivity and co-expression. They are expected or not, please briefly explain(hint: google the related pathway of the genes).(**1 sentence**)

8th In *Plots* menu, for horizontal and vertical axis, please plot TP53 putative copy number alteration vs TP53 mRNA expression by changing Gene to TP53.
9th Under *Mutations* menu,you'll see different hotspot domains of different elements related to the selected cancer type based on somatic mutation rate.
10th Under *Co-expression* menu,based on Pearson and Spearman correlation values, which gene/genes are highly correlated to **TP53**?

11th Under *Survival* menu, you'll see the Kaplan-Meier Estimation graph for the patients of breast cancer. Please ,comment on this figure briefly(no longer than **1 sentence**).

12th For *Network* menu, when you change the option of Hide Drugs to **Show FDA Approved Drugs**, please write name of one of ERB2 targeting drug shown in the network

13th Under *CN* menu, you can see the Copy Number of specific genes on the chromosome via Integrative Genomics Viewer (IGV)
For a detailed view, we need to download Cytoscope. If you wish,you can do it at your home.

CROSS CANCER STUDY
When you click on **Modify Query** on top-left of the page, you'll return to the page you start, which includes main options: **Query** and Download.

14th Now, we'll select **All** studies recorded by cBioPortal from the *Query* section.

15th **Only Mutation** will be selected as *Select Data Type Priority*

16th We need to clear the *Enter Gene Set*

17th We'll look for BRCA1 expression across all tumor types.So just enter **BRCA1** on *Enter Gene Set* part, then click to *Submit*.

A new page will open. In this new page, you'll see the alteration frequency of BRCA1 across all tumor samples.
18th So, Is breast cancer the only one having higher alteration of BRCA1 gene?(Please,briefly explain)
19th Under *Expression* menu,please check which **mutation type**(eg. missense, frameshift) is **dominated** in **BRCA1 for breast cancer**, please write it down?

SINGLE TUMOR-INDIVIDUAL PATIENT VIEW

When you click **Modify Query** on top-left of the page, please select just *one type of dataset* **20th** (eg. Breast Invasive Carcinoma,TCGA Cell,2015), and click on the ***Summary*** (rightmost side of the dataset)). Play with that menu. **What kind of information can you able to get from this section, please explain it briefly(1 sentence)**.

In the Summary Page of the tumor dataset you've selected, on top of the graphs, you'll see the number of patients/samples selected.

21th Plesae click on *View the Selected Patients*.

22th Now, you have a chance to analyze each patient individually. Please, play with that page, and write **which type of information you can get from this menu(in 1 sentence)**.

Finally

Please, briefly explain the possible usage/purpose of this site in your research in two or three sentences.

Please,explain the possible usage of this site for precision medicine in one short sentence. (hint:use info from Individual Patient Analysis Part)

----------Happly Ever After End----------

!!!OPTIONAL!!!

We have an optional section,now.It is time to combine R and cBioPortal.
Don'to worry, to play with coding part of this portal, I'll give all the codes required. You'll just copy and paste, and enjoy :)

CGDS-R package usage

23th
CRAN package of R has already included this **CGDS** package, you just need to **install** it from CRAN library of R(inside of R studio-install part).Let's start:

> **library(cgdsr)**

We need to connect with CBioPortal datasets by using this library:

> **cgdsObj= CGDS("http://www.cbioportal.org/")**

To make sure that you connect,you can test some of elements of cdgsObj object:

> **test(cgdsObj)**

If it says OK for every test it has done, then your object is ok.

Now, we wanted to get the studies cBİoPortal has:

>**getCancerStudies(cgdsObj)[,c(1,2)]**

We just want to 1st and 2nd column of cancer studies: first column shows the id of each study(datasets), and second column is the expanded version of this study.

You need to select correct and unique id of the study you want to continue with. (e.g Bladder Urothelial Carcinoma,TCGA,Provisional)

>**getGeneticProfiles(cgdsObj,"blca_tcga")[,c(1:2)]**

We just want to see fisrt two cloumns of genetic profile which'll show genetic profile id and genetic profile name,respectively).

To get the case lists(what we want to anayze: all tumors or tumor samples with cna data as we did in the Select Patient/Case Part) from this profile.

>**getCaseLists(cgdsObj,"blca_tcga")[,c(1:2)]**

We selected first two cloumns which include case list ids and case list names. For a detailed explanation, you may want to see third column.

Now, we want to get a graph like 8th step, plots part. We'll. Let's say you want to plot CNA alteration of BRCA1 and mRNA expression of BRCA1 on R-studio.

>getProfileData(cgdsObj,"BRCA1",c("blca_tcga_gistic","blca_tcga_rna_seq_v2_mrna"),"blca_tcga_all")[c(1:5),]

> brcaProfile<-
getProfileData(cgdsObj,”BRCA1”,c(“blca_tcga_gistic”,”blca_tcga_rna_seq_v2_mrna”),”blca_tcga_all”)
>boxplot(brcaProfile[,2]~brcaProfile[,1],main=”CNA status and mRNA Expression of BRCA1”, xlab=”CNA”, ylab=”mRNA”)
>stripchart(brcaProfile[,2]~brcaProfile[,1],vertical=T,add=T,method=”jitter”,pch=1,col=”yellow”)

Or you may want to see the expression profile of different genes. You can do it.
>getProfileData(cgdsObj,c(“BRCA1”,”TP53”),”blca_tcga_rna_seq_v2_mrna”,”blca_tcga_all”)[c(1:5),]
> plot (brcaProfile, main=”BRCA1 and TP53 mRNA Expression”, xlab=”BRCA1 mRNA”, ylab=”TP53 mRNA”)

You can also retrieve clinical data by using the function below
>getClinicalData(cgdsObj,”blca_all”)[c(1:3),]