cBioPortal Hands-on Introduction

Version 02/December/2019

What is cBioPortal for Cancer Genomics?

The tool cBioPortal is an open-source software, which is hosted by Health-RI as a data integration platform where processed data from different research studies have been made available for view and query.

What is the purpose of this hand-out?

In this training material, we aim to guide you with your first cBioPortal experience. The examples are using available public studies to introduce how to use the basic features of this tool. The estimated time to go through these examples is 30 minutes.
Let's get started!

1. Logging in

In order to become a user of cBioPortal, you will need two accounts for the following systems:
https://trait.topdesk.net/
https://cbioportal.health-ri.nl/

Requesting a topdesk account:
Send an Email to servicedesk@health-ri.nl where you request to have an account for trait.topdesk.net to become a user of cBioPortal on Health-RI server. You will receive account details within one working day.

Requesting a cbioportal account:
Go to https://trait.topdesk.net/.
Click on "Use Self-Service Portal".
Login with your account details.
Click on 'Services and Request Forms' --> 'cBioPortal' --> 'cBioPortal request new user'.
Finally, fill in the form where you will be asked to provide a Google-affiliated account. Optionally you can already request access to studies that you might want to have access to.
Following the submission of this request, the Servicedesk will contact you within one working day to confirm your access to the tool.
Go to https://cbioportal.health-ri.nl/ and log-in using your Google Account.

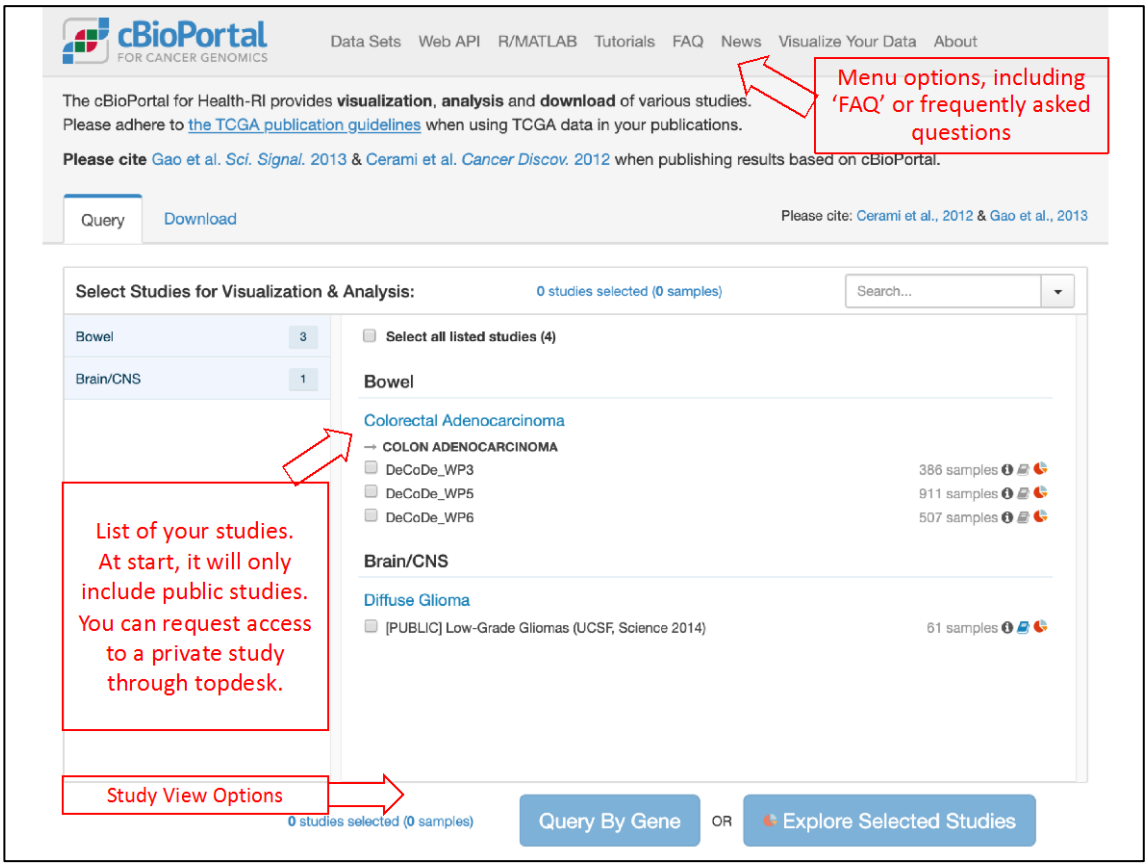
How does a Google account work?

When using your GMAIL for logging-in into cBioPortal, the server will only use your Google Identity to see who is logged-in. cBioPortal will not get to know your Google Account password or receive access to your email and information. No patient data will be stored on your Google Account.

Note

In case you are following this training in a workshop session, you will be provided with separate instructions to login with a test account.

2. cBioPortal interface



In case you need to access a private study (which may be hidden to your account at the moment), please go to trait.topdesk.net.

How to get access to a private study in cBioPortal:
Click on "Use Self-Service Portal".
Login with your account details.
Click on 'Services and Request Forms' --> 'cBioPortal' --> 'cBioPortal give user access to existing study'.
Finally, fill in the form where you will be asked to provide the name of the study.
Upon receiving a confirmation of access from the Servicedesk, refresh the page and you will be able to see and use the study.

3. Study view

There are two ways in how you can explore a study:

1) Study exploration = 'Explore Selected Studies'
2) Gene panel-based view = 'Query By Gene'

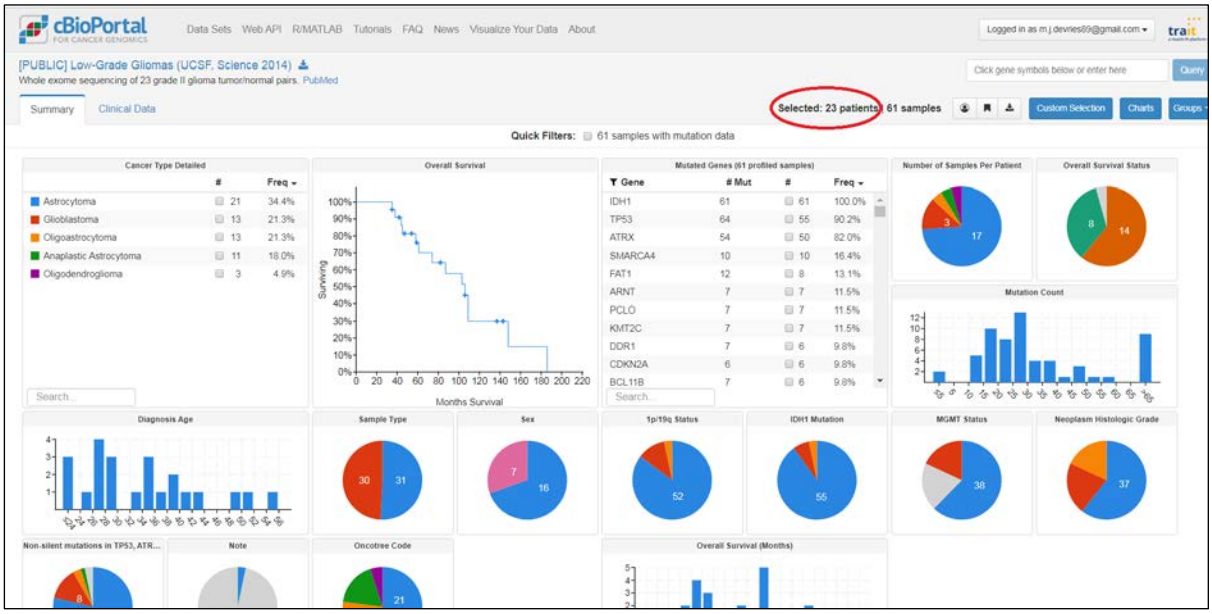
In this handout, we will use the public study **'Low-Grade Gliomas (UCSF, Science 2014)'.**

Once you select the study name, you should be able to clink on one of the following options:

4. Study exploration

Select the study and click on 'Explore Selected Studies'

You should now see the following screen, where clinical and mutation data are visualized in charts (e.g., pie-charts, tables and graphs).



You can use these interactive charts to filter the data based on the different available details.

Example
Select 'Astrocytoma' in the 'Cancer Type Detailed' section. Note how all other charts adapt to this selected criterion. Next, play around by extending your selection. In addition to 'astrocytoma', you can click on 'IDH1' in 'Mutated Genes' section and 'Male'.

Clearing your selection can be done at the top by using the button 'Clear All Filters'.

Having learned this, you should now be able to answer the following: *(the answers are at the end of this handout)*

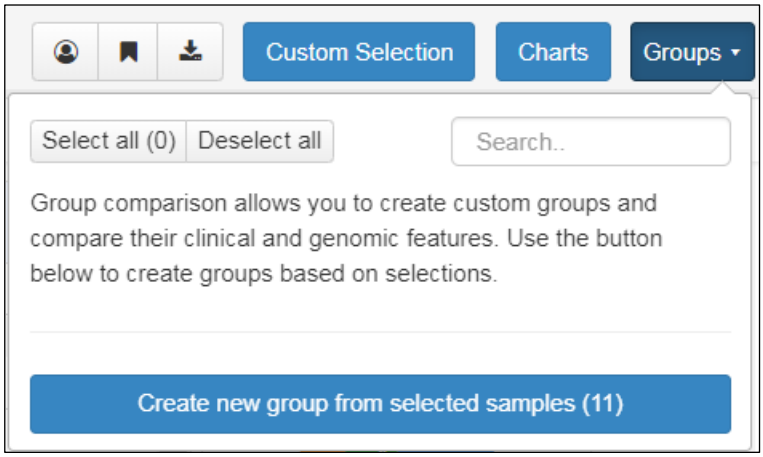
> **Question 1:** Are there any patients in this dataset: stated as 'alive' and have a glioblastoma with 19q deletion?

For further information, you can check **Tutorial #1: Single Study Exploration** on the publically available Tutorials webpage or download the tutorial via the following link.

5. Study exploration: Group comparison

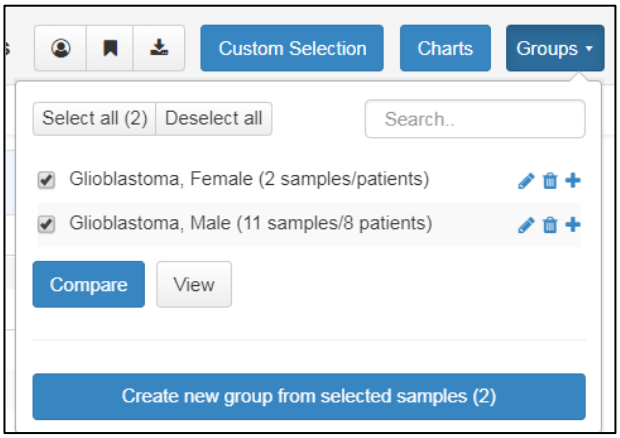
In the previous paragraph you have made a first cohort selection. However, it might be handy to compare this selected group to another one. This is possible in cBioPortal through the 'Group Comparison' functionality.

Example
First, **clear all filters and select all males with a glioblastoma.**
Next, go to the right corner, select 'Groups' and click on 'Create new group from selected samples'.



Give it a name (e.g. "Glioblastoma, Male") and click on 'Create'.

Next, clear all filters and make a new selection: all females with a glioblastoma. Again, go to 'Groups' and create another group. Now, having both groups selected in the group selection view. Click on 'compare' and a new tab will open in your browser.



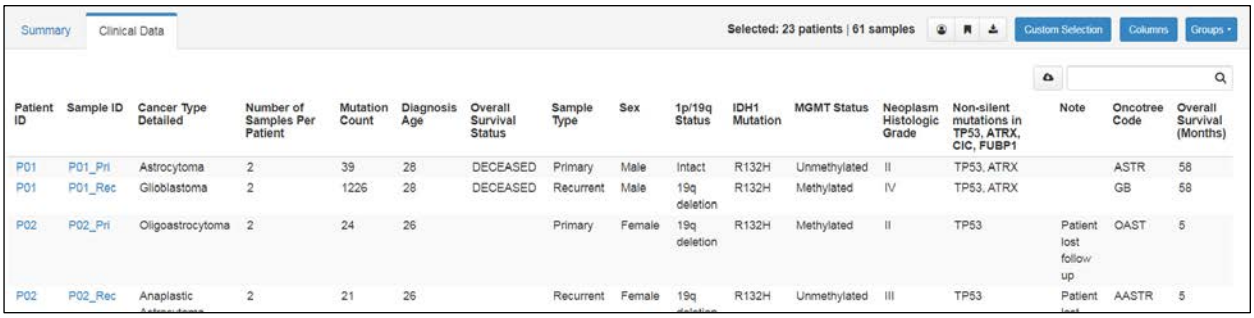
In the new screen, you have new tabs for: 'Overlap', 'Survival', 'Clinical' and 'Mutations'. Go through each tab and see which information is displayed. Now, you should now be able to answer question 2.

> **Question 2:** Which group displays a better survival: Astrocytoma patients with a methylated or unmethylated MGMT status? Is the result statistically significant?

For further information, you can check **Tutorial #6: Group Comparison** on the Tutorials webpage or download the tutorial via the following link.

6. Study exploration: Patient level view

Next, we would like to see the data of an individual patient. **Close the group comparison view, clear all filters again and click on 'Clinical Data'.** You should now be able to see a tabular view of the dataset with each row representing one sample per one patient.



Patient ID	Sample ID	Cancer Type Detailed	Number of Samples Per Patient	Mutation Count	Diagnosis Age	Overall Survival Status	Sample Type	Sex	1p/19q Status	IDH1 Mutation	MGMT Status	Neoplasm Histologic Grade	Non-silent mutations in TP53, ATRX, CIC, FUBP1	Note	Oncotree Code	Overall Survival (Months)
P01	P01_Pri	Astrocytoma	2	39	28	DECEASED	Primary	Male	Intact	R132H	Unmethylated	II	TP53, ATRX		ASTR	58
P01	P01_Rec	Glioblastoma	2	1226	28	DECEASED	Recurrent	Male	19q deletion	R132H	Methylated	IV	TP53, ATRX		GB	58
P02	P02_Pri	Oligoastrocytoma	2	24	26		Primary	Female	19q deletion	R132H	Methylated	II	TP53	Patient lost follow up	OAST	5
P02	P02_Rec	Anaplastic	2	21	26		Recurrent	Female	19q	R132H	Unmethylated	III	TP53	Patient	AASTR	5

At the left, **click on patient number 1: P01.** A new tab will open in your browser with the Patient ID and basic information and overview of all samples available for this patient.

At the top a timeline is being displayed, in the middle a genetic profile and at the bottom a tabular overview of which mutations have been found.

Hover over with your mouse on everything you see to notice where additional information is being displayed, for example on the colored dots in the timeline. You may Answer the following two questions.

> **Question 3:** What was the last treatment that was given to the patient?
>
> **Question 4:** What hotspot mutation seemed to have developed over time since it was not detected in the first biopsy?

For further information, you can check **Tutorial #3: Patient View** on the Tutorials webpage or download the tutorial via the following link.

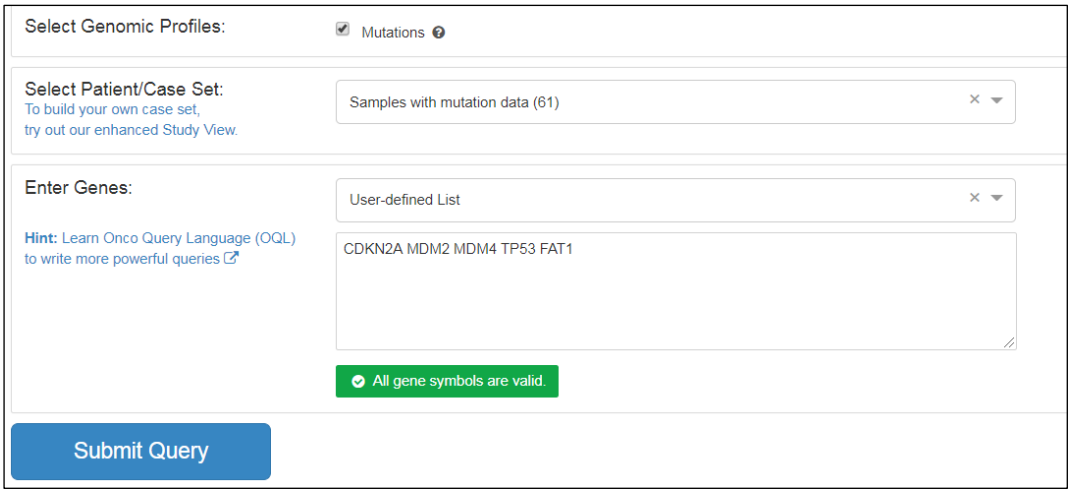
7. Gene panel-based view

Go back to the start page of www.cbioportal.health-ri.nl and select the same study again: *Low-Grade Gliomas (UCSF, Science 2014)*. However, this time we will visualize the same study in a genetic way. Therefore, **click on 'Query By Gene'.**

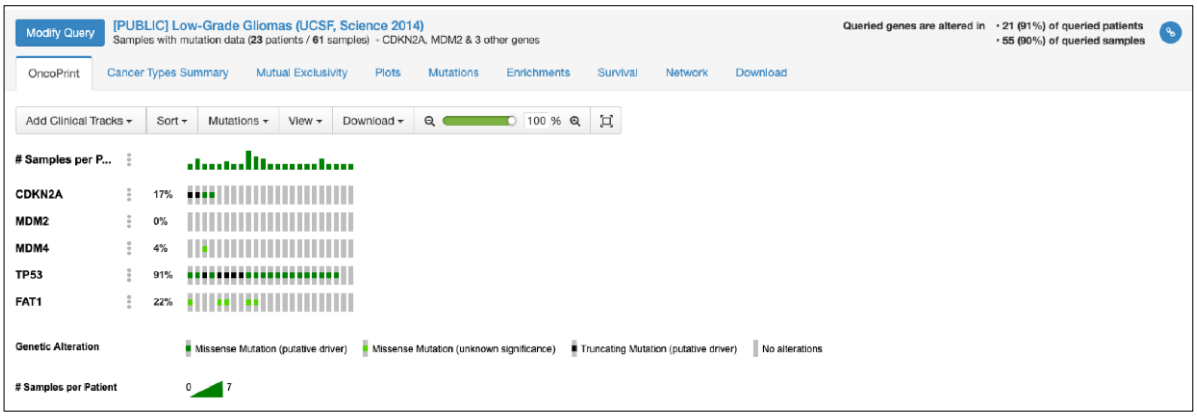
A new screen appears that wants you to define/modify your query request with four options: 'Selected Studies', 'Select Genomic Profiles', 'Select Patient/Case Set', and 'Enter Genes'.

Next to 'Select Genomic Profiles' you will be able to see the genetic data types that are available. For this study, we have mutation data only but when a study contains RNA sequencing, proteomics or anything else it will display this information here.

Next to 'Enter Genes' we will have to define your gene panel list. For example, **click on the drop-down menu of 'Enter Genes'** and you will see a list of popular panels. **Click on 'Glioblastoma TP53 Pathway'** and you will see the names of the four genes that are involved in this pathway. Let's also add another gene manually 'FAT1', by typing the name next to TP53 (note: the dropdown menu will now indicate 'user-defined list'). In case you mistype FAT1, an error message will appear which will indicate which gene symbol is incorrect.

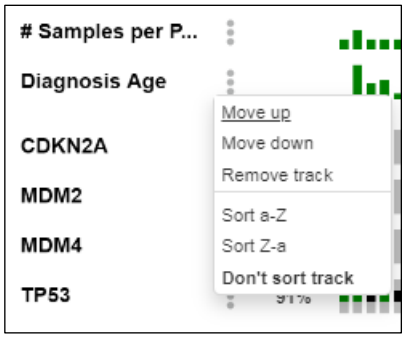


Next, **click on 'Submit Query'** and a new screen appears with several tabs (OncoPrint, Cancer Types Summary, Mutual Exclusivity, Plots, Mutations, Enrichments, Survival, Network, and Download). In this training, we will make use of two of them but please feel free to explore the rest as well.



8. Gene panel-based view: OncoPrint

Go to the OncoPrint tab and notice how each gene from the selected panel is visualized. Each grey rectangle represents a patient in which the mutation (when present) is shown. **Click on 'Add Clinical Tracks' and select 'Diagnosis Age'**. Click on the three grey dots that are next to 'Diagnosis Age' (see screenshot), **move the diagnosis age to top (Move up) and order them from low to high (Sort a-Z)**. Notice how all the other bars are re-organized as well.



Next, **add another clinical track: 1p/19q status.** Order this one as well and **hover over the OncoPrint with your mouse** in order to answer the following:

Question 5: What is the (diagnosis) age of the patient with a 1p and 19q co-deletion?

Finally, **download the view in PNG format** using the 'download' button.

9. Gene panel-based view: Plots

Next, click on the tab 'Plots' where you can plot variables on 2 axes. For each axis, you can specify a 'Data Type' and a set of features depending on the data type selected.

For example, on the horizontal axis you can choose "Data Type = Clinical Attribute", and below the sub-head of "Clinical Attribute", you can use the drop-down list to choose **'overall survival (months)'**. As you can see, the plot changes according to the changed parameter.

Next, you can make further changes to the vertical axis:
"Data Type = Mutation"
"Group Mutations by = Mutated vs Wild-type"
"Gene = FAT1"

Next, **hover over with your mouse** in the boxplot that has been created.

> **Question 6:** Do the two patients with FAT1 mutation and relative long overall survival have the same protein change?

For further information, you can check **Tutorial #2: Single Study Query** on the Tutorials webpage or download the tutorial via the following link.

Summary

Having learned all of this, you have become familiar with the basics of cBioPortal. You have learned how to explore a study in multiple ways: from overall view, to an individual view down to a single gene. The answers and explanations to the questions are on the next page. It is recommended to read the explanations when your answer was incorrect.

Please note that we have covered cBioPortal only partially but feel free to explore the rest of the publicly available datasets as well. In case you want to upload your own data to cBioPortal, please send an email to servicedesk@health-ri.nl and we will help you out.

Further online resources

There are two additional tutorials on Tutorials webpage:

- Tutorial #4: Virtual Studies
- Tutorial #5: Onco Query Language (OQL)

The FAQs on cBioPortal include:

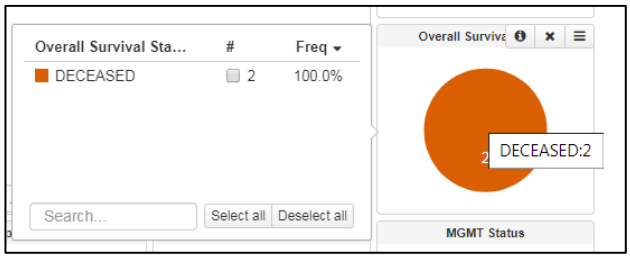
- General Questions
- Data Questions
    - General
    - TCGA
    - DNA (Mutations, Copy Number & Fusions)
    - RNA
    - DNA Methylation, and
    - Clinical Data
- Analysis Questions
- Results View

Answers to included questions & brief explanations

Question 1: Are there any patients in this dataset stated as 'alive', who have a glioblastoma with 19q deletion?

Short answer: No

Explanation: Make a cohort selection by selecting 'Glioblastoma' and '19q deletion' at 1p/19q status, resulting in only two patients. Hover your mouse over the pie chart of Overall Survival Status to see that they have both deceased.



Question 2: Which group displays a better survival: Astrocytoma patients with a methylated or unmethylated MGMT status? Is the result statistically significant?

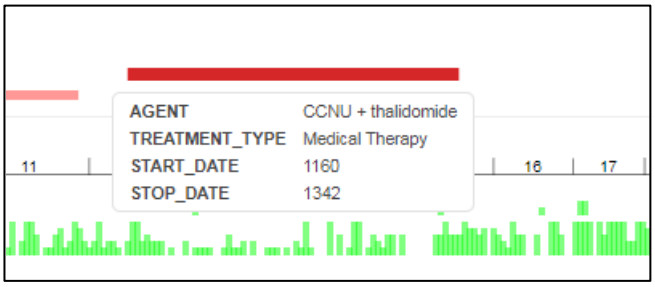
Short answer: practically the same, with statistically insignificant p-value of 0.978

Explanation: Make a first cohort of astrocytoma patients with methylated MGMT status (10 patients) and a second cohort of astrocytoma patients with unmethylated MGMT status (3 patients). Save them both separately as a group. Click on 'compare' and go to the 'Survival' tab. Look at the legenda to see what the colored lines correspond to. The p-value is been given in this legend as well.

Question 3: What was the last treatment that was given to the patient?

Short answer: CCNU + Thalodomide

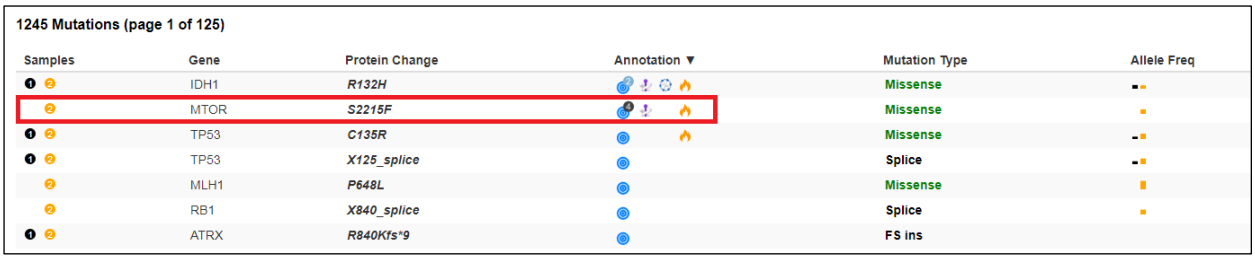
Explanation: In the patient view of PO1, watch in the timeline. The latest treatment given is the red bar, which will present a legend when you hover over your mouse.



Question 4: What hotspot mutation seemed to have developed over time since it was not detected in the first biopsy?

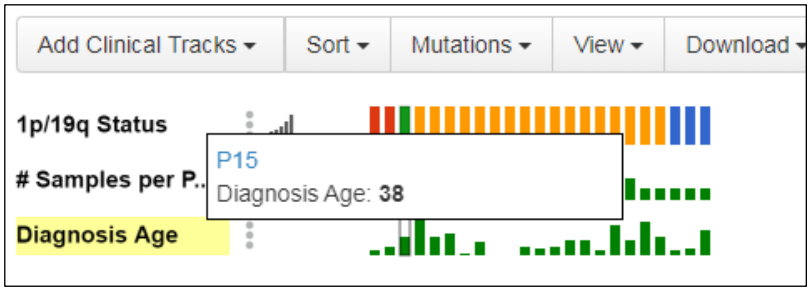
Short answer: S2215F

Explanation: Look below the timeline and CNA overview to see a table with mutations. At the most left you see a column named 'Sample Dots'. In the timeline you see that the black dots represent the first biopsy and the yellow dot the second biopsy. In the mutation table, the 'Annotation' column gives 'fire' symbols for hotspot mutations (hover over your mouse). Notice that there are only three hotspot mutations from which only one was detected in the second biopsy. The protein change states this to be S2215F.



Question 5: What is the (diagnosis) age of the patient with a 1p and 19q co-deletion?

Short answer: 38
Explanation: Look at the legend of 1p/19q status to see that the green-colored rectangle represents the 1p/19q co-deletion. Next, hover over your mouse to the green bar in 'Diagnosis Age' that is below the green rectangle. A legend appears stating the age of this patient.



Question 6: Do the two patients with FAT1 mutation and relative long overall survival have the same protein change?

Short answer: No
Explanation: In the boxplot given, two patients at the 'Mutated' bar can be observed with relative long overall survival compared to the other. Hover over your mouse to see an information box appearing. In here, you can see the protein change of the FAT1 gene. One patient has A1229T protein change while the other has G1687E.