cBioPortal Tutorial #6: Group Comparison

Compare clinical and genomic features of user-defined groups of samples/patients

Last update: June 20, 2019

Tutorial Objectives

- Explain what the new group comparison functionality enables
- Delineate the different ways to define groups and enter the group comparison view
- Highlight potential use cases for group comparison functionality

Group Comparison Overview

Group Comparison is a new suite of analysis features which allow you to compare the clinical and genomic features of user-defined groups of samples.

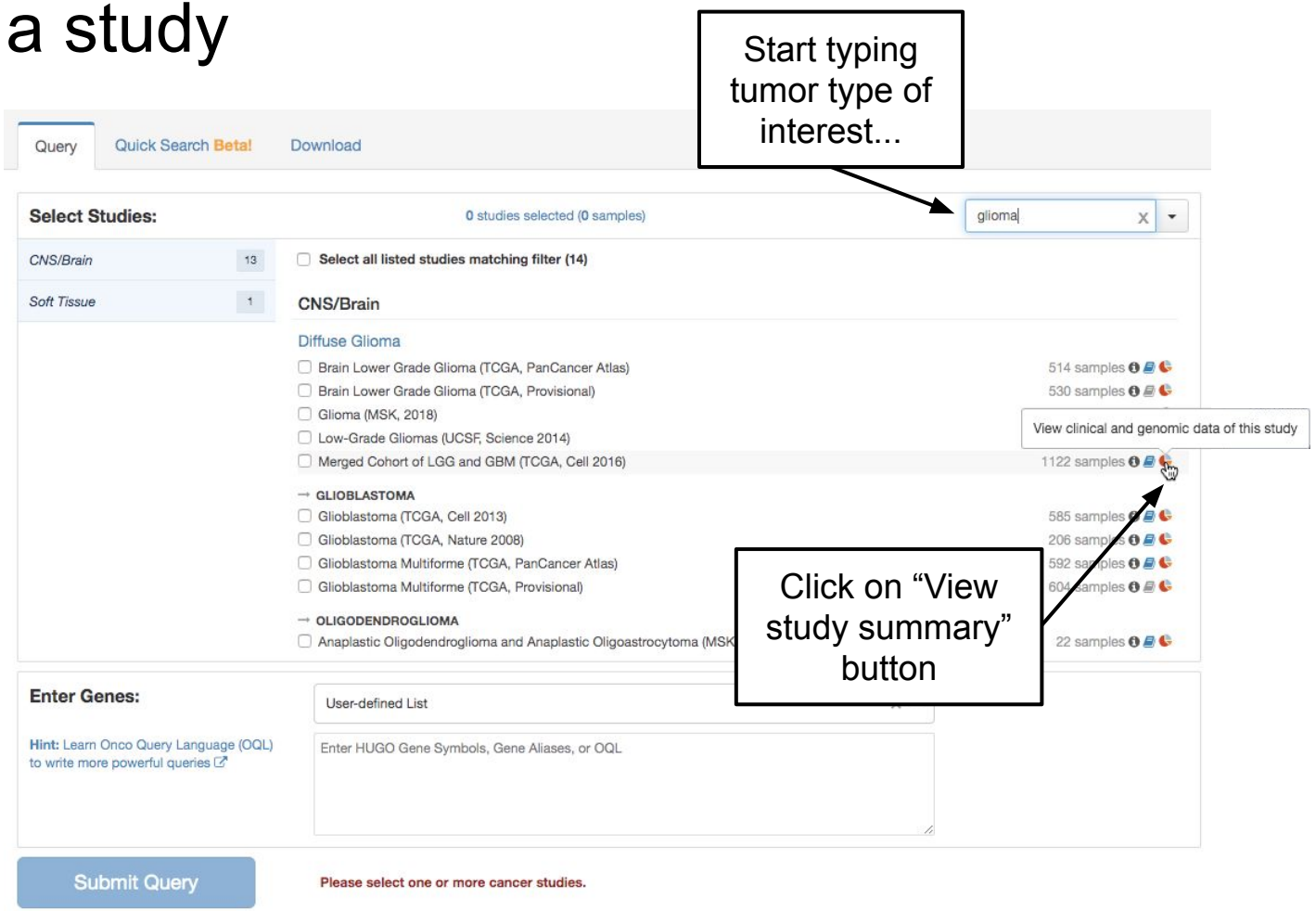
Groups can be defined in Study View based on any clinical or genomic features.

Groups can also be defined within the Group Comparison page based on the union/intersection of the compared groups.

In this tutorial, we will demonstrate two different ways to enter the group comparison view.

We begin by selecting a study of interest and viewing it in the Study Summary page.

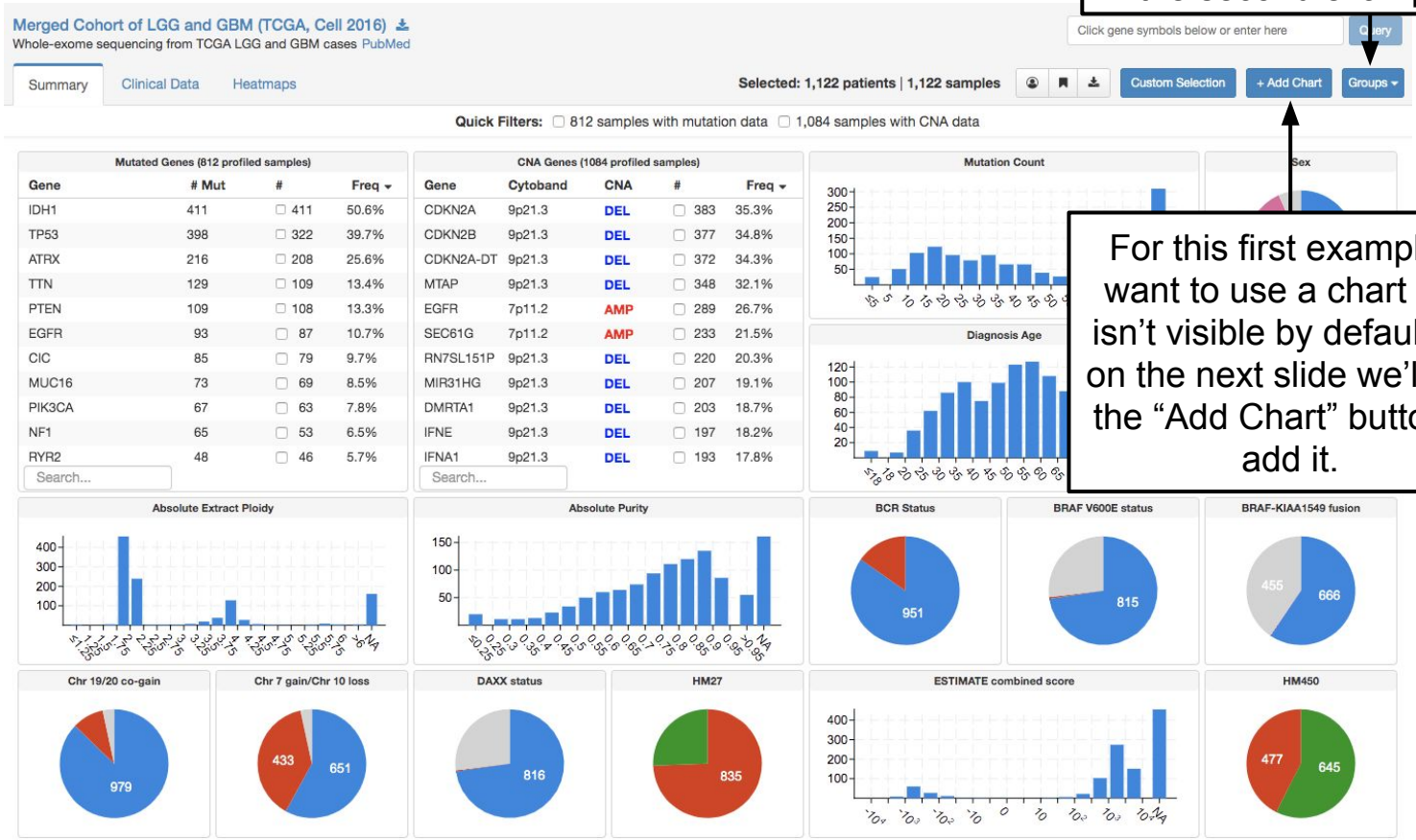
Select a study



Link to this page

Study View

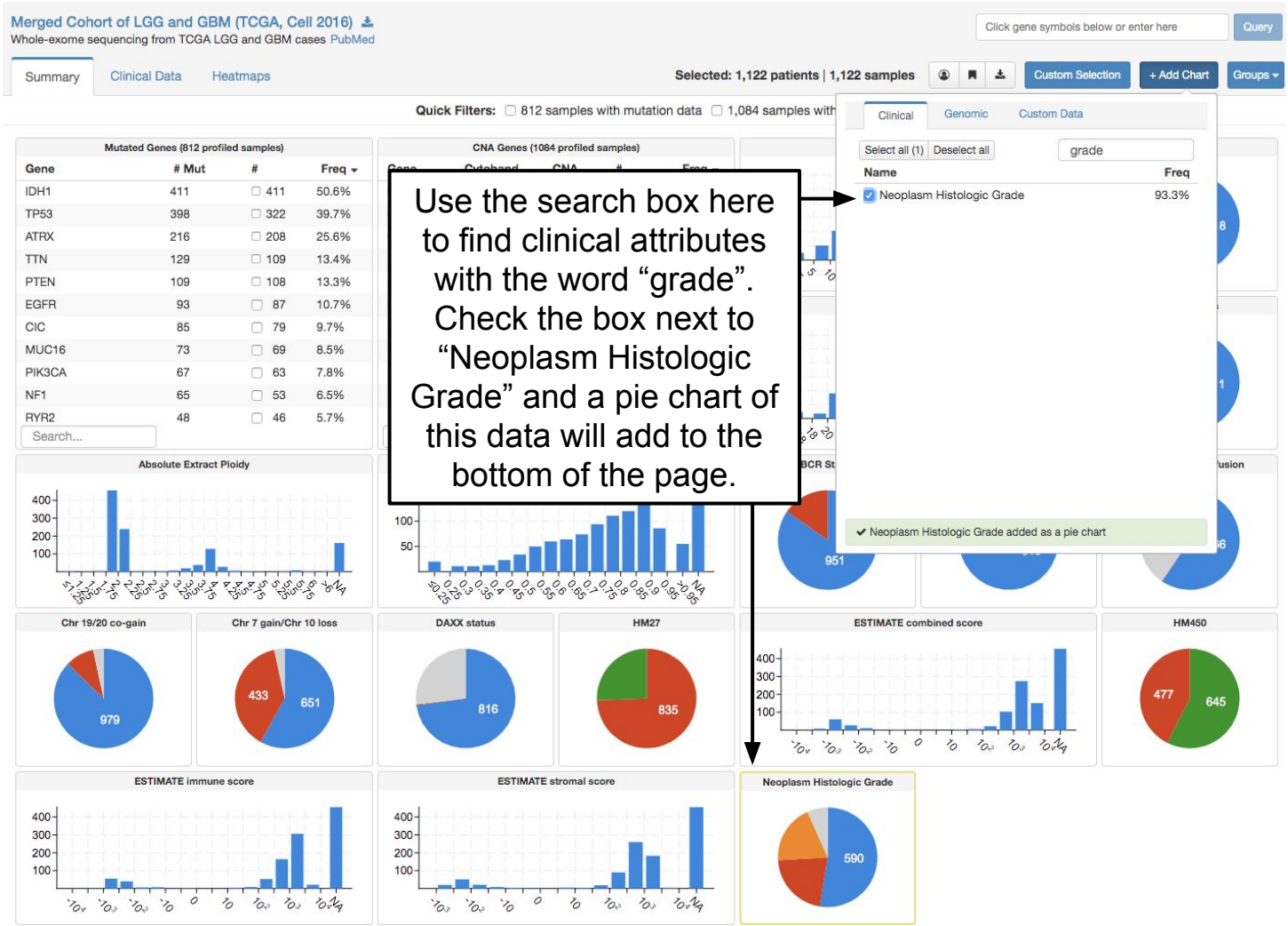
Notice this new “Groups” button. We’ll use this in the second example.



For this first example, I want to use a chart that isn’t visible by default, so on the next slide we’ll use the “Add Chart” button to add it.

[Link to this page]

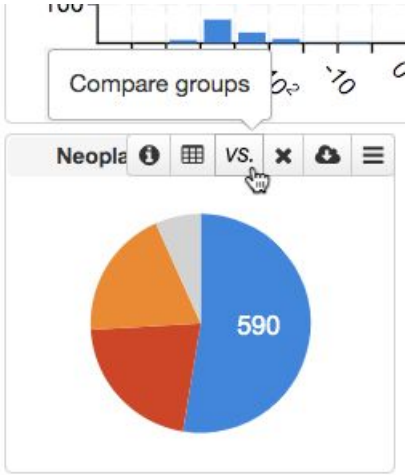
Study View



[Link to this page]

Study View

Let’s compare samples of different histologic grades. Hover over the “Neoplasm Histologic Grade” pie chart and notice the new vs. button. We’re going to click on this, and it will bring us to the new group comparison page where we can compare the clinical & genomic features of samples/patients by grade.



Link to this page

Group Comparison: Header

All group comparison pages share the same header:

The attribute used to create the groups.

Each tab has specific functionality. We'll go through these one-by-one over the next few slides.

The original study. Click to return to study view.

The available groups. Click on a group name to include or exclude it from analysis. Click the "x" to remove the group from the comparison session. Groups can also be reordered by dragging the group name.

Link to this page

Group Comparison: Overlap

The Overlap tab shows which samples or patients may overlap among the selected groups. In this example, we can see that there is no overlap in samples or patients. In the next example, we'll look at how overlapping samples/patients are managed.

This view can also be used to create additional groups. We'll do this later.

Link to this page

Group Comparison: Survival

The Survival tab shows a Kaplan-Meier plot of Overall Survival or Disease/Progression-free Survival based on the selected groups. This tab will only be visible when the original study contains survival data.

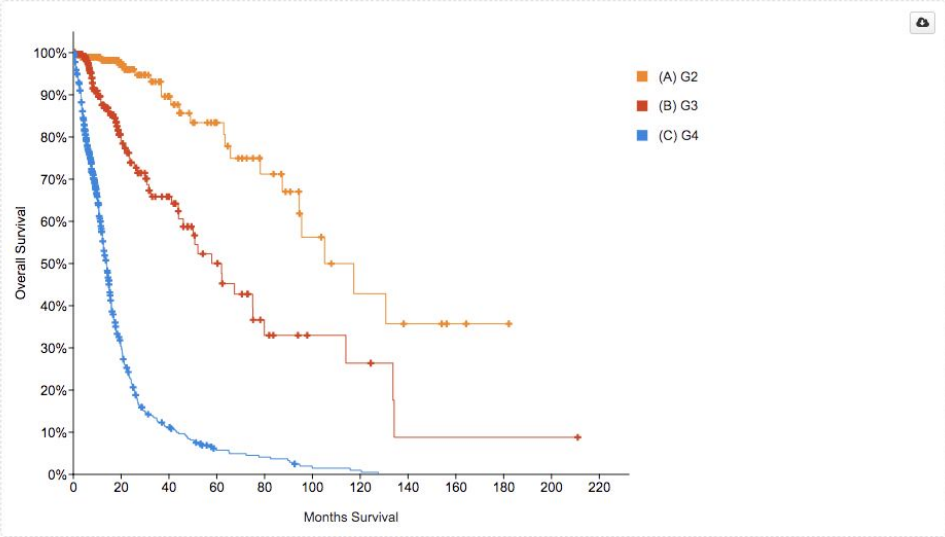
Merged Cohort of LGG and GBM (TCGA, Cell 2016)

Groups from ***Neoplasm Histologic Grade***

Groups: (drag to reorder) (A) G2 (216) (B) G3 (241) (C) G4 (590) Select all | Deselect all

Overlap | Survival | Clinical | Mutations | Copy-number

Overall Survival Kaplan-Meier Estimate (Overall patient survival status.)

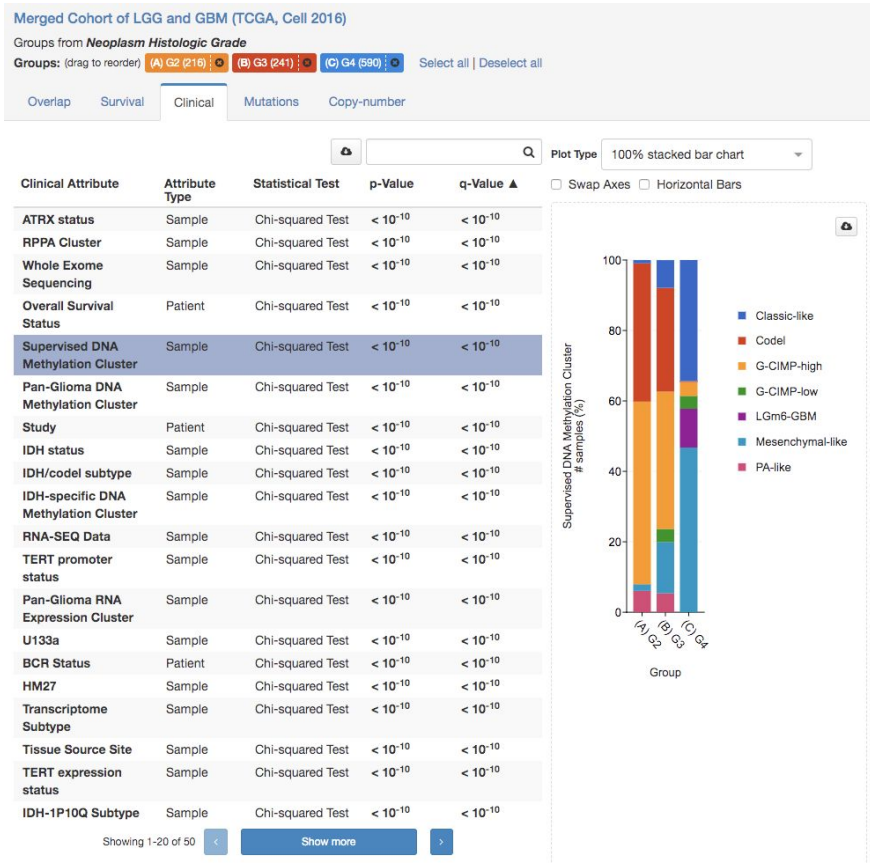


	Number of Cases, Total	Number of Cases, Deceased	Median Months Survival
(A) G2	**216**	**22**	**105.1**
(B) G3	**241**	**58**	**62**
(C) G4	**589**	**444**	**13.9**

Link to this page

Group Comparison: Clinical

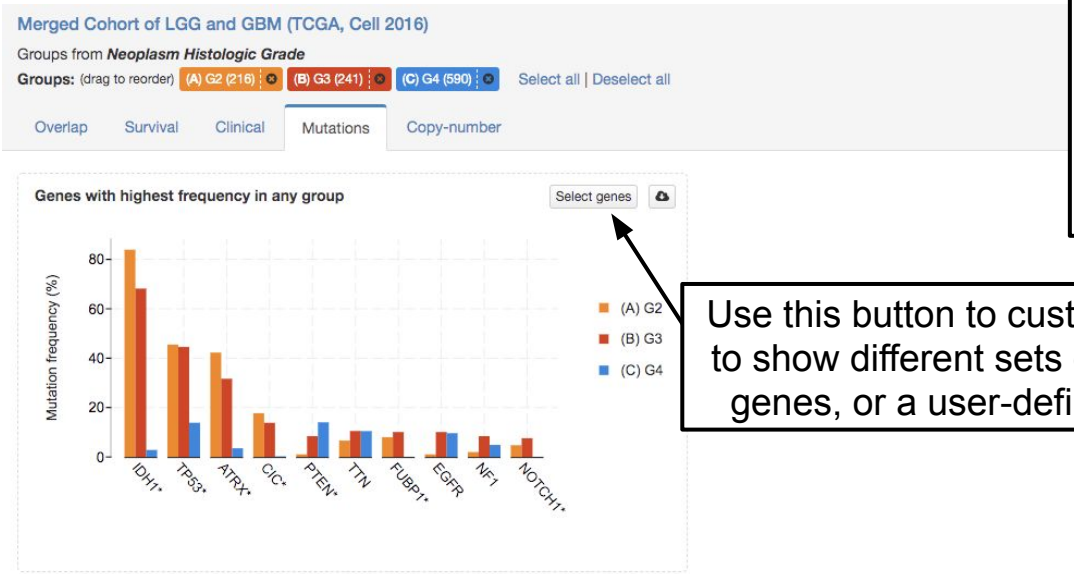
The Clinical tab shows all the same clinical attributes that are present in Study View. Select a clinical attribute in the table (Supervised DNA Methylation Cluster is selected here) and a plot will appear to the right with the distribution of that clinical attribute across the selected groups.



Link to this page

Group Comparison: Mutations

The Mutations tab compares the frequency of mutations in genes across the selected groups. The visible plots change depending on how many groups are selected. This screenshot shows the view with 3 or more groups selected.



Use this button to customize the plot to show different sets or numbers of genes, or a user-defined gene list

Mutations

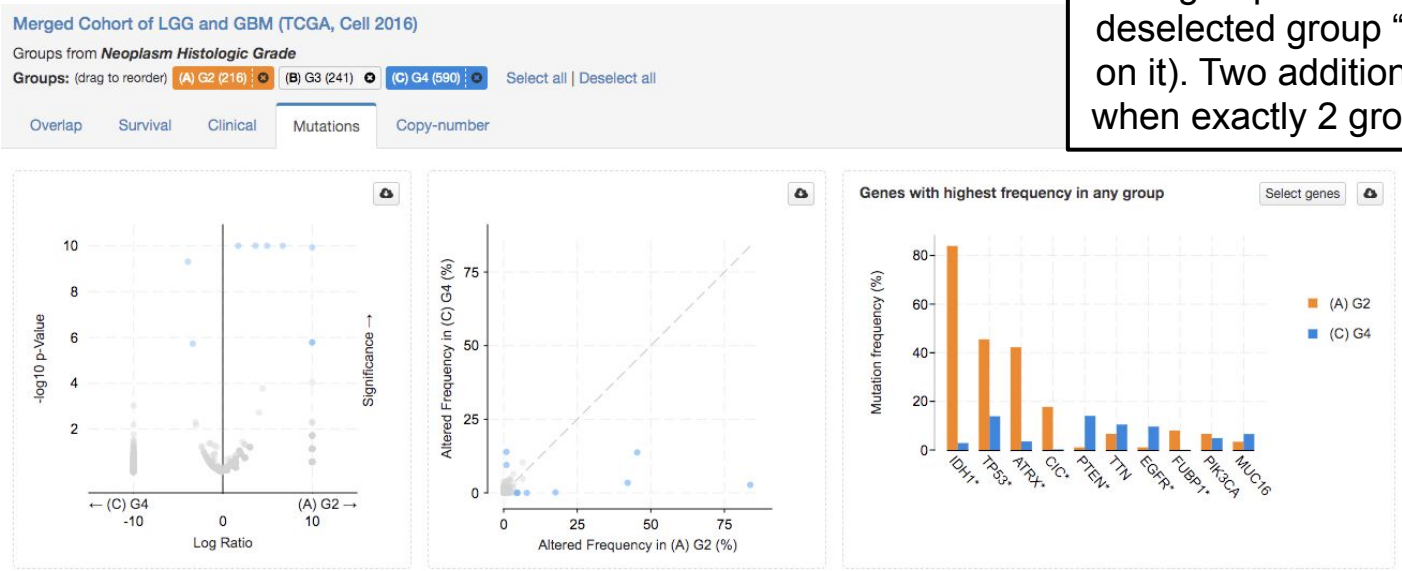
Select enriched groups | Significant only | Columns

Gene	Cytoband	(A) G2	(B) G3	(C) G4	p-Value	q-Value ▲	Most enriched in
ATRX	Xq21.1	91 (42.13%)	76 (31.54%)	20 (3.39%)	0.00	0.00	(A) G2
CIC	19q13.2	38 (17.59%)	33 (13.69%)	1 (0.17%)	0.00	0.00	(A) G2
IDH1	2q34	181 (83.80%)	164 (68.05%)	16 (2.71%)	0.00	0.00	(A) G2
TP53	17p13.1	98 (45.37%)	107 (44.40%)	81 (13.73%)	0.00	0.00	(A) G2
FUBP1	1p31.1	17 (7.87%)	24 (9.96%)	0 (0.00%)	5.63e-13	1.04e-9	(B) G3
NOTCH1	9q34.3	10 (4.63%)	18 (7.47%)	0 (0.00%)	1.47e-9	2.277e-6	(B) G3

Link to this page

Group Comparison: Mutations

This screenshot shows the view with 2 groups selected (notice that I deselected group “(B) G3” by clicking on it). Two additional plots are shown when exactly 2 groups are compared.



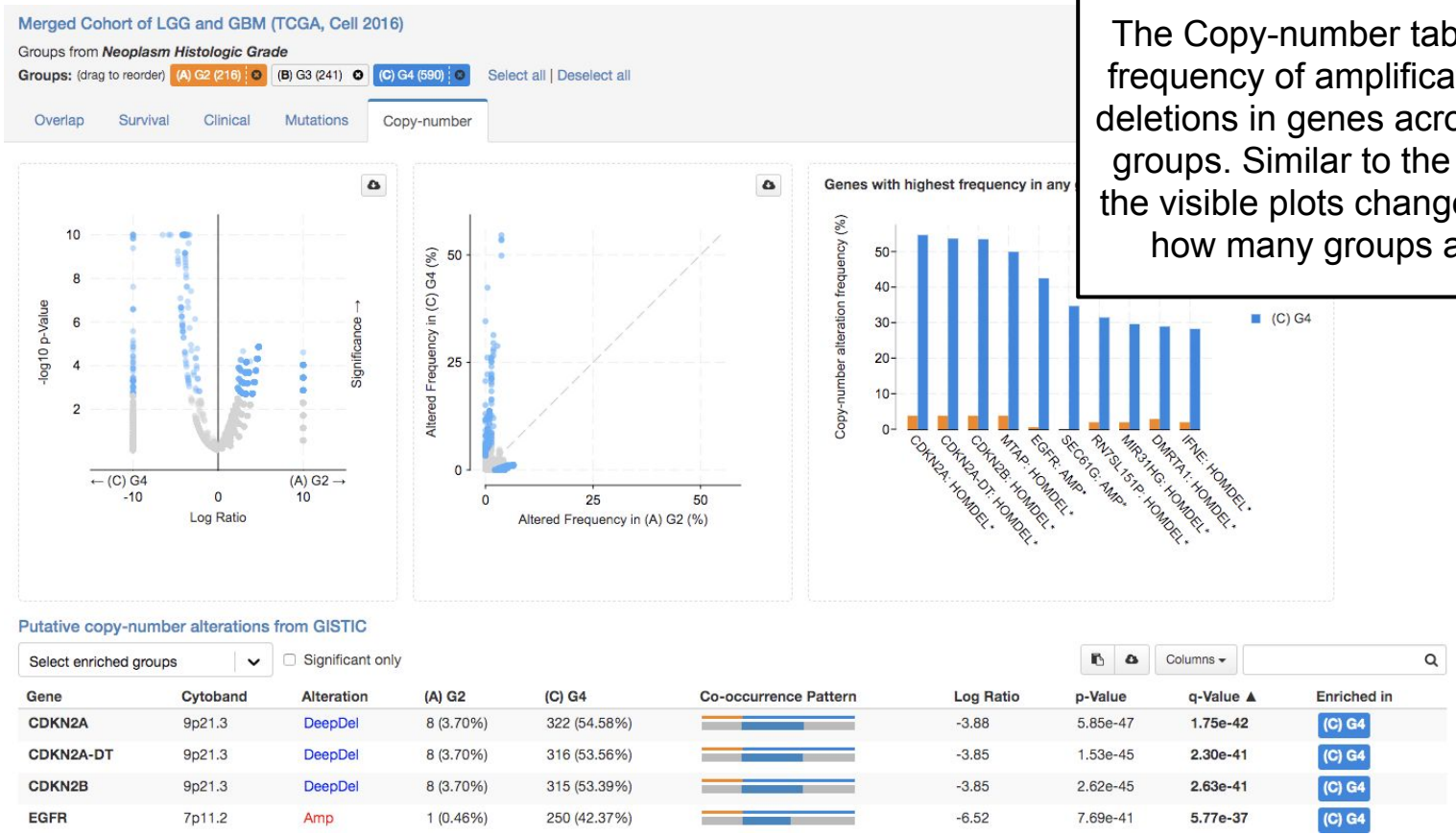
Mutations

Select enriched groups | Significant only | Columns

Gene	Cytoband	(A) G2	(C) G4	Co-occurrence Pattern	Log Ratio	p-Value	q-Value ▲	Enriched in
IDH1	2q34	181 (83.80%)	16 (2.71%)		4.95	1.49e-122	**1.13e-118**	(A) G2
ATRX	Xq21.1	91 (42.13%)	20 (3.39%)		3.64	3.98e-40	**1.51e-36**	(A) G2
CIC	19q13.2	38 (17.59%)	1 (0.17%)		6.70	4.27e-22	**1.08e-18**	(A) G2
TP53	17p13.1	98 (45.37%)	81 (13.73%)		1.72	5.41e-20	**1.03e-16**	(A) G2
FUBP1	1p31.1	17 (7.87%)	0 (0.00%)		>10	1.18e-10	**1.79e-7**	(A) G2
PTEN	10q23.31	2 (0.93%)	82 (13.90%)		-3.91	4.93e-10	**6.24e-7**	(C) G4

Link to this page

Group Comparison: Copy-Number



The Copy-number tab compares the frequency of amplifications and deep deletions in genes across the selected groups. Similar to the Mutations tab, the visible plots change depending on how many groups are selected.

Putative copy-number alterations from GISTIC

Select enriched groups | Significant only | Columns

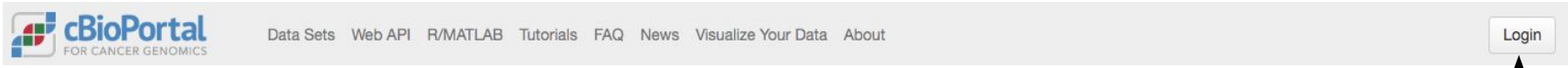
Gene	Cytoband	Alteration	(A) G2	(C) G4	Co-occurrence Pattern	Log Ratio	p-Value	q-Value ▲	Enriched in
CDKN2A	9p21.3	DeepDel	8 (3.70%)	322 (54.58%)		-3.88	5.85e-47	1.75e-42	(C) G4
CDKN2A-DT	9p21.3	DeepDel	8 (3.70%)	316 (53.56%)		-3.85	1.53e-45	2.30e-41	(C) G4
CDKN2B	9p21.3	DeepDel	8 (3.70%)	315 (53.39%)		-3.85	2.62e-45	2.63e-41	(C) G4
EGFR	7p11.2	Amp	1 (0.46%)	250 (42.37%)		-6.52	7.69e-41	5.77e-37	(C) G4

Link to this page

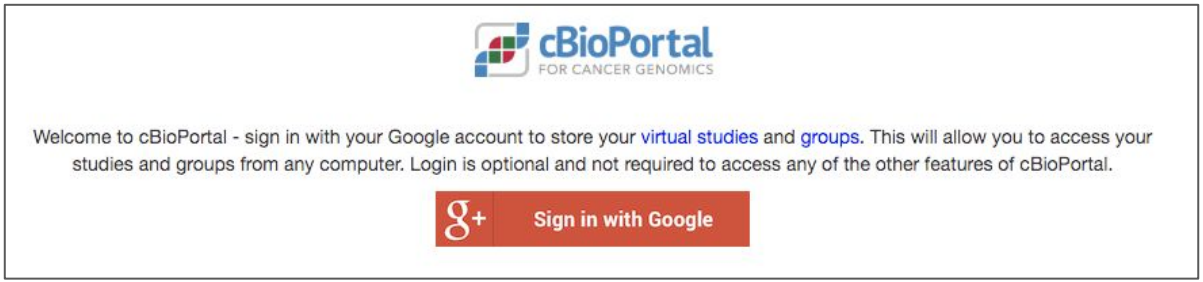
For the second approach to group comparison, we will define our own groups.

To do this, you must be logged in.

Logging in



This button is in the header on all pages. Click on it. It will bring you to a page where you can sign in (or register) using any google account.

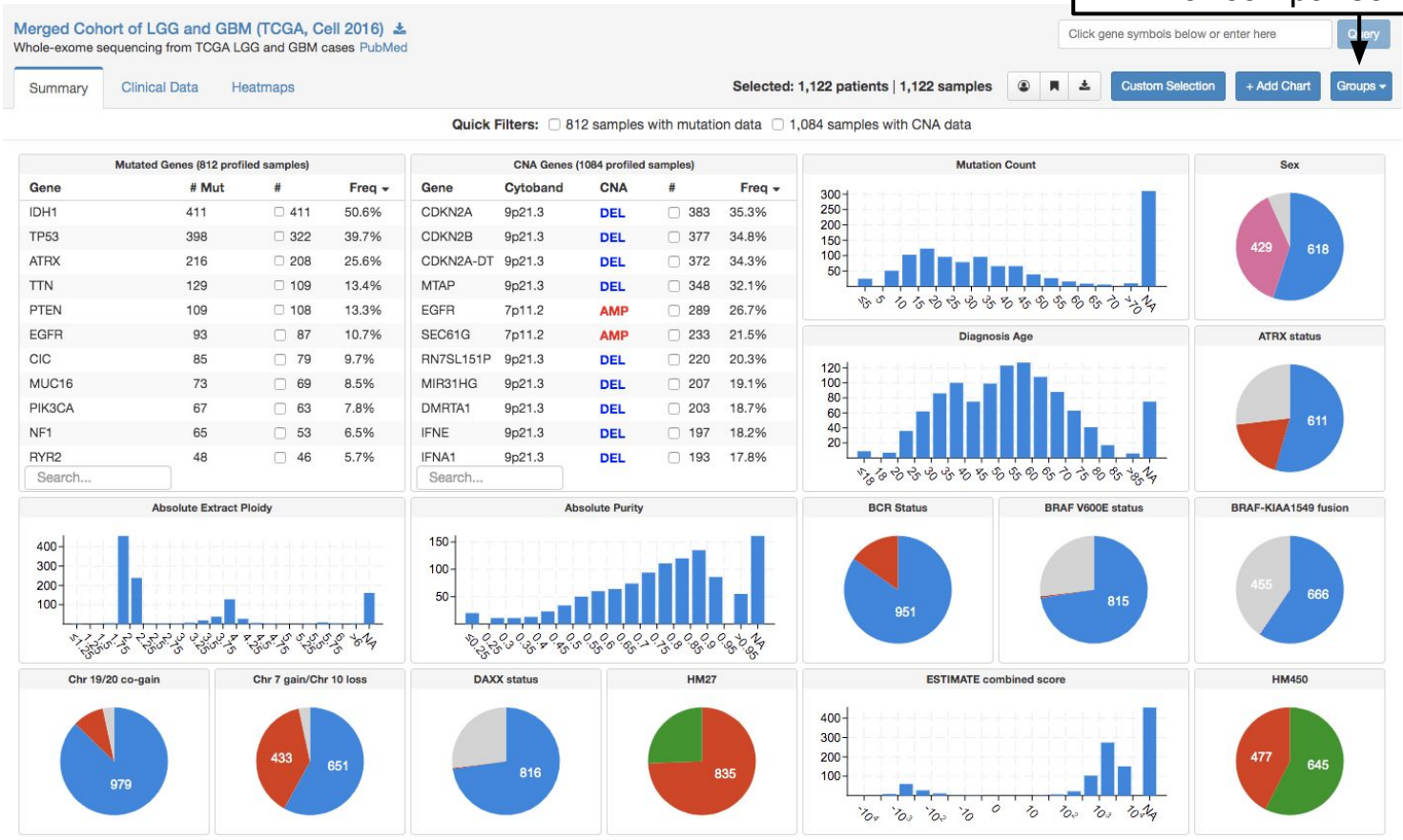


Link to this page

Now that we’re logged in,
let’s go back to study view.

Study View: Defining Groups

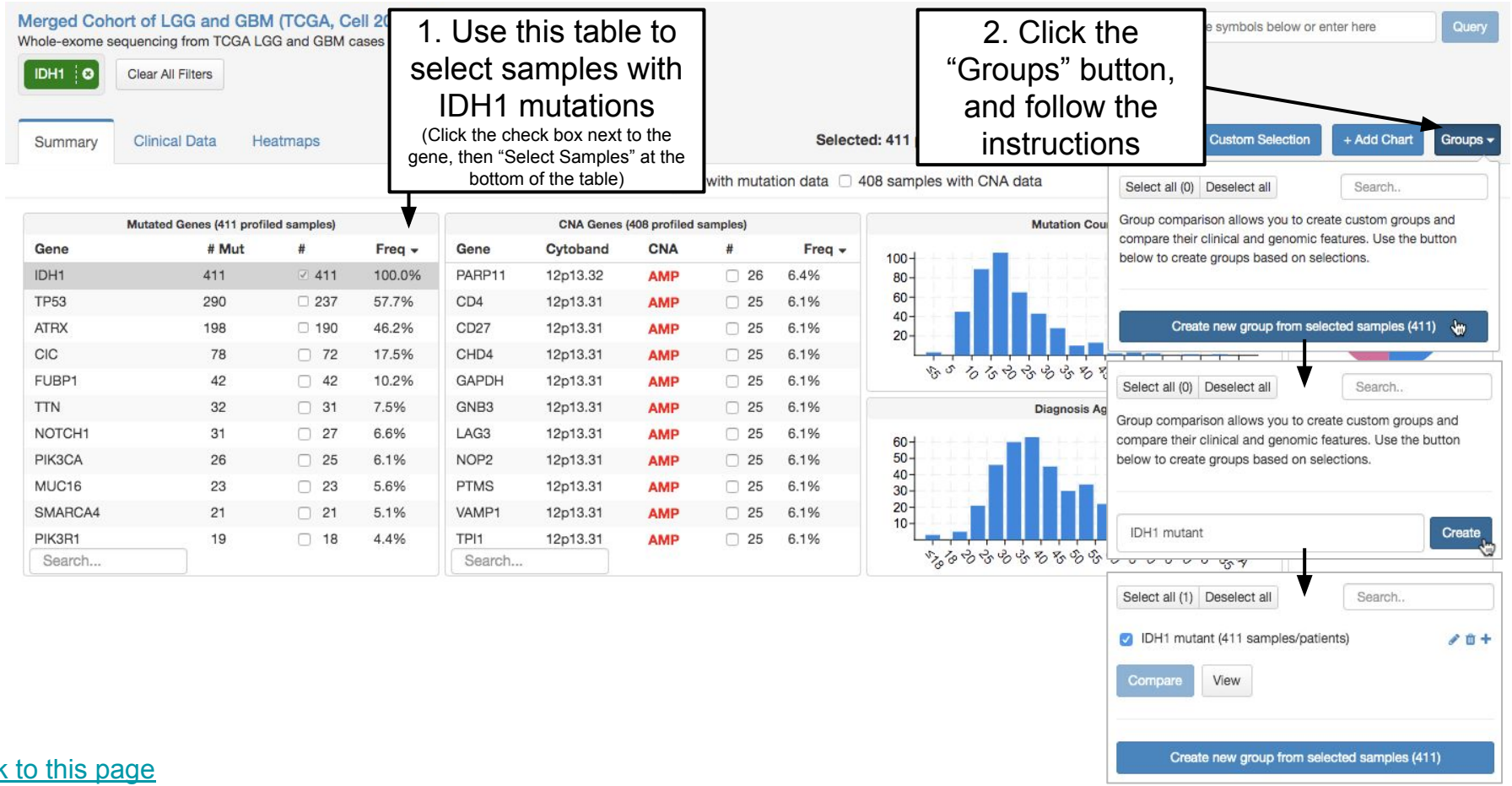
Let's use this "Groups" button to define groups for comparison.



[Link to this page]

Define Groups by Filtering

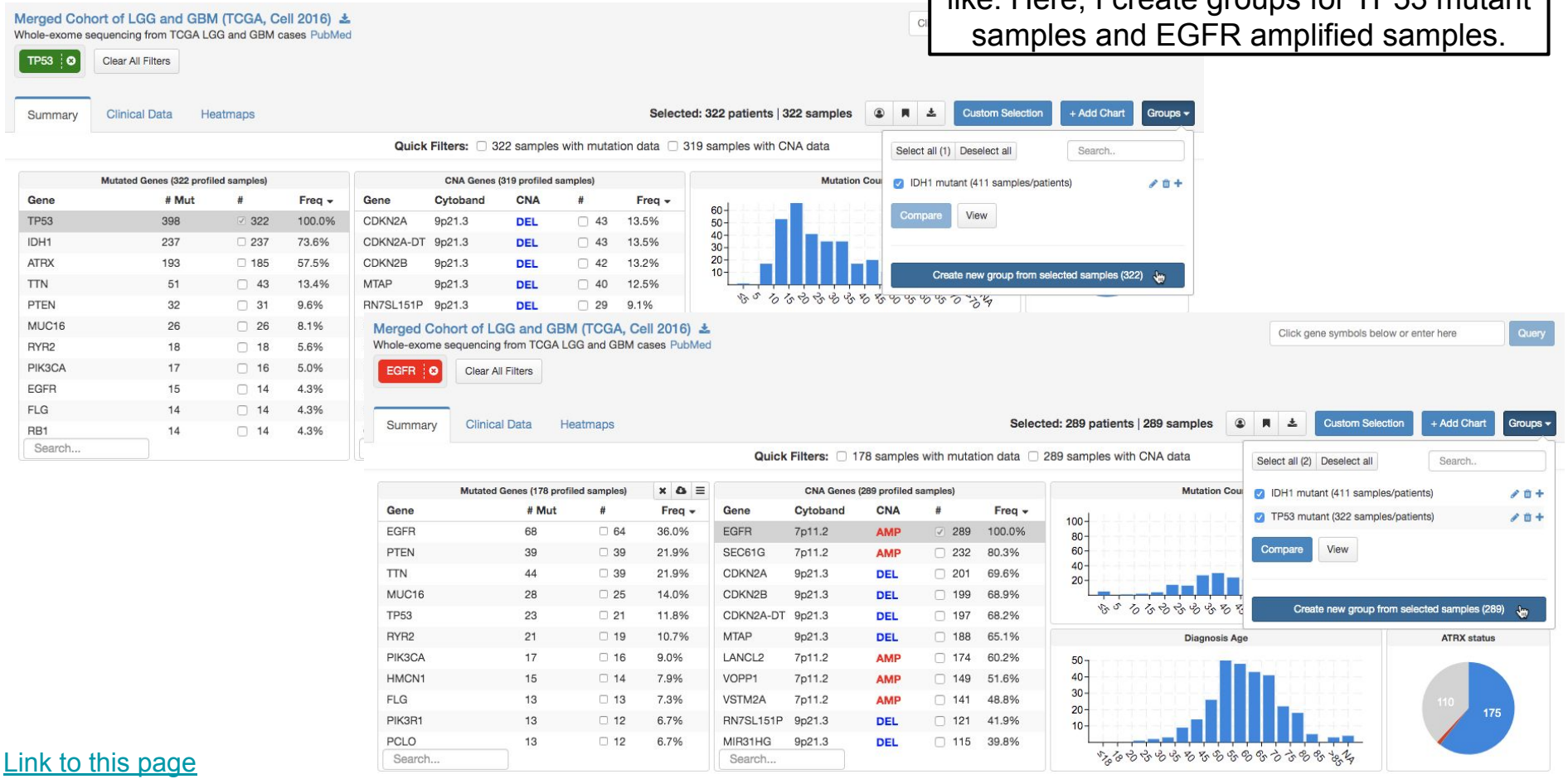
We define groups by applying filters in study view. Here, I select samples with IDH1 mutations:



[Link to this page]

Define Groups by Filtering

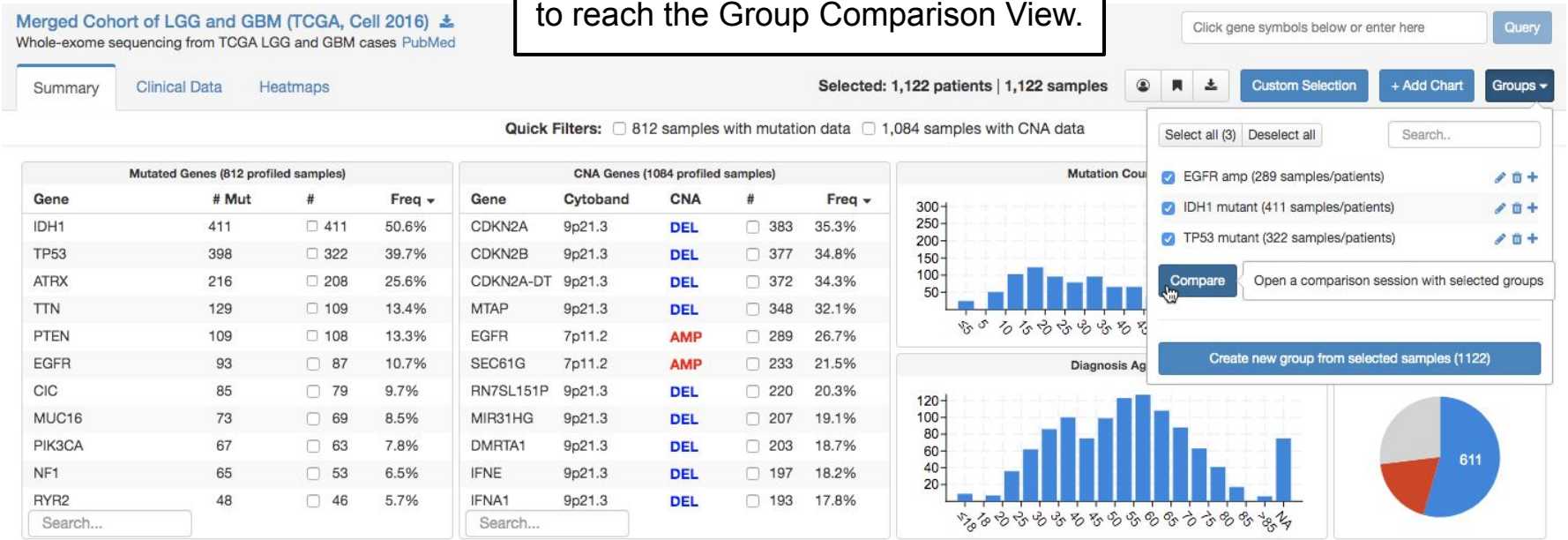
Clear that filter, and continue to apply new filters to define as many groups as you like. Here, I create groups for TP53 mutant samples and EGFR amplified samples.



Link to this page

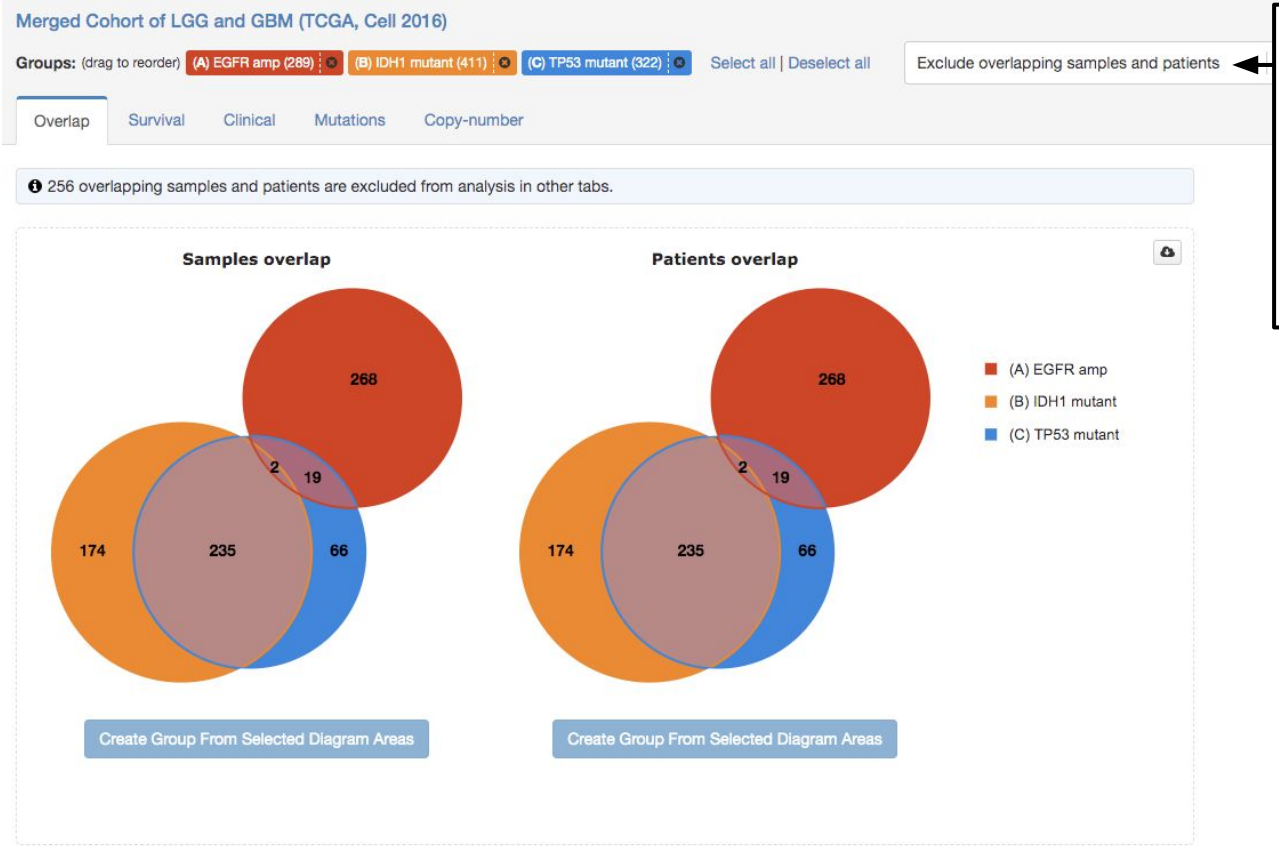
Compare User-Defined Groups

Once all groups are defined, select them and click the “Compare” button to reach the Group Comparison View.



[Link to this page]

Group Comparison: Overlapping Samples/Patients

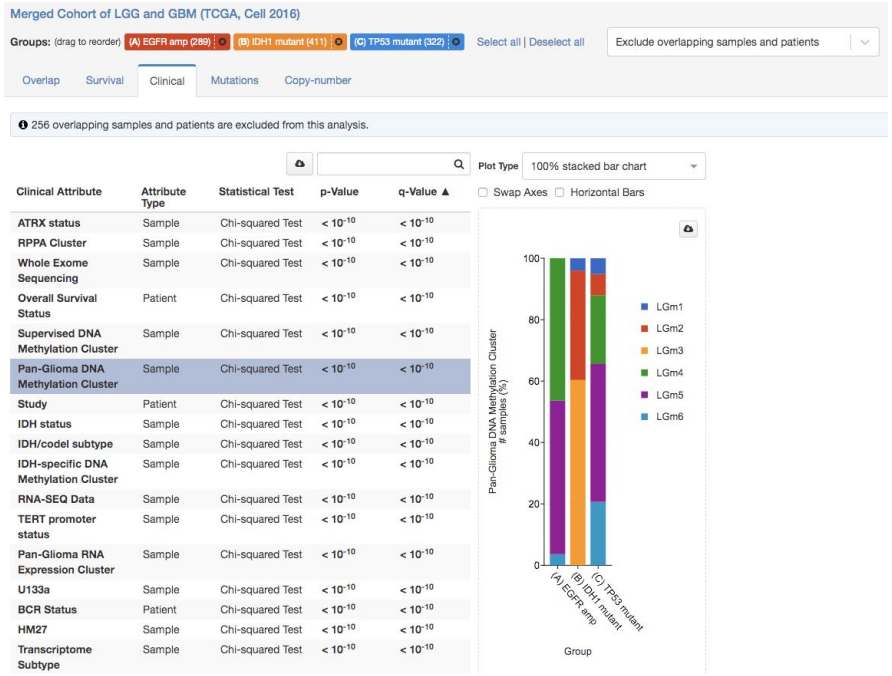


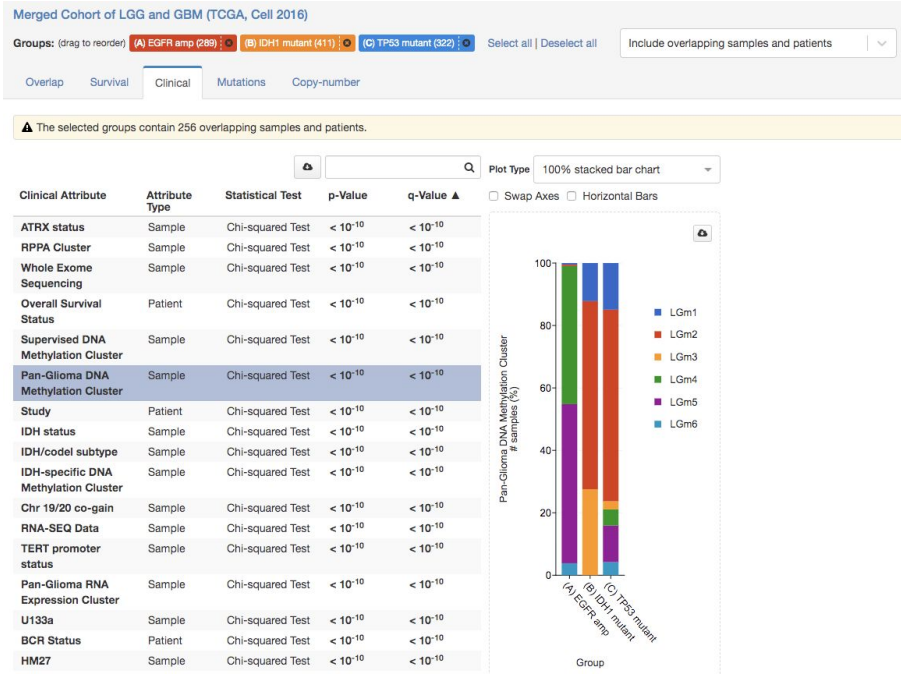
When samples/patients overlap among groups, a drop-down appears in the header which allows you to decide to exclude (default) or include those overlapping samples/patients in the analysis.

Link to this page

Group Comparison: Overlapping Samples/Patients

When samples/patients overlap among groups, each tab will include a warning message to make clear how those samples/patients are handled. This can be changed at any time by using the drop-down menu in the header of the page.

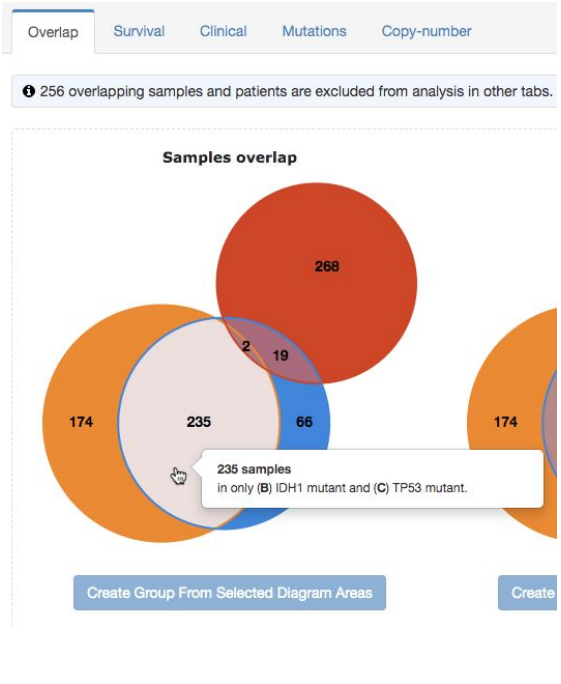




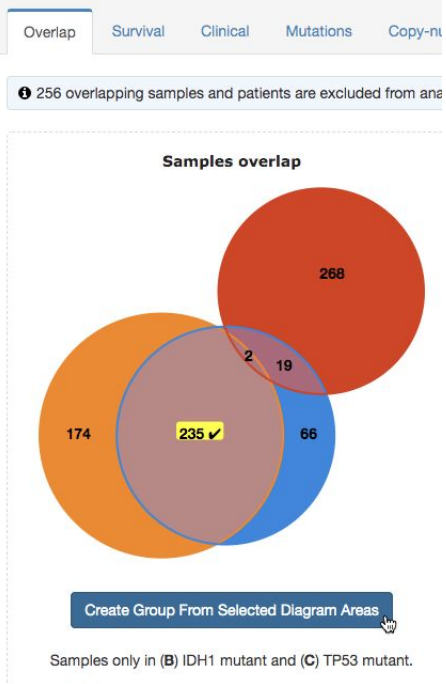
[Link to this page]

The Overlap tab can also be used to create new groups. For example, let’s say I want to create a group of samples with mutations in both IDH1 and TP53, and without amplification of EGFR.

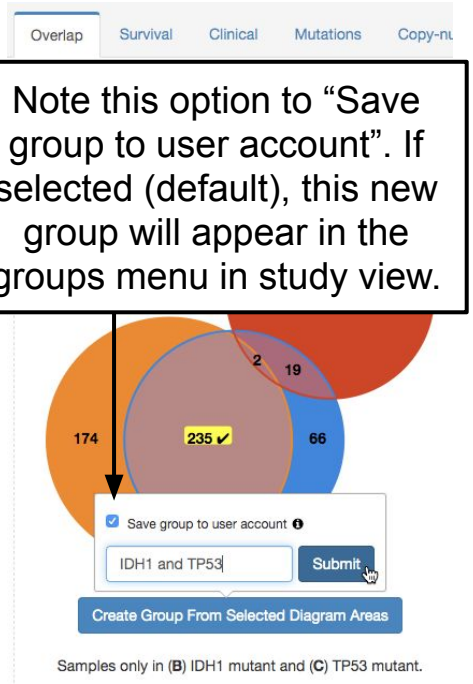
Custom Groups from Overlap Tab



Hover over the venn diagram to find the segment of interest. Click on it.



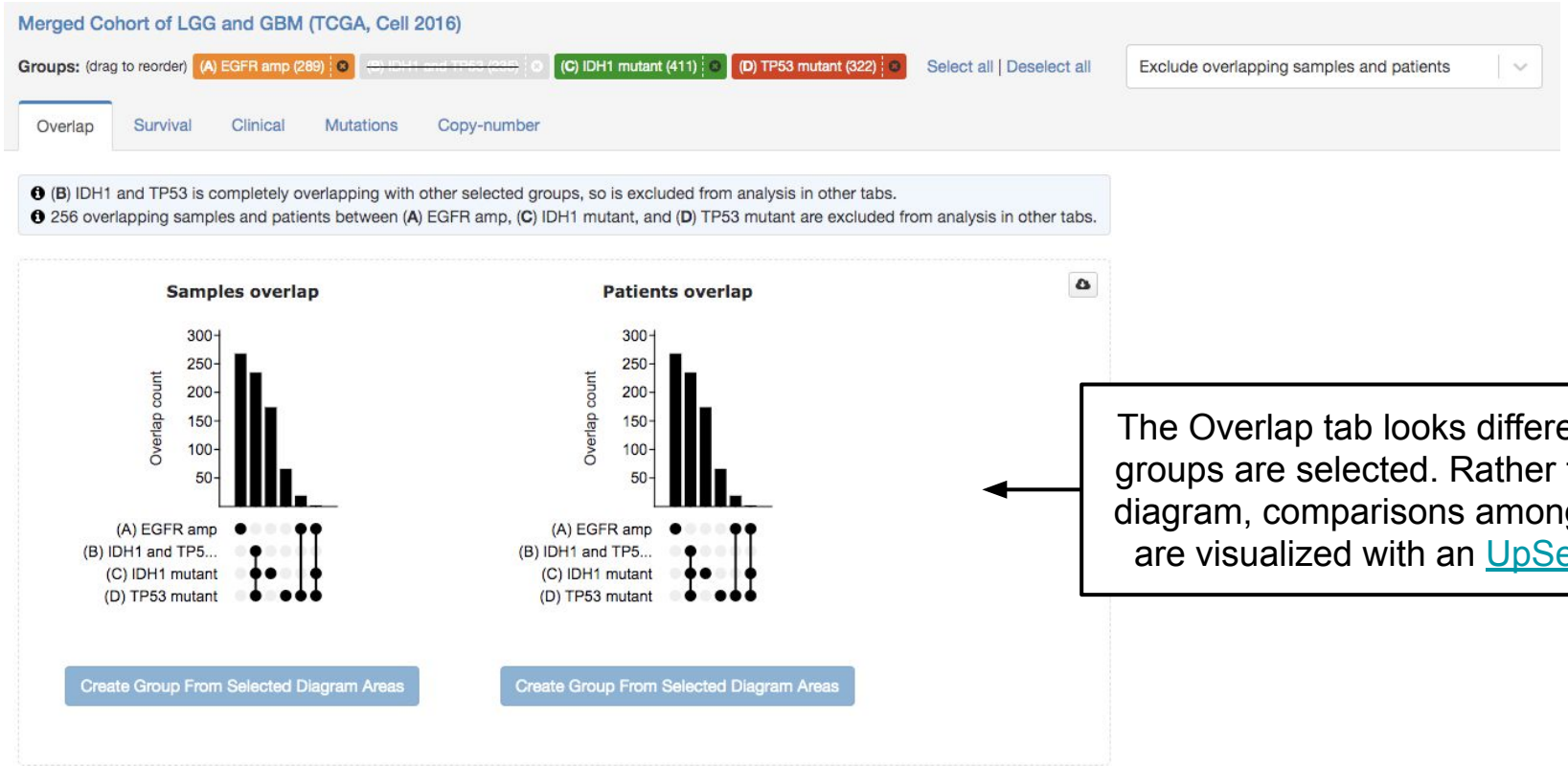
The “Create Group…” button is now enabled. Click on it.



Enter a name for this new group and click “Submit”.

[Link to this page]

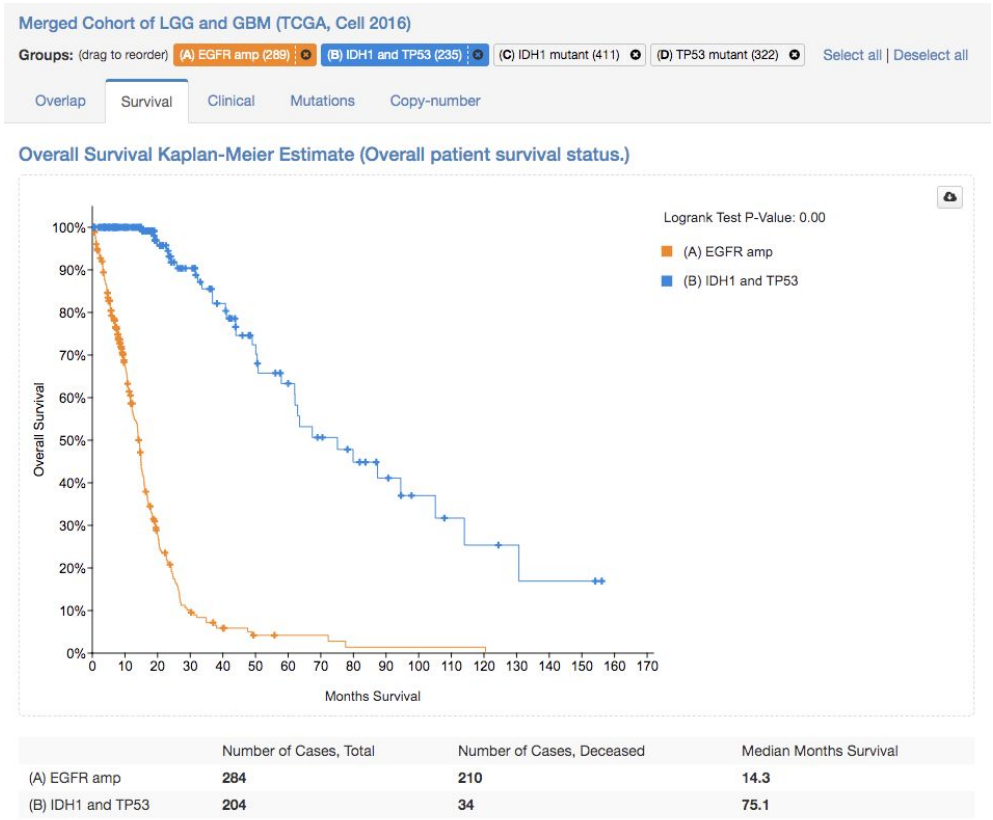
Custom Groups from Overlap Tab



The Overlap tab looks different when 4+ groups are selected. Rather than a venn diagram, comparisons among 4+ groups are visualized with an UpSet diagram.

Link to this page

Custom Groups from Overlap Tab



	Number of Cases, Total	Number of Cases, Deceased	Median Months Survival
(A) EGFR amp	**284**	**210**	**14.3**
(B) IDH1 and TP53	**204**	**34**	**75.1**

Now I can compare, for example, overall survival for patients with EGFR amplification vs IDH1 and TP53 mutations.

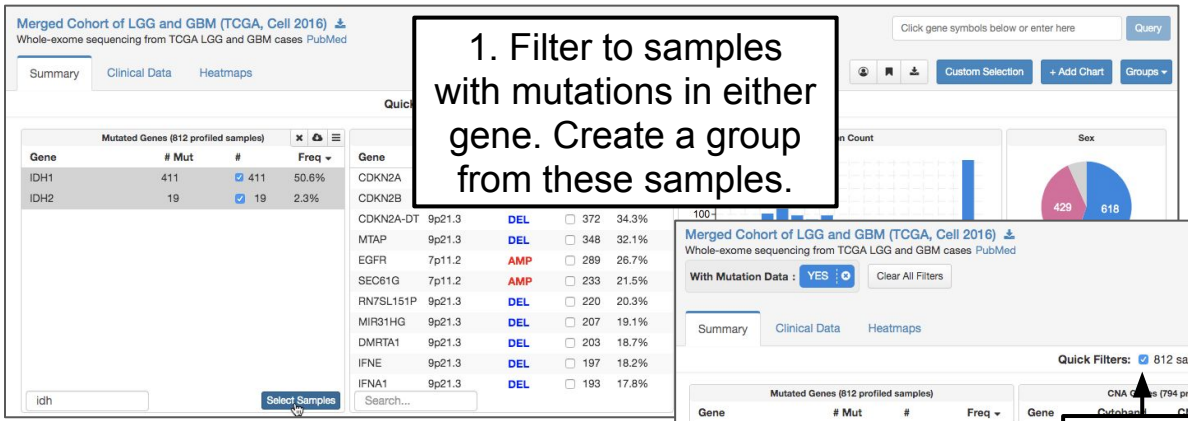
Link to this page

There are many powerful ways to take advantage of this new functionality.

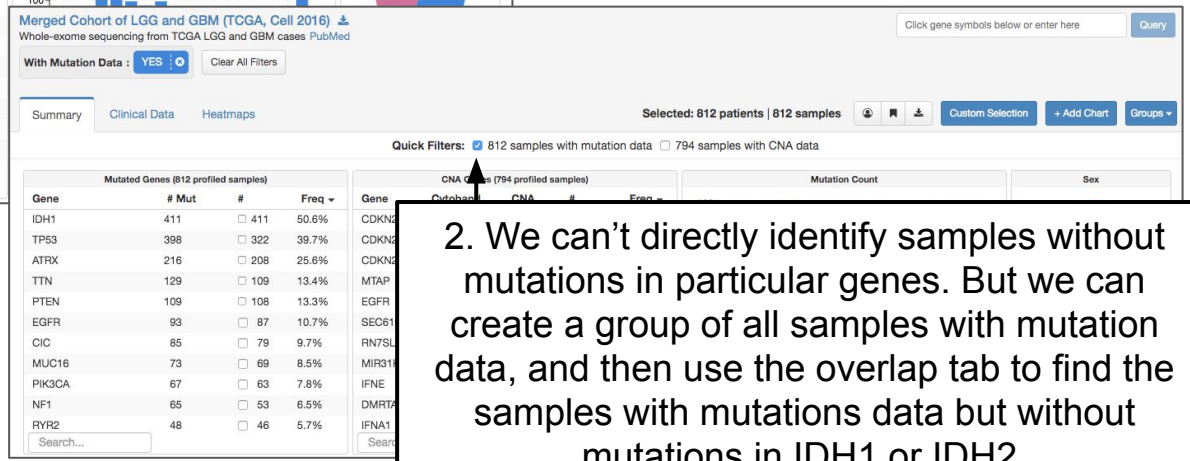
What follows are a few possibilities...

Identify/Compare Samples Without a Specific Alteration

How can we compare samples with or without specific genomic features? As an example, let's take samples with IDH1 or IDH2 mutations and compare to samples without mutations in either gene.



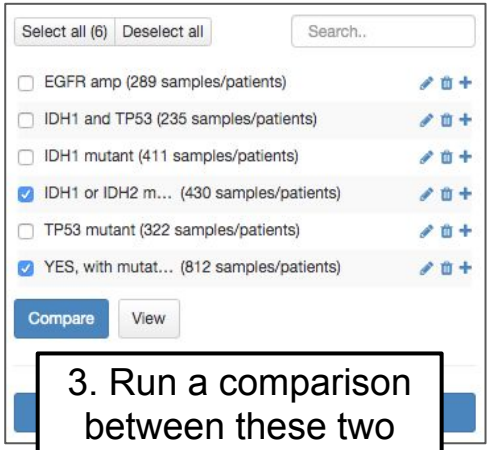
1. Filter to samples with mutations in either gene. Create a group from these samples.



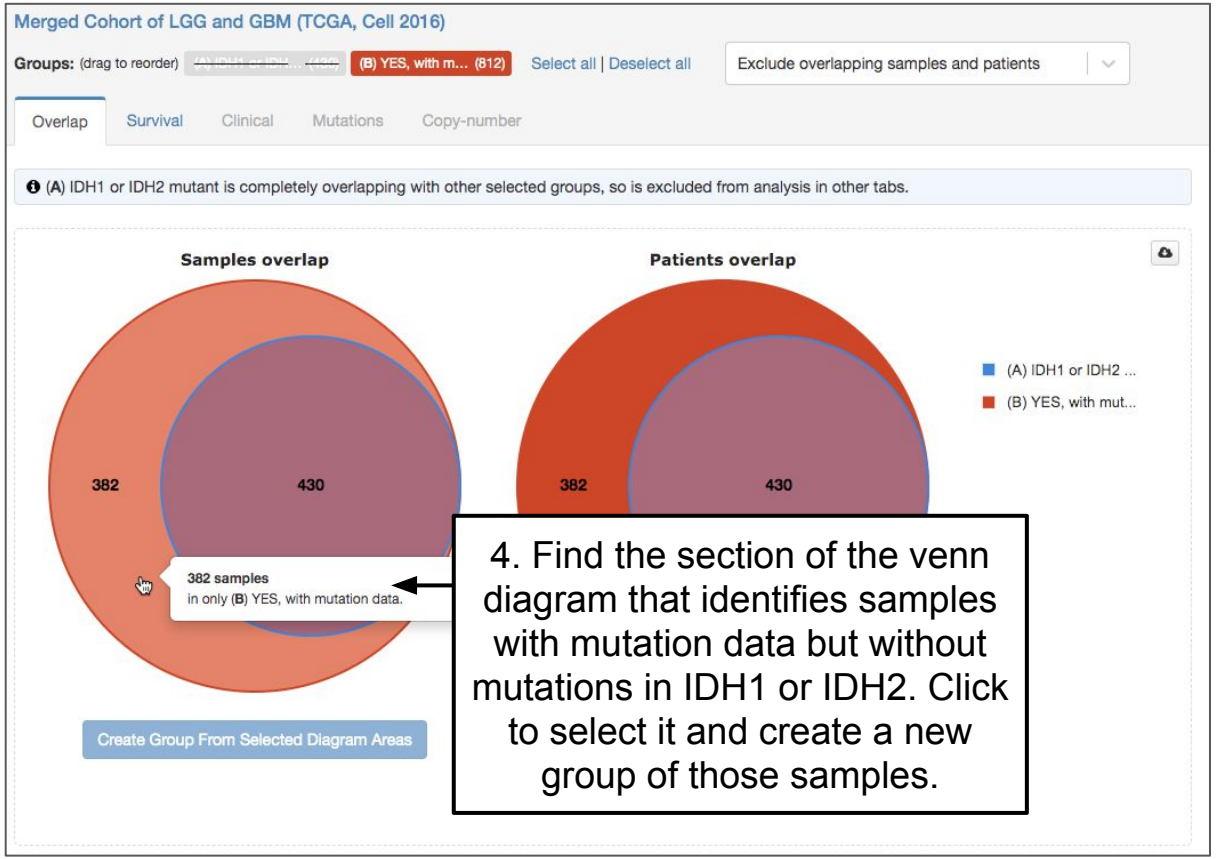
2. We can't directly identify samples without mutations in particular genes. But we can create a group of all samples with mutation data, and then use the overlap tab to find the samples with mutations data but without mutations in IDH1 or IDH2.

Link to this page

Identify/Compare Samples Without a Specific Alteration



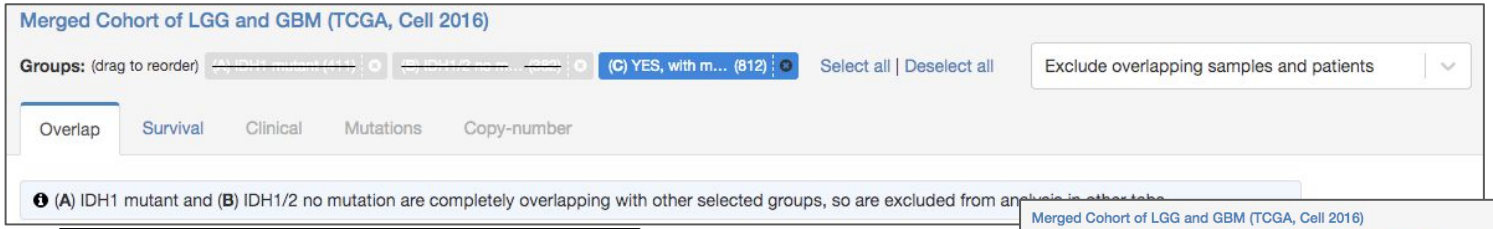
3. Run a comparison between these two groups.



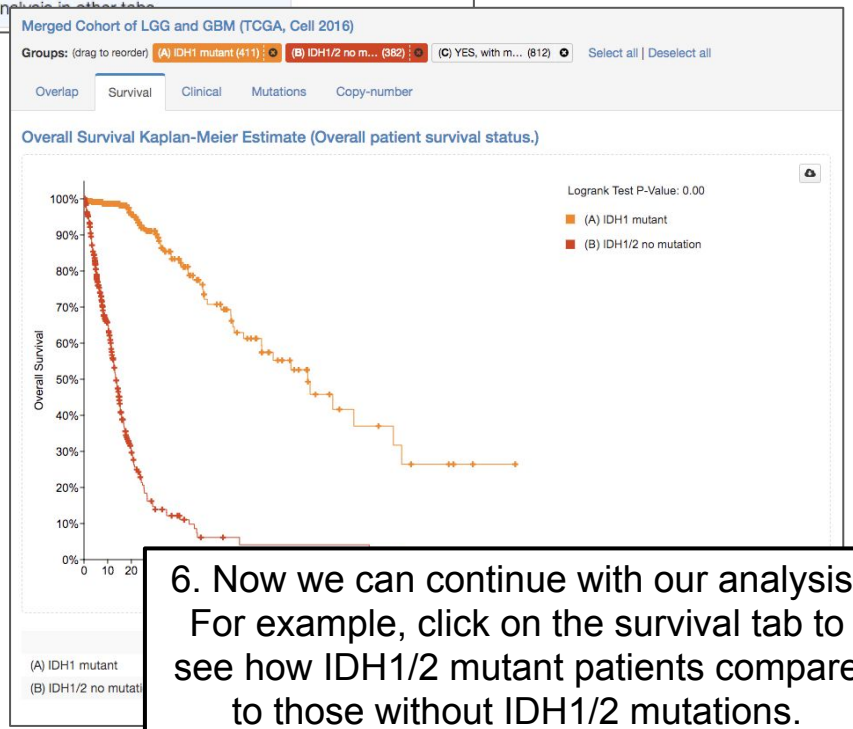
4. Find the section of the venn diagram that identifies samples with mutation data but without mutations in IDH1 or IDH2. Click to select it and create a new group of those samples.

[Link to this page]

Identify/Compare Samples Without a Specific Alteration



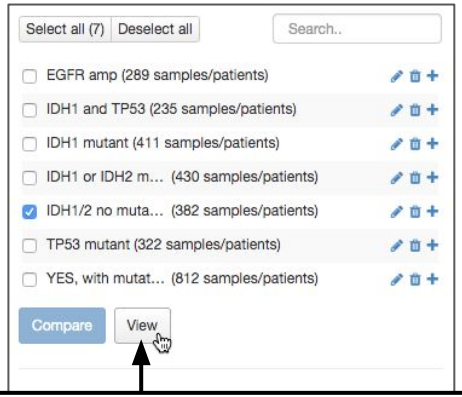
5. When the page reloads, we see the newly created group and a warning that some groups are excluded from analysis due to all samples overlapping with another group. Deselect the “(C) YES, with m…” group.



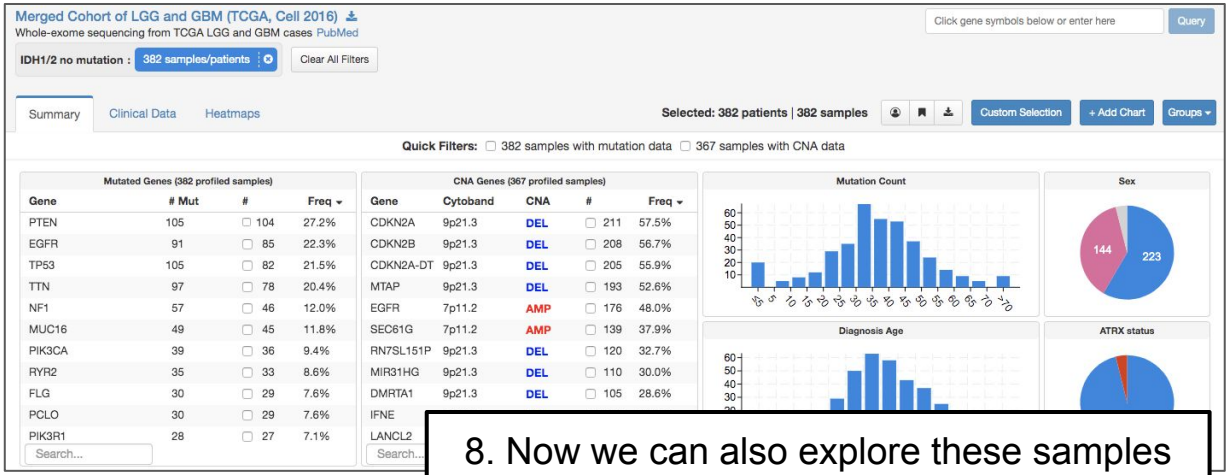
6. Now we can continue with our analysis. For example, click on the survival tab to see how IDH1/2 mutant patients compare to those without IDH1/2 mutations.

Link to this page

Identify/Compare Samples Without a Specific Alteration



7. Remember that these new groups get saved back to Study View. You can use the Groups dropdown to filter study view to this newly defined group.

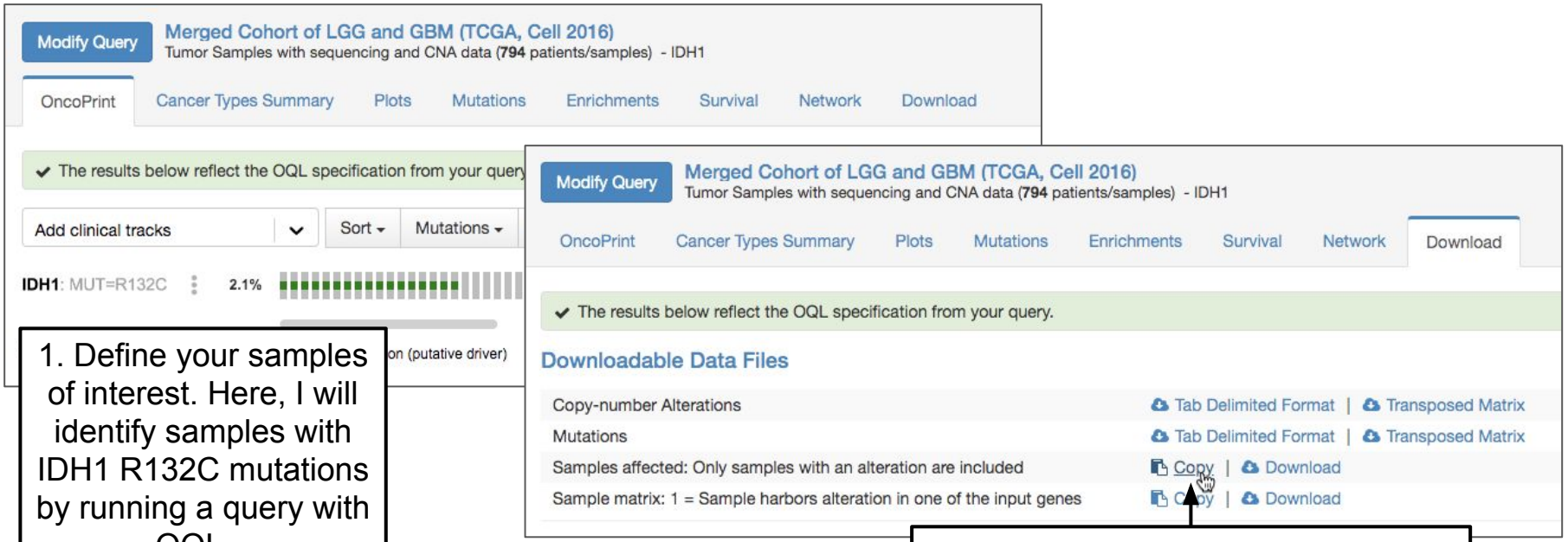


8. Now we can also explore these samples without IDH1/2 mutations in study view.

[Link to this page]

Create a Group from a User-Defined List

What if you have identified a set of samples of interest based on your own analysis? You can create a group from that list in study view.

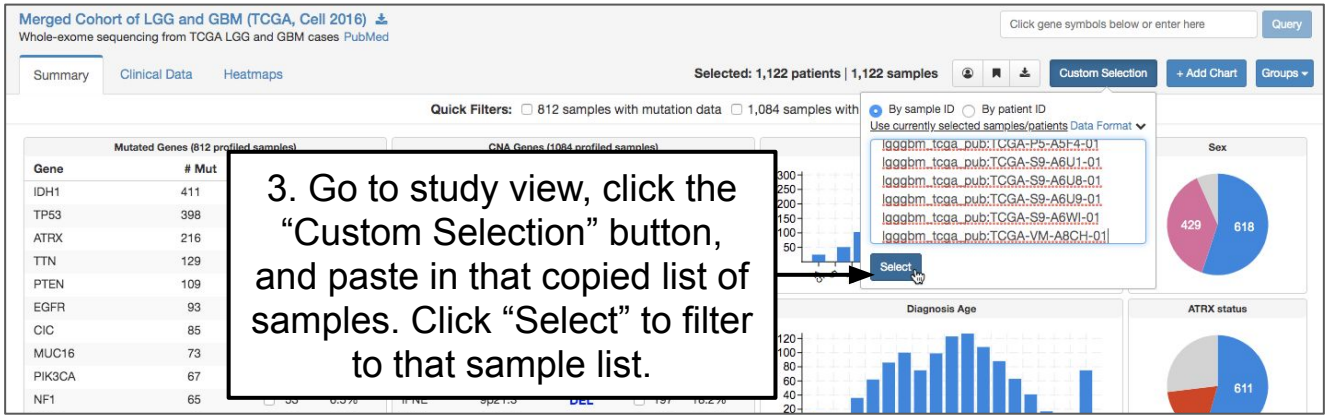


1. Define your samples of interest. Here, I will identify samples with IDH1 R132C mutations by running a query with OQL...

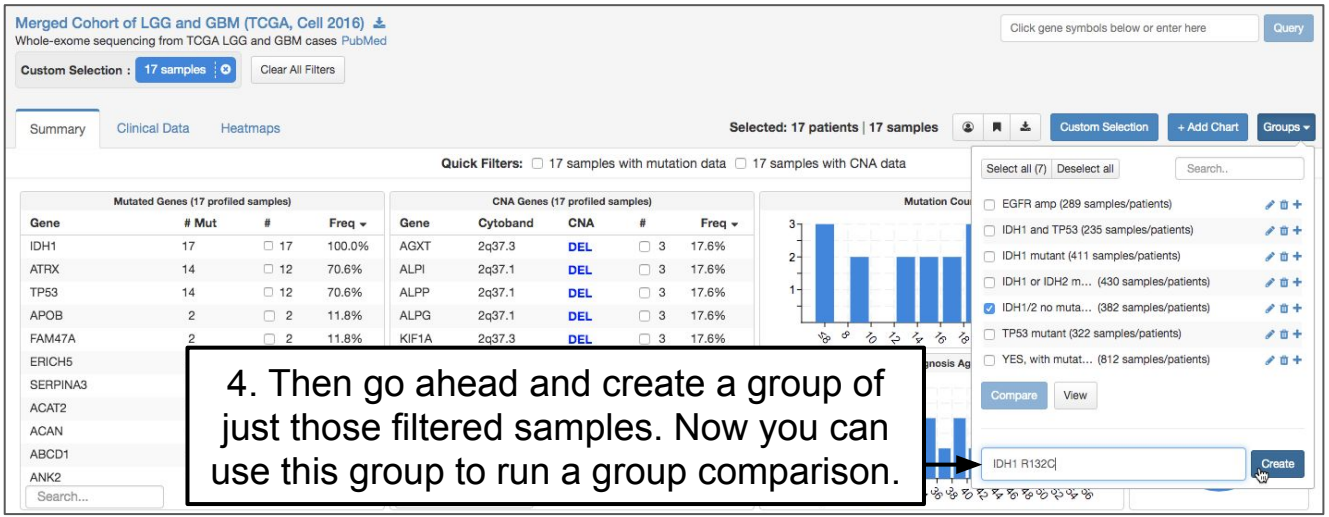
2. ...and then use the Download tab to copy a list of the altered samples.

Link to this page

Create a Group from a User-Defined List



3. Go to study view, click the "Custom Selection" button, and paste in that copied list of samples. Click "Select" to filter to that sample list.



4. Then go ahead and create a group of just those filtered samples. Now you can use this group to run a group comparison.

Link to this page

Questions?
Check out our other tutorials
or email us at:
cbioportal@googlegroups.com