

cBioPortal Tutorial #2: Single Study Query

Query one or multiple genes in a single dataset

Tutorial Objectives

- Show how to run a single-study query from the main page
- Walk through each of the data/analysis tabs in a single-study query
 - OncoPrint
 - Cancer Types Summary
 - Mutual Exclusivity
 - Plots
 - Mutations
 - Co-expression
 - Comparison/Survival
 - CN Segments
 - Pathways
 - Download
- Show how to modify and re-run a query

In this tutorial, blue boxes provide an overview of each tab on cBioPortal ...

... while green boxes ask a biological question that we can answer using cBioPortal.

Overview of Tabs in a Single Study Query

Note that depending on the query run and the data available for a particular study, not all of these will be present (e.g. a study without mRNA expression data will not have a Co-expression tab)

- **OncoPrint:** Overview of genetic alterations per sample in each query gene
- **Cancer Types Summary:** Frequency of alteration in each query gene in the detailed cancer types included in this study
- **Mutual Exclusivity:** Statistical analysis to determine if query genes are mutually exclusively altered
- **Plots:** explore the relationships among genetic alterations, gene expression, protein levels, DNA methylation and available clinical features
- **Mutations:** Details about mutations called in each query gene
- **Co-expression:** Explore which genes have mRNA/protein levels correlated with query genes
- **Comparison/Survival:** Explore overlaps, outcomes, clinical attributes and genomic data comparisons among groups of samples as defined by the query
- **CN Segments:** Explore copy number changes with the Integrated Genomics Viewer (IGV)
- **Pathways:** Explore queried genes in TCGA-defined pathways
- **Download:** Download data or copy sample lists

We're going to run a query in a TCGA Lower-Grade Glioma study. The next few slides will show how to run this query from the Query page. You can also run the same query from a Single Study Exploration, as we did in Tutorial #1.

Query overview

Browse available datasets and select studies to explore or query

Number of studies for each tissue of origin (click to filter)

Search studies

List of all studies, organized by organ system

Please cite: Cerami et al., 2012 & Gao et al., 2013

What's New @cbiportal

If you are running your own local / private installation of cBioPortal, please consider adding it to our new map. Use this form to register your instance: bit.ly/2XW3wNG. We will count all entries!

for low-volume email news alerts

Subscribe

queries

- Primary vs. metastatic prostate cancer
- RAS/RAF alterations in colorectal cancer
- BRCA1 and BRCA2 mutations in ovarian cancer
- POLE hotspot mutations in endometrial cancer
- TP53 and MDM2/4 alterations in GBM
- PTEN mutations in GBM in text format
- Patient view of an endometrial cancer case
- All TCGA Pan-Cancer
- MSK-IMPACT clinical cohort, Zehir et al. 2017
- Histone mutations across cancer types

Local Installations Host your own

Data Sets Web API R/MATLAB Tutorials/Webinars FAQ News Visualize Your Data About cBio Portal Login

Query Quick Search Beta Download

Select Studies for Visualization & Analysis: 0 studies selected (0 samples)

Search...

PanCancer Studies 7

PanCancer Studies

- MSK-IMPACT Clinical Sequencing Cohort (MSKCC, Nat Med 2017)
- Metastatic Solid Cancers (UMich, Nature 2017)
- MSS Mixed Solid Tumors (Broad/Dana-Farber, Nat Genet 2018)
- SUMMIT - Neratinib Basket Study (Multi-Institute, Nature 2018)
- TMB and Immunotherapy (MSKCC, Nat Genet 2019)
- Tumors with TRK fusions (MSK, Clin Cancer Res 2020)
- Cancer Therapy and Clonal Hematopoiesis (MSK, Nat Genet 2020)

10945 samples

Pediatric Cancer Studies

- Pediatric Preclinical Testing Consortium (CHOP, Cell Rep 2019)
- Pediatric Acute Lymphoid Leukemia - Phase II (TARGET, 2018)
- Pediatric Rhabdoid Tumor (TARGET, 2018)
- Pediatric Wilms' Tumor (TARGET, 2018)
- Pediatric Acute Myeloid Leukemia (TARGET, 2018)
- Pediatric Neuroblastoma (TARGET, 2018)
- Pediatric Pan-Cancer (DKFZ, Nature 2017)
- Pediatric Pan-cancer (Columbia U, Genome Med 2016)
- Acute Lymphoblastic Leukemia (St Jude, Nat Genet 2016)
- Acute Lymphoblastic Leukemia (St Jude, Nat Genet 2015)
- Pediatric Ewing Sarcoma (DFCI, Cancer Discov 2014)

261 samples

1978 samples

72 samples

657 samples

1025 samples

1089 samples

961 samples

103 samples

73 samples

93 samples

107 samples

112 samples

0 studies selected (0 samples)

Query By Gene OR Explore Selected Studies

[Link to this page](#)

Single study query

Please cite: Cerami et al., 2012 & Gao et al., 2013

glioma

Select Studies for Visualization & Analysis: 1 study selected (514 samples) Deselect all

Immunogenomic Studies 17

Select all listed studies matching filter (17)

CNS/Brain 16

Immunogenomic Studies

Glioblastoma (Columbia, Nat Med. 2019) 42 samples

Soft Tissue 1

CNS/Brain

Diffuse Glioma

Brain Lower Grade Glioma (TCGA, Firehose Legacy)

Brain Lower Grade Glioma (TCGA, PanCancer Atlas) 530 samples

Glioma (MSK, Nature 2019) 514 samples

Glioma (MSKCC, Clin Cancer Res 2019) 91 samples

Low-Grade Gliomas (UCSF, Science 2014) 1004 samples

Merged Cohort of LGG and GBM (TCGA, Cell 2016) 61 samples

→ GLIOBLASTOMA

Brain Tumor PDXs (Mayo Clinic, 2019) 1102 samples

Glioblastoma (Columbia, Nat Med. 2019) 530 samples

Glioblastoma (TCGA, Cell 2013) 97 samples

Glioblastoma (TCGA, Nature 2008) 42 samples

Glioblastoma Multiforme (TCGA, Firehose Legacy) 543 samples

Glioblastoma Multiforme (TCGA, PanCancer Atlas) 206 samples

→ OLIGODENDROGLIOMA

Anaplastic Oligodendrogloma and Anaplastic Oligoastrocytoma (MSK...) 604 samples

22 samples

1 study selected (514 samples) Deselect all

Query By Gene OR Explore Selected Studies

1. Filter the list of studies (optional)

Login

What's New @cbioportal

cBioPortal @cbioportal

If you are running your own local / private installation of cBioPortal, please consider adding it to our new map. Use this form to register your instance: bit.ly/2XW3wNG. We will count all entries!

Sign up for low-volume email news alerts

Subscribe

Example Queries

- Primary vs. metastatic prostate cancer
- RAS/RAF alterations in colorectal cancer
- BRCA1 and BRCA2 mutations in ovarian cancer
- POLE hotspot mutations in endometrial cancer
- TP53 and MDM2/4 alterations in GBM
- PTEN mutations in GBM in text format
- Patient view of an endometrial cancer case
- All TCGA Pan-Cancer
- MSK-IMPACT clinical cohort, Zehir et al. 2017
- Histone mutations across cancer types

Local Installations Host your own

Link to this page

3. Select "Query By Gene"

Single study query

4. This section lists all data types available for the selected study. Select data types to query. By default, Mutations and CNA will be selected (if available).

5. Select sample set. For most studies, an appropriate sample set will be automatically selected given the data types selected in Step 4.

Query Quick Search Beta! Download Please cite: Cerami et al., 2012 & Gao et al., 2013

Selected Studies: [Modify](#) Brain Lower Grade Glioma (TCGA, PanCancer Atlas) (514 total samples)

Select Genomic Profiles:

Mutations [?](#)
 Putative copy-number alterations from GISTIC [?](#)
 mRNA Expression. Select one of the profiles below:
 mRNA expression z-scores relative to diploid samples (RNA Seq V2 RSEM) [?](#)
 mRNA expression z-scores relative to all samples (log RNA Seq V2 RSEM) [?](#)
 Protein expression z-scores (RPPA) [?](#)

Select Patient/Case Set:
To build your own case set, try out our enhanced Study View.

Samples with mutation and CNA data (507)

Enter Genes:
Hint: Learn Onco Query Language (OQL) to write more powerful queries [?](#)

User-defined List
IDH1 EGFR ←
All gene symbols are valid

Submit Query 

6. Submit query

Refine your query: You can use Onco Query Language (OQL) to define which specific alterations to include. See [specifications](#) or [OQL tutorial](#).

Performing a query as shown in the previous slides or as shown in Tutorial #1 will both bring you to Results View, shown on the next slide.

Results View is made up of multiple tabs, each with specific functionality, which all share a header.

Results View Header: General Information

The name of the study.
Click to view the full study in Study View.

The number (percentage) of samples/patients with an alteration in any of the query genes

Modify Query



Brain Lower Grade Glioma (TCGA, PanCancer Atlas)

Samples with mutation and CNA data (507 patients/samples) - IDH1 & EGFR

OncoPrint

Cancer Types Summary

Mutual Exclusivity

Plots

Mutations

Co-expression

Comparison/Survival

Queried genes are altered in 444 (88%) of queried patients/samples



CN Segments

Pathways

Download

The number of samples and patients included in the query. Note that these numbers can differ from each other if some patients have more than one tumor sample profiled.

Click on the number of patients/samples to go to Study View for just the queried samples.

Save a link to the current session. Useful for sharing with others or returning to a query at a later date.

Results View Header: Variant Settings

Use this menu to control how alterations are visualized. Changes made here are immediately reflected across Results View. However over the  to confirm how individual tabs reflect these selections.

↓

Brain Lower Grade Glioma (TCGA, PanCancer Atlas)
Samples with mutation and CNA data (507 patients/samples) - IDH1 & EGFR 

Queried genes are altered in 444 (88%) of queried patients/samples 

Annotate Data 

Putative drivers vs VUS:
 OncoKB driver annotation
 Hotspots 🔥
 cBioPortal ≥ 10
 COSMIC ≥ 10

Filter Data

Exclude mutations and copy number alterations of unknown significance
 Exclude germline mutations
 Exclude samples that are not profiled for all queried genes in all queried profiles

Plots Mutations Co-expression Comparison/Survival CN Segments Pathways Download

Mutations

Missense

Set the definition of a putative driver vs variant of unknown significance (VUS).

Check boxes to exclude VUS (as defined above) or germline alterations. When checked, VUS or germline alterations are considered not present, so a sample with only VUS or germline alterations will be treated as an unaltered sample.

Check box to exclude samples where queried genes are not profiled or genomic profiles are not available.

[Link to this page](#)

But wait! What if I changed my mind?
Can we modify a query?

Modify Query

Click on “Modify Query”. This button is available on all tabs and can be used at any time. This will bring up the query interface from the homepage (see next slide for a screenshot).

OncoPrint

Brain Lower Grade Glioma (TCGA, PanCancer Atlas)
Samples with mutation and CNA data (507 patients/samples) - IDH1 & EGFR

Queried genes are altered in 444 (88%) of queried patients/samples

Cancer Types Summary Mutual Exclusivity Plots Mutations Co-expression Comparison/Survival CN Segments Pathways Download

Add Clinical Tracks 97 ▾ Add Heatmap Tracks 4 ▾ Sort ▾ Mutations ▾ View ▾ Download

IDH1 78%

EGFR 11%

Genetic Alteration

Legend: Missense Mutation (putative driver) Missense Mutation (unknown significance) Fusion Amplification No alterations

You can also click the for a quick edit of the queried genes, including OQL edits. To change other query settings, use the “Modify Query” button.



Modify Query

The existing query is pre-populated for your convenience. You can change the study, the genomic profiles, the patient/case set or the gene set. Simply hit “Submit” when you are happy with the modified query.

Select Studies for Visualization & Analysis: 1 study selected (514 samples) Deselect all

Quick select: TCGA PanCancer Atlas Studies Curated set of non-redundant studies

PanCancer Studies

- 7 MSK-IMPACT Clinical Sequencing Cohort (MSKCC, Nat Med 2017) 10945 samples
- 8 Metastatic Solid Cancers (UMich, Nature 2017) 500 samples
- 3 MSS Mixed Solid Tumors (Broad/Dana-Farber, Nat Genet 2018) 249 samples
- 3 SUMMIT - Neratinib Basket Study (Multi-Institute, Nature 2018) 141 samples
- 1 TMB and Immunotherapy (MSK, Nat Genet 2019) 1661 samples
- 1 Tumors with TRK fusions (MSK, Clin Cancer Res 2020) 106 samples
- 1 Cancer Therapy and Clonal Hematopoiesis (MSK, Nat Genet 2020) 24146 samples

Pediatric Cancer Studies

- 16 Bone 261 samples
- 2 Bowel 1978 samples
- 10 Pediatric Preclinical Testing Consortium (CHOP, Cell Rep 2019) 72 samples
- 1 Pediatric Acute Lymphoid Leukemia - Phase II (TARGET, 2018) 657 samples
- 1 Pediatric Rhabdoid Tumor (TARGET, 2018) 1025 samples
- 1 Pediatric Wilms' Tumor (TARGET, 2018) 1025 samples
- 1 Pediatric Anaplastic Meningeal Leiomysia (TARGET, 2018) 1025 samples

Select Genomic Profiles:

- Mutations
- Putative copy-number alterations from GISTIC
- mRNA Expression. Select one of the profiles below:
 - mRNA expression z-scores relative to diploid samples (RNA Seq V2 RSEM)
 - mRNA expression z-scores relative to all samples (log RNA Seq V2 RSEM)
- Protein expression z-scores (RPPA)

Select Patient/Case Set:
To build your own case set, try our enhanced Study View.

Samples with mutation and CNA data (507)

Enter Genes:

Hint: Learn Onco Query Language (OQL) to write more powerful queries

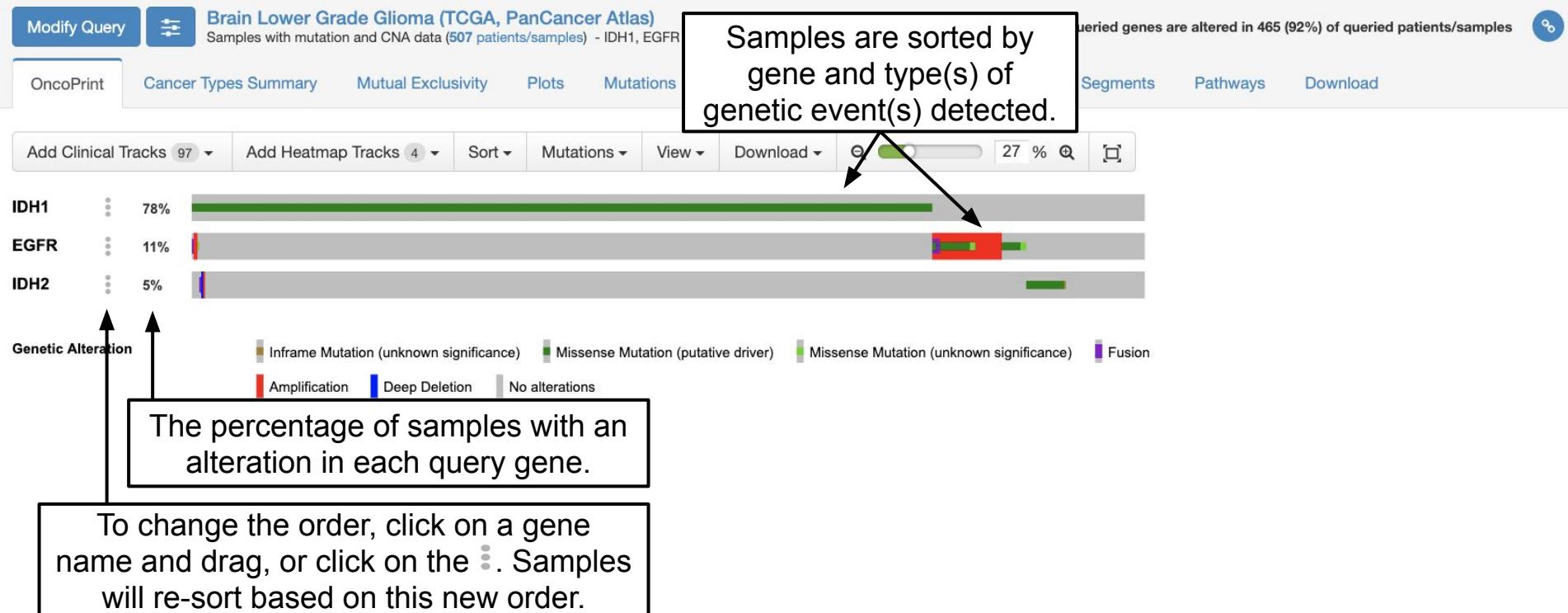
User-defined List
IDH1 EGFR IDH2

All gene symbols are valid.

In this case, I've added a third gene (IDH2) to the query.

OncoPrint

Summary of alterations per sample. Each sample is a column. Each gene is a row. Different kinds of genetic alterations are highlighted with different colors.



OncoPrint: Features

Add clinical tracks (options will vary depending on the data available for each study)

Add heatmaps with RNA/protein levels, DNA methylation, or other data, e.g. treatment response (available data varies by study)

Change the sample sorting order

Customize visualization

The screenshot shows the OncoPrint interface with several features highlighted:

- OncoPrint**: The main navigation tab.
- Cancer Types Summary**: A secondary tab.
- Mutual Exclusivity**: A secondary tab.
- Plots**: A secondary tab.
- Mutations**: A secondary tab.
- Co-expression**: A secondary tab.
- Comparison/Survival**: A secondary tab.
- CN Segments**: A secondary tab.
- Pathways**: A secondary tab.
- Download**: A secondary tab.

Below the tabs, there are dropdown menus and search/filter options:

- Add Clinical Tracks**: Shows 97 items, with "age" selected.
- Add Heatmap Tracks**: Shows 4 items.
- Sort**: A dropdown menu.
- Mutations**: A dropdown menu.
- View**: A dropdown menu.
- Download**: A dropdown menu.
- Search**: A search bar with a toggle switch for case sensitivity and a progress bar.
- Download**: A download icon.

A detailed view of the mutations section is shown on the right, displaying a list of clinical tracks and a visualization of mutation patterns across samples.

Annotations on the interface:

- Add Clinical Tracks (97)**: Points to the "Add Clinical Tracks" dropdown.
- Add Heatmap Tracks (4)**: Points to the "Add Heatmap Tracks" dropdown.
- Sort**: Points to the "Sort" dropdown.
- Mutations**: Points to the "Mutations" dropdown.
- View**: Points to the "View" dropdown.
- Download**: Points to the "Download" dropdown.
- Search**: Points to the search bar.
- Download**: Points to the download icon.
- Change the rules by which mutations are colored**: Points to the color scale legend in the mutation visualization.
- Download figure as PNG, PDF or SVG. Download patient/sample IDs in same order as OncoPrint**: Points to the download icon in the mutation visualization.

OncoPrint: Zoom

There may be more samples hiding off-screen. Scroll to the right or zoom out or use minimap to see them.

Change the zoom by clicking the zoom in/out icons or moving the slider or typing a value

OncoPrint Cancer Types Summary Mutual Exclusivity Plots Mutations Co-expression Comparison/Survival CN Segments Pathways Download

Add Clinical Tracks 97

Add Heatmap Tracks 4

Sort

Mutations

View

Download

Q

55 %

Q



Click here to open "minimap" (see below)

Diagnosis Age

IDH1

EGFR

IDH2

Genetic Alteration

Diagnosis Age

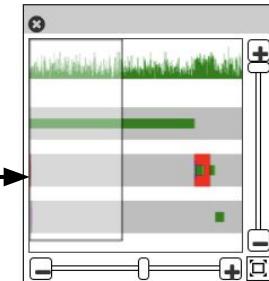


Inframe Mutation (unknown significance) Missense Mutation (putative driver) Missense Mutation (unknown significance) Fusion

Amplification Deep Deletion

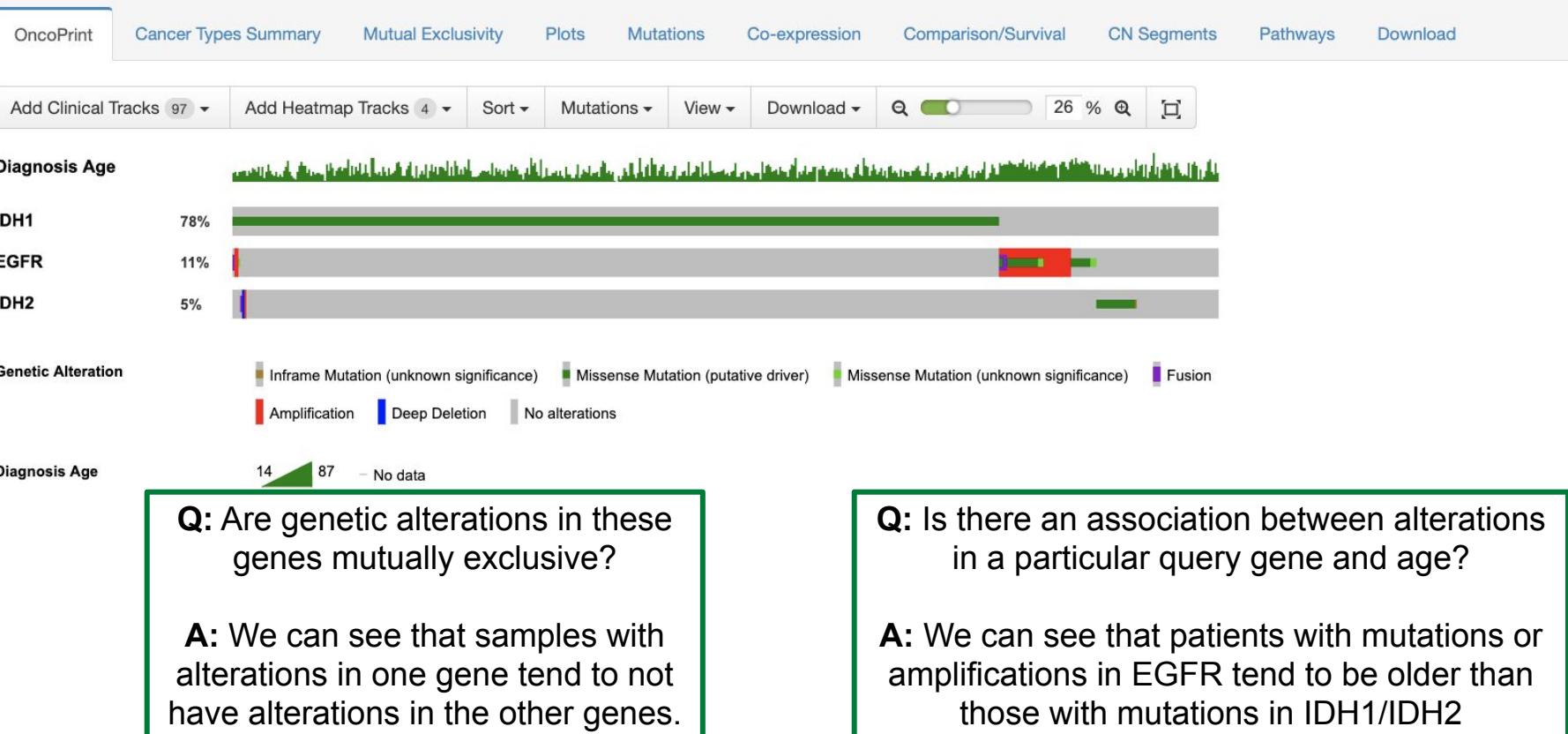
14 87 - No data

Minimap shows a small version of the full OncoPrint and allows you to zoom in each direction independently. The rectangle can be dragged to move around OncoPrint or resized to change the zoom.



This button zooms OncoPrint to show all samples with alterations

OncoPrint: What can we learn?



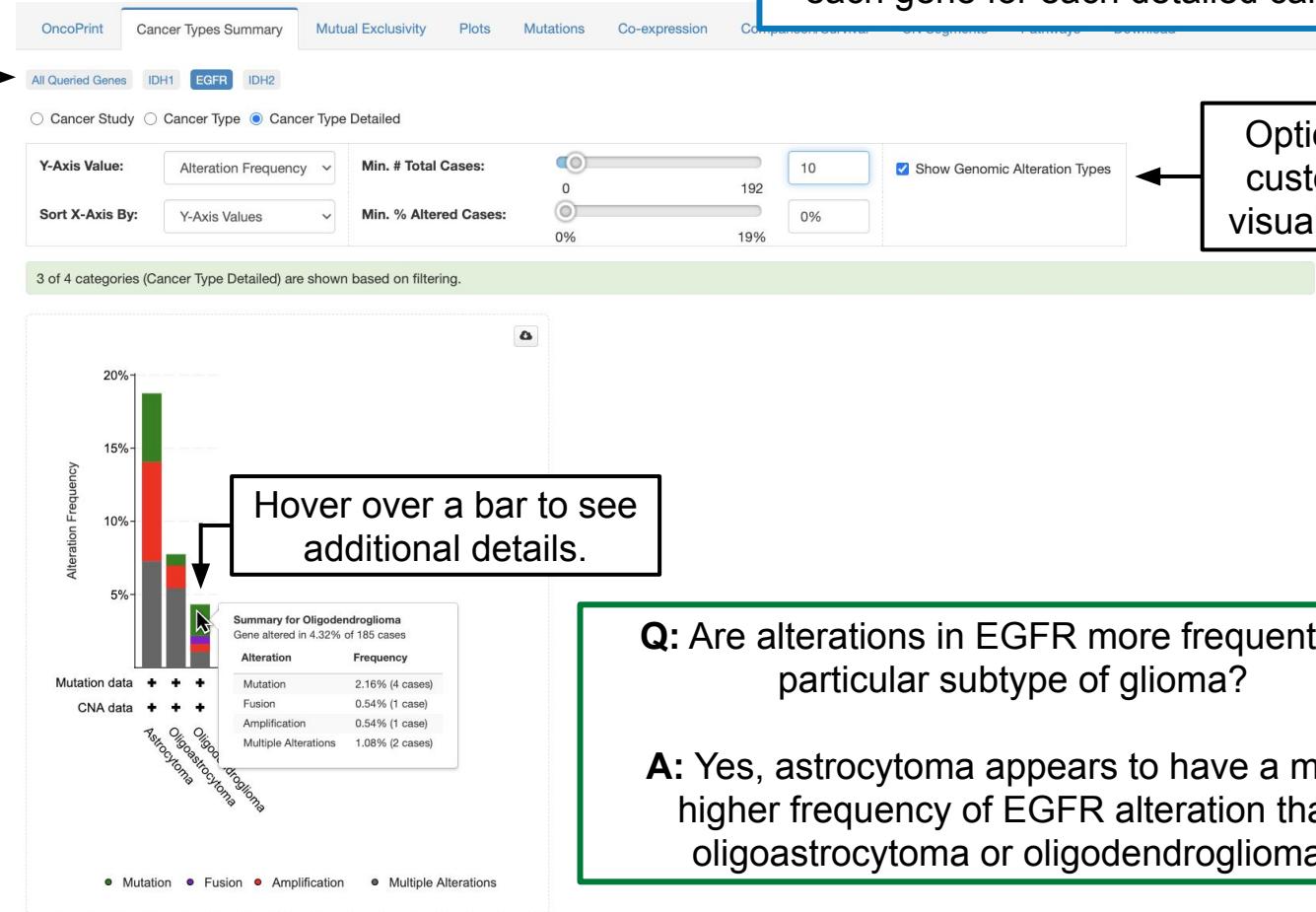
[Link to this page](#)

Now we're going to go through all the other tabs and ask some questions about alterations in *IDH1*, *IDH2* and *EGFR* in the TCGA Lower-Grade Glioma study.

Note: Depending on the data available for a particular study, not all of the following tabs will be present (e.g. a study without expression data will not have a Co-expression tab)

Cancer Types Summary

Plots for all queried genes together and each individual gene are available as separate tabs.



Q: Are alterations in EGFR more frequent in a particular subtype of glioma?

A: Yes, astrocytoma appears to have a much higher frequency of EGFR alteration than oligoastrocytoma or oligodendrogloma.

Mutual Exclusivity

All pairwise combinations of query genes analyzed for mutual exclusivity or co-occurrence in the queried samples.

On the OncoPrint tab we could see visually that alterations in these three query genes tended to be mutually exclusive. Here we can address that same question with a statistical analysis.

OncoPrint Cancer Types Summary Mutual Exclusivity Plots Mutations Gene Expression Comparison Survival Other Segments Pathways Domains

The analysis tested **3** pairs between the **3** tracks in the OncoPrint.

Mutual exclusivity Co-occurrence Significant only

Columns ▾  

Q

A	B	Neither	A Not B	B Not A	Both	Log2 Odds Ratio	p-Value	q-Value ▲	Tendency	←
IDH1	EGFR	63	390	50	4	<-3	<0.001	<0.001	Mutual exclusivity	
IDH1	IDH2	92	391	21	3	<-3	<0.001	<0.001	Mutual exclusivity	
EGFR	IDH2	429	54	24	0	<-3	0.063	0.063	Mutual exclusivity	

Showing 1-3 of 3

A positive value here suggests that alterations in these genes co-occur in the same samples, while a negative value suggests that alterations in these genes are mutually exclusive and occur in different samples.

$$\log_2 \left(\frac{\text{odds of alteration in B given alteration in A}}{\text{odds of alteration in B given lack of alteration in A}} \right)$$

Click on any column header to sort. Hover over the column names for more details about how values are calculated.

[Link to this page](#)

Plots

Example plot settings

Choose type of data

Select a query gene

Swap horizontal & vertical axis

OncoPrint Cancer Types Summary Mutual Exclusivity Plots Mutations

Examples: Mut# vs Dx FGA vs Dx Mut# vs FGA mRNA vs Dx mRNA vs mut type mRN

Data Type: Copy Number
Copy Number Profile: Putative copy-number alteration...
Gene: EGFR
Filter categories: Select...
 Sort Categories by Median
Horizontal Axis: Swap Axes
Data Type: mRNA
mRNA Profile: mRNA expression z-scores relat...
Gene: EGFR
Search Case(s): Case ID...
Search Mutation(s): Protein Change...



Depending on available data types for a given study, this tab allows for plots comparing mutations, copy number, mRNA expression, protein levels and DNA methylation of query genes, along with any available clinical attributes.

Select color scheme

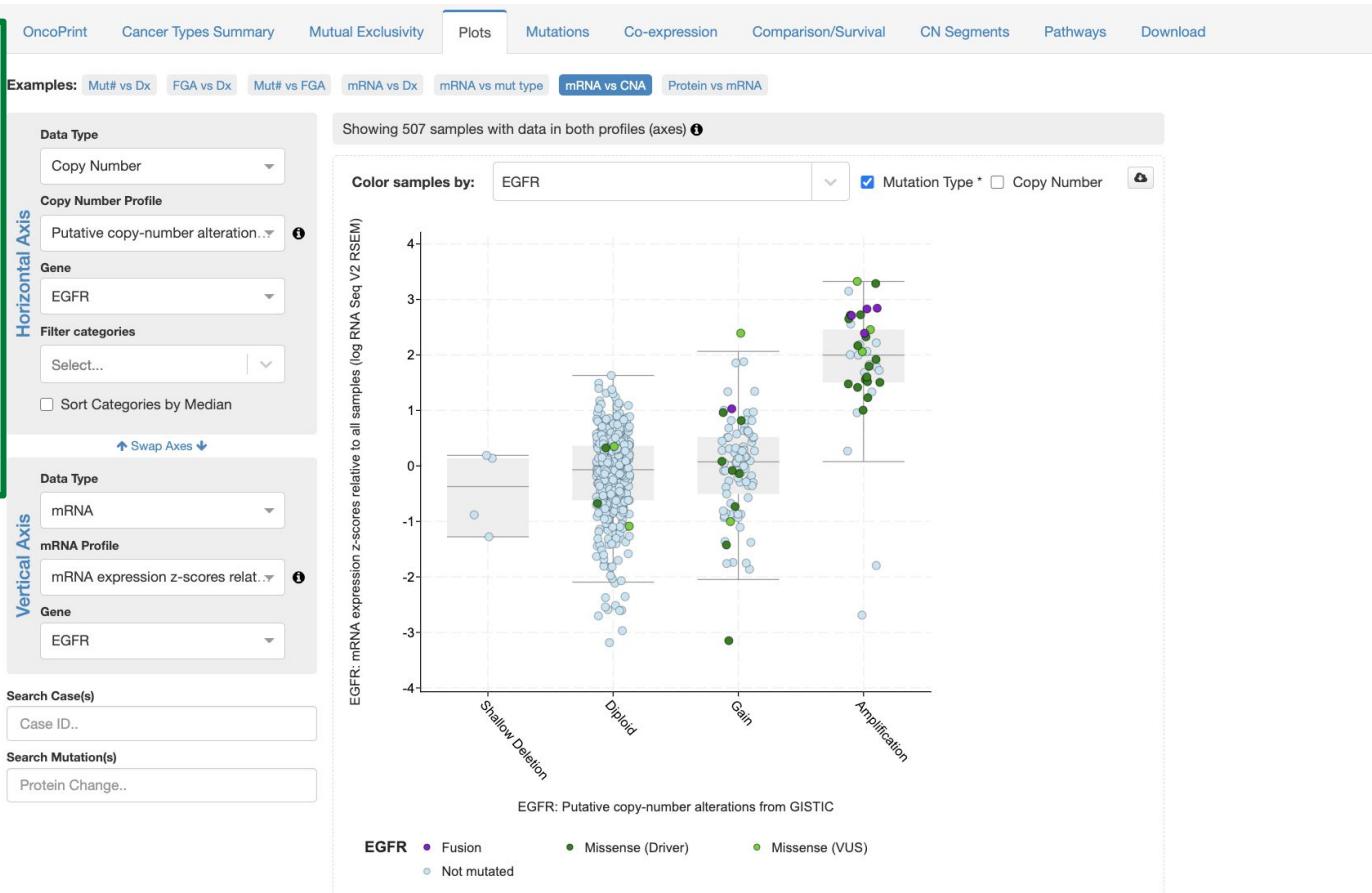
Driver vs. VUS annotation settings are in the ☰ menu in the header of the page.

[Link to this page](#)

Plots

Q: Does amplification of EGFR alter gene expression?

A: Yes, we can see that higher copy number of EGFR (x-axis) is associated with increased expression (y-axis).



[Link to this page](#)

Mutations

This tab shows details about all mutations called in each query gene.

Each gene appears on a separate tab



Table of all mutations with annotations

47 Mutations (page 1 of 2)

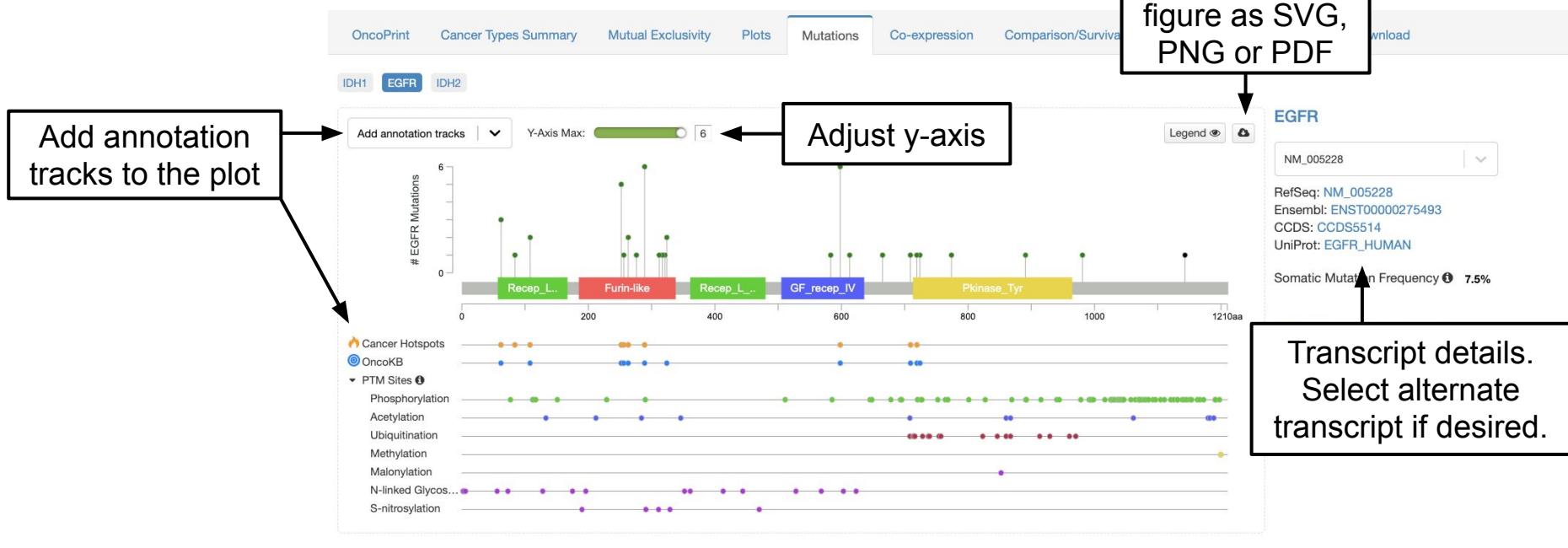
Sample ID	Cancer Type	Protein Change	Annotation ▾	Mutation Type	Copy #	COSMIC	Allele Freq (T)	# Mut in Sample
TCGA-KT-A7W1-01	Astrocytoma	G719D	● ○ ● ○ 🔥	Missense	Amp	125	0.18	39
TCGA-DU-A5TT-01	Oligodendrogloma	A289V	● ○ 🔥	Missense	Gain			
TCGA-E1-A7YM-01	Astrocytoma	A289V	● ○ 🔥	Missense	Gain			
TCGA-FG-6692-01	Oligodendrogloma	A289V	● ○ 🔥	Missense	Amp			
TCGA-FG-A70Z-01	Oligoastrocytoma	A289V	● ○ 🔥	Missense	Amp			
TCGA-KT-A7W1-01	Astrocytoma	A289V	● ○ 🔥	Missense	Amp			
TCGA-TM-A7C3-01	Astrocytoma	A289V	● ○ 🔥	Missense	Amp	50	0.71	
TCGA-HT-8107-01	Oligodendrogloma	R108K	● ○ 🔥	Missense	Diploid	17	0.15	
TCGA-HT-8110-01	Astrocytoma	R108K	● ○ 🔥	Missense	Amp	17	0.94	

Show additional columns

Filter based on any visible text column

Link to this page

Mutations



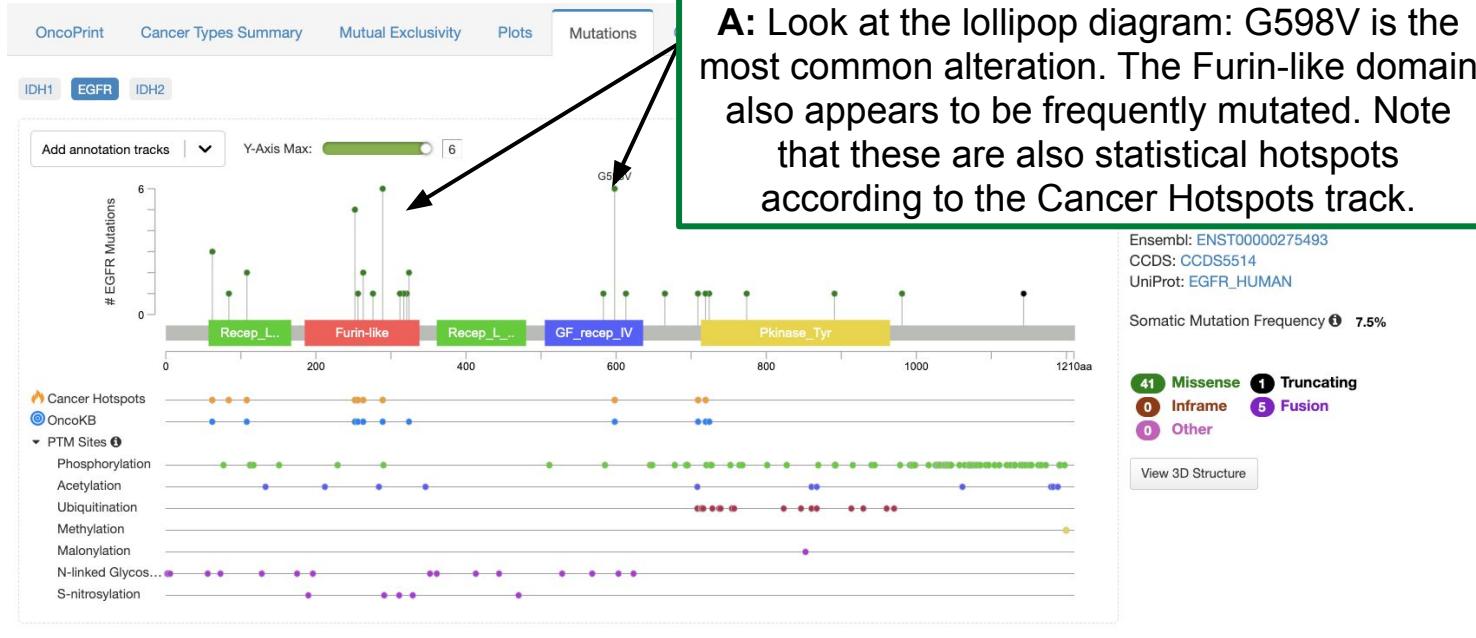
Sample ID	Cancer Type	Protein Change	Annotation ▾	Mutation Type	Copy #	COSMIC	Allele Freq (T)	# Mut in Sample
TCGA-KT-A7W1-01	Astrocytoma	G719D	● ● ○ ○	Missense	Amp	125	0.18	39
TCGA-DU-A5TT-01	Oligodendrogloma	A289V	● ○ ○ ○	Missense	Gain	50	0.45	45
TCGA-E1-A7YM-01	Astrocytoma	A289V	● ○ ○ ○	Missense	Gain	50	0.18	35
TCGA-FG-6692-01	Oligodendrogloma	A289V	● ○ ○ ○	Missense	Amp	50	0.95	63
TCGA-FG-A70Z-01	Oligoastrocytoma	A289V	● ○ ○ ○	Missense	Amp	50	0.88	47
TCGA-KT-A7W1-01	Astrocytoma	A289V	● ○ ○ ○	Missense	Amp	50	0.02	39

Link to this page

Mutations

Q: What are the hotspots for EGFR mutation in glioma?

A: Look at the lollipop diagram: G598V is the most common alteration. The Furin-like domain also appears to be frequently mutated. Note that these are also statistical hotspots according to the Cancer Hotspots track.

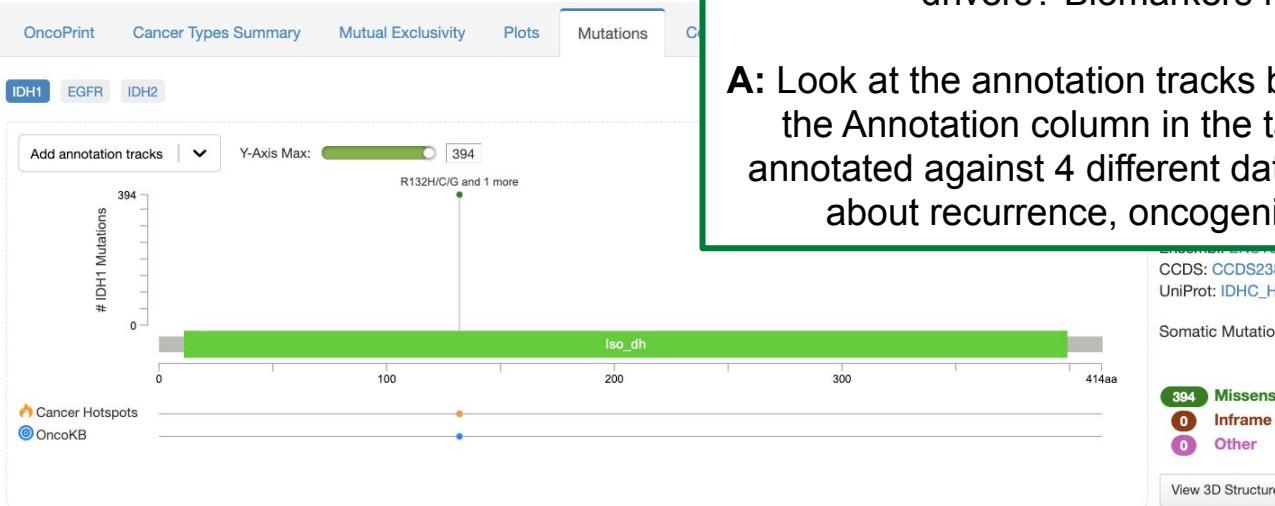


47 Mutations (page 1 of 2)

Sample ID	Cancer Type	Protein Change	Annotation ▾	Mutation Type	Copy #	COSMIC	Allele Freq (T)	# Mut in Sample
TCGA-KT-A7W1-01	Astrocytoma	G719D	•, 🔥, ⚡, 🔥	Missense	Amp	125	0.18	39
TCGA-DU-A5TT-01	Oligodendrogloma	A289V	•, 🔥, ⚡, 🔥	Missense	Gain	50	0.45	45
TCGA-E1-A7YM-01	Astrocytoma	A289V	•, 🔥, ⚡, 🔥	Missense	Gain	50	0.18	35
TCGA-FG-6692-01	Oligodendrogloma	A289V	•, 🔥, ⚡, 🔥	Missense	Amp	50	0.95	63
TCGA-FG-A70Z-01	Oligoastrocytoma	A289V	•, 🔥, ⚡, 🔥	Missense	Amp	50	0.88	47
TCGA-KT-A7W1-01	Astrocytoma	A289V	•, 🔥, ⚡, 🔥	Missense	Amp	50	0.02	39

Link to this page

Mutations



3 This mutation is in [OncoKB](#) as a Level 3 variant. Hover over this symbol to see additional information, including that this is a known oncogenic mutation.

TCGA-DB-A4XF-01 Astrocytoma *R132C*
TCGA-DB-A64S-01 Oligodendroglioma *R132C*
TCGA-DB-A7
TCGA-FG-811
TCGA-HT-747

This mutation is annotated in [CIViC](#). Hover over this symbol for additional information.

Q: The mutations in IDH1 appear to be highly recurrent. Are these mutations known hotspots? Known oncogenic drivers? Biomarkers for any drugs?

A: Look at the annotation tracks below the lollipop plot and the Annotation column in the table. Each mutation is annotated against 4 different databases with information about recurrence, oncogenicity and drugability.

Annotation ▾

Mutation Type Copy # COSMIC Allele # Mut in

Missense Diploid 4964 0.40 28

Missense Diploid 4964 0.32 12

Missense Diploid 4964 0.20 25

Missense Diploid 4964 0.38 36

Missense Diploid 4964 0.25 22

This mutation is a [recurrent hotspot](#) based on a [statistical analysis of mutation frequency](#).

This mutation is in [My Cancer Genome](#).

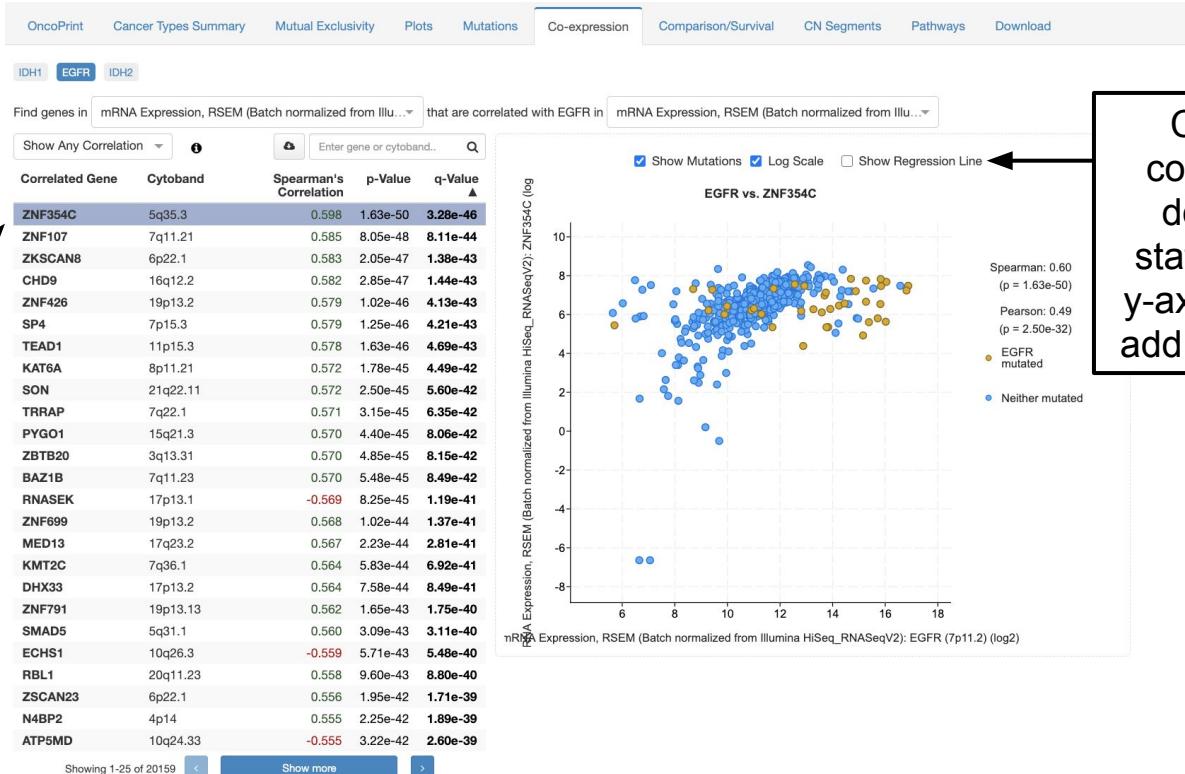
Co-Expression

Each gene appears on a separate tab

Select from available data types

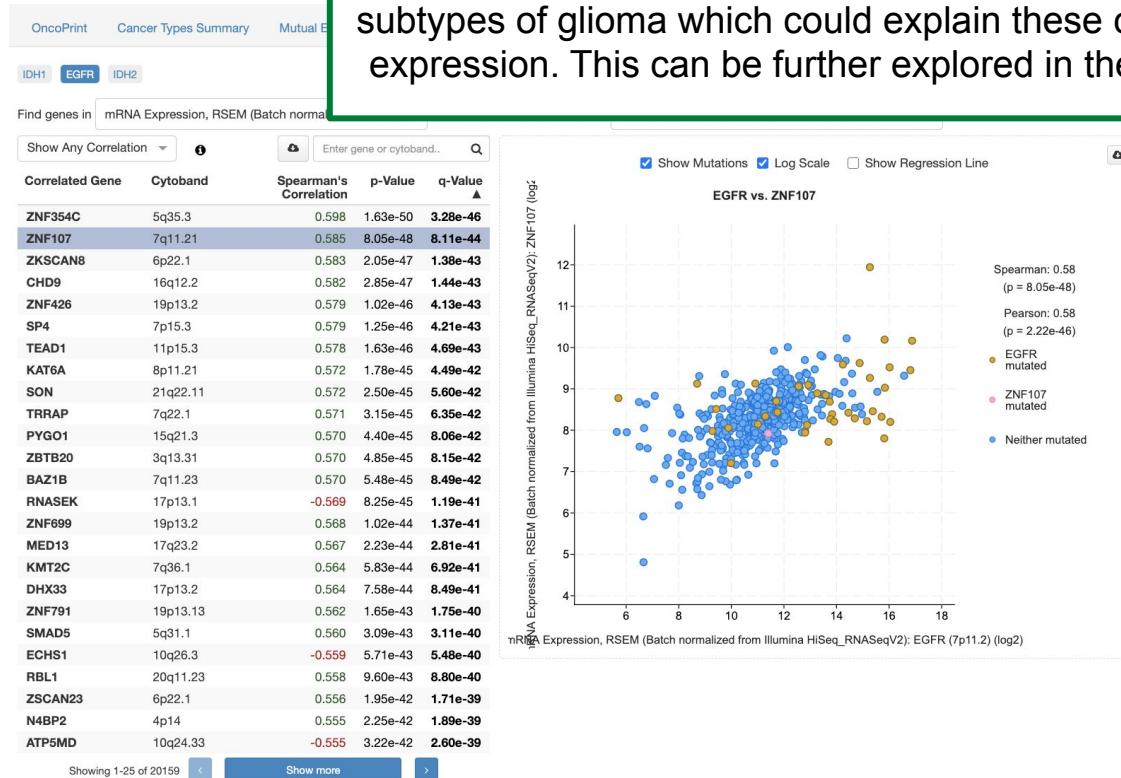
Click on a gene name to see correlation plot

Compares mRNA/protein level expression of your query genes against all other genes.



Check boxes to color-code sample dots by mutation status, change x- or y-axis to log scale, or add a regression line.

Co-Expression



Q: Several genes on chr7 show high expression correlation with EGFR within this cohort (see table on the left). Why might that be?

A: EGFR is also located on chr7 and is frequently gained in some subtypes of glioma which could explain these correlated increases in expression. This can be further explored in the “CN Segments” tab.

Comparison

This tab enables the comparison of all available data types between samples with or without alterations in the query genes. This tab replaces and enhances the old “Enrichments” tab.

The Comparison tab is the same as the Group Comparison functionality that is accessible from Study View. See the [Group Comparison Tutorial](#) for more details about the functionality of this tab.

Oncoprint Cancer Types Summary Mutual Exclusivity Plots Mutations Co-expression Comparison/Survival CN Segments Pathways Download

Groups: (drag to reorder) Altered group (465) Unaltered group (42) IDH1 (394) EGFR (54) IDH2 (24) Select all | Deselect all

Overlap Survival Clinical Mutations Copy-number mRNA Protein Microbiome Signature

By default, the “Altered” (one or more alterations in one or more query genes) and “Unaltered” (no alterations in any query gene) groups are selected.

Groups can be toggled on or off by clicking on them them. Analyses will update as the selections change.

Additional groups (deselected by default) correspond to each track shown in OncoPrint.

Comparison: Overlap

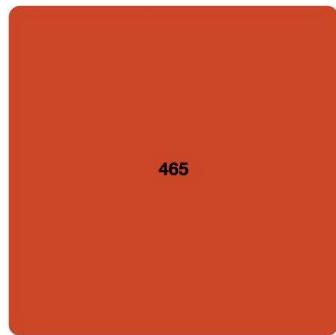
The Overlap subtab shows samples or patients that may overlap among the selected groups.

[OncoPrint](#)[Cancer Types Summary](#)[Mutual Exclusivity](#)[Plots](#)[Mutations](#)[Co-expression](#)[Comparison/Survival](#)[CN Segments](#)[Pathways](#)[Download](#)

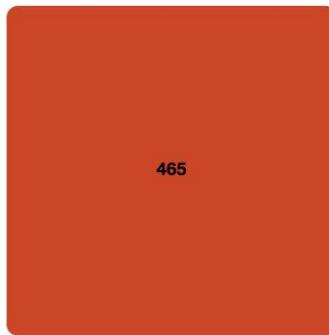
Groups: (drag to reorder) Altered group (465) Unaltered group (42) IDH1 (394) EGFR (54) IDH2 (24) Select all | Deselect all

Overlap Survival Clinical Mutations Copy-number mRNA Protein Microbiome Signature

Samples overlap



Patients overlap



- Altered group
- Unaltered group

[Create Group From Selected Diagram Areas](#)

[Create Group From Selected Diagram Areas](#)

Select one or more regions in the overlap plot to create a new group. If you are logged in, this new group can be saved to your profile for future use. Saved groups can be found under the “Groups” button in Study View.



Comparison: Survival

Select among different outcome measures. Options here depend on data availability for the study.

Oncoprint Cancer Types Summary Mutual Exclusivity Plots Mut

Groups: (drag to reorder) Altered group (465) Unaltered group (42) IDH1 (394) EGFR (54) IDH2 (24) Select all | Deselect all

Overlap Survival Clinical Mutations Copy-number mRNA Protein Microbiome Signature

ⓘ Interpret all outcome results with caution, as they can be confounded by many different variables that are not controlled for in these analyses. Overall

Survival Type	# Patients With Data	p-Value ▲	q-Value
Progression Free	505	< 10 ⁻¹⁰	1.21e-10
Disease-specific	497	5.79e-8	1.16e-7
Overall	505	4.27e-7	5.69e-7
Disease Free	129	0.458	0.458

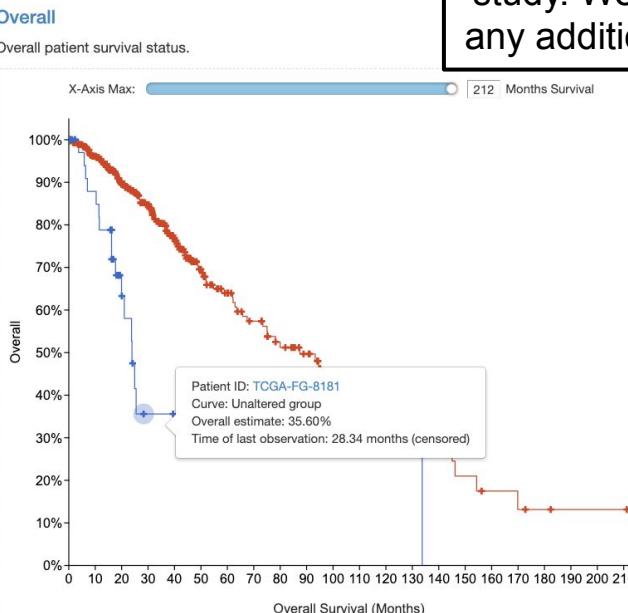
Showing 1-4 of 4

Q: Do patients with alterations in IDH1, IDH2 or EGFR have different outcomes compared to patients without alterations in any of those genes?

A: Patients with alterations in IDH1, IDH2 or EGFR have significantly better OS and DFS than patients without those alterations.

The Survival subtab replaces the old “Survival” tab. This subtab will only be visible if outcome data is available for the selected study.

Note: These plots reflect data as provided by the study. We do not perform any additional processing.



Comparison: Clinical

The Clinical subtab compares all available clinical data among the selected groups.

OncoPrint Cancer Types Summary Mutual Exclusivity Plots Mutations Co-expression Comparison/Survival CN Segments Pathways Download

Groups: (drag to reorder) Altered group (465) Unaltered group (42) IDH1 (394) EGFR (54) IDH2 (24) Select all | Deselect all

Overlap Survival Clinical Mutations Copy-number mRNA Protein Microbiome Signature

Clinical Attribute Attribute Type Statistical Test p-Value q-Value ▲ Swap Axes Horizontal Bars

Clinical Attribute	Attribute Type	Statistical Test	p-Value	q-Value ▲
Subtype	Patient	Chi-squared Test	< 10 ⁻¹⁰	< 10 ⁻¹⁰
10p Status	Sample	Chi-squared Test	< 10 ⁻¹⁰	< 10 ⁻¹⁰
10q Status	Sample	Chi-squared Test	5.52e-10	1.42e-8
11q Status	Sample	Chi-squared Test	2.25e-9	4.33e-8
22 (22q) Status	Sample	Chi-squared Test	3.58e-8	5.52e-7
20q Status	Sample	Chi-squared Test	1.15e-7	1.479e-6
19q Status	Sample	Chi-squared Test	1.97e-7	2.169e-6
6p Status	Sample	Chi-squared Test	2.64e-7	2.544e-6
1p Status	Sample	Chi-squared Test	8.04e-7	6.883e-6
7p Status	Sample	Chi-squared Test	9.21e-7	7.095e-6
Progress Free Survival (Months)	Patient	Kruskal Wallis Test	1.092e-5	7.646e-5
7q Status	Sample	Chi-squared Test	1.494e-5	9.586e-5
14 (14q) Status	Sample	Chi-squared Test	3.814e-5	2.259e-4
Aneuploidy Score	Sample	Kruskal Wallis Test	4.561e-5	2.509e-4
19p Status	Sample	Chi-squared Test	6.237e-5	3.201e-4
8q Status	Sample	Chi-squared Test	1.934e-4	9.305e-4
20p Status	Sample	Chi-squared Test	2.128e-4	9.640e-4
1q Status	Sample	Chi-squared Test	5.575e-4	2.385e-3
Months of disease-specific survival	Patient	Kruskal Wallis Test	1.118e-3	4.306e-3
Overall Survival (Months)	Patient	Kruskal Wallis Test	1.118e-3	4.306e-3

Showing 1-20 of 77 < Show more >

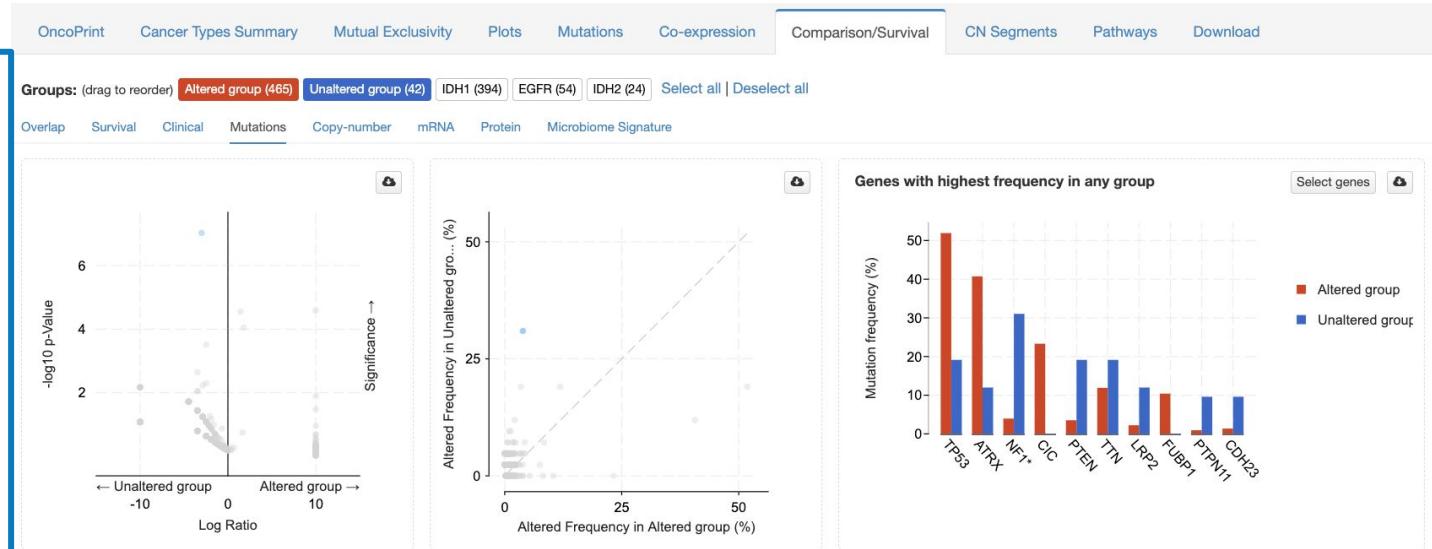
Plot Type: 100% stacked bar chart

The chart displays the distribution of IDH status across different clinical groups. The Y-axis represents the percentage of samples (# samples %) from 0 to 100. The X-axis shows the clinical groups: Altered group and Unaltered group. For each group, there are three bars representing IDH status: LGG_IDHmut-codel (blue), LGG_IDHmut-non-codel (pink), and LGG_IDHwt (green). In the Altered group, the LGG_IDHmut-codel bar reaches approximately 100%. In the Unaltered group, the LGG_IDHwt bar reaches approximately 100%, while the LGG_IDHmut-codel bar is near zero.

Click on a clinical attribute to visualize the data in the plot on the right.

Comparison: Genomic Profiles

The Mutations, Copy-number, mRNA and Protein subtabs replace the old “Enrichments” tab. These analyses ask whether mutations (or copy-number alterations, mRNA expression or protein expression) in a particular gene is enriched in one of the selected groups. These, and additional subtabs like Microbiome Signature, will be visible depending on the data available for each study.



Mutations

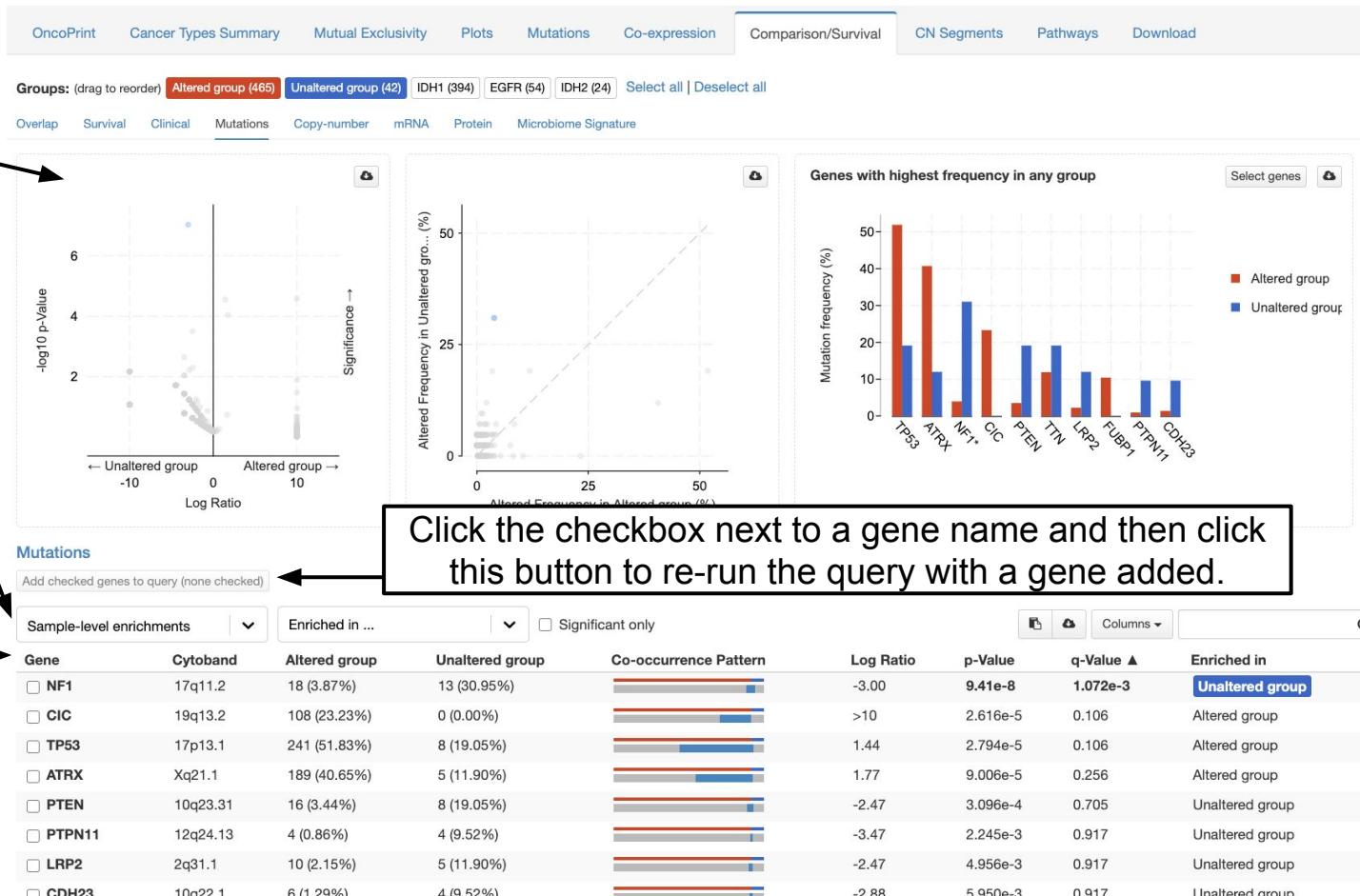
Add checked genes to query (none checked)

Sample-level enrichments | Enriched in ... | Significant only

Gene	Cytoband	Altered group	Unaltered group	Co-occurrence Pattern	Log Ratio	p-Value	q-Value ▲	Enriched in
<input type="checkbox"/> NF1	17q11.2	18 (3.87%)	13 (30.95%)		-3.00	9.41e-8	1.072e-3	Unaltered group
<input type="checkbox"/> CIC	19q13.2	108 (23.23%)	0 (0.00%)		>10	2.616e-5	0.106	Altered group
<input type="checkbox"/> TP53	17p13.1	241 (51.83%)	8 (19.05%)		1.44	2.794e-5	0.106	Altered group
<input type="checkbox"/> ATRX	Xq21.1	189 (40.65%)	5 (11.90%)		1.77	9.006e-5	0.256	Altered group
<input type="checkbox"/> PTEN	10q23.31	16 (3.44%)	8 (19.05%)		-2.47	3.096e-4	0.705	Unaltered group
<input type="checkbox"/> PTPN11	12q24.13	4 (0.86%)	4 (9.52%)		-3.47	2.245e-3	0.917	Unaltered group
<input type="checkbox"/> LRP2	2q31.1	10 (2.15%)	5 (11.90%)		-2.47	4.956e-3	0.917	Unaltered group
<input type="checkbox"/> CDH23	10q22.1	6 (1.29%)	4 (9.52%)		-2.88	5.950e-3	0.917	Unaltered group

Link to this page

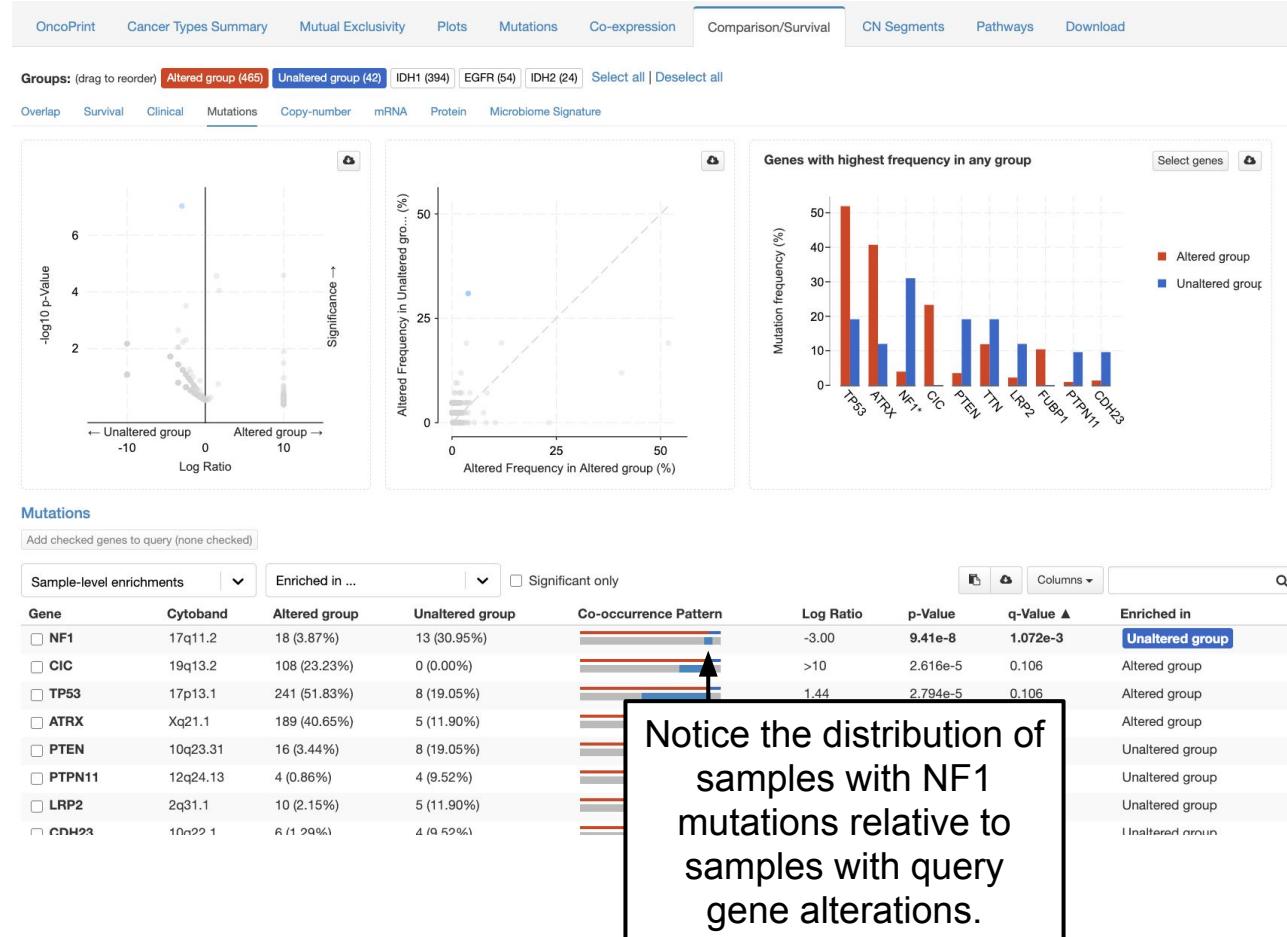
Comparison: Genomic Profiles



Comparison

Q: Alterations in IDH1, IDH2 and EGFR are mutually exclusive but some samples have alterations in none of these genes. Do samples without IDH1, IDH2 or EGFR alterations commonly have mutations in one or more other genes?

A: Mutations in NF1 are significantly mutually exclusive with alterations in IDH1, IDH2 and EGFR (see table). Try adding NF1 to the query (check the box next to NF1 and then click “Add checked genes to query”) and examine the OncoPrint and the Mutual Exclusivity tabs.



CN Segments

View copy number for each sample at each query gene via the [Integrated Genomics Viewer](#) (IGV).



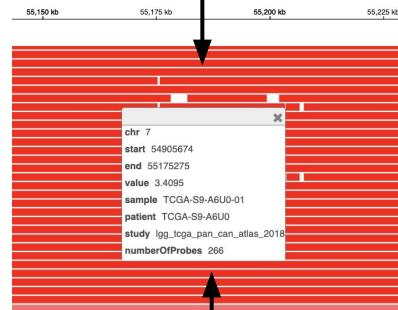
Plots for each gene appear on a separate tab.

Toggle track labels, a vertical line marking the center of the viewing screen, and a vertical line that moves with your cursor. Use to zoom in or out.

Click for track settings, including expanding the height of each sample (see below)

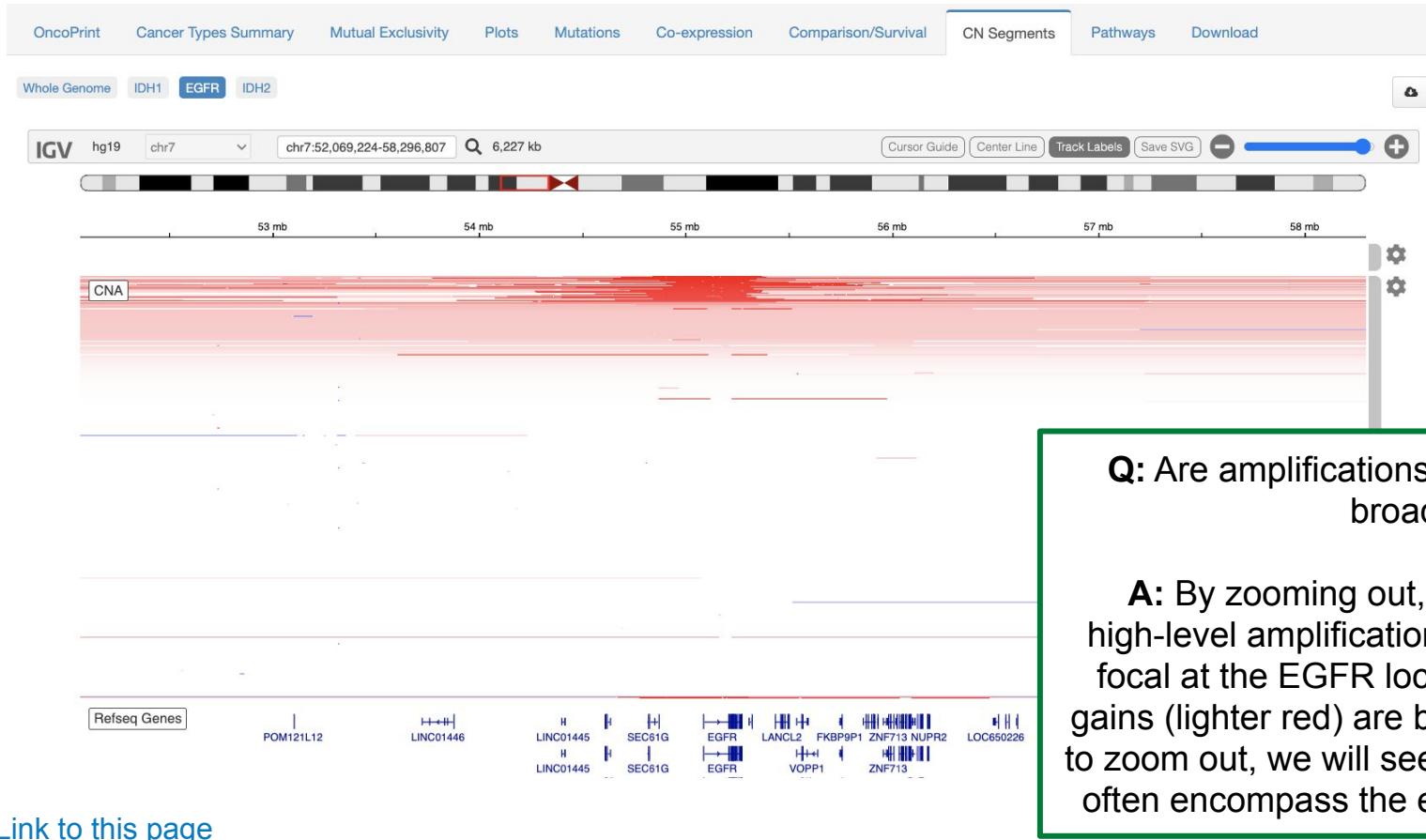
Each row is a single sample

Gene structures



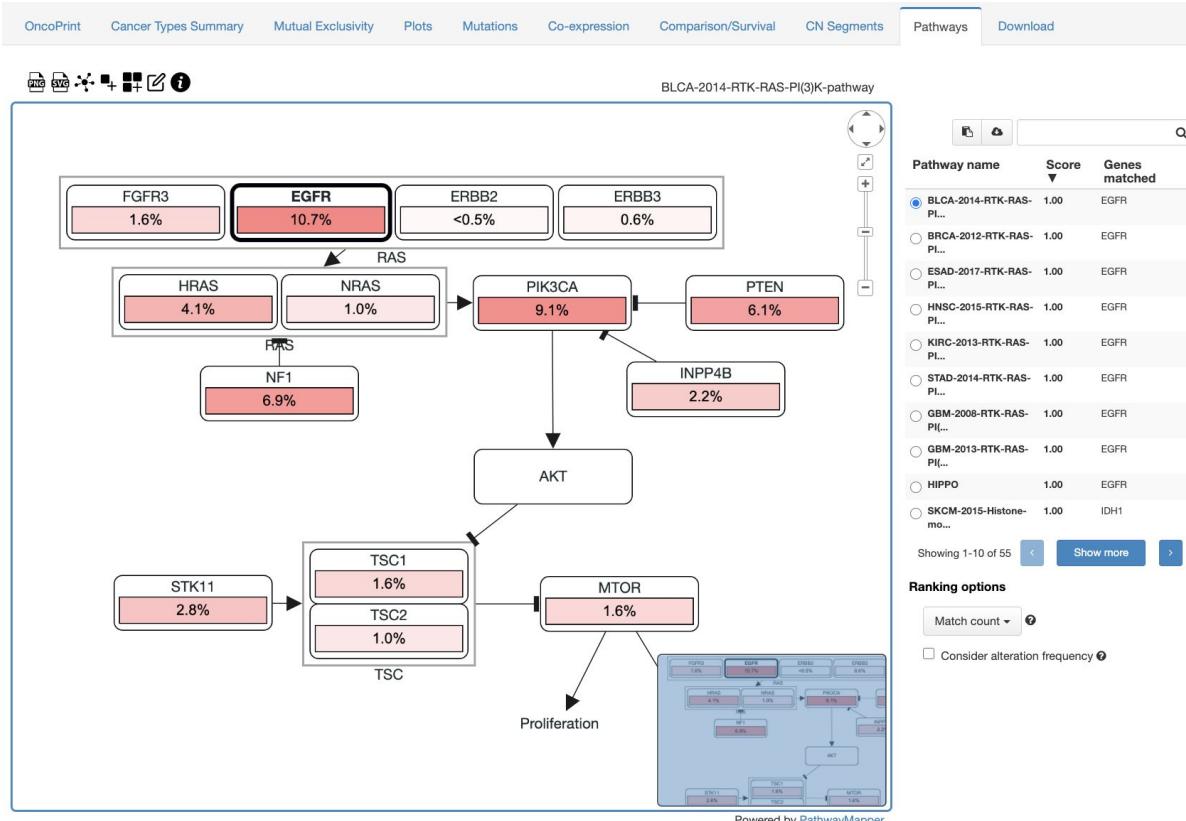
Click on a read for details

CN Segments



Pathways

The Pathways tab replaces the now retired “Network” tab. This tab in an integration with [PathwayMapper](#). The tab enables exploration of the queried genes in the context of Pathways defined by TCGA. For more detail on this tab, refer to the [Pathways Tutorial](#).



[Link to this page](#)

Download

Download data or copy lists of samples.

OncoPrint Cancer Types Summary Mutual Exclusivity Plots Mutations Co-expression Comparison/Survival CN Segments Pathways Download

Downloadable Data Files

Copy-number Alterations (OQL is not in effect)
Mutations (OQL is not in effect)
Altered samples: List of samples with alterations
Unaltered samples: List of samples without any alteration
Sample matrix: List of all samples where 1=altered and 0=unaltered
Log2 copy-number values ⓘ
mRNA Expression, RSEM (Batch normalized from Illumina HiSeq_RNASeqV2)
mRNA expression z-scores relative to diploid samples (RNA Seq V2 RSEM)
mRNA expression z-scores relative to all samples (log RNA Seq V2 RSEM)
Protein expression (RPPA) ⓘ
Protein expression z-scores (RPPA) ⓘ
Microbiome Signatures (log RNA Seq CPM) ⓘ

Tab Delimited Format | Transposed Matrix

Tab Delimited Format | Transposed Matrix

Copy | Download | Query | Virtual Study

Copy | Download | Query | Virtual Study

Copy | Download

Tab Delimited Format | Transposed Matrix

Heatmap tracks added in the OncoPrint tab can be downloaded here.

Download queried data types for the queried genes.

Download all other data types for the queried genes.

Gene Alteration Frequency

Gene Symbol	Num Samples Altered	Percent Samples Altered ▾
IDH1	394	78%
EGFR	54	11%
IDH2	24	5%

Columns q

Frequency of gene alteration for each gene in the query

Type of Genetic Alterations Across All Samples

Study ID	Sample ID	Patient ID	Altered ▾	IDH1	EGFR	IDH2
lgg_tcga_pan_can_atlas_2018	TCGA-CS-4938-01	TCGA-CS-4938	1	R132H	no alteration	no alteration
lgg_tcga_pan_can_atlas_2018	TCGA-CS-4941-01	TCGA-CS-4941	1	no alteration	AMP	no alteration
lgg_tcga_pan_can_atlas_2018	TCGA-CS-4942-01	TCGA-CS-4942	1	R132H	no alteration	no alteration
lgg_tcga_pan_can_atlas_2018	TCGA-CS-4943-01	TCGA-CS-4943	1	R132H	no alteration	no alteration
lgg_tcga_pan_can_atlas_2018	TCGA-CS-4944-01	TCGA-CS-4944	1	R132H	no alteration	no alteration
lgg_tcga_pan_can_atlas_2018	TCGA-CS-5390-01	TCGA-CS-5390	1	R132H	no alteration	no alteration
lgg_tcga_pan_can_atlas_2018	TCGA-CS-5393-01	TCGA-CS-5393	1	R132H	AMP	no alteration

Columns q

List of all samples with status of each query gene for each type of alteration queried (see [OQL tutorial](#) for explanation of alteration types included in a query)

Download

Download data or copy lists of samples.

OncoPrint Cancer Types Summary Mutual Exclusivity Plots Mutations Co-expression Comparison/Survival CN Segments Pathways Download

Downloadable Data Files

- Copy-number Alterations (OQL is not in effect)
- Mutations (OQL is not in effect)
- Altered samples: List of samples with alterations
- Unaltered samples: List of samples without any alteration
- Sample matrix: List of all samples where 1=altered and 0=unaltered
- Log2 copy-number values ⓘ
- mRNA Expression, RSEM (Batch normalized from Illumina HiSeq_RNASeqV2)
- mRNA expression z-scores relative to diploid samples (RNA Seq V2 RSEM)
- mRNA expression z-scores relative to all samples (log RNA Seq V2 RSEM)
- Protein expression (RPPA) ⓘ
- Protein expression z-scores (RPPA) ⓘ
- Microbiome Signatures (log RNA Seq CPM) ⓘ

Gene Alteration Frequency

Gene Symbol	Num Samples Altered	Percent Samples Altered ▾
IDH1	394	78%
EGFR	54	11%
IDH2	24	5%

Showing 1-3 of 3

Type of Genetic Alterations Across All Samples

Study ID	Sample ID	Patient ID	Altered ▾	IDH1	EGFR	IDH2
lgg_tcga_pan_can_atlas_2018	TCGA-CS-4938-01	TCGA-CS-4938	1	R132H	no alteration	no alteration
lgg_tcga_pan_can_atlas_2018	TCGA-CS-4941-01	TCGA-CS-4941	1	no alteration	AMP	no alteration
lgg_tcga_pan_can_atlas_2018	TCGA-CS-4942-01	TCGA-CS-4942	1	R132H	no alteration	no alteration
lgg_tcga_pan_can_atlas_2018	TCGA-CS-4943-01	TCGA-CS-4943	1	R132H	no alteration	no alteration
lgg_tcga_pan_can_atlas_2018	TCGA-CS-4944-01	TCGA-CS-4944	1	R132H	no alteration	no alteration
lgg_tcga_pan_can_atlas_2018	TCGA-CS-5390-01	TCGA-CS-5390	1	R132H	no alteration	no alteration
lgg_tcga_pan_can_atlas_2018	TCGA-CS-5393-01	TCGA-CS-5393	1	R132H	AMP	no alteration

List of samples that have an alteration in one or more query genes

List of samples that have no alterations in any query genes

List of all samples with summary classification:
0 = no alteration in any query gene
1 = alteration in one or more query genes

Advanced feature: use these lists to build a custom sample list to run a new query, to create [virtual studies](#) or to build [custom groups](#).

Questions?

Check out our other tutorials
or email us at:

cboportal@googlegroups.com