

# 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 (includes Survival, formerly a separate tab)
  - CN Segments
  - Pathways (replaces the Network tab)
  - 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:** 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.

Browse available  
datasets and  
select studies to  
explore or query

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

List of all studies,  
organized by  
organ system



# Single study query

**1. Filter the list of studies (optional)**

**2. Check the box for study of interest.**

**3. Select "Query By Gene"**

The screenshot shows the cBioPortal interface for a single study query. The search bar at the top contains the term 'glioma'. Below the search bar, the 'Select Studies for Visualization & Analysis' section displays a list of studies. The 'CNS/Brain' category is selected, showing 15 studies. The 'Soft Tissue' category shows 1 study. The 'Diffuse Glioma' section lists several studies, including 'Brain Lower Grade Glioma (TCGA, Firehose Legacy)' and 'Brain Lower Grade Glioma (TCGA, PanCancer Atlas)'. The 'Encapsulated Glioma' section lists 'Pilocytic Astrocytoma (ICGC, Nature Genetics 2013)'. The 'Miscellaneous Neuroepithelial Tumor' section is also visible. The 'Query By Gene' button is highlighted with a red box and an arrow pointing to it. The 'Explore Selected Studies' button is also visible. The 'What's New' section on the right shows a tweet from cBioPortal and a 'Subscribe' button for low-volume email news alerts. The 'Cancer Studies' section shows a bar chart of cases by top 20 primary sites, with 'Breast' and 'Prostate' being the most common.

**What's New** @cbiportal

cBioPortal @cbiportal

It's a small addition, but the new pencil icon allows you to quickly edit and update genes and query criteria (via OQL), in your cBioPortal queries.

Sign up for low-volume email news alerts

Subscribe

**Cancer Studies**

The portal contains 282 cancer studies (details)

Cases by Top 20 Primary Sites

Breast  
Prostate  
CNS/Brain  
Lung  
Lymphoid  
Bowel  
Kidney  
Stomach  
Myeloid  
Bladder  
Skin  
Uterus  
Head/Neck

# Single study query

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

Selected Studies: [Modify](#) Brain Lower Grade Glioma (TCGA, Firehose Legacy) (530 total samples)

Select Genomic Profiles:

- ☒ Mutations [?](#)
- ☒ Putative copy-number alterations from GISTIC [?](#)
- ☐ mRNA Expression. Select one of the profiles below:
  - ☐ mRNA Expression z-Scores (microarray) [?](#)
  - ☐ mRNA Expression z-Scores (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 (283) [×](#)

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](#)

4. This section lists all data types available for the selected study. Select data types to query.

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

5. Type gene(s) or select from pre-defined gene lists. cBioPortal will confirm that all entries are valid gene symbols.

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

6. Submit query

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

The screenshot shows the header of the OncoPrint interface. On the left, there is a 'Modify Query' button and a list of tabs: 'OncoPrint' (selected), 'Cancer Types Summary', 'Mutual Exclusivity', 'Plots', 'Mutations', 'Co-expression', 'Comparison', 'Survival', 'CN Segments', 'Pathways', and 'Download'. The main header area contains the study name 'Brain Lower Grade Glioma (TCGA, Firehose Legacy)' in blue, followed by the text 'Samples with mutation and CNA data (283 patients/samples) - IDH1 & EGFR'. To the right, it states 'Queried genes are altered in 243 (86%) of queried patients/samples' next to a circular icon with a link symbol.

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.

The screenshot shows the OncoPrint interface for a query titled "Brain Lower Grade Glioma (TCGA, Firehose Legacy)". The interface includes a "Modify Query" button, a list of tabs (OncoPrint, Annotate Data, Plots, Mutations, Co-expression, Comparison, Survival, CN Segments, Pathways, Download), and a "Genetic Alteration" section. The "Annotate Data" tab is active, showing options for "Putative drivers vs VUS" and "Filter Data". The "Putative drivers vs VUS" section has checkboxes for "OncoKB driver annotation", "Hotspots", "cBioPortal >= 10", and "COSMIC >= 10". The "Filter Data" section has checkboxes for "Exclude mutations and copy number alterations of unknown significance" and "Exclude germline mutations". A "Mutations" table is visible in the background, showing a list of mutations with a "View" dropdown and a "Download" button. A "Missense" label is also present.

**Brain Lower Grade Glioma (TCGA, Firehose Legacy)**  
Samples with mutation and CNA data (283 patients/samples) - IDH1 & EGFR

Queried genes are altered in 243 (86%) of queried patients/samples ⓘ

**Annotate Data ⓘ**

☒ Putative drivers vs VUS:

- ☒ OncoKB driver annotation
- ☒ Hotspots 🔥
- ☐ cBioPortal  $\geq 10$
- ☐ COSMIC  $\geq 10$

**Filter Data ⓘ**

- ☐ Exclude mutations and copy number alterations of unknown significance
- ☐ Exclude germline mutations

**Mutations** View Download 🔍 52 % 🔍

Missense

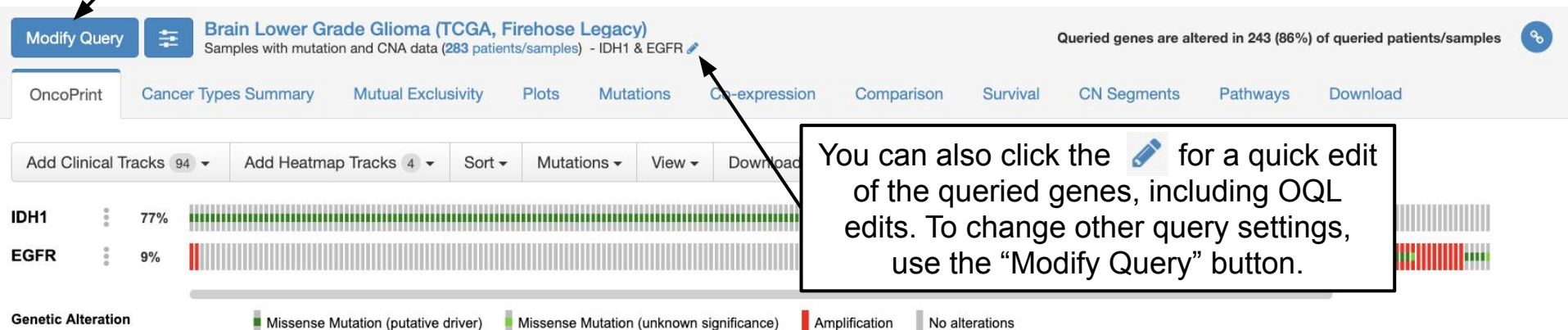
Set the definition of a putative driver vs variant of unknown significance (VUS).

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.

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).



# 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.

Cancel Modify Query

Brain Lower Grade Glioma (TCGA, Firehose Legacy)  
Samples with mutation and CNA data (283 patients/samples) - IDH1 & EGFR

Queried genes are altered in 243 (86%) of queried patients/samples

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

Query

Visualization & Analysis: 1 study selected (530 samples) Deselect all

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

3

PanCancer Studies

☐ MSK-IMPACT Clinical Sequencing Cohort (MSKCC, Nat Med 2017)

☐ Pan-Lung Cancer (TCGA, Nat Genet 2016)

☐ Pediatric Pan-cancer (Columbia U, Genome Med 2016)

10945 samples

1144 samples

103 samples

9

Cell lines

☐ Cancer Cell Line Encyclopedia (Broad, 2019)

☐ Cancer Cell Line Encyclopedia (Novartis/Broad, Nature 2012)

☐ NCI-60 Cell Lines (NCI, Cancer Res 2012)

1739 samples

1020 samples

67 samples

2

Adrenal Gland

☐ Adrenocortical Carcinoma Project (2019)

☐ Adrenocortical Carcinoma (TCGA, Firehose Legacy)

1049 samples

20 samples

Select Genomic Profiles:

☒ Mutations

☒ Putative copy-number alterations from GISTIC

☐ mRNA Expression. Select one of the profiles below:

☐ mRNA Expression z-Scores (microarray)

☐ mRNA Expression z-Scores (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 (283)

Enter Genes:

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

User-defined List

IDH1 EGFR IDH2

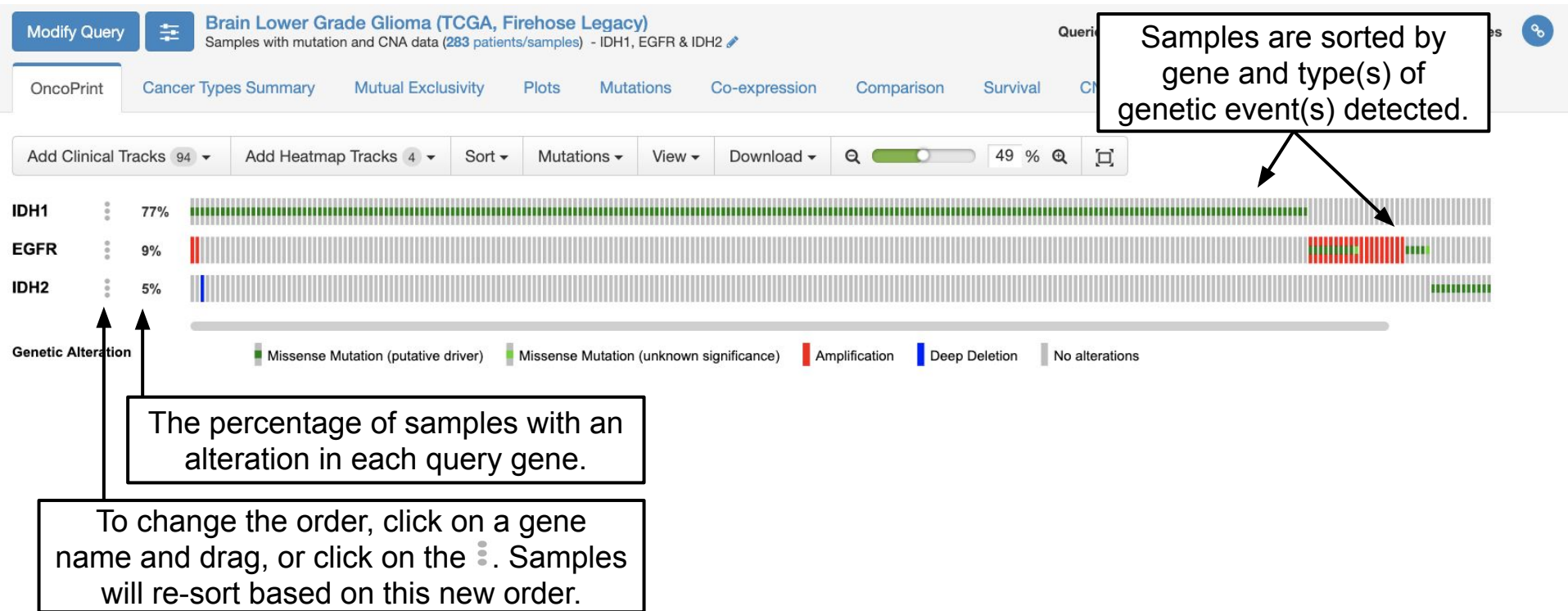
All gene symbols are valid.

Submit Query

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 a heatmap with RNA or protein levels or treatment response (when available)

Change the sample sorting order

Customize visualization

OncoPrint

Cancer Types Summary

Mutual Exclusivity

Plots

Mutations

Co-expression

Comparison

Survival

CN Segments

Pathways

Download

Add Clinical Tracks 94

Add Heatmap Tracks 4

Sort

Mutations

View

Download

Q

49 %

Q

Q

Select all (3) Deselect all

age

Name

Freq

- ☒ Diagnosis Age
- ☒ Animal Insect Allergy Age
- ☐ Age of Food Allergy

100.0%  
2.8%  
1.4%

Change the rules by which mutations are colored.

Download figure as PNG, PDF or SVG. Download patient/sample IDs in same order as OncoPrint.

# OncoPrint: Zoom

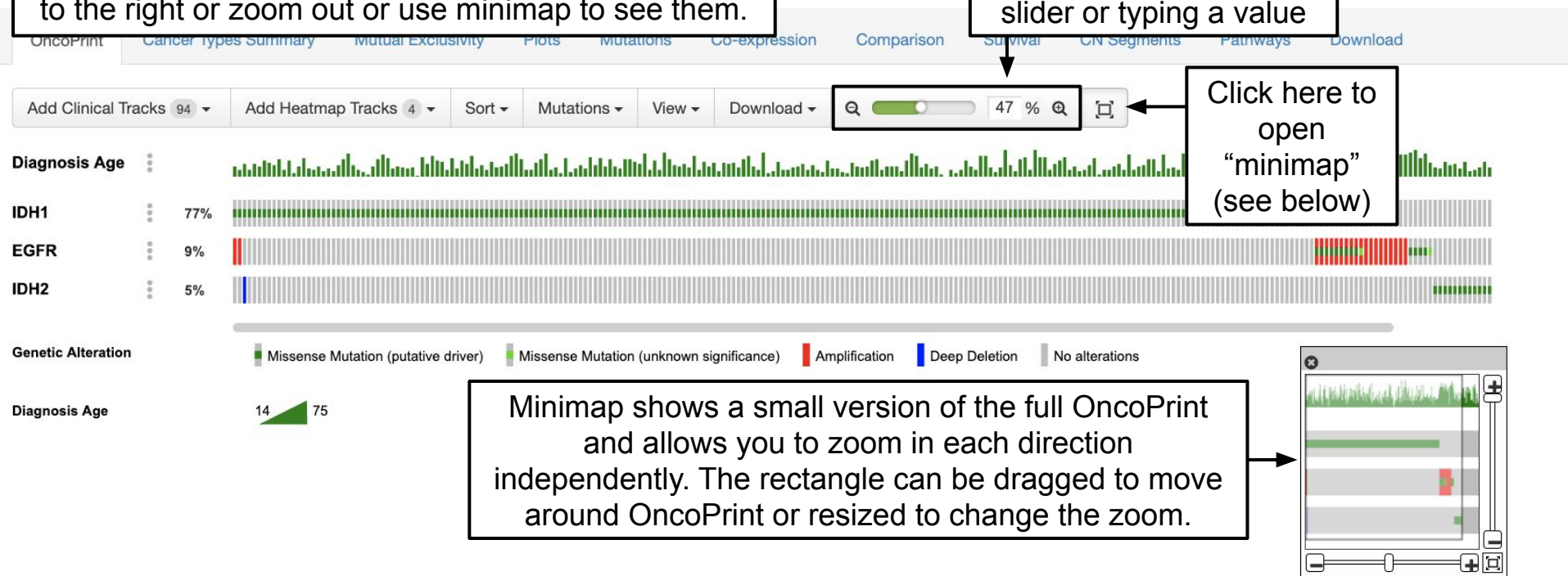
There may be more samples hiding off-screen. Scroll to the right or zoom out or use minimap to see them.

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

Click here to open "minimap" (see below)

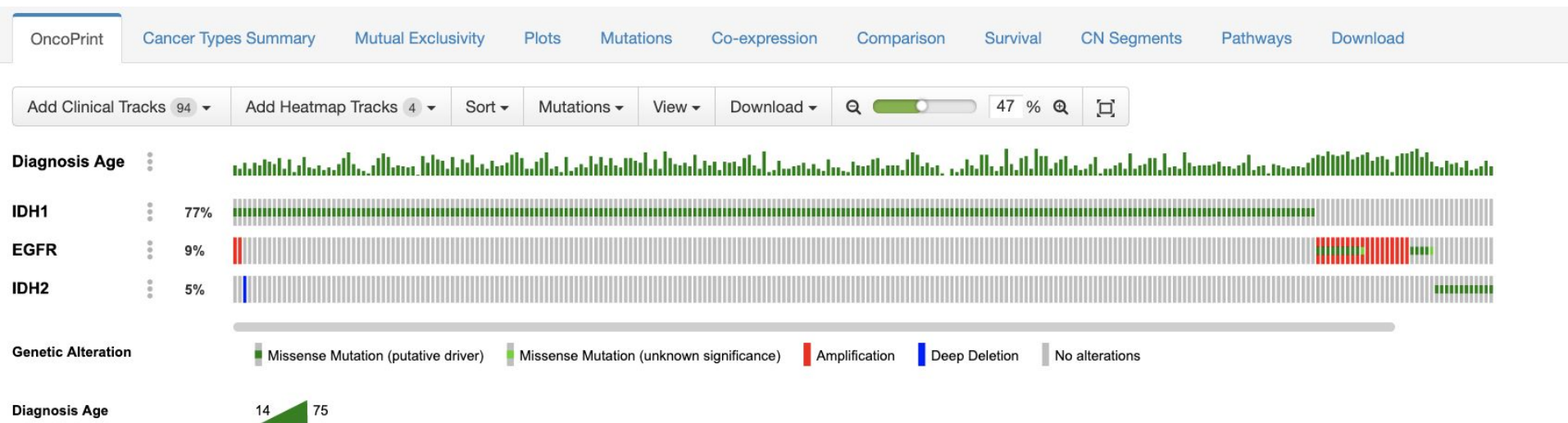
Minimap shows a small version of the full OncoPrint and allows you to zoom in each direction independently. The rectangle can be dragged to move around OncoPrint or resized to change the zoom.

This button zooms OncoPrint to show all samples with alterations





# OncoPrint: What can we learn?



**Q:** Are genetic alterations in these genes mutually exclusive?

**A:** We can see that samples with alterations in one gene tend to not have alterations in the other genes.

**Q:** Is there an association between alterations in a particular query gene and age?

**A:** We can see that patients with mutations or amplifications in EGFR tend to be older than those with mutations in IDH1/IDH2

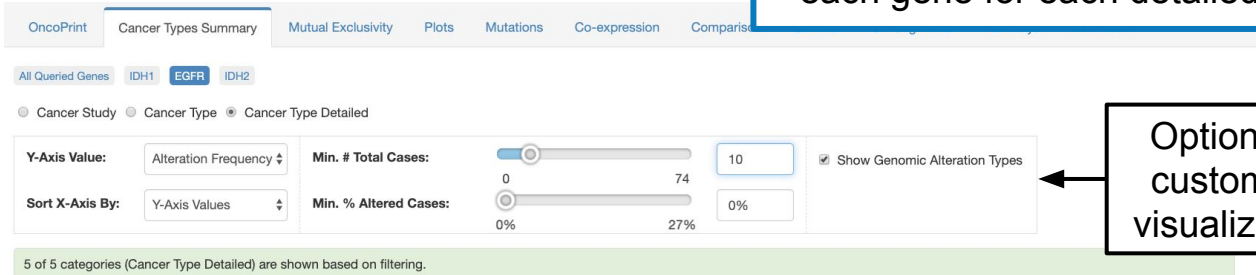
Now we're going to go through all the other tabs and ask some questions about alterations in *IDH1*, *IDH2* and *EGFR* in the TCGA Lower-Grade Glioma study.

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)

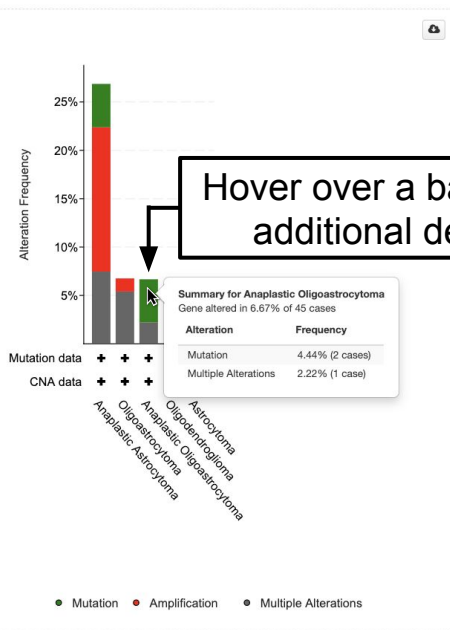
# Cancer Types Summary

Histogram of the frequency of alterations in each gene for each detailed cancer type.

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



Options to customize visualization



Hover over a bar to see additional details.

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

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

# 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 Mutat Survival

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

☒ Mutual exclusivity ☒ Co-occurrence ☐ Significant only

Columns ▾

A	B	Neither	A Not B	B Not A	Both	Log2 Odds Ratio	p-Value	q-Value ▲	Tendency
IDH1	EGFR	40	217	24	2	<-3	<0.001	<0.001	Mutual exclusivity
IDH1	IDH2	52	218	12	1	<-3	<0.001	<0.001	Mutual exclusivity
EGFR	IDH2	244	26	13	0	<-3	0.278	0.278	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.

# Plots


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.

Example plot settings

Choose type of data

Select a query gene

Swap horizontal & vertical axis

Modify Query  Brain Lower Grade Glioma (TCGA, Firehose Legacy) Samples with mutation and CNA data (283 patients/samples) - IDH1

OncoPrint Cancer Types Summary Mutual Exclusivity **Plots** Mut

Examples: Mut# vs Dx FGA vs Dx Mut# vs FGA mRNA vs Dx mRNA vs mut type mRNA vs CNA mRNA vs methyl Protein vs mRNA

**Horizontal Axis**

Data Type  
Copy Number

Copy Number Profile  
Putative copy-number alteration..

Gene  
EGFR

**Vertical Axis**

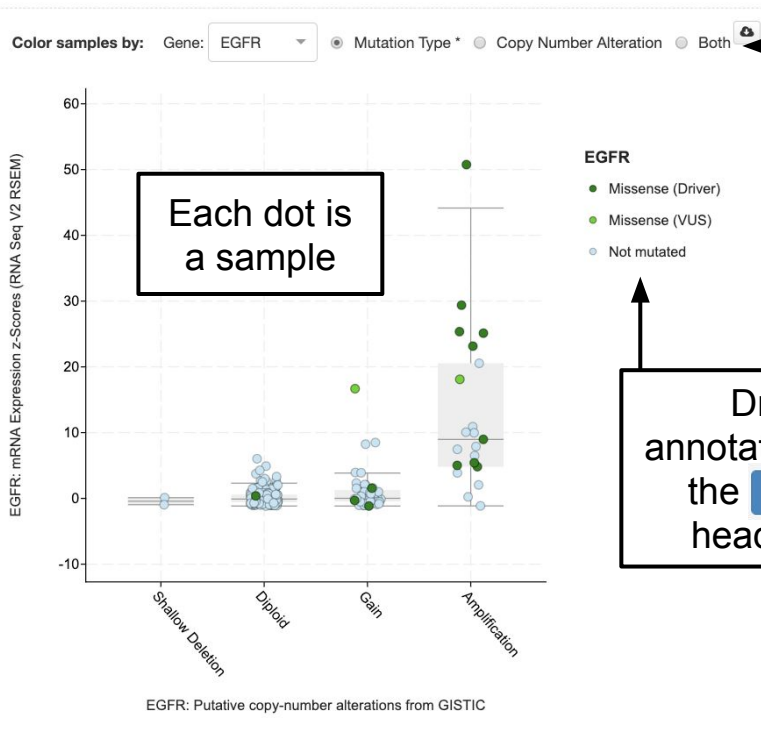
Data Type  
mRNA

mRNA Profile  
mRNA Expression z-Scores (RN..


Gene  
Same gene (EGFR)

Search Case(s)  
Case ID..

Search Mutation(s)  
Protein Change..



Select color scheme

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

\* Driver annotation settings are located in the settings menu  at the top of the 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).



\* Driver annotation settings are located in the settings menu ☰ at the top of the page.

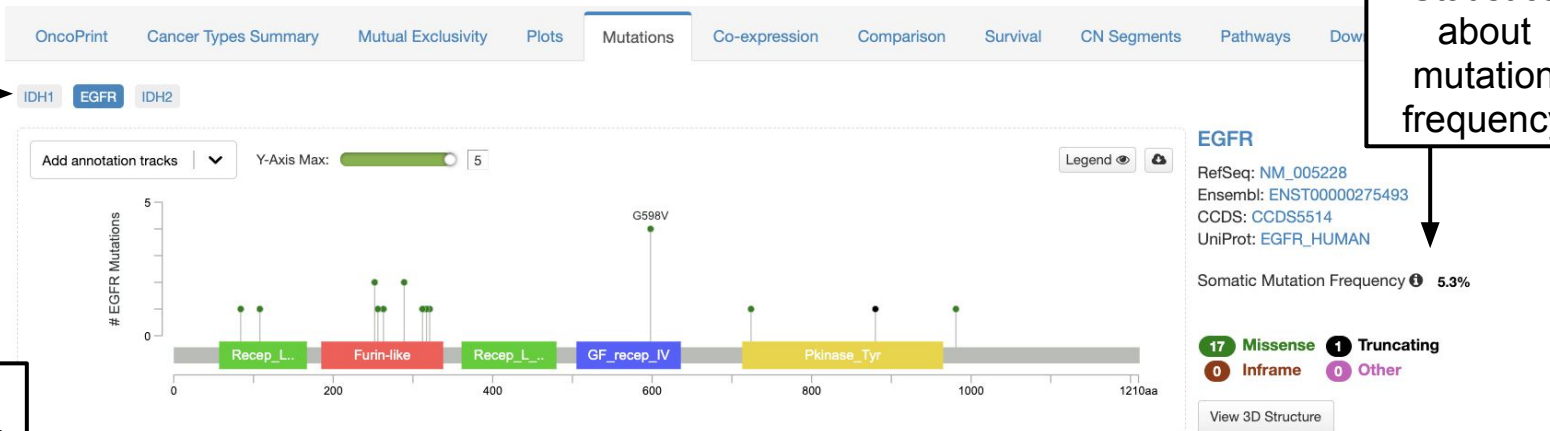
# Mutations

This tab shows details about all mutations called in a particular gene.

Each gene appears on a separate tab

Table of all mutations with annotations

Statistics about mutation frequency



8 Mutations (page 1 of 1)

Sample ID	Cancer Type	Protein Change	Annotation ▼	Mutation Type	Copy #	COSMIC	Allele Freq (T)	# Mut in Sample
TCGA-FG-6692-01	Anaplastic Oligoastrocytoma	A289V		Missense	Amp	50	0.95	59
TCGA-DU-A5TT-01	Anaplastic Oligoastrocytoma	A289V		Missense			0.0	
TCGA-HT-8110-01	Anaplastic Astrocytoma	R108K		Missense			0.0	
TCGA-HT-A61C-01	Anaplastic Oligoastrocytoma	T263P		Missense			0.0	
TCGA-DU-7013-01	Anaplastic Astrocytoma	G598V		Missense			0.0	
TCGA-DU-8162-01	Oligoastrocytoma	G598V		Missense			0.0	
TCGA-FG-A4MU-01	Oligoastrocytoma	G598V		Missense	Amp	36	0.0	
TCGA-HT-A5RC-01	Anaplastic Astrocytoma	G598V		Missense	Amp	36	0.41	46
TCGA-HT-8104-01	Anaplastic Astrocytoma	G724S		Missense	Amp	7	0.06	49
TCGA-DU-7292-01	Anaplastic Astrocytoma	R252P		Missense	Amp	12	0.79	34
TCGA-DU-7012-01	Anaplastic Astrocytoma	D256G		Missense	Gain	5	0.38	39

Show additional columns

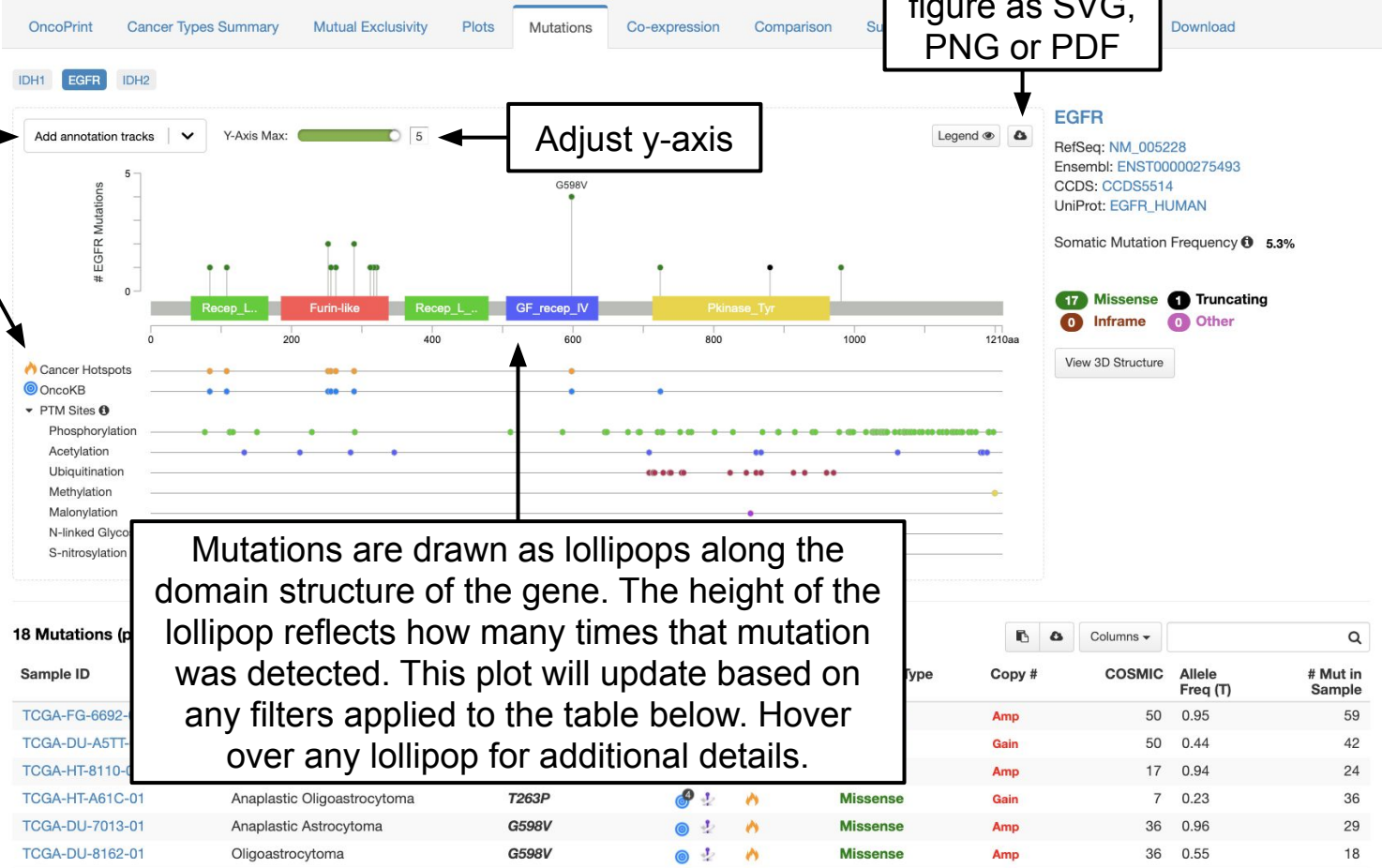
Filter based on any visible text column

# Mutations

Add annotation tracks to the plot

Adjust y-axis

Download figure as SVG, PNG or PDF



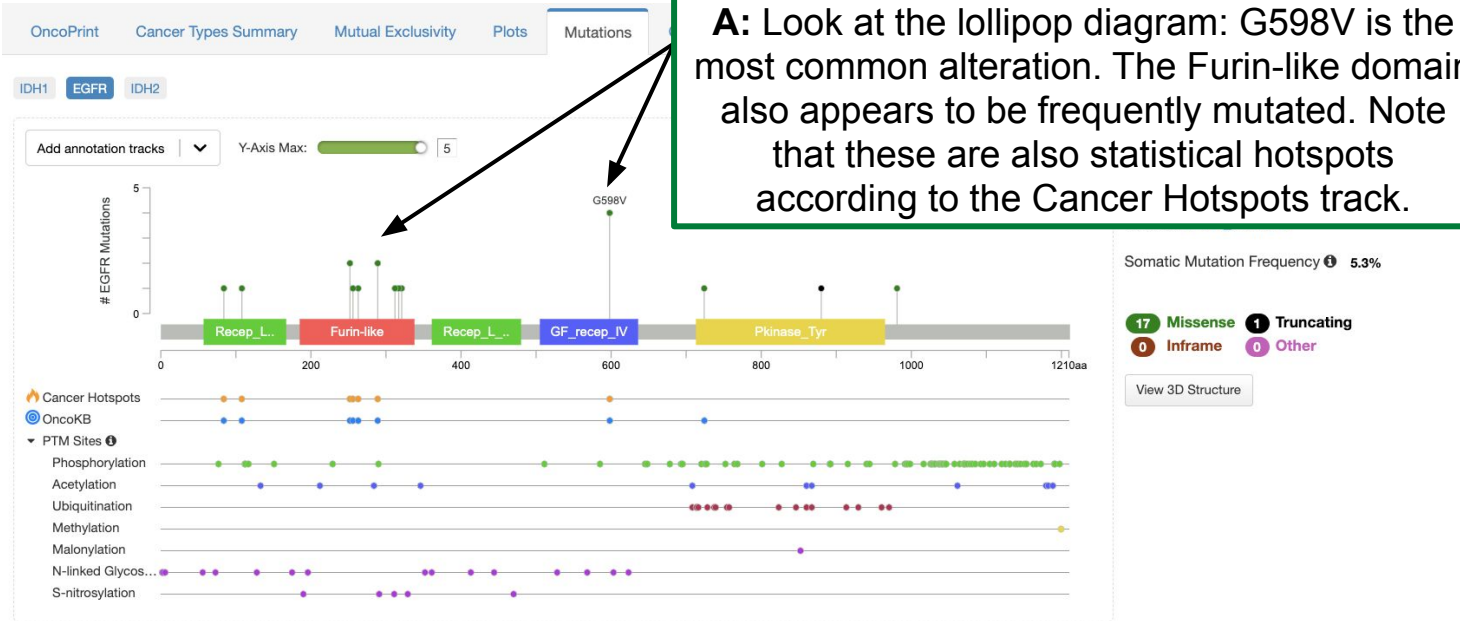
Mutations are drawn as lollipops along the domain structure of the gene. The height of the lollipop reflects how many times that mutation was detected. This plot will update based on any filters applied to the table below. Hover over any lollipop for additional details.





















# 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.



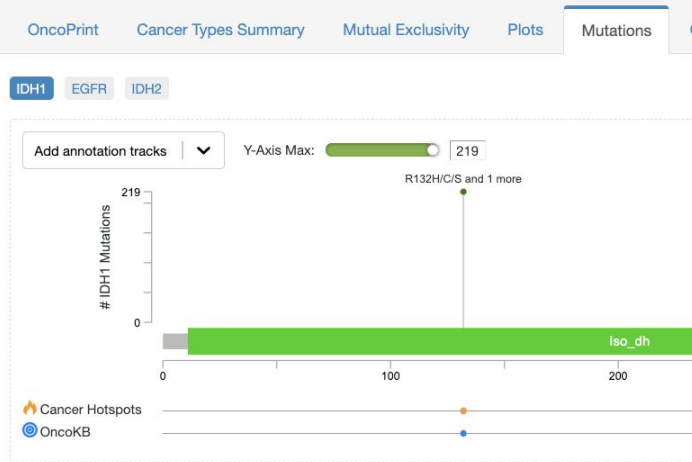
18 Mutations (page 1 of 1)

Sample ID	Cancer Type	Protein Change	Annotation ▼	Mutation Type	Copy #	COSMIC	Allele Freq (T)	# Mut in Sample
TCGA-FG-6692-01	Anaplastic Oligoastrocytoma	A289V	  	Missense	Amp	50	0.95	59
TCGA-DU-A5TT-01	Anaplastic Oligoastrocytoma	A289V	  	Missense	Gain	50	0.44	42
TCGA-HT-8110-01	Anaplastic Astrocytoma	R108K	  	Missense	Amp	17	0.94	24
TCGA-HT-A61C-01	Anaplastic Oligoastrocytoma	T263P	  	Missense	Gain	7	0.23	36
TCGA-DU-7013-01	Anaplastic Astrocytoma	G598V	  	Missense	Amp	36	0.96	29
TCGA-DU-8162-01	Oligoastrocytoma	G598V	  	Missense	Amp	36	0.55	18

# Mutations


**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.





Somatic Mutation Frequency ⓘ 77.4%

219 Missense 0 Truncating  
0 Inframe 0 Other

 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.

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

You may also see this symbol  which means the mutation is a [recurrent hotspot](#) based on a [statistical analysis of 3D protein conformation](#).

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

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

TCGA-HT-7693-01	Oligodendroglioma	R132C	Annotation	Missense	Diploid	4964	0.39	29
TCGA-HT-7855-01	Anaplastic Astrocytoma	R132C	Annotation	Missense	ShallowDel	4964	0.48	24
TCGA-DB-A4XD-01			Annotation	Missense		4964	0.46	30
TCGA-DB-A4XF-01			Annotation	Missense		4964	0.39	28
TCGA-P5-A5EZ-01			Annotation	Missense		4964	0.40	17
TCGA-P5-A5F4-01			Annotation	Missense		4964	0.24	25
TCGA-P5-A5F4-01			Annotation	Missense		4964	0.31	24
TCGA-P5-A5F4-01			Annotation	Missense		4964	0.31	24

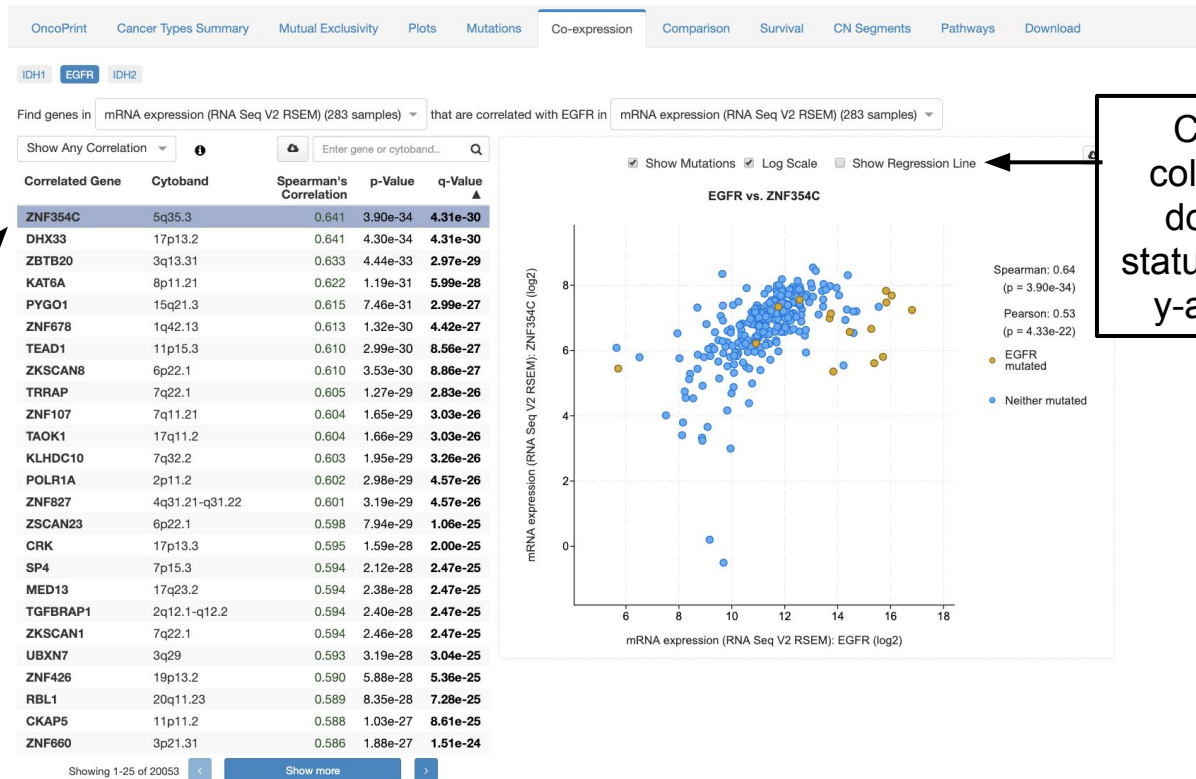
# Co-Expression

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

Each gene appears on a separate tab

Select from available data types

Click on a gene name to see correlation plot



Check boxes to color-code sample dots by mutation status or change x- or y-axis to log scale

# Co-Expression

**Q:** Which genes have expression that correlates with EGFR expression across the cohort?

**A:** EGFR is on chr7 and many other genes located on chr7 have expression levels correlated with EGFR expression (see table on the left). Chr7 is frequently gained in some subtypes of glioma which could explain these correlated increases in expression.

OncoPrint Cancer Types Summary Mutual Exclusion

IDH1 EGFR IDH2

Find genes in mRNA expression (RNA Seq V2 RSEM) (283)

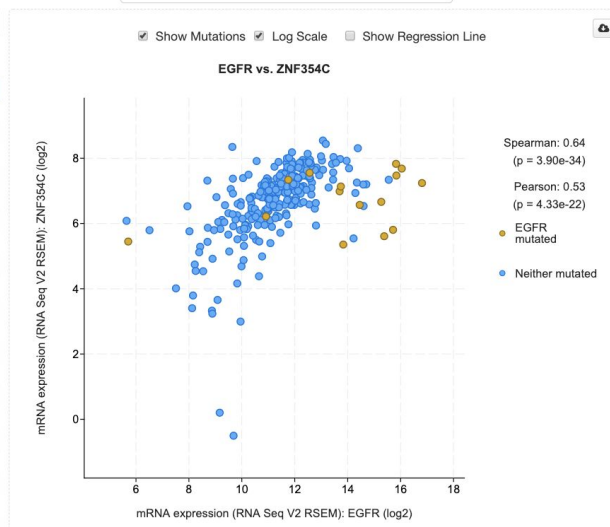
Show Any Correlation

Enter gene or cytoband..

Correlated Gene	Cytoband	Spearman's Correlation	p-Value	q-Value
ZNF354C	5q35.3	0.641	3.90e-34	4.31e-30
DHX33	17p13.2	0.641	4.30e-34	4.31e-30
ZBTB20	3q13.31	0.633	4.44e-33	2.97e-29
KAT6A	8p11.21	0.622	1.19e-31	5.99e-28
PYGO1	15q21.3	0.615	7.46e-31	2.99e-27
ZNF678	1q42.13	0.613	1.32e-30	4.42e-27
TEAD1	11p15.3	0.610	2.99e-30	8.56e-27
ZKSCAN8	6p22.1	0.610	3.53e-30	8.86e-27
TRRAP	7q22.1	0.605	1.27e-29	2.83e-26
ZNF107	7q11.21	0.604	1.65e-29	3.03e-26
TAOK1	17q11.2	0.604	1.66e-29	3.03e-26
KLHDC10	7q32.2	0.603	1.95e-29	3.26e-26
POLR1A	2p11.2	0.602	2.98e-29	4.57e-26
ZNF827	4q31.21-q31.22	0.601	3.19e-29	4.57e-26
ZSCAN23	6p22.1	0.598	7.94e-29	1.06e-25
CRK	17p13.3	0.595	1.59e-28	2.00e-25
SP4	7p15.3	0.594	2.12e-28	2.47e-25
MED13	17q23.2	0.594	2.38e-28	2.47e-25
TGFBRAP1	2q12.1-q12.2	0.594	2.40e-28	2.47e-25
ZKSCAN1	7q22.1	0.594	2.46e-28	2.47e-25
UBXN7	3q29	0.593	3.19e-28	3.04e-25
ZNF426	19p13.2	0.590	5.88e-28	5.36e-25
RBL1	20q11.23	0.589	8.35e-28	7.28e-25
CKAP5	11p11.2	0.588	1.03e-27	8.61e-25
ZNF660	3p21.31	0.586	1.88e-27	1.51e-24

Showing 1-25 of 20053

Show more



# 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.

The screenshot shows the OncoPrint web application interface. The top navigation bar includes tabs for OncoPrint, Cancer Types Summary, Mutual Exclusivity, Plots, Mutations, Co-expression, Comparison (which is the active tab), Survival, CN Segments, Pathways, and Download. Below this, the 'Groups' section displays a list of selected groups: 'Altered group (255)' in a red box and 'Unaltered group (28)' in a blue box. To their right are three deselected groups: 'IDH1 (219)', 'EGFR (26)', and 'IDH2 (13)', each in a light blue box. Further right are links for 'Select all' and 'Deselect all'. Below the groups, a row of data type tabs is visible: Overlap, Survival, Clinical, Mutations, Copy-number, mRNA, and Protein. Three callout boxes provide additional information: 1. A box on the left points to the 'Altered' and 'Unaltered' groups, stating that these are the default selections. 2. A box in the middle points to the 'IDH1', 'EGFR', and 'IDH2' groups, stating that these represent additional groups deselected by default, corresponding to tracks in OncoPrint. 3. A box on the right points to the 'Select all' and 'Deselect all' links, stating that groups can be toggled on or off by clicking on them, and analyses will update as selections change.

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

Groups: (drag to reorder) **Altered group (255)** **Unaltered group (28)** IDH1 (219) EGFR (26) IDH2 (13) [Select all](#) [Deselect all](#)

Overlap Survival Clinical Mutations Copy-number mRNA Protein

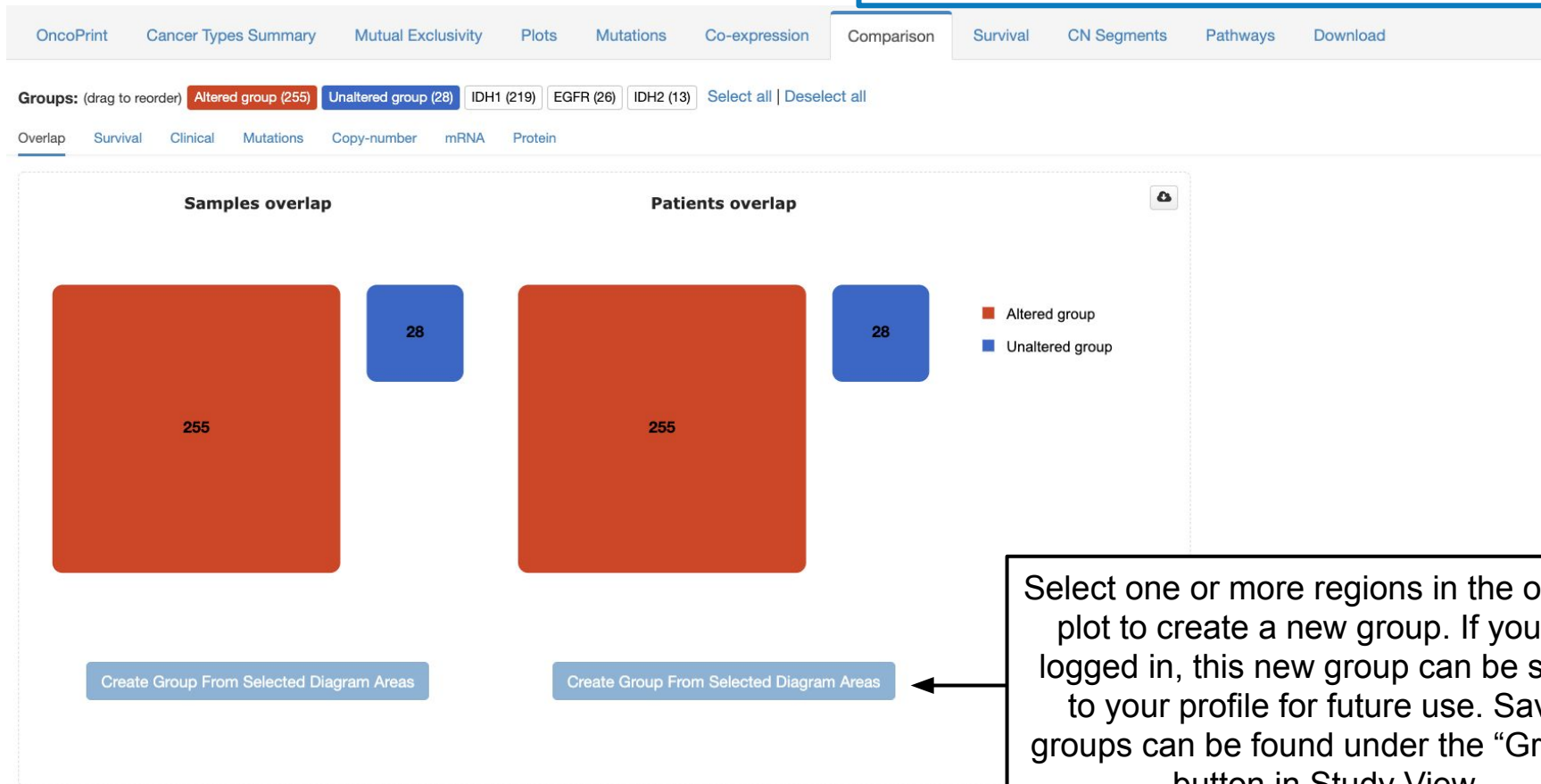
By default, the “Altered” (one or more alterations in one or more query genes) and “Unaltered” (no alterations in any query gene) groups are selected.

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

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

# Comparison: Overlap

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



# Comparison: Survival

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

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



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

**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.



# Comparison: Clinical

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

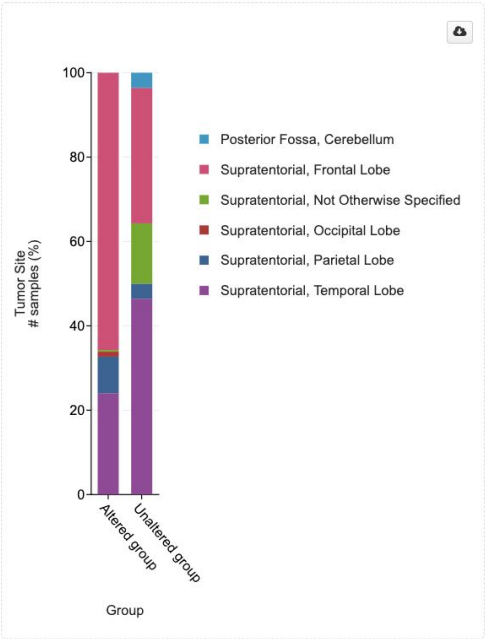
Groups: (drag to reorder) **Altered group (255)** **Unaltered group (28)** IDH1 (219) EGFR (26) IDH2 (13) [Select all](#) [Deselect all](#)

Overlap Survival **Clinical** Mutations Copy-number mRNA Protein

Plot Type 100% stacked bar chart

☐ Swap Axes ☐ Horizontal Bars

Clinical Attribute	Attribute Type	Statistical Test	p-Value	q-Value ▲
Tumor Site	Patient	Chi-squared Test	5.56e-9	4.23e-7
Disease Free (Months)	Patient	Kruskal Wallis Test	1.308e-4	4.971e-3
New Neoplasm Event Post Initial Therapy Indicator	Patient	Chi-squared Test	3.976e-4	0.0101
Overall Survival (Months)	Patient	Kruskal Wallis Test	7.763e-4	0.0147
Fraction Genome Altered	Sample	Kruskal Wallis Test	0.0255	0.350
Race Category	Patient	Chi-squared Test	0.0276	0.350
Overall Survival Status	Patient	Chi-squared Test	0.0516	0.490
Patient's Vital Status	Patient	Chi-squared Test	0.0516	0.490
Mutation Count	Sample	Kruskal Wallis Test	0.0738	0.623
Neoplasm Histologic Type Name	Patient	Chi-squared Test	0.0848	0.644
Inherited genetic syndrome found	Patient	Chi-squared Test	0.0998	0.661
Disease Free Status	Patient	Chi-squared Test	0.104	0.661
Ldh1 mutation test method	Patient	Chi-squared Test	0.114	0.667
Primary Therapy Outcome Success Type	Patient	Chi-squared Test	0.138	0.746
Supratentorial Localization	Patient	Chi-squared Test	0.161	0.810
Neoadjuvant Therapy Type Administered Prior To Resection	Patient	Chi-squared Test	0.178	0.810



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.

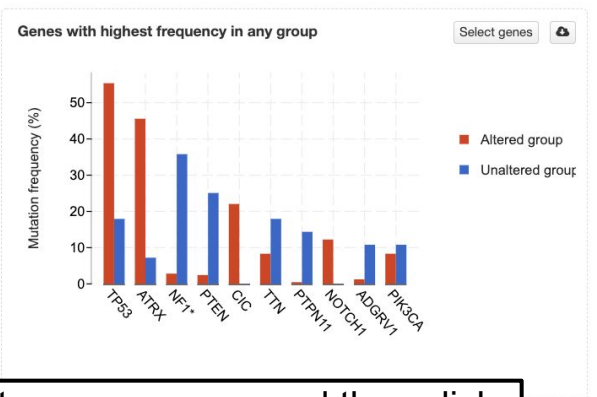
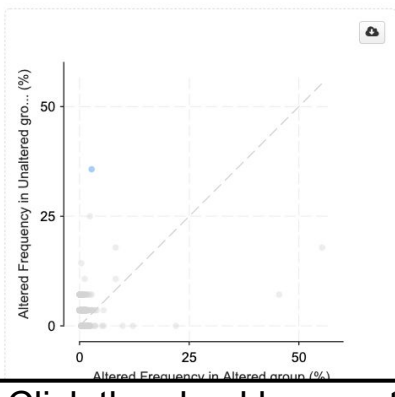
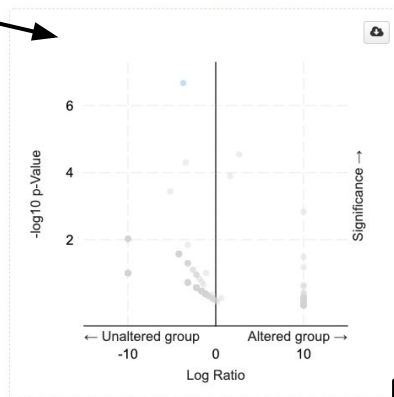


# Comparison: Genomic Profiles

Groups: (drag to reorder) **Altered group (255)** **Unaltered group (28)** IDH1 (219) EGFR (26) IDH2 (13) [Select all](#) [Deselect all](#)

[Overlap](#)
[Survival](#)
[Clinical](#)
[Mutations](#)
[Copy-number](#)
[mRNA](#)
[Protein](#)

Hover over a dot to see the gene name



Select sample-level or patient-level analysis

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

Click the checkbox next to a gene name and then click this button to re-run the query with a gene added.

Mutations

Add checked genes to query (none checked)

Sample-level enrichments

Select enriched groups

☐ Significant only

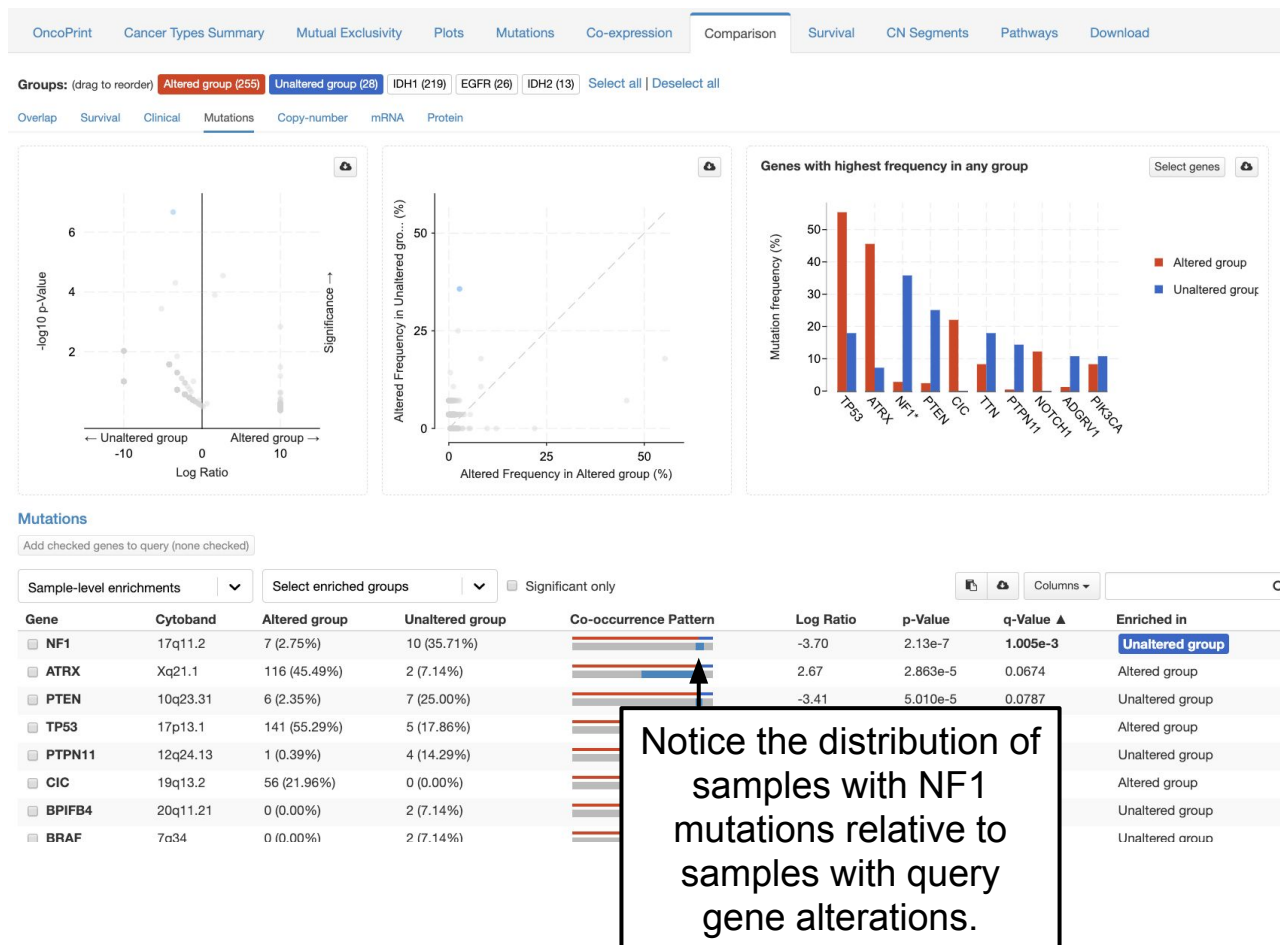
Columns

Gene	Cytoband	Altered group	Unaltered group	Co-occurrence Pattern	Log Ratio	p-Value	q-Value ▲	Enriched in
<input type="checkbox"/> NF1	17q11.2	7 (2.75%)	10 (35.71%)		-3.70	2.13e-7	1.005e-3	<b>Unaltered group</b>
<input type="checkbox"/> ATRX	Xq21.1	116 (45.49%)	2 (7.14%)		2.67	2.863e-5	0.0674	Altered group
<input type="checkbox"/> PTEN	10q23.31	6 (2.35%)	7 (25.00%)		-3.41	5.010e-5	0.0787	Unaltered group
<input type="checkbox"/> TP53	17p13.1	141 (55.29%)	5 (17.86%)		1.63	1.272e-4	0.150	Altered group
<input type="checkbox"/> PTPN11	12q24.13	1 (0.39%)	4 (14.29%)		-5.19	3.644e-4	0.343	Unaltered group
<input type="checkbox"/> CIC	19q13.2	56 (21.96%)	0 (0.00%)		>10	1.460e-3	0.767	Altered group
<input type="checkbox"/> BPIFB4	20q11.21	0 (0.00%)	2 (7.14%)		<-10	9.473e-3	0.767	Unaltered group
<input type="checkbox"/> BRAF	7q34	0 (0.00%)	2 (7.14%)		<-10	9.473e-3	0.767	Unaltered group

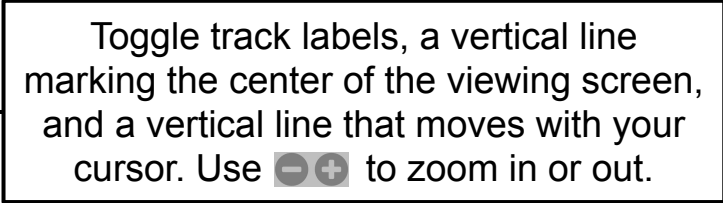
# Comparison


**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.



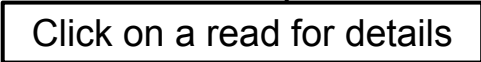
View copy number for each sample at each query gene via the [Integrated Genomics Viewer](#) (IGV).



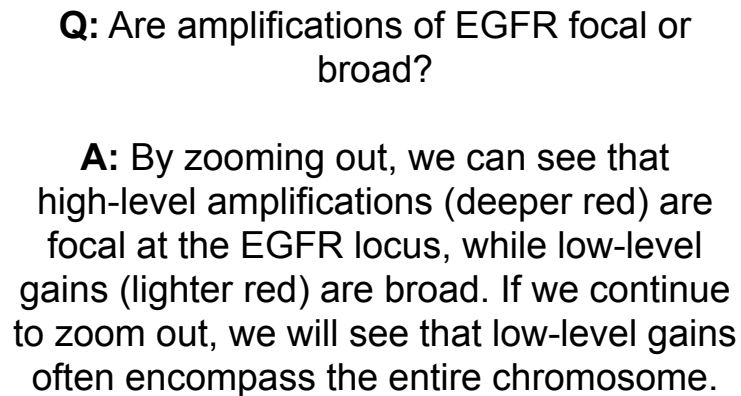
Click  for track settings, including expanding the height of each sample (see below)

Each row is a single sample

Gene  
structures



[Link to this page](#)



**A:** By zooming out, we can see that high-level amplifications (deeper red) are focal at the EGFR locus, while low-level gains (lighter red) are broad. If we continue to zoom out, we will see that low-level gains often encompass the entire chromosome.

# Network

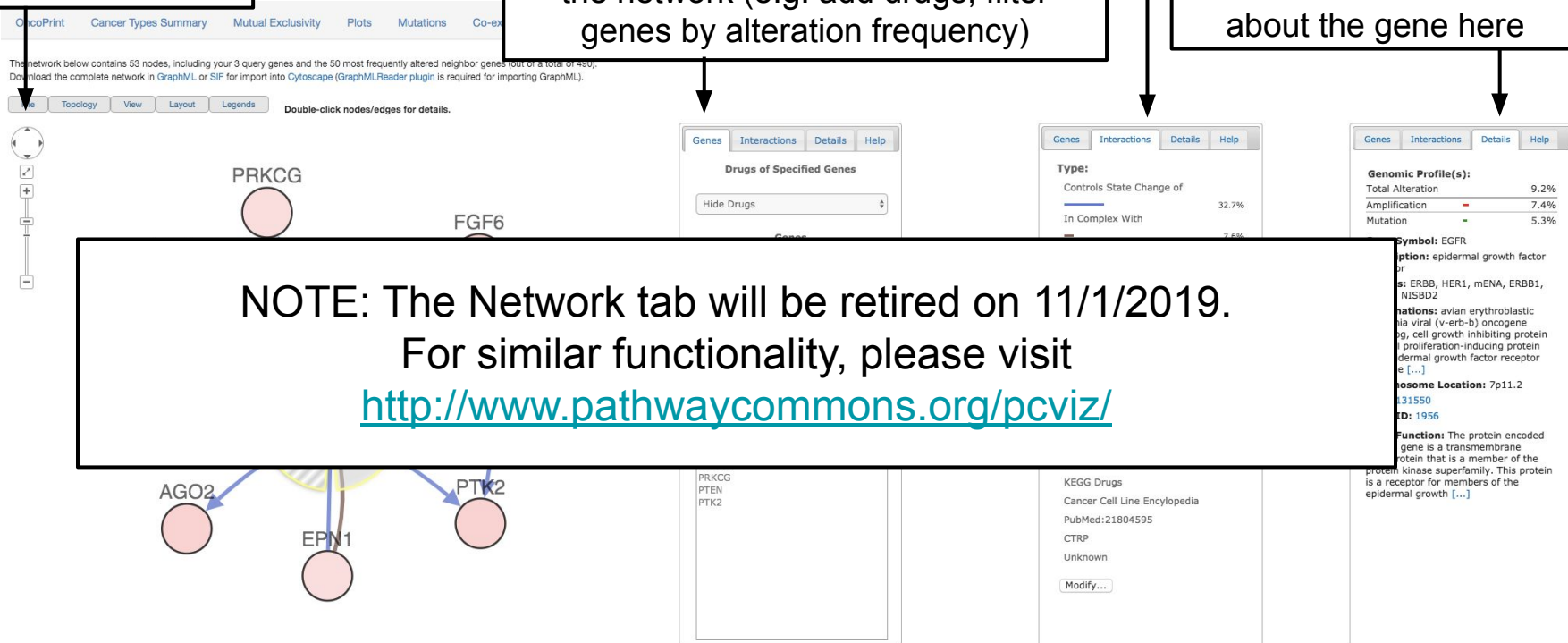
Visualize biological interaction networks centered on your query genes, with color-coding and filter options based on the frequency of genomic alterations in each gene. Click on the “Help” tab for a more detailed explanation.

Change zoom and move around network

View or modify the nodes included in the network (e.g. add drugs, filter genes by alteration frequency)

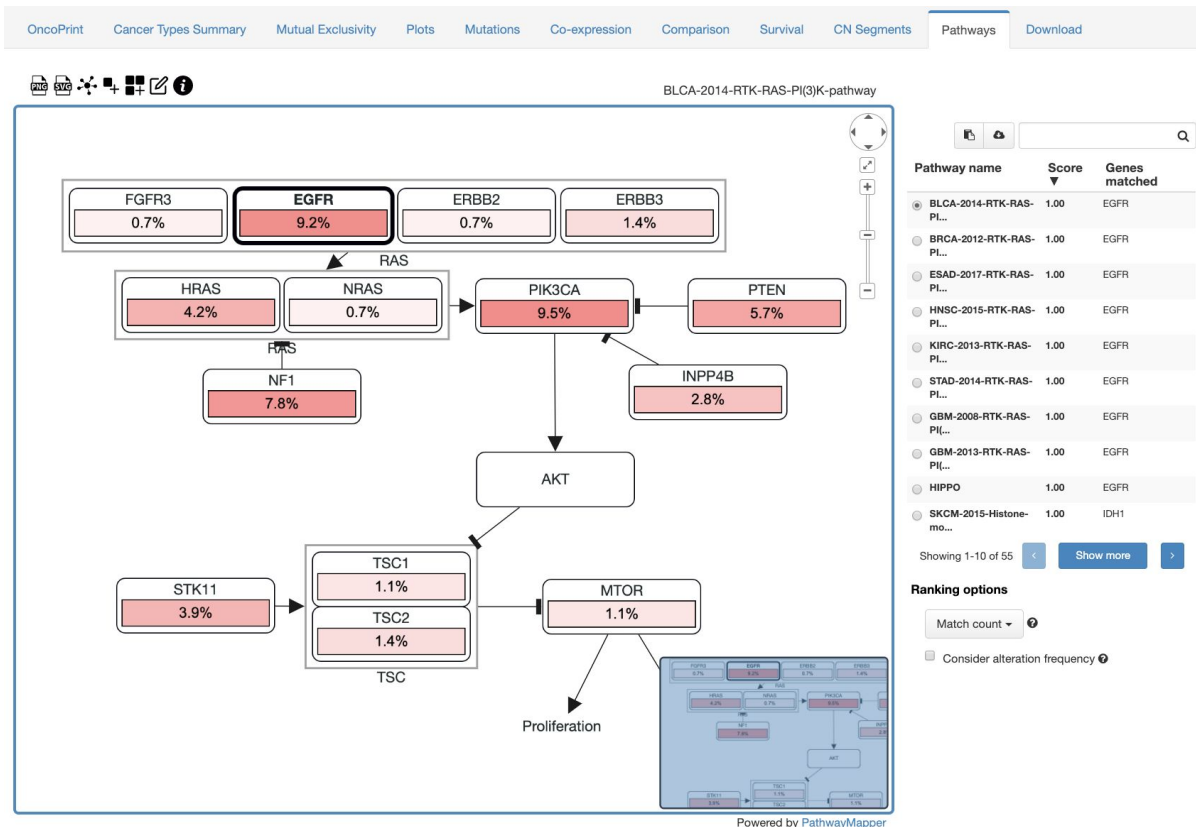
View or modify the types of interactions (edges) utilized in the plot

Click on any node to see detailed information about the gene here



# 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](#).





# 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)	<a href="#">Tab Delimited Format</a>   <a href="#">Transposed Matrix</a>
Mutations (OQL is not in effect)	<a href="#">Tab Delimited Format</a>   <a href="#">Transposed Matrix</a>
Altered samples: List of samples with alterations	<a href="#">Copy</a>   <a href="#">Download</a>   <a href="#">Query</a>   <a href="#">Virtual Study</a>
Unaltered samples: List of samples without any alteration	<a href="#">Copy</a>   <a href="#">Download</a>   <a href="#">Query</a>   <a href="#">Virtual Study</a>
Sample matrix: List of all samples where 1=altered and 0=unaltered	<a href="#">Copy</a>   <a href="#">Download</a>
Relative linear copy-number values ⓘ	<a href="#">Tab Delimited Format</a>   <a href="#">Transposed Matrix</a>
mRNA expression (microarray) ⓘ	<a href="#">Tab Delimited Format</a>   <a href="#">Transposed Matrix</a>
mRNA Expression z-Scores (microarray) ⓘ	<a href="#">Tab Delimited Format</a>   <a href="#">Transposed Matrix</a>
mRNA expression (RNA Seq V2 RSEM) ⓘ	<a href="#">Tab Delimited Format</a>   <a href="#">Transposed Matrix</a>
mRNA Expression z-Scores (RNA Seq V2 RSEM) ⓘ	<a href="#">Tab Delimited Format</a>   <a href="#">Transposed Matrix</a>
Methylation (HM450) ⓘ	<a href="#">Tab Delimited Format</a>   <a href="#">Transposed Matrix</a>
Protein expression (RPPA) ⓘ	<a href="#">Tab Delimited Format</a>   <a href="#">Transposed Matrix</a>
Protein expression Z-scores (RPPA) ⓘ	<a href="#">Tab Delimited Format</a>   <a href="#">Transposed Matrix</a>

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	219	77.4%
EGFR	26	9.2%
IDH2	13	4.6%

Showing 1-3 of 3

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	TCGA-CS-6290-01	TCGA-CS-6290	1	R132H	no alteration	no alteration
lgg_tcga	TCGA-DU-5849-01	TCGA-DU-5849	1	no alteration	no alteration	R172K
lgg_tcga	TCGA-DU-5852-01	TCGA-DU-5852	1	no alteration	AMP, R252C, I981F	no alteration
lgg_tcga	TCGA-DU-5854-01	TCGA-DU-5854	1	no alteration	AMP	no alteration
lgg_tcga	TCGA-DU-5855-01	TCGA-DU-5855	1	R132H	no alteration	no alteration
lgg_tcga	TCGA-DU-5870-01	TCGA-DU-5870	1	R132H	no alteration	no alteration
lgg_tcga	TCGA-DU-5871-01	TCGA-DU-5871	1	R132G	no alteration	no alteration
lgg_tcga	TCGA-DU-5872-01	TCGA-DU-5872	1	R132H	no alteration	no alteration

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)	<a href="#">Tab Delimited Format</a>	<a href="#">Transposed Matrix</a>
Mutations (OQL is not in effect)	<a href="#">Tab Delimited Format</a>	<a href="#">Transposed Matrix</a>
Altered samples: List of samples with alterations	<a href="#">Copy</a>	<a href="#">Download</a>   <a href="#">Query</a>   <a href="#">Virtual Study</a>
Unaltered samples: List of samples without any alteration	<a href="#">Copy</a>	<a href="#">Download</a>   <a href="#">Query</a>   <a href="#">Virtual Study</a>
Sample matrix: List of all samples where 1=altered and 0=unaltered	<a href="#">Copy</a>	<a href="#">Download</a>
Relative linear copy-number values ⓘ	<a href="#">Tab Delimited Format</a>	<a href="#">Transposed Matrix</a>
mRNA expression (microarray) ⓘ	<a href="#">Tab Delimited Format</a>	<a href="#">Transposed Matrix</a>
mRNA Expression z-Scores (microarray) ⓘ	<a href="#">Tab Delimited Format</a>	<a href="#">Transposed Matrix</a>
mRNA expression (RNA Seq V2 RSEM) ⓘ	<a href="#">Tab Delimited Format</a>	<a href="#">Transposed Matrix</a>
mRNA Expression z-Scores (RNA Seq V2 RSEM) ⓘ	<a href="#">Tab Delimited Format</a>	<a href="#">Transposed Matrix</a>
Methylation (HM450) ⓘ	<a href="#">Tab Delimited Format</a>	<a href="#">Transposed Matrix</a>
Protein expression (RPPA) ⓘ	<a href="#">Tab Delimited Format</a>	<a href="#">Transposed Matrix</a>
Protein expression Z-scores (RPPA) ⓘ	<a href="#">Tab Delimited Format</a>	<a href="#">Transposed Matrix</a>

### Gene Alteration Frequency

Gene Symbol	Num Samples Altered	Percent Samples Altered ▼
IDH1	219	77.4%
EGFR	26	9.2%
IDH2	13	4.6%

Showing 1-3 of 3

### Type of Genetic Alterations Across All Samples

Study ID	Sample ID	Patient ID	Altered ▼	IDH1	EGFR	IDH2
lgg_tcga	TCGA-CS-6290-01	TCGA-CS-6290	1	<b>R132H</b>	no alteration	no alteration
lgg_tcga	TCGA-DU-5849-01	TCGA-DU-5849	1	no alteration	no alteration	<b>R172K</b>
lgg_tcga	TCGA-DU-5852-01	TCGA-DU-5852	1	no alteration	<b>AMP, R252C, I981F</b>	no alteration
lgg_tcga	TCGA-DU-5854-01	TCGA-DU-5854	1	no alteration	<b>AMP</b>	no alteration
lgg_tcga	TCGA-DU-5855-01	TCGA-DU-5855	1	<b>R132H</b>	no alteration	no alteration
lgg_tcga	TCGA-DU-5870-01	TCGA-DU-5870	1	<b>R132H</b>	no alteration	no alteration
lgg_tcga	TCGA-DU-5871-01	TCGA-DU-5871	1	<b>R132G</b>	no alteration	no alteration
lgg_tcga	TCGA-DU-5872-01	TCGA-DU-5872	1	<b>R132H</b>	no alteration	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:

[cbioportal@googlegroups.com](mailto:cbioportal@googlegroups.com)