

# cBioPortal Tutorial #4: Virtual Studies

Create and save your own custom study comprised  
of samples from existing studies

# Tutorial Objectives

- Explain what virtual studies are and how to create them
- Delineate the different features available when a user is or is not logged in
- Show how to log in

# Virtual Study Overview

## **What is a Virtual Study?**

A virtual study is a custom study comprised of samples from one or more existing studies.

## **Why is a virtual study useful?**

The virtual study feature allows you to define a custom cohort of samples that fit your specific genomic or clinical criteria of interest. These samples can be a subset of the data available in an existing study, or result from the combination of multiple existing studies. This cohort of samples can then be queried or explored just like a traditional study, and can be returned to at a later date or shared with a collaborator.

In this tutorial, we will create a virtual study comprised of a subset of samples from two existing TCGA glioma studies.

We begin by selecting the studies of interest and viewing them in the Study Summary page.

# Creating a virtual study: Selecting studies

3. Click "View summary"

1. Use the search box to find the studies of interest

2. Select the studies of interest

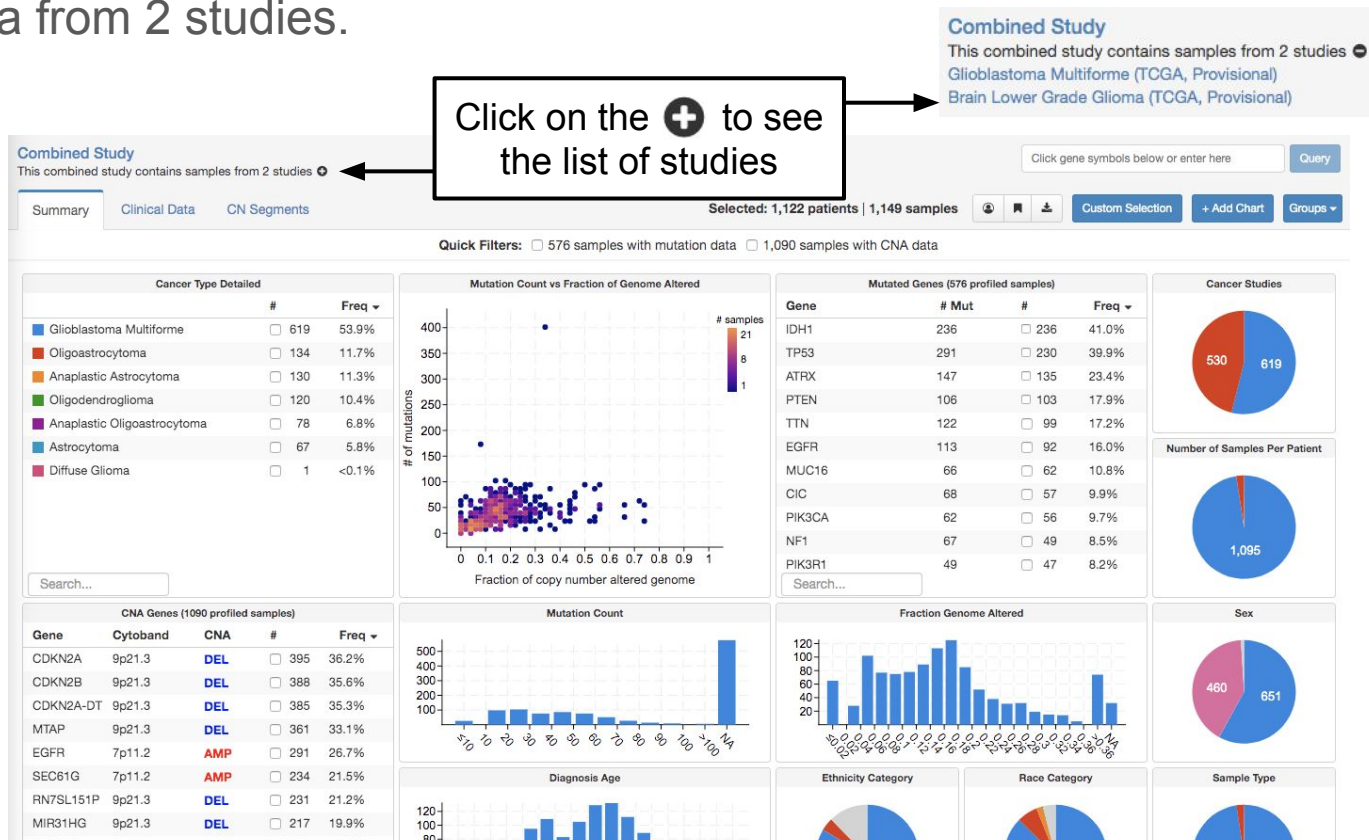
The screenshot shows the OncoPrint 'QUERY' interface. At the top, there are tabs for 'QUERY' and 'DOWNLOAD DATA'. Below these, a message says 'Open summary of selected studies in a new window.' and a 'View summary' button is highlighted. The 'Select Studies:' section shows '2 studies selected (1136 samples)' and a 'Deselect all' button. A search box contains 'glioma provisional'. The 'CNS/Brain' filter is selected, showing a list of cancer types with checkboxes and sample counts:

Cancer Type	Selection	Sample Count
Diffuse Glioma		
Brain Lower Grade Glioma (TCGA, Provisional)	<input checked="" type="checkbox"/>	532 samples
GLIOBLASTOMA		
Glioblastoma Multiforme (TCGA, Provisional)	<input checked="" type="checkbox"/>	604 samples
Miscellaneous Neuroepithelial Tumor		
Pheochromocytoma and Paraganglioma (TCGA, Provisional)	<input type="checkbox"/>	184 samples

Below the study list, there are sections for 'Select Molecular Profiles' (Mutation, Copy number alterations), 'Select Patient/Case Set' (All (1136)), and 'Enter Genes' (User-defined List, Enter HUGO Gene Symbols or Gene Aliases). A 'Submit Query' button is at the bottom, with a note: 'Please enter one or more gene symbols.'

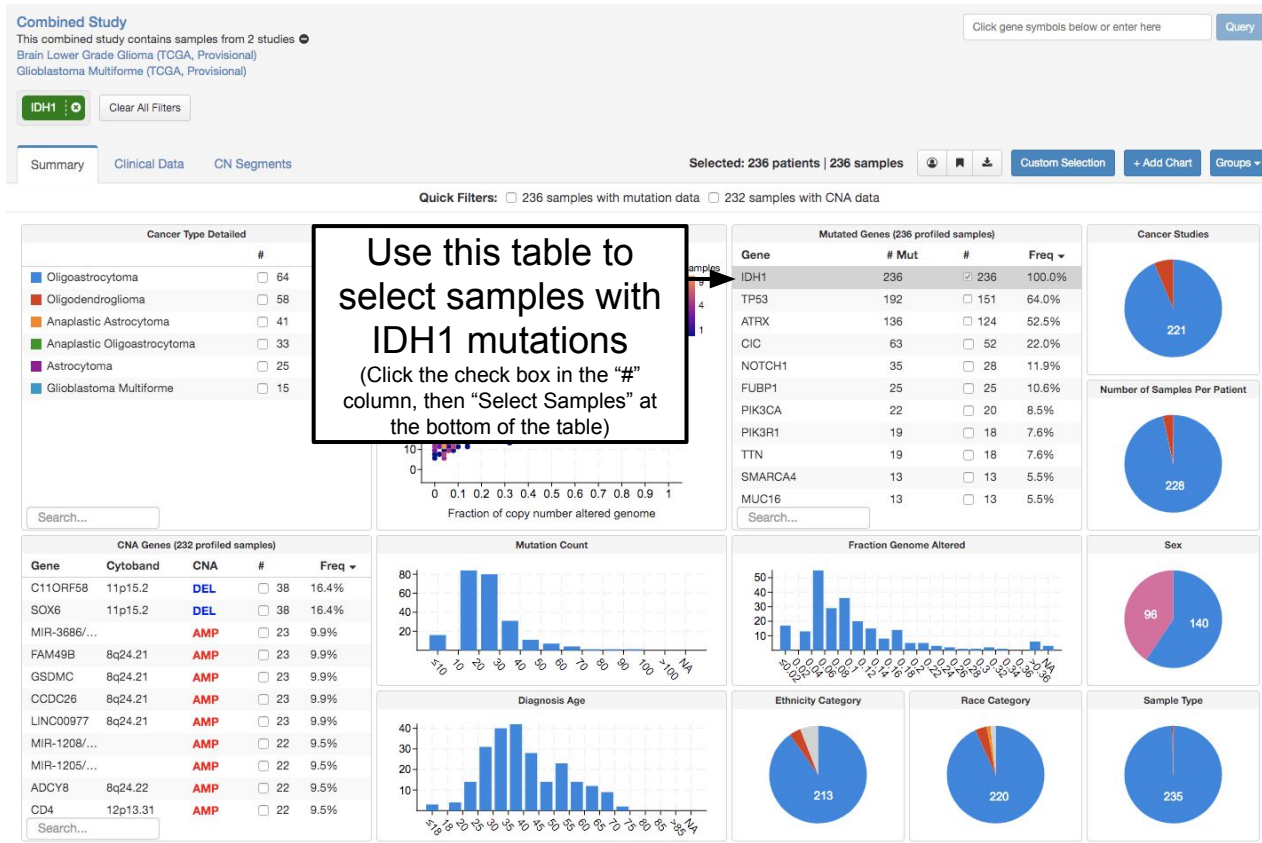
# Creating a virtual study: Study summary page

This is the same Study Summary page that you see for a single study, but now with data from 2 studies.



# Creating a virtual study: Study summary page

In this example, we will additionally filter to samples with mutations in IDH1.



We are now ready to create our virtual study.

What you see will differ slightly depending on if you are logged in or not. The next two slides show what the virtual study button looks like in each scenario.



# Virtual study button (not logged in)

**Combined Study**  
This combined study contains samples from 2 studies  
Brain Lower Grade Glioma (TCGA, Provisional)  
Glioblastoma Multiforme (TCGA, Provisional)

IDH1 ☐ Clear All Filters

Summary Clinical Data CN Segments

Quick Filters: 236 samples with mutation data 232 samples

Cancer Type Detailed

Cancer Type	#
Oligoastrocytoma	64
Oligodendroglioma	58
Anaplastic Astrocytoma	41
Anaplastic Oligoastrocytoma	36
Astrocytoma	25
Glioblastoma Multiforme	15

Search...

Click here to create and share your virtual study

Enter a name for your virtual study (optional)

Text box pre-filled with a description of the studies contributing samples and filters applied to the samples. You can edit this text.

List of studies contributing to samples with links to the study summary for each

Share Virtual Study

Selected: 236 patients | 236 samples

Custom Selection + Add Chart Groups

Selected samples (2019-06-11)

Share

236 samples from 2 studies:  
- Brain Lower Grade Glioma (TCGA, Provisional) (221 samples)  
- Glioblastoma Multiforme (TCGA, Provisional) (15 samples)

Filters:  
- Mutated Genes:  
- IDH1

This virtual study was derived from:  
- Brain Lower Grade Glioma (TCGA, Provisional)  
- Glioblastoma Multiforme (TCGA, Provisional)

Click on the link to open your virtual study, or click "Copy" to copy the URL to your clipboard.

My Virtual Study - GBM LGG IDH1 [mut](#) Share

236 samples from 2 studies:  
- Brain Lower Grade Glioma (TCGA, Provisional) (221 samples)  
- Glioblastoma Multiforme (TCGA, Provisional) (15 samples)

Filters:  
- Mutated Genes:  
- IDH1

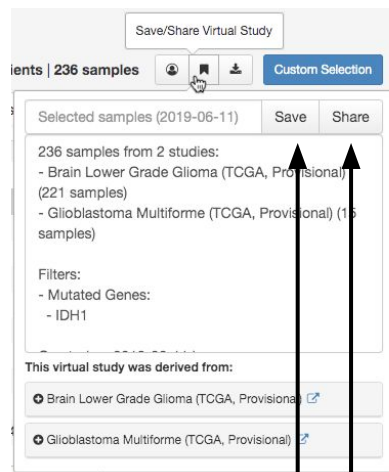
This virtual study was derived from:  
- Brain Lower Grade Glioma (TCGA, Provisional)  
- Glioblastoma Multiforme (TCGA, Provisional)

ents | 236 samples Custom Selection

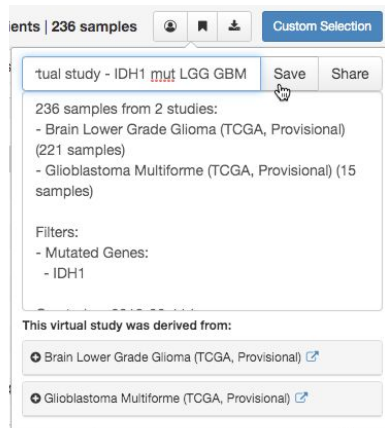
<https://www.cbioportal.org/study?id=5cfff55ae4b0ab4137874d6a> Copy View

[Link to this page](#)

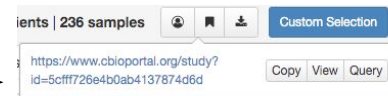
# Virtual study button (logged in)



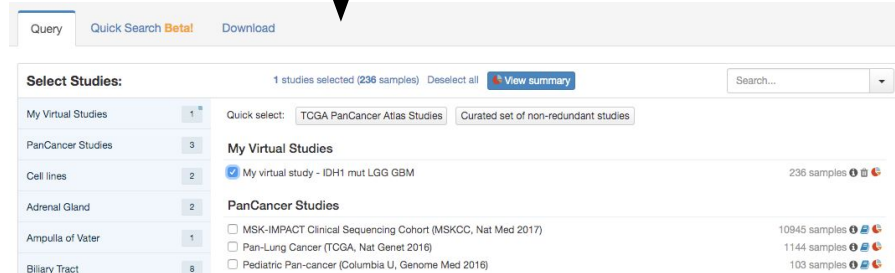
The “Save” button only appears when you are logged in.



The “Share” button works exactly the same as when not logged in



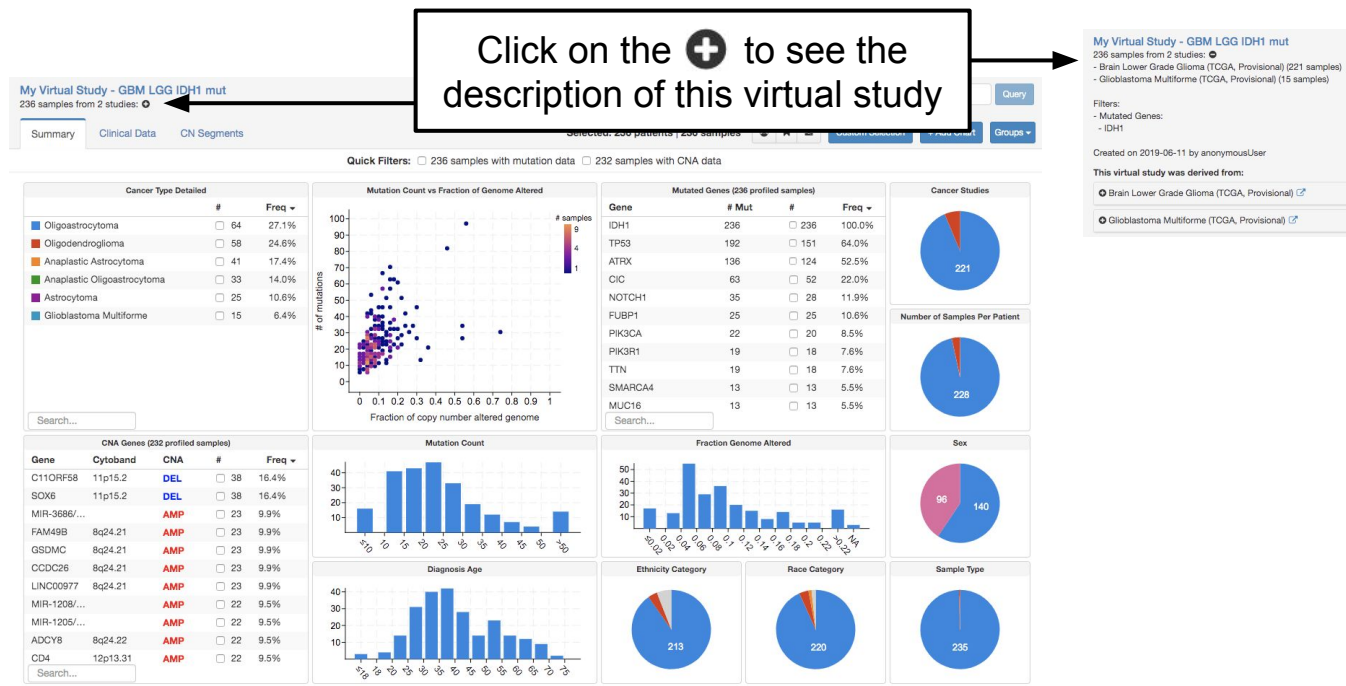
When you save a study, it is added to the homepage, at the top of the study list under “My Virtual Studies”. Clicking “Query” brings you to the query selector with your new virtual study pre-selected.



Whether or not you are logged in, let's continue on to view this new virtual study.

# Virtual study: Study summary page

We can now see the study summary page for our newly created virtual study. You can now explore these samples or run a query, just like with a traditional study. Remember that while making this study we filtered down to IDH1 mutant samples, but the non-mutant samples were still present. Now, those samples without mutations are completely gone.

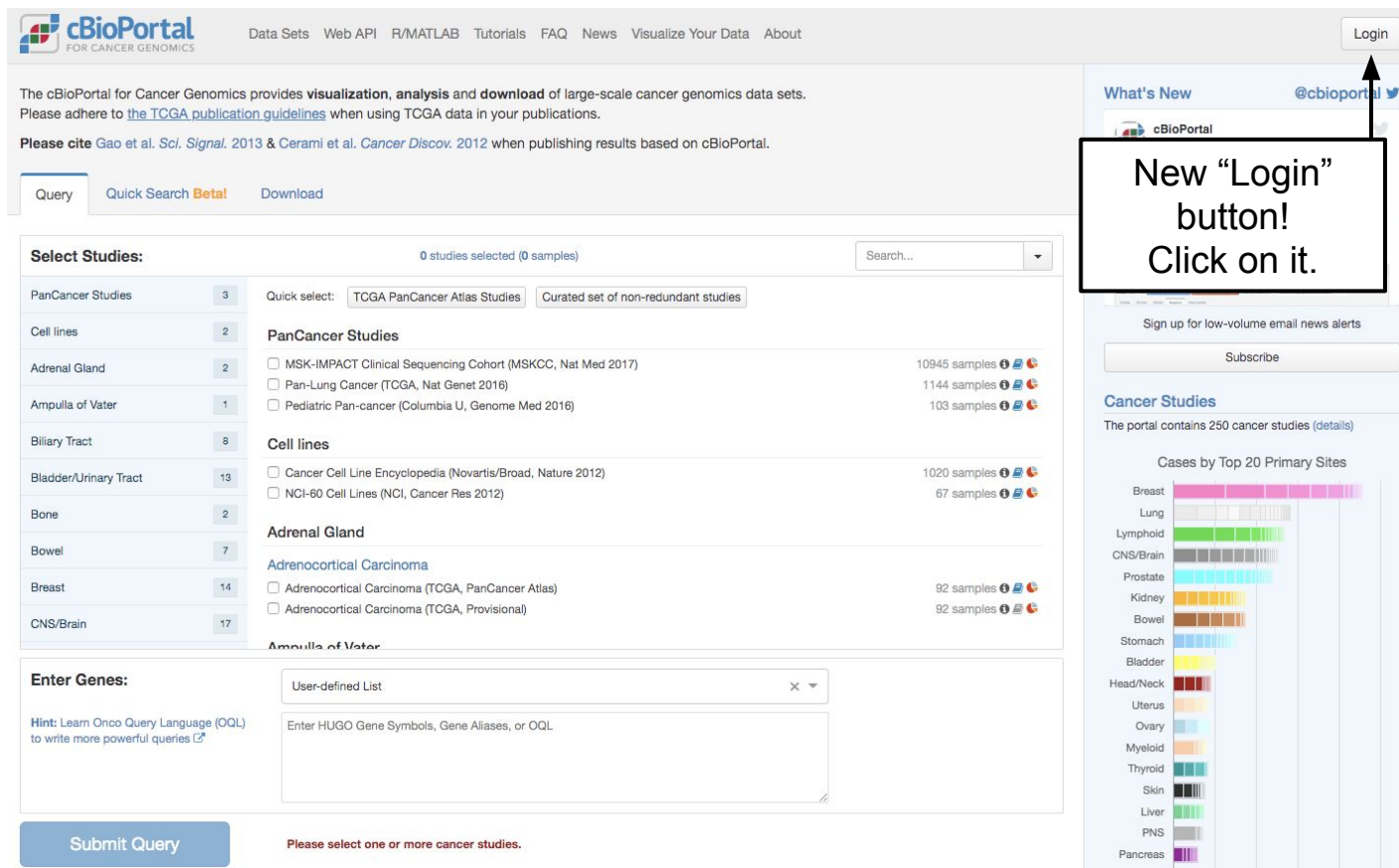


Logging in is a new feature for cBioPortal.

If you use a private instance of cBioPortal, you might already go through an authentication process (through your university, for example). If so, there is nothing else you need to do.

But if you don't normally log in, then let's go through how this works.

# Logging in



The screenshot shows the cBioPortal website interface. At the top right, there is a 'Login' button. A red box highlights this button, and a red arrow points to it from a text box that says 'New "Login" button! Click on it.' The website header includes the cBioPortal logo and navigation links: Data Sets, Web API, R/MATLAB, Tutorials, FAQ, News, Visualize Your Data, and About. Below the header, there is a message about the portal's purpose and a citation request. The main content area is divided into sections for 'Select Studies' and 'Enter Genes'. The 'Select Studies' section includes a list of cancer types with sample counts and a 'Quick select' dropdown. The 'Enter Genes' section has a text input field and a 'Submit Query' button. On the right side, there is a 'What's New' section with a cBioPortal logo and a 'Subscribe' button. Below that is a 'Cancer Studies' section with a bar chart showing cases by top 20 primary sites.

**Header:** cBioPortal FOR CANCER GENOMICS | Data Sets | Web API | R/MATLAB | Tutorials | FAQ | News | Visualize Your Data | About | **Login**

The cBioPortal for Cancer Genomics provides **visualization, analysis** and **download** of large-scale cancer genomics data sets. Please adhere to [the TCGA publication guidelines](#) when using TCGA data in your publications.

Please cite [Gao et al. \*Sci. Signal.\* 2013](#) & [Cerami et al. \*Cancer Discov.\* 2012](#) when publishing results based on cBioPortal.

**Navigation:** Query | Quick Search **Beta!** | Download

**Select Studies:** 0 studies selected (0 samples) Search...

**PanCancer Studies** 3

**Cell lines** 2

**Adrenal Gland** 2

**Ampulla of Vater** 1

**Biliary Tract** 8

**Bladder/Urinary Tract** 13

**Bone** 2

**Bowel** 7

**Breast** 14

**CNS/Brain** 17

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

**PanCancer Studies**

- ☐ MSK-IMPACT Clinical Sequencing Cohort (MSKCC, Nat Med 2017) 10945 samples
- ☐ Pan-Lung Cancer (TCGA, Nat Genet 2016) 1144 samples
- ☐ Pediatric Pan-cancer (Columbia U, Genome Med 2016) 103 samples

**Cell lines**

- ☐ Cancer Cell Line Encyclopedia (Novartis/Broad, Nature 2012) 1020 samples
- ☐ NCI-60 Cell Lines (NCI, Cancer Res 2012) 67 samples

**Adrenal Gland**

**Adrenocortical Carcinoma**

- ☐ Adrenocortical Carcinoma (TCGA, PanCancer Atlas) 92 samples
- ☐ Adrenocortical Carcinoma (TCGA, Provisional) 92 samples

**Ampulla of Vater**

**Enter Genes:** User-defined List x

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

**Submit Query**

Please select one or more cancer studies.

**What's New** @cbioportal

Sign up for low-volume email news alerts

**Subscribe**

**Cancer Studies**

The portal contains 250 cancer studies (details)

**Cases by Top 20 Primary Sites**

Primary Site	Cases
Breast	10945
Lung	1144
Lymphoid	103
CNS/Brain	1020
Prostate	67
Kidney	92
Bowel	92
Stomach	
Bladder	
Head/Neck	
Uterus	
Ovary	
Myeloid	
Thyroid	
Skin	
Liver	
PNS	
Pancreas	

[Link to this page](#)

# Logging in

We use Google for authentication. Sign in with any google account.



Welcome to cBioPortal - sign in with your Google account to store your [virtual studies](#) and [groups](#). This will allow you to access your studies and groups from any computer. Login is optional and not required to access any of the other features of cBioPortal.



# Logging in

The screenshot shows the cBioPortal website interface. At the top, the cBioPortal logo is on the left, and navigation links (Data Sets, Web API, R/MATLAB, Tutorials, FAQ, News, Visualize Your Data, About) are in the center. On the right, it says "Logged in as your email address" with a "Sign out" link. Below the header, a message states: "The cBioPortal for Cancer Genomics provides **visualization, analysis and download** of large-scale cancer genomics data sets. Please adhere to [the TCGA publication guidelines](#) when using TCGA data in your publications. Please cite Gao et al. *Sci. Signal.* 2013 & Cerami et al. *Cancer Discov.* 2012 when publishing results based on cBioPortal."

Below the message are three tabs: "Query", "Quick Search **Beta!**", and "Download". The "Query" tab is active. The main content area is titled "Select Studies:" and shows "0 studies selected (0 samples)". It has a search bar and a "Quick select:" dropdown with options "TCGA PanCancer Atlas Studies" and "Curated set of non-redundant studies".

The "My Virtual Studies" section shows a list of virtual studies. One study is selected: "My virtual study - IDH1 mut LGG GBM" with 236 samples. Below this, the "PanCancer Studies" section lists several studies with sample counts: "MSK-IMPACT Clinical Sequencing Cohort (MSKCC, Nat Med 2017)" (10945 samples), "Pan-Lung Cancer (TCGA, Nat Genet 2016)" (1144 samples), and "Pediatric Pan-cancer (Columbia U, Genome Med 2016)" (103 samples). The "Cell lines" section lists "Cancer Cell Line Encyclopedia (Novartis/Broad, Nature 2012)" (1020 samples) and "NCI-60 Cell Lines (NCI, Cancer Res 2012)" (67 samples). The "Adrenal Gland" section is also visible.

At the bottom, the "Enter Genes:" section has a "User-defined List" dropdown and a text input field for "Enter HUGO Gene Symbols, Gene Aliases, or OQL". A "Submit Query" button is at the bottom left. A red message at the bottom center says "Please select one or more cancer studies."

On the right side, there is a "What's New" section with a "cBioPortal" link and a "Sign out" link. Below this is a "Cases by Top 20 Primary Sites" bar chart showing the number of cases for various cancer types: Breast, Lung, Lymphoid, CNS/Brain, Prostate, Kidney, Bowel, Stomach, Bladder, Head/Neck, Uterus, Ovary, Myeloid, Thyroid, Skin, Liver, PNS, Pancreas, and Glioma.

Two callout boxes are present. The first box, with an arrow pointing to the "Sign out" link, contains the text: "This is what you see when you log in". The second box, with an arrow pointing to the "My virtual study - IDH1 mut LGG GBM" entry, contains the text: "Here's the virtual study I made".



# Virtual Study Authentication Questions

## **Is it necessary to log in to use virtual studies?**

No. A user that has not logged in can create virtual studies and run queries in those studies (by using the query box on the study summary page). Links to virtual studies are permanent, so you can save the link on your computer and come back to it anytime, or share it with others.

## **If I do log in, what additional functionality do I gain?**

If you log in, you gain the ability to save your virtual study to the list of existing studies on the homepage. This makes a virtual study functionally the same as any other study: you can access your virtual studies in the query builder and you can combine an existing virtual study with any other study to create a new virtual study.

Questions?

Check out our other tutorials  
or email us at:

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