

Tutorial: CGC

i) What is CGC?

The Seven Bridges Cancer Genomics Cloud (CGC; www.cancergenomicscloud.org) is scalable data access and analysis platform that connects researchers to a wide range of public datasets, for example, The Cancer Genomics Atlas (TCGA), Therapeutically Applicable Research to Generate Effective Treatments (TARGET). The CGC enables collaborative, reproducible analysis across both public and private teams. To reduce the unnecessary burden of installing bioinformatics software, the CGC provides a public toolkit and bioinformatics tools and workflows.

ii) Using CGC to answer an oncology question

Background: By identifying the key information points (gene mutation) at the beginning of tumor development, the diagnosis can be made at its early stage, and the effectiveness of treatment is relatively high. Barrett's esophagus is an important precancerous lesion of esophageal carcinoma, but not all patients will develop esophageal cancer. The discovery of specific gene changes of precancerous esophageal lesion helps to identify the high-risk population that will develop into esophageal cancer.

In this tutorial, a Gene Mutation Search will be performed to answer the following question:

What are the top mutated genes in the male in esophageal cancer and its case information?

Step-by-step instructions:

1. Visit the CGC using browser and log in to query the public datasets.
2. Once logged in, access the **Case Browser** from the top navigation bar on the landing page. Click on the "Data" tab, and then click **Case Browser**. (Figure 1)

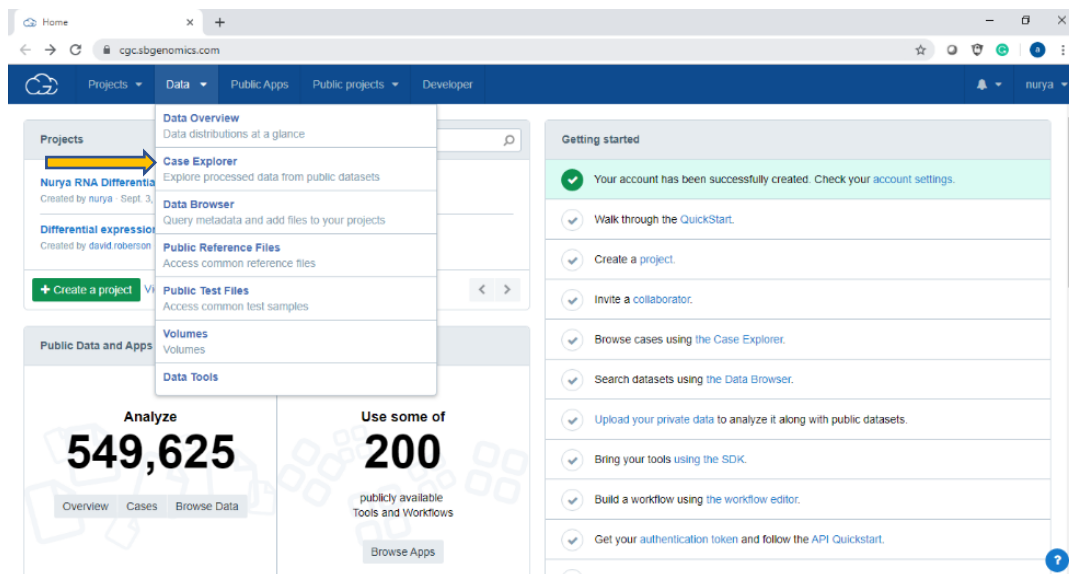


Figure 1 access the Case Browser

3. In order to query for the top mutated genes in the male in esophageal carcinoma. For this

4. To find esophageal carcinoma (ESCA) cases detail information. Scroll down the **Case Browser** page, under the bottom of this page, there is a table (Figure 3) listed the basic disease information, ->click on one of the cases ->and click on “Continue To Data Browser” ->**Navigate to a details page**(Figure 4), [which includes the Simple, Treatment, Diagnosis, Exposure et al.](#) (answer to the proposed question).



Figure 3 Esophageal Carcinoma basic information

CGC follows the FAIR (Findable, Accessible, Interoperable, and Reusable) principles. Researchers will be able to perform queries and computational analyses across datasets. Suggestions to improve CGC include: Besides projects, data, public projects, and developer, CGC can include links to other external databases. For example, the latest research progress of diseases, clinical and animal experimental reports, and published sources.