

Exercise 1A: Explore TCGA Ovarian Cancer Data Set in GDC

1. Go to Genomic Data Commons (GDC) at <https://gdc-portal.nci.nih.gov/>.

Can't find your data? Click here for more information.

NIH NATIONAL CANCER INSTITUTE GDC Data Portal

Home Projects Data Analysis

Quick Search Login Cart GDC Apps

Harmonized Cancer Datasets

Genomic Data Commons Data Portal

Get Started by Exploring:

Projects Data

Perform Advanced Search Queries, such as:

Cases of kidney cancer diagnosed at the age of 20 and below182 Cases1,514 Files

ChV data of female brain cancer cases459 Cases1,788 Files

Gene expression quantification data in TCGA-GBM project166 Cases522 Files

Cases by Primary Site

DATA PORTAL SUMMARY

Data Release 4.0 - October 31, 2016

PROJECTS39

PRIMARY SITE29

CASES14,531

FILES274,821

Infrastructure

Data is continuously being processed and harmonized by the GDC.

View GDC system statistics:

Compute Infrastructure12,800 Cores87.96 TB RAM

Storage Infrastructure4.98 PB Used5.42 PB Total

View Data Download Statistics Report >

Documentation

Learn how to use the GDC Data Portal to its full potential with common topics such as:

Browse Data using Facet Search

Search Data with Advanced Search Technology

Project Based Data Availability

Controlled Access Data

Visit the Documentation Website >

GDC Applications

The GDC Data Portal is a robust data-driven platform that allows cancer researchers and bioinformaticians to search and download cancer data for analysis. The GDC applications include:

Data PortalWebsiteData Transfer ToolAPI

Data Submission PortalDocumentationLegacy ArchiveGDC cBio Portal

Site Home | Policies | Accessibility | FOIA

U.S. Department of Health and Human Services | National Institutes of Health | National Cancer Institute | USA.gov

NIH... Turning Discovery into Health ®

UI 1.4.1 © 1416e89, API 1.5.0 © 9f568ba

2. Click on "Projects." This shows a summary of all datasets.

Can't find your data? Click here for more information.

NIH NATIONAL CANCER INSTITUTE GDC Data Portal

Home Projects Data Analysis

Quick Search Login Cart GDC Apps

Start searching by selecting a facet or try the Data Search

Table Graph

Projects (39)

ID	Disease Type	Primary Site	Program	Cases	Available Cases per Data Category						Files
					Seq	Exp	SNV	CNV	Clinical	Bio	
TARGET-NBL	Neuroblastoma	Nervous System	TARGET	1,120	270	151	216	0	1,120	0	2,803
TCGA-BRCA	Breast Invasive Carcin...	Breast	TCGA	1,099	1,099	1,097	1,044	1,099	1,097	1,099	27,207
TARGET-AML	Acute Myeloid Leukemia	Blood	TARGET	923	299	272	8	0	447	0	1,870
TARGET-WT	High-Risk Wilms Tumor	Kidney	TARGET	663	128	128	34	0	128	0	1,321
TCGA-GBM	Glioblastoma Multiforme	Brain	TCGA	617	406	166	396	593	596	617	9,657
TCGA-OV	Ovarian Serous Cysta...	Ovary	TCGA	608	575	492	443	573	587	608	13,054
TCGA-LUAD	Lung Adenocarcinoma	Lung	TCGA	585	582	519	569	518	522	585	14,804
TCGA-UCEC	Uterine Corpus Endo...	Uterus	TCGA	560	559	559	542	547	548	560	13,604
TCGA-KIRC	Kidney Renal Clear Ce...	Kidney	TCGA	537	535	534	339	532	537	537	12,272
TCGA-HNSC	Head and Neck Squa...	Head and Neck	TCGA	528	528	528	510	521	528	528	12,895
TCGA-LGG	Brain Lower Grade Gli...	Brain	TCGA	516	516	516	513	514	515	516	12,603
TCGA-THCA	Thyroid Carcinoma	Thyroid	TCGA	507	507	507	496	505	507	507	12,703
TCGA-LUSC	Lung Squamous Cell ...	Lung	TCGA	504	504	504	497	504	504	504	13,124
TCGA-PRAD	Prostate Adenocarcino...	Prostate	TCGA	500	498	498	498	498	500	500	12,568
TCGA-STAD	Stomach Adenocarcino...	Stomach	TCGA	478	443	439	441	443	443	478	10,835
TCGA-SKCM	Skin Cutaneous Melan...	Skin	TCGA	470	470	469	470	470	470	470	11,265
TCGA-COAD	Colon Adenocarcinoma	Colorectal	TCGA	463	460	459	433	458	459	463	11,827
TCGA-BLCA	Bladder Urothelial Car...	Bladder	TCGA	412	412	412	412	412	412	412	10,193
TARGET-OS	Osteosarcoma	Bone	TARGET	384	0	0	0	0	73	0	3
TCGA-LIHC	Liver Hepatocellular C...	Liver	TCGA	377	377	376	375	376	377	377	9,511

3. Click on "TCGA-OV." This shows the summary of the TCGA-OV (TCGA Ovarian Cancer) project.

Can't find your data? Click here for more information.

NIH NATIONAL CANCER INSTITUTE GDC Data Portal

Home Projects Data Analysis

Quick Search Login Cart GDC Apps

TCGA-OV

Download ManifestDownload ClinicalDownload Biospecimen

Summary

Project IDTCGA-OV

Disease NameOvarian Serous Cystadenocarcinoma

Disease TypeOvarian Serous Cystadenocarcinoma

Primary SiteOvary

ProgramTCGA

CASES608

FILES13,054

ANNOTATIONS361

Case and File Counts by Experimental Strategy

Experimental Strategy	Cases	Files
Genotyping Array	573	2,292
Methylation Array	602	623
WXS	460	5,931
RNA-Seq	376	1,516
miRNA-Seq	489	1,497

Case and File Counts by Data Category

Data Category	Cases	Files
Raw Sequencing Data	575	1,929
Transcriptome Profiling	492	2,135
Simple Nucleotide Variation	443	4,880
Copy Number Variation	573	2,292
Clinical	587	587
Biospecimen	608	608