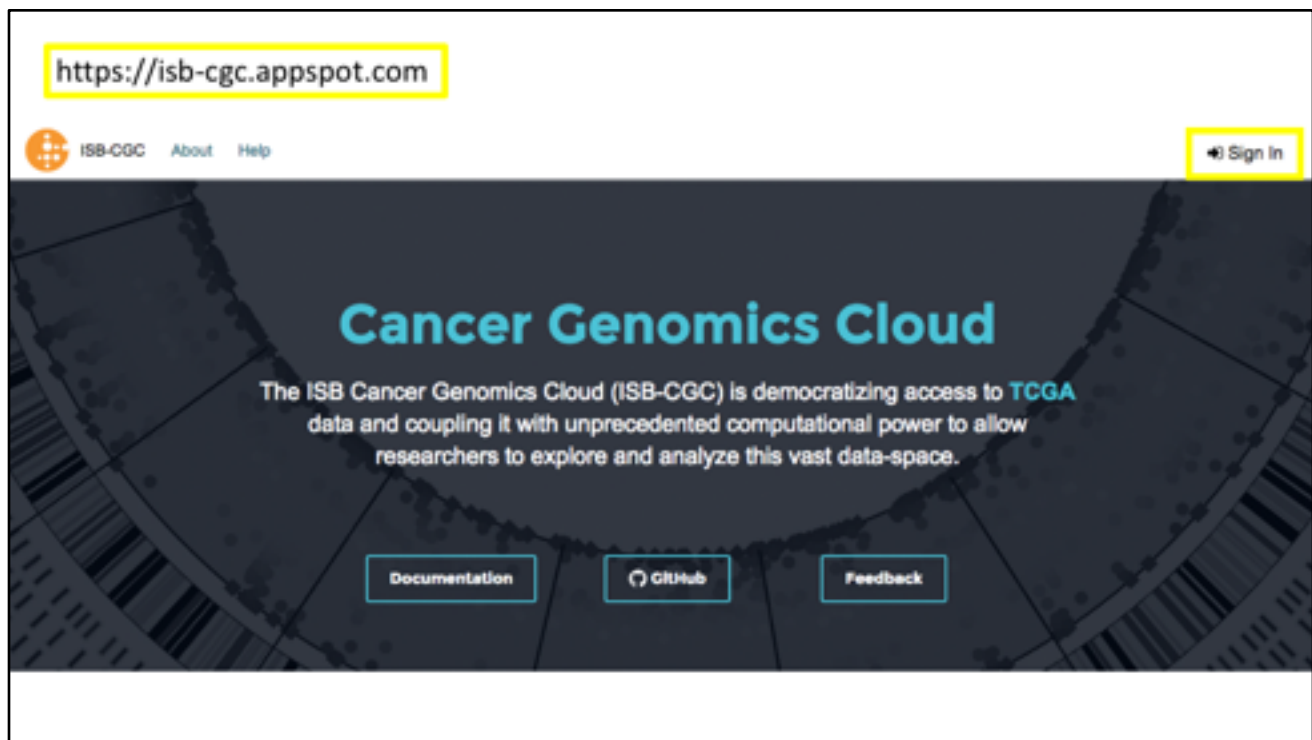


# An Introduction to the ISB-CGC Web App IGV

brought to you by

## The ISB Cancer Genomics Cloud





- This is our main landing page.
- There are a few links here that you can use to get to documentation, code, and send us feedback.
- You may only log in using a Google managed identity by clicking the Sign In button.

Your Dashboard > Saved Cohorts >

## CCLE

[New Workbook](#) | [Edit](#) | [Comments](#) | [Duplicate](#) | [Delete](#) | [View Files](#) | [Download IDs](#) | [Share](#) | Shared With (0)

### Selected Filters

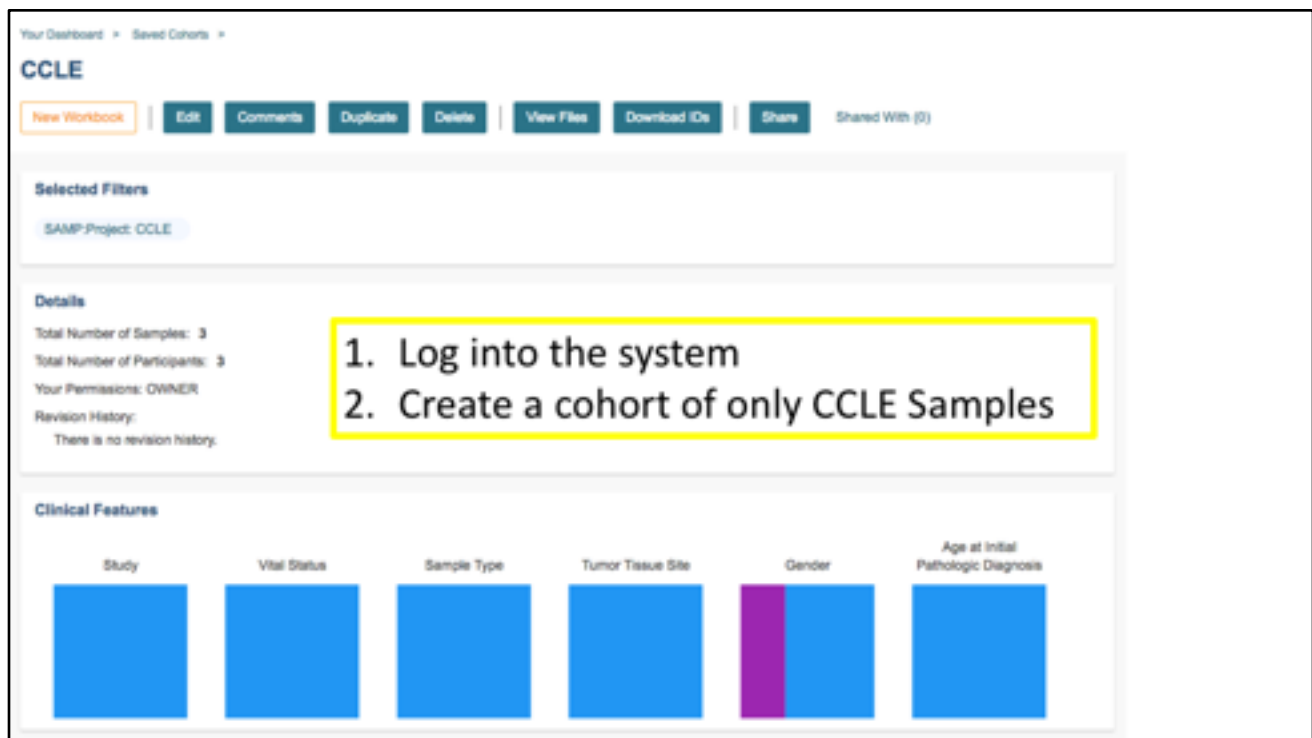
SAMP-Project: CCLE

### Details

Total Number of Samples: 3  
Total Number of Participants: 3  
Your Permissions: OWNER  
Revision History:  
There is no revision history.

### Clinical Features

Study	Vital Status	Sample Type	Tumor Tissue Site	Gender	Age at Initial Pathologic Diagnosis



- Log into the system and create a cohort of CCLE Data.
- Note that the screenshots in this lesson is not indicative of real data and uses a smaller sample set of fake data.

Your Dashboard > Saved Cohorts >

## CCLE

[New Workbook](#) | [Edit](#) [Comments](#) [Duplicate](#) [Delete](#) [View Files](#) [Download IDs](#) | [Share](#) Shared With (0)

### Selected Filters

SAMP-Project: CCLE

### Details

Total Number of Samples: 3  
Total Number of Participants: 3  
Your Permissions: OWNER  
Revision History:  
There is no revision history.

### Clinical Features

Study	Vital Status	Sample Type	Tumor Tissue Site	Gender	Age at Initial Pathologic Diagnosis

### 3. Click View Files Button

## 4. File Listing Page

< CCLE

Download File List as CSV

### Platforms

- ☐ IlluminaHiSeq DNaseq (1)
- ☐ SNP6 (1)
- ☐ HiSeq/BCGSC (2)

### File Listing

Showing 20 files of 4, Page: 1

Previous Page

Next Page

Launch IGV

Sample Barcode	Pipeline	Platform	Data Level	Data Type	IGV
CCLE-LUDLU-1-DN A-08	broad.mit.edu__DNA Seq	IlluminaHiSeq DNaseq	Level_1	DNA Sequence-Alignment	<input checked="" type="checkbox"/> GA4GH
CCLE-LUDLU-1-DN A-08	broad.mit.edu__snp_ cnv	SNP6	Level_1	Copy Number Results-SNP	
CCLE-NCI-H2087-R NA-08	broad.mit.edu__RNA Seq	HiSeq/BCGSC	Level_1	RNA Sequence-Alignment	<input checked="" type="checkbox"/> GA4GH
CCLE-NCI-H727-RN A-08	broad.mit.edu__RNA Seq	HiSeq/BCGSC	Level_1	RNA Sequence-Alignment	<input type="checkbox"/> GA4GH

- This is the file listing page.
- You can see what kind of files are available for the cohort you are looking at based on sample barcode.
- You can also download the file list as a CSV. Note: that if the list of files is too large, you will get an error. Try making the cohort smaller to trim down the file list size.
- If the file is available in Google Genomics, then the IGV column will be populated with a GA4GH checkbox. Selecting one or more will activate the Launch IGV button.

## 4. File Listing Page

← CCLE



### Platforms

- ☐ IlluminaHiSeq Dnaseq (1)
- ☐ SNP6 (1)
- ☐ HiSeq/BCGSC (2)

### File Listing

Showing 20 files of 4. Page: 1

[Previous Page](#)

[Next Page](#)

[Launch IGV](#)

Sample Barcode	Pipeline	Platform	Data Level	Data Type	IGV
CCLE-LUDLU-1-DN A-08	broad.mit.edu__DNA Seq	IlluminaHiSeq DNAS eq	Level_1	DNA Sequence-Align ment	<input checked="" type="checkbox"/> GA4GH
CCLE-LUDLU-1-DN A-08	broad.mit.edu__snp_ cnv	SNP6	Level_1	Copy Number Result s-SNP	
CCLE-NCI-H2087-R NA-08	broad.mit.edu__RNA Seq	HiSeq/BCGSC	Level_1	RNA Sequence-Align ment	<input checked="" type="checkbox"/> GA4GH
CCLE-NCI-H727-RN A-08	broad.mit.edu__RNA Seq	HiSeq/BCGSC	Level_1	RNA Sequence-Align ment	<input type="checkbox"/> GA4GH

## 5. IGV Browser



- You may be prompted to log in again to authorize access of the information required to display. Please be sure to allow pop-ups on this page for that, otherwise you may receive an access denied error.
- Here, we have the two samples that were previously selected displayed and the IGV browser is accessing the reads from Google Genomics using the GA4GH API.

Your Dashboard > Saved Cohorts >

## All TCGA Data

Owner: [New Workbook](#) | [Comments](#) | [Duplicate](#) | [Remove](#) | [View Files](#) | [Download IDs](#) | Shared With (12)

**Selected Filters**  
There were no filters used to create this cohort.

**Details**  
Total Number of Samples: 1378  
Total Number of Participants: 594  
Your Permissions: READER  
Revision History:  
There is no revision history.

### 6. All TCGA Cohort

**Clinical Features**

The 'Clinical Features' section displays six trellis plots representing different clinical variables for the cohort. Each plot is a square divided into colored rectangles, where the color represents a category and the size of the rectangle represents the frequency of that category. The features shown are: Study (a complex multi-colored plot), Vital Status (purple and red), Sample Type (purple, pink, yellow, and red), Tumor Tissue Site (a complex multi-colored plot), Gender (purple and blue), and Age at Initial Pathologic Diagnosis (purple, green, red, yellow, and blue).

- We are now going to look at another example of the IGV Viewer using the All TCGA Cohort
- Again, we click the View Files button.



## 7. All TCGA Cohort File Listing

### All TCGA Data

Platforms

- ☐ Mixed DnaSeq Cont (4)
- ☐ GA (406)
- ☐ 2% (210)
- ☐ GA/BCSC (206)
- ☐ GA/UNC V2 (200)
- ☒ IlluminaHiSeq TotalRNAseqV2 (28)
- ☐ HiSeq/BCSC (348)
- ☐ IlluminaHiSeq None (177)
- ☐ Illumina DnaSeq Curated (206)
- ☐ IlluminaHiSeq DnaSeq Cont Automated (86)
- ☐ IlluminaHiSeq DnaSeq (1016)
- ☐ HiSeq (1096)
- ☐ HiSeq/UNC V2 (1106)
- ☐ SNP (3818)
- ☐ RPPA (381)
- ☐ IlluminaHiSeq None (122)
- ☐ Illumina DnaSeq Automated (2470)
- ☐ Illumina DnaSeq Cont Automated (432)
- ☐ 450k (548)
- ☐ Mixed DnaSeq Curated (53)
- ☐ IlluminaHiSeq DnaSeq Automated (793)

### File Listing

Showing 20 files of 28. Page: 1

Previous Page Next Page

Sample Barcode	Pipeline	Platform	Data Level	Data Type	IGV
TCGA-BL-A13J-01A	unc.edu__TotalRNASeqV2	IlluminaHiSeq TotalRNASeqV2	Level 3	RSEM_isotforms_normalized	
TCGA-BL-A13J-01A	unc.edu__TotalRNASeqV2	IlluminaHiSeq TotalRNASeqV2	Level 3	RSEM_genes_normalized	
TCGA-BL-A13J-01A	unc.edu__TotalRNASeqV2	IlluminaHiSeq TotalRNASeqV2	Level 3	RSEM_isotforms	
TCGA-BL-A13J-01B	unc.edu__TotalRNASeqV2	IlluminaHiSeq TotalRNASeqV2	Level 1	RNA Sequence-Alignment	
TCGA-BL-A13J-01A	unc.edu__TotalRNASeqV2	IlluminaHiSeq TotalRNASeqV2	Level 1	RNA Sequence-Alignment	
TCGA-BL-A13J-01A	unc.edu__TotalRNASeqV2	IlluminaHiSeq TotalRNASeqV2	Level 3	RSEM_genes	
TCGA-BL-A13J-01A	unc.edu__TotalRNASeqV2	IlluminaHiSeq TotalRNASeqV2	Level 3	junction_quantification	
TCGA-BL-A13J-01A	unc.edu__TotalRNASeqV2	IlluminaHiSeq TotalRNASeqV2	Level 3	exon_quantification	
TCGA-A8-267T-01A	unc.edu__TotalRNASeqV2	IlluminaHiSeq TotalRNASeqV2	Level 1	RNA Sequence-Alignment	
TCGA-A8-267T-01A	unc.edu__TotalRNASeqV2	IlluminaHiSeq TotalRNASeqV2	Level 3	RSEM_genes_normalized	
TCGA-A8-267T-01A	unc.edu__TotalRNASeqV2	IlluminaHiSeq TotalRNASeqV2	Level 3	RSEM_isotforms_normalized	
TCGA-A8-267T-01A	unc.edu__TotalRNASeqV2	IlluminaHiSeq TotalRNASeqV2	Level 3	exon_quantification	
TCGA-A8-267T-01B	unc.edu__TotalRNASeqV2	IlluminaHiSeq TotalRNASeqV2	Level 1	RNA Sequence-Alignment	

- Selected are all files that were produced using the Illumina HiSeq Total RNA Seq v2 Platform.
- Note that the entire IGV column is empty right now. This is because the user is not currently authorized to view Level 1 .bam files.
- To gain access, follow the instructions in this documentation:
  - Short version: <http://bit.ly/1TJErfN>
  - Long version: <http://isb-cancer-genomics-cloud.readthedocs.io/en/latest/sections/webapp/Gaining-Access-To-TCGA-Contolled-Access-Data.html>

## 8. Associating your Google Identity with a valid NIH login.

- Documentation can be found here:
  - Short Version: <http://bit.ly/1TJErfN>
  - Long Version: <http://isb-cancer-genomics-cloud.readthedocs.io/en/latest/sections/webapp/Gaining-Access-To-TCGA-Contolled-Access-Data.html>
- Note: To access TCGA controlled-access data, you must also be authorized by dbGaP.

## 9. All TCGA Cohort File Listing – After NIH authentication

← All TCGA Data

**Platforms**

- ☐ Mixed OncoSeq Cont (9)
- ☐ GA (405)
- ☐ 27k (210)
- ☐ GABRGSC (226)
- ☐ GABRG V2 (263)
- ☒ IlluminaHiSeq TotalRNAseqV2 (28)
- ☐ HBBGBCSC (348)
- ☐ IlluminaHiSeq None (177)
- ☐ Illumina OncoSeq Curated (206)
- ☐ IlluminaHiSeq OncoSeq Cont Automated (98)
- ☐ IlluminaHiSeq OncoSeq (1016)
- ☐ HBBG V2 (4108)
- ☐ HBBG V2 (4108)
- ☐ BNG (3818)
- ☐ RPPA (391)
- ☐ IlluminaHiSeq None (122)
- ☐ Illumina OncoSeq Automated (2470)
- ☐ Illumina OncoSeq Cont Automated (452)
- ☐ ABB (348)
- ☐ Mixed OncoSeq Curated (53)
- ☐ IlluminaHiSeq OncoSeq Automated (790)

**File Listing**

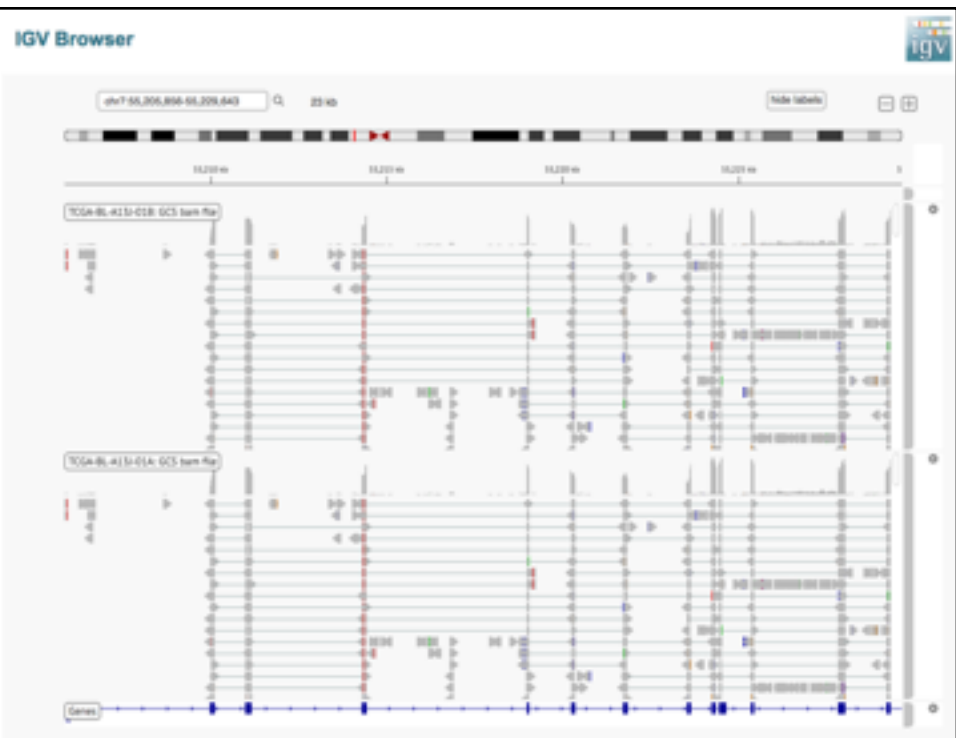
Showing 20 files of 26. Page: 1

[Previous Page](#) [Next Page](#) [Launch IGV](#)

Sample Barcode	Pipeline	Platform	Data Level	Data Type	IGV
TCGA-BL-A13J-01A	unc.edu__TotalRNASeqV2	IlluminaHiSeq TotalRNAseqV2	Level 3	RSEM_isotoms_normalized	
TCGA-BL-A13J-01A	unc.edu__TotalRNASeqV2	IlluminaHiSeq TotalRNAseqV2	Level 3	RSEM_genes_normalized	
TCGA-BL-A13J-01A	unc.edu__TotalRNASeqV2	IlluminaHiSeq TotalRNAseqV2	Level 3	RSEM_isotoms	
TCGA-BL-A13J-01B	unc.edu__TotalRNASeqV2	IlluminaHiSeq TotalRNAseqV2	Level 1	RNA Sequence-Alignment	<input type="checkbox"/> Cloud Storage
TCGA-BL-A13J-01A	unc.edu__TotalRNASeqV2	IlluminaHiSeq TotalRNAseqV2	Level 1	RNA Sequence-Alignment	<input type="checkbox"/> Cloud Storage
TCGA-BL-A13J-01A	unc.edu__TotalRNASeqV2	IlluminaHiSeq TotalRNAseqV2	Level 3	RSEM_genes	
TCGA-BL-A13J-01A	unc.edu__TotalRNASeqV2	IlluminaHiSeq TotalRNAseqV2	Level 3	junction_quantification	
TCGA-BL-A13J-01A	unc.edu__TotalRNASeqV2	IlluminaHiSeq TotalRNAseqV2	Level 3	exon_quantification	
TCGA-A6-2677-01A	unc.edu__TotalRNASeqV2	IlluminaHiSeq TotalRNAseqV2	Level 1	RNA Sequence-Alignment	<input type="checkbox"/> Cloud Storage
TCGA-A6-2677-01A	unc.edu__TotalRNASeqV2	IlluminaHiSeq TotalRNAseqV2	Level 3	RSEM_genes_normalized	
TCGA-A6-2677-01A	unc.edu__TotalRNASeqV2	IlluminaHiSeq TotalRNAseqV2	Level 3	RSEM_isotoms_normalized	
TCGA-A6-2677-01A	unc.edu__TotalRNASeqV2	IlluminaHiSeq TotalRNAseqV2	Level 3	exon_quantification	
TCGA-A6-2677-01B	unc.edu__TotalRNASeqV2	IlluminaHiSeq TotalRNAseqV2	Level 1	RNA Sequence-Alignment	<input type="checkbox"/> Cloud Storage

- Again, selected are all files that were produced using the Illumina HiSeq Total RNA Seq v2 Platform.
- Now, you can see that there are entries in the IGV column.
- Cloud Storage indicates that these are .bam files that the ISB-CGC pilot is hosting in Google Cloud Storage.
- Selecting one or more of those will activate the Launch IGV button.

## 10. IGV Browser



- You may be prompted to log in again to authorize access of the information required to display. Please be sure to allow pop-ups on this page for that, otherwise you may receive an access denied error.
- Here, we have two samples and we are looking at. The labels on the two tracks show that they are TCGA samples and GCS bam files.
- Note: this is a mock bam file that has been associated with a real TCGA sample identifier.

# Acknowledgements

The copyright to the Integrative Genomics Viewer is held by the Broad Institute, and the software has been released under the MIT License. For more information about IGV please see the [IGV Home Page](#) or the [IGV Github Repository](#).

We are grateful to the IGV team for their assistance in integrating the IGV into the ISB-CGC web application.

- <http://www.broadinstitute.org/software/igv/home>
- <https://github.com/igvteam/igv>