

An Introduction to the ISB-CGC Web App IGV

brought to you by

The ISB Cancer Genomics Cloud



Main Landing Page

- Handy place to access documentation, code, and send feedback
- You may only log in using a Google Managed identity by clicking the signin button

<https://isb-cgc.appspot.com>

Cancer Genomics Cloud

The ISB Cancer Genomics Cloud (ISB-CGC) is democratizing access to **TCGA** data and coupling it with unprecedented computational power to allow researchers to explore and analyze this vast data-space.

[Documentation](#)

[GitHub](#)

[Feedback](#)

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.

1. Log into the system
2. Create a cohort of only CCLE Samples

Clinical Features

Study



Vital Status



Sample Type



Tumor Tissue Site



Gender



Age at Initial
Pathologic Diagnosis



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.

3. Click View Files Button

Clinical Features

Study



Vital Status



Sample Type



Tumor Tissue Site



Gender



Age at Initial
Pathologic Diagnosis



4. 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.
- If the file is available in Google Cloud Storage as a .bam file, then the IGV column will also have that option available.
- The checkboxes will be disabled if you do not have the correct permissions to view the data.
- Selecting one or more will activate the Launch IGV button.

[← 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 DNAS eq	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

4. File Listing Page

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 will 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.



All TCGA Data

Owner:

[New Workbook](#)

[Comments](#)

[Duplicate](#)

[Remove](#)

[View Files](#)

[Download IDs](#)

Shared With (13)

Selected Filters

There were no filters used to create this cohort.

Details

Total Number of Samples: 1378

Total Number of Participants: 594

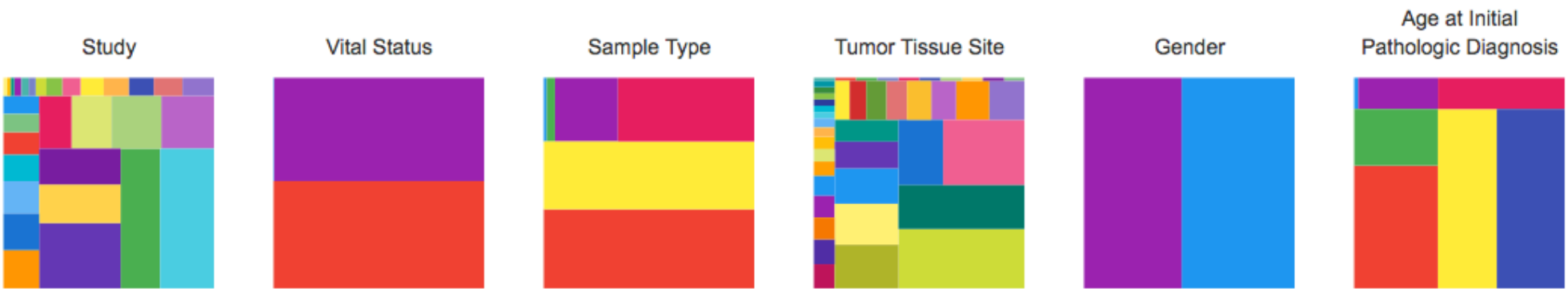
Your Permissions: READER

Revision History:

There is no revision history.

6. All TCGA Cohort

Clinical Features



7. All TCGA Cohort File Listing

- Selected are all files that were produced using the Illumina HiSeq
- Note that the entire IGV column has Cloud Storage data available. However, the checkboxes are disabled because this user has not authenticated and authorized their ERA Commons ID.
- After authenticating, these checkboxes will be active.

< All TCGA Data

Platforms

- Mixed Dnaseq Automated (809)
-
- Illuminahiseq Dnaseq Cont Automated (920)**
 - 27k (2591)
 - GA (5317)
 - GA/BCGSC (2030)
 - GA/UNC V2 (4352)
 - Illuminahiseq Totalrnaseqv2 (369)
 - Mixed Dnaseq Cont (1447)
 - Illuminamiseq Dnaseq (16)
 - Illuminaga Dnaseq (9785)
 - Mixed Dnaseq Cont Curated (212)
 - Illuminaga Dnaseq Curated (5476)
 - Illuminamiseq None (3857)
 - HiSeq/BCGSC (15298)
 - Mixed Dnaseq Curated (107)
 - Illuminahiseq Dnaseq Automated (6433)
 - HiSeq (29566)
 - Illuminaga Dnaseq Automated (20573)
 - SNP6 (66560)
 - RPPA (7668)
 - Illuminahiseq Dnaseq (20277)**
 -
- Illuminaga Dnaseq Cont Automated (8786)**
 - HiSeq/UNC V2 (71786)
 - 450k (9635)

Selected Files

Launch IGV

0 file(s) selected (limit 5)

Select files

File Listing

Download File List as CSV

Showing 20 files of 20277. Page: 1

Previous Page Next Page

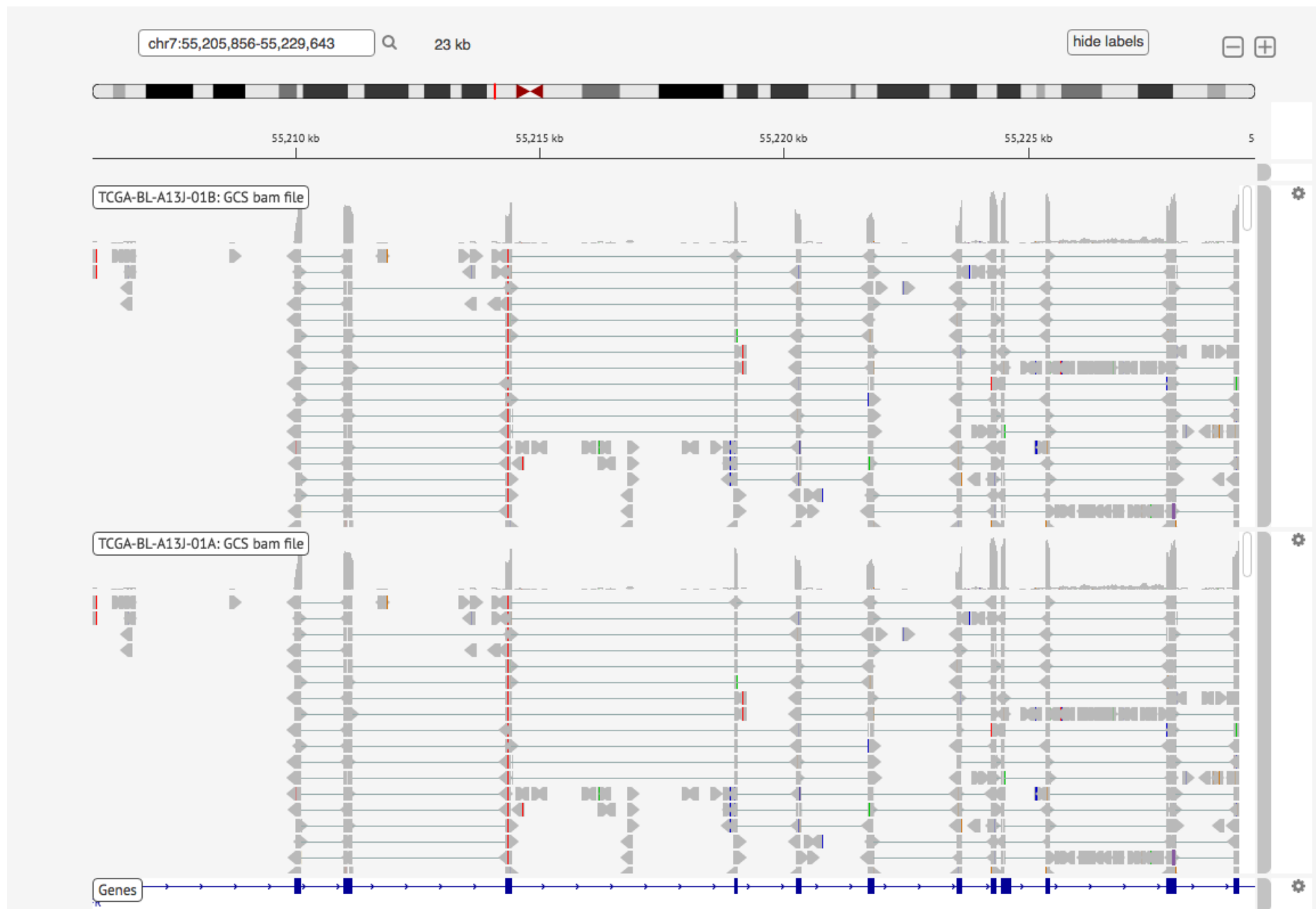
Sample Barcode	Pipeline	Platform	Data Level	Data Type	IGV
TCGA-W7-A93P-11A	genome.wustl.edu___DNASEq	IlluminaHiSeq DNASEq	Level 1	DNA Sequence-Alignment	<input type="checkbox"/> Cloud Storage
TCGA-W5-AA21-11A	genome.wustl.edu___DNASEq	IlluminaHiSeq DNASEq	Level 1	DNA Sequence-Alignment	<input type="checkbox"/> Cloud Storage
TCGA-W5-AA30-11A	genome.wustl.edu___DNASEq	IlluminaHiSeq DNASEq	Level 1	DNA Sequence-Alignment	<input type="checkbox"/> Cloud Storage
TCGA-4G-AAZT-11A	genome.wustl.edu___DNASEq	IlluminaHiSeq DNASEq	Level 1	DNA Sequence-Alignment	<input type="checkbox"/> Cloud Storage
TCGA-W7-A93P-01A	genome.wustl.edu___DNASEq	IlluminaHiSeq DNASEq	Level 1	DNA Sequence-Alignment	<input type="checkbox"/> Cloud Storage
TCGA-W6-AA0T-01A	genome.wustl.edu___DNASEq	IlluminaHiSeq DNASEq	Level 1	DNA Sequence-Alignment	<input type="checkbox"/> Cloud Storage

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. IGV Browser

- You will 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.