



Oct 06, 2020

# Protocol for IGV

1

1UCSC

1

Works for me

This document is published without a DOI.

UCSC BME 22L



Alyssa Ayala

## DOCUMENT CITATION

2020. Protocol for IGV. **protocols.io**  
<https://protocols.io/view/protocol-for-igv-bmtxk6pn>

## LICENSE

————— This is an open access document distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited

## CREATED

Sep 28, 2020

## LAST MODIFIED

Oct 06, 2020

## DOCUMENT INTEGER ID

42583

## DISCLAIMER:

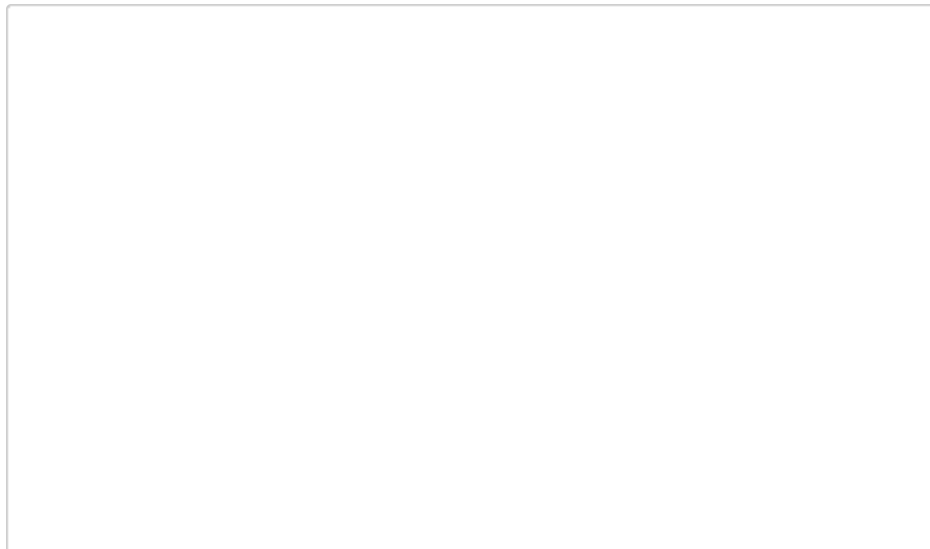
### DISCLAIMER – FOR INFORMATIONAL PURPOSES ONLY; USE AT YOUR OWN RISK

The protocol content here is for informational purposes only and does not constitute legal, medical, clinical, or safety advice, or otherwise; content added to [protocols.io](#) is not peer reviewed and may not have undergone a formal approval of any kind. Information presented in this protocol should not substitute for independent professional judgment, advice, diagnosis, or treatment. Any action you take or refrain from taking using or relying upon the information presented here is strictly at your own risk. You agree that neither the Company nor any of the authors, contributors, administrators, or anyone else associated with [protocols.io](#), can be held responsible for your use of the information contained in or linked to this protocol or any of our Sites/Apps and Services.

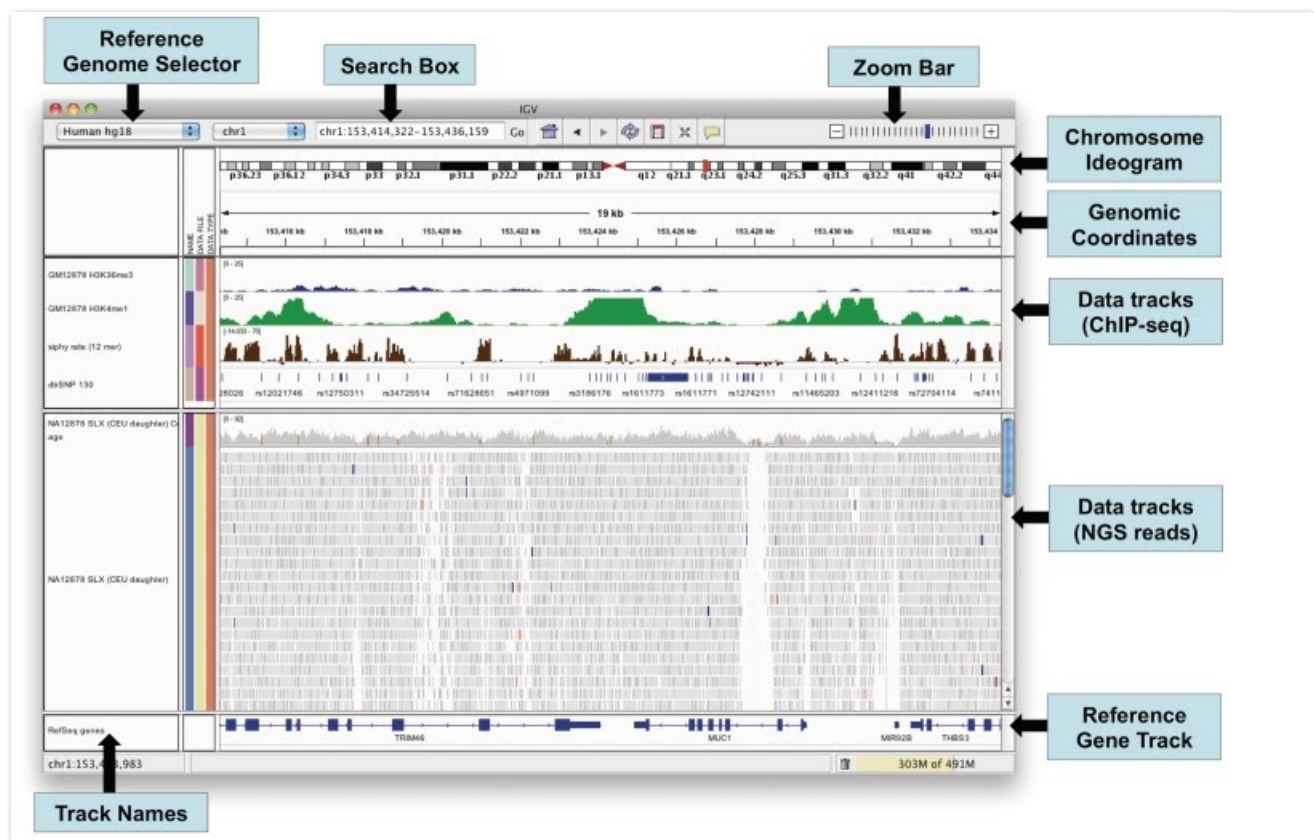
## Interactive Genome Viewer (IGV)

The interactive Genome Viewer is an interactive tool that allows one to visually analyze genomic data. With the advancement of next-generation sequencing and array-based profiling, there has come a need to analyze these diverse data-sets. IGV is a tool that allows us to analyze genomic information in real-time with great resolution and detail, allowing you to adjust the view from the entire genome down to individual base pairs. IGV is also useful because it gives you the ability to view, group, and filter different types of data sets simultaneously.

IGV first began development in 2007 in response to the need for The Cancer Genome Atlas project, where they were in demand of a tool to visualize integrated copy number data, expression, mutation, and clinical data. When trying to analyze this data, they realized the vast size of the data sets was a challenge. IGV addresses this by indexing the data files where they allow you to view several hundred samples



## Features



### Reference Genome:

In order to view the data in IGV, it is necessary that you upload a reference genome first. There are several pre-loaded reference genomes that IGV allows you to select from. You also have the ability to upload your own genome by uploading a FASTA file. When the viewer zoom is adjusted sufficiently enough, one can view the reference genome as a separate track. You can also see the individual nucleotides as colored bars or their letters depending on the zoom level. On further adjusting, one can even see the three letter translation of amino acids.

### Loading Data:

IGV has been designed to accept data that can be mapped to genomic coordinates. It can accept over 30 different file formats, namely the formats that support genome annotations, sequence alignments, variant calls, and microarray data. There are 4 ways that data files can be loaded into IGV. These are:

1. Using the built-in file browser to select a file on the local file system.
2. Entering the URL of a file accessible over a network via HTTP or FTP.
3. Entering the URL of a Distributed Annotation System (DAS) feature source
4. Selecting entries from the 'File > Load from Server' menu.

### Viewing Data:

IGV gives users the ability to view multiple data sets simultaneously, using the same or different types of data. For NGS data, it is very useful for analyzing sequence alignments as it is good at taking SAM, BAM, and Goby files. Due to the magnitude of the data stored in these reads, IGV varies the level of detail of the reads based on the zoom settings. There is a user-settable visibility threshold (default 30kb) that once the user passes that threshold, the individual reads will become visible as horizontal bars. Zooming further than this will reveal the individual bases. IGV also uses color and transparency to highlight important events. For instance, some of the gray tracks will be colored to indicate that there are mismatches within the alignment, which is helpful when identifying SNPs. The size and the color of these bars indicate the allele frequencies at these locations. Reads that match the reference genome will match the color of the reference genome tracks.

### Instructions

**We have provided you with Sars-CoV 2 sequencing data that was acquired using the nanopore. Download this data and view it using IGV. You can download the file in the Lab 6 file manager. Use the data to complete the lab.**

**Instructions:** Now that you have downloaded the sequencing data and know how IGV works, analyze the data and answer the following questions. You will need to use the Sars-Cov 2 genome as the reference genome the data will be aligned to. Load the file into IGV. In the location box search this specific region NC\_045512.2:11,577-11,615.

[UCSC Genome Browser for SARS-CoV-2](#) (wuhCor1)

**Then,** go to the UCSC Human Genome Browser and look at the Sars-CoV 2 browser. Enter the location of the variant you saw in the data in the browser and answer the following questions. Go to the section of the comparative genome and select the full display mode of Human CoV.

