Genome Browsers Guide

Take a Class

This guide supports the Galter Library class called Genome Browsers. See our Classes schedule for the next available offering. If this class is not on our upcoming schedule, it is still available to you or your group by request.

Background

Genome browsers allow users to view gene information for a species in relation to other genes on a chromosome. They also facilitate comparison between species to investigate syntenic relationships.

This guide provides a summary of three major genome browsers: Ensembl Genomes, NCBI’s Map Viewer and the University of California Santa Cruz (UCSC) Genome Browser. Each of these browsers provide a graphical interface and features to aid users in finding gene information as well as information on specific features of genes such as exons, non-coding regions and variation.

The Ensembl Genome Browser

The Ensembl genome browser interface can be accessed at:

www.ensembl.org
On the Ensembl home page you can:

- Begin a search in the search box for a gene, chromosomal region or disease name
- Select a specific genome to browse or search using the **popular genomes** or using the **drop down menu** of other species
- Go directly to genome **BLAST or BLAT** from the upper right corner
- Use **BioMart** to select specific data within genomes
- Browse the **Docs and FAQs**
- Use the **help documents and videos** on the right to learn more about using Ensembl

**Species Home Pages**

If you click on any of the popular genomes, or choose a genome from the pull down menu, you will be taken to an entry page for that species.
On this page, you can see a summary description of the genome, including number and length of contigs. Click on the Assembly and Genebuild button in the upper right to see more details about the genome, including number of genes, pseudogenes, SNPs and other data.

To get back to the Ensembl home page, click the Ensembl logo in the upper left corner.

Searching Ensembl

Enter your search terms in the box. To search for a gene, you can use the word "gene" in your search. This guide uses the gene ACE (angiotensin converting enzyme). To search across all genomes, just click the Go button. To search a specific species, use the drop down menu on the home page.

Your results will be returned by feature and by species. On the results page, click the term Gene to expand the gene menu. You can see numbers of results for the gene in each species.
Click on the species of choice to see the individual results.

Result in Detail

1 Gene matches your query

**ACE** [Ensembl/Havana merge gene: ENSG00000159840]

**Description** angiotensin I converting enzyme (peptidyl-dipeptidase A) 1 [Source:HGNC Symbol;Acc:2707] [Type: protein coding Ensembl/Havana merge gene];

**Location** 17:61554432-61599209

**Source:** E51; **Feature type:** Gene; **Species:** Homo sapiens;

Now click on the blue text link for the gene name to see the individual results in the genome browser.

**Ensembl Gene View**

You are now in the Ensembl gene view pages.
Note the tabs at the top of the page. You can navigate between gene and chromosome views using these tabs. Use the menus at the left to see specific features of the gene. Click on any blue text link (such as the transcript identifiers) to view detailed pages for each linked feature.

Further down the page, you will see the contig and transcripts for your gene.

Havana genes are manually-curated genes from the Havana gene project (human genes only). Ensembl genes are products of the Ensembl genomes computed gene features. The blue line on the diagram is the contig. Genes above the blue contig line are coded on the forward DNA strand. Genes shown under the blue line are coded on the reverse strand. All splice variants are depicted in the diagram, including pseudogenes.

To change any of the features shown in the graphic, click on the Configure this page link in the left menu bar.
The transcript configuration panel will open.

- Click any category in the **left menu** to open available feature tracks
- Click the boxes next to any feature you wish to view
- Use the **Show info** button to see descriptions of the feature tracks
- You can change several feature categories by opening new feature menus
- When you are done adding features, click the **checkmark** in the upper right
- Your browser view will reload with the new features
- If you don’t like the changes you’ve made, you can click the **reset text link** near the bottom of the configuration window

Click on a transcript ID in the table at the top of the page.

**Transcript View**

Your view will change and you will now have a Transcripts tab at the top of the page, as well as a transcript-based menu on the left.
You can now view specific features of the transcript using the menu to access sequence, variation, protein, oligo probes, ontologies and evidence.

- **cDNA** view will show you the cDNA plus the encoded amino acid sequence, plus any variation consequences
- **Exon** view shows the exons highlighted in the sequence
  - Use the **Configure this page** button to change the amount of sequence flanking the features

You can also click the tabs at the top of the page to go back to the gene view or the chromosomal view. Go back to the **Gene** view.

**Viewing Variation in Gene View**

To see SNPs and other variation features for a gene, click on either the **Variation table** or **Variation image** in the left menu in Ensembl's **gene view tab**.
Depending on your gene, there may be such a great number of variation features that the image is too busy to read. You can adjust the image and the table to reflect only the variation features you desire by using the Configure this page button:

- **Configure this page** is context-dependent, so you will only see variation options when you are viewing variation features in the gene view tab
- Select the variation features you want to see
  - For example, to view only non-synonymous variations, click on **Consequence type** in the left pane, then uncheck all of the boxes except **Non-synonymous**, and click **checkmark** in the upper right corner

Other Features in Ensembl Genome Browser

- View **synteny**, **multi-species views** and **alignment** in the **Location** tab
- View an interactive Java-based **gene tree**, **orthologs**, **paralogs** and **alignments** in the **Gene** tab
- View **regulation** features in the **Gene** tab
- Export sequence of gene and flanking regions with the **Export data** button

There are many features and functions in all of the pages at Ensembl. Users are strongly encouraged to spend time using the browser and the support documents and videos available at the Ensembl web pages.

---

**NCBI Genomes and MapViewer**

Searching genomic sequences for gene features at the NCBI is somewhat different than finding these features at Ensembl. Gene information in a genomic context is best found through links in the Entrez Gene database records to your gene of choice in MapViewer, the NCBI's genome browser.

Go to the NCBI home page:


8 of 17
Search for your gene of interest by typing it in the text box at the top of the page (this guide uses the gene ACE as an example). Click **Search**

- Click on the results in the Gene category

In the Entrez Gene database, click on the **gene name symbol** to access the record for your gene

- Scroll down the gene record until you see the section called **Genomic context**
- Click on the link to see your gene in MapViewer

You are now viewing your gene in MapViewer, NCBI's genome browser and viewer.
Features in MapViewer

There are many features and options in MapViewer to choose from:

- You can see the **genome build number** at the top of the page.
- An **ideogram** of the chromosome is located in the left menu.
- You can **zoom in or out** using the left menu bar or by clicking anywhere on a map and using the pop up menu to zoom.
- **Remove any map** from the view by clicking the X next to the map name.
- Read descriptions of maps by clicking on the **map name**.
- Your gene in sequence is depicted by a **vertical red line** running parallel to the maps.
- Link out to other NCBI databases using the **blue text links** in the pinkish menu on the right.
- Customize your map by clicking on the **Maps & Options** button.

Click the **Maps & Options** button (near the upper right of your MapViewer window).
See the region shown on the maps

Change the organism and assembly—organisms available for full assembly are currently human, chimp, mouse and rat

Select maps to add to your view by clicking them in the Available Maps menu—use control key (command key on a Mac) to select more than one map

Click the ADD button to add selected maps to your view

Move maps up and down using the Move keys

When you are done adding or removing maps, click the Apply or OK button

The maps in your viewer window will now update

Links in MapViewer

Links to other NCBI databases can be accessed using the links in the pink rectangle in MapViewer.

- The gene symbol links you back to the Entrez Gene record
- The arrow indicates the DNA strand on which your gene is localized. Down is the forward strand.
- OMIM links to your gene’s record in the Online Mendelian Inheritance in Man database
- HGNC links to the Hugo Gene Naming Consortium page for your gene (HUGO is not an NCBI database)
- sv links to the sequence viewer where you can view your gene's sequence and coding properties
- pr links to the Entrez Protein record for your gene's protein product
- dl links to a page where you can set the sequence to download in FASTA format
- ev links to the Evidence Viewer, where you can view all the sequence data that contributed to your gene's build
- mm links to the MapViewer Model Maker, where you can build different splice variants of your gene by linking exons together
- sts links to the Sequence Tagged Sites database, which will show you primers that are likely to produce the
sequence for your gene in any available species' models

- **CCDS** links to the Consensus Coding DNA Sequence project database, which identifies a core set of human protein coding regions that are consistently annotated by multiple public resources
- **SNP** links to the NCBI's SNP database, dbSNP

**Other Ways to Access MapViewer**

You can also access MapViewer without starting from an Entrez Gene record.

- On the [NCBI home page](https://www.ncbi.nlm.nih.gov), click on the left side blue menu item **Genomes and Maps**

- Scroll down the page to the **Tools** section and click on **MapViewer**
- You will now be at the MapViewer home page, which looks very different from the view you access from an Entrez Gene record
NCBI's **Entrez Genomes database** is not very useful on its own. It supplies the data that is used in MapViewer, but the Genomes database is not as easily searched from the home page as some other NCBI databases. If you do want to do a gene name search in NCBI's Entrez Genomes database, you should use the **Limits tab** from the [Entrez Genome home](#) page to limit your search to the Gene Name field. Then you can further limit by species, if you wish.

---

**Summary and More Information**

Every user will develop preferences for genome browsers, and there is no "right" or "wrong" browser to use. While preference is subjective, here are some strengths of each of the browsers:

- Ensembl contains more genomes than any other browser, with access to annotation features for each species represented by the genome
- NCBI's MapViewer provides the quickest access to a gene's sequence
- UCSC allows more complete and detailed information to be accessed on gene and transcription features

### Tutorials and Guides

There are excellent help documents for each of these browsers. There are even some demonstration videos for Ensembl and UCSC genome browsers.

- Ensembl has a [help page of video tutorials and pdf guides](#)
- NCBI has some [examples of how to use MapViewer in the NCBI Handbook](#)

As always, the [Biosciences & Bioinformatics Librarian](#) at Galter is available to help you.

---

**University of California Santa Cruz (UCSC) Genome Browser**
The UCSC Genome Browser is the most dense and complex of the three genome browsers shown in this guide. Despite the complicated look of the browser, users are encouraged to try it, because it provides a number of tracks and views that are available nowhere else.

The UCSC Genome Browser is available at:

http://genome.ucsc.edu/

There are many tools and features available directly from this home page. They are all listed both along the top dark blue menu bar and along the left side lighter blue menu. The central text area briefly describes each tool or database. This guide only refers to the Genomes and Tables section of the UCSC Genome Bioinformatics pages, but the other tools are worth exploring, as well.

Click on the Genome Browser link in the left menu or on Genomes in the top menu.

You can search by chromosome and position, by keyword, or by gene symbol. You can configure the tracks that display in your browser results in this window, but you can also add tracks to the display after your search. You can change the clade, genome or assembly and adjust the width of the map in pixels.

The browser will retain your search from the same computer from session to session, so to reset the browser settings to their defaults, click the text link to reset.
When you've entered your search terms, click submit.

You now see the browser maps in the top half of the page and the settings for maps and display that you can change in the bottom half of the page. The size of your view in base pairs, is displayed right above the chromosome view above your genome tracks window. The default tracks displayed in the graphic are currently:

- Base position mapping
- UCSC predicted genes
- RefSeq genes (NCBI)
- Human mRNA
- Spliced ESTs
- Species conservation alignments
- SNPs

Each map can be hidden or displayed in a number of condensed or expanded forms. You can hide all tracks (using the button below the map display), then display only the maps you want using the drop down menus in each section. There are dozens of choices of maps and tracks.

- You can tell the direction the gene is coded by the small arrows on the Gene and RNA tracks: arrows pointing right mean that the gene is coded on the forward strand, and arrows pointing left show that the gene is coded on the reverse strand
- You can zoom in or out or move in any direction in the map
- Click on any track in the display to expand that section to view all features available
  - For example, if you click on the SNP track, you will see all of the snips displayed with their reference numbers from the NCBI dbSNP database
- When viewing a gene's region, click on DNA in the top blue menu bar to retrieve the cDNA sequence in FASTA format
- Click on a transcript to get detailed information, from sequence details to translated protein structures to ontology details
- Click on the blue or gray bars at the left of the track window to see detailed information about the tracks, evidence and data that was used to create them and citations to literature supporting the methods used to create the track

UCSC's Tables

UCSC Genome Browsers Tables are the easiest and fastest way to find transcription start and end sites, specific genomic coordinates of exon boundaries and many other details not visible in the graphical view of the map.
When viewing a gene or chromosomal region in the UCSC Genome Browser, click on the Tables link in the top blue menu bar to access this detailed information on features of the sequence.

In the tables set-up window, you have many options to set your output. The tables are in a MySQL database, and the Tables set-up window allows users to see the details of the table set-up schema, create intersections between tables, filter data and much more.

- From the tables set up page, you will see your gene’s chromosomal position in the text box, but you must select the radio button next to position. Otherwise, your table output will include the entire genome.
- Choose the type of track groups you want to view. Genes and Gene Prediction Tracks will give you features of exons and transcription sites.
- You choose the tracks to view: UCSC Genes is the default.
- You can set output format to sequence, hyperlinks, all fields, selected fields.
- Save the file by typing a file location in the output file text box, or leave this box blank to view the data as tab-formatted text in your browser window.
- Click get output when you are ready to view your data.

![Table Browser](image)