

Accessing the Human Genome

Three websites provide free, online access to the Human Genome:

National Center for Biotechnology Information:

MapViewer

<http://www.ncbi.nlm.nih.gov>

European Bioinformatics Institute: Ensembl

<http://www.ensembl.org>

Univ. of California – Santa Cruz: Genome Browser

<http://genome.ucsc.edu>

Primary assembly of the human genome sequence is done at NCBI; additional annotation and display features are conducted individually at each site

I. NCBI's Human Genome Resources

<http://www.ncbi.nlm.nih.gov/genome/guide/human/>

The Organism Home Page provides 6 categories of links:

- Genes and Human Health (OMIM, Entrez Gene, Ref Seq)
- The Genomic Sequence (FTP, BLAST)
- Maps and Markers (MapViewer, Types of Maps)
- Transcribed Sequences (Unigene, GEO)
- Cytogenetics (BACs, SKY/CGH)
- Comparative Genomics (Homologene, TaxPlot)

Genome View: Search Page for the Map Viewer

Simple Search Options

- Keyword
- Keyword + chromosome #
- Assembly: NCBI reference, Celera, MHC haplotypes.

Advanced Search Options

- Fields
- Object types
- Chromosome numbers
- Types of map
- Record properties
- Show linked entries

The MapViewer provides a graphical display of multiple maps of the human genome that can be easily manipulated. The maps are displayed vertically, left to right, across the screen. Up to ten maps can be simultaneously displayed.

There are four types of maps available:

- Cytogenetic Maps

Map elements (genes, markers, clones) are localized on chromosomes using fluorescent or radioactive labeled probes. The probe location is identified relative to the chromosomal banding pattern (karyotype) created by a chromosomal stain such as Giemsa.

- Genetic Linkage Maps

These maps are based on recombination frequency studies of genetic linkage in inheritance. The frequency of recombination between two sites is inversely correlated with the distance between them. The standard unit of measurement for these maps is the cM (centimorgan): 1cM represents a 1% probability that a recombination has occurred between two positions on a chromosome.

- Radiation Hybrid Maps

These maps are similar to genetic linkage maps, except that the chromosome breakage and recombination is artificially-induced with radiation rather than measured through inheritance. The pattern of break frequency between markers is correlated with distance between the markers. These maps are static (will not be updated with more markers).

- Sequence-Based Maps

These maps are based on the human genome sequence in basepairs as generated by the Human Genome Project. Position of elements on the maps is identified by BLAST comparisons to the genomic sequence or computer prediction. There are multiple types of sequence maps that show genes, markers, variation sites, ESTs, and other annotated features.

Manipulation the Map Viewer

- SideBar
- Top of the Display
- Map Graphics
- Maps & Options
- Zoom Options
- Map Summaries

Blue SideBar includes

•**Links**

- Help Documents; FTP Site
- Data As Table View
- Maps and Options (Add and Remove Maps)

•**Compress Map** – squeeze maps to left

•**Region shown** – select region shown in Master Map units

•**Zoom graphic** – zoom in and out by clicking on graphic

•**Ideogram** – graphically demonstrates which region of chromosome is being shown

Top of page displays:

- Human Genome Build # (links to statistics)
- Option to BLAST against the genome
- Quick links to each human chromosome
- Search Query
- Identification of Master Map
- View Summary links to further information on bottom of page
- Maps and Options button (add and remove maps)
- Region Displayed – region of chromosome shown; units of the Master Map
- Download/ViewSequence/Evidence –Download the genomic sequence

Graphical Map Display

- Maps aligned left to right
- Right-most map is Master Map (name in red)
- Click on a map name to get more information.
- Blue arrow – make a displayed map become the Master Map.
- Blue “x” - eliminate a map from the display.
- Click on the identifier for a map element for more information.
- Queried map element marked with pink background on Master Map

Maps and Options Window

Allows selection of the maps from the menu. Use buttons to order the maps for the display, to identify the MasterMap (automatically the one on the bottom of the selected list), and add/remove rulers for each map.

MapView: Zoom Options

Zoom Menu

Hold mouse over center of MasterMap until cursor becomes a small hand, then left click once to view Zoom Menu.

Summary of Zoom Options

- Use zoom menu: Click on your desired zoom level and the MapViewer will adjust the display accordingly.
- Use the floating icon (thumbnail zoom box) on the left sidebar. Click on desired level.
- Type in the exact region in the Region Shown text boxes on the blue left hand side bar and click on 'go'.

The Map Summary

Shows statistical information on each map in the display:

- region of chromosome shown in the specific map units
- total number of map elements in the region
- total number of map elements actually displayed

Help Pages

MapView Help (<http://www.ncbi.nlm.nih.gov/mapview/static/MapViewHelp.html>)

Provides help with searching techniques and methods for modifying the display.

Human Maps Help (<http://www.ncbi.nlm.nih.gov/mapview/static/humansearch.html>)

Provides descriptions of each individual map and the data represented by that map.

The Genes_Seq Map: Genes Located in the Genomic Sequence

This map shows sequence locations for genes based on BLAST alignments of mRNAs to the genomic contigs. If multiple splice variants are known for a gene, a flattened view that shows all possible exons is provided.

Exons are thick lines while **introns** are the thin connecting lines between the exons. Genes shown on the **left of the grey line** are transcribed in the - orientation (from bottom up), and those on the **right** in the + orientation (from top down).

Colors and Evidence Codes:

Gene Color on Map	Evidence Code	Evidence Type
Blue	C	Confirmed; BLAST alignment of mRNA plus ESTs
Light Green	E	EST evidence only
Dark Brown	PE	Predicted by Gnomon Plus EST evidence
Light Brown	P	Predicted by Gnomon only
Orange	?	Uncertain: There is some discrepancy in the available evidence

Annotation Links from the Genes_Seq Map

OMIM: Display the OMIM record for the gene

SV: Sequence Viewer - display the genomic sequence at the base pair level

PR: Display protein records in Entrez Protein for that gene

DL: Download the genomic sequence

EV: Evidence Viewer - display the evidence for the exon/intron structure of the gene

MM: Model Maker - manipulate the evidence shown in the evidence viewer to create and download different gene structure models

HM: Homologene - identify homologous genes in other organisms

Viewing Human/Mouse/Rat Maps Simultaneously

The human Maps & Options dialog box allows the user to select maps from two other organisms besides Homo Sapiens:

- Mus Musculus – the mouse
- Rattus Norvegicus – the rat

The maps from these two organisms can be displayed in the human MapViewer, side-by-side with the human maps.

- Select organism to see available maps
- Choose maps and add to display list
- Display list will show organism and map name
- Select Master Map/map order and toggle rulers
- Click Apply

The maps from different organisms are shown on differently-colored backgrounds.

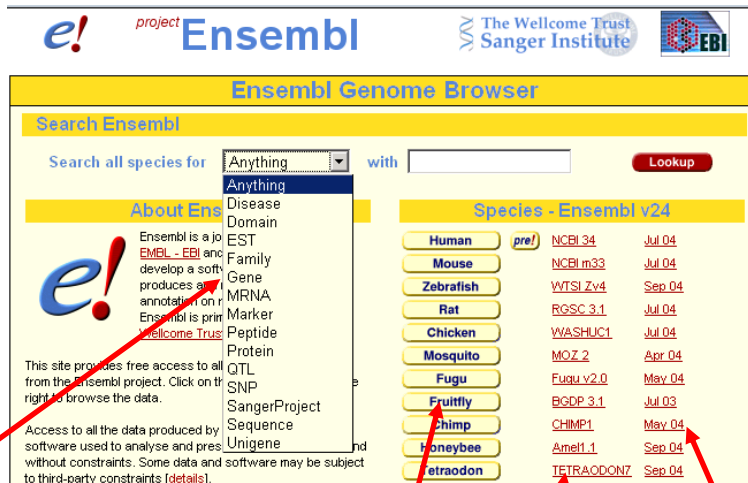
The mouse and rat maps show the chromosome next to the map title. The positioning of corresponding elements such as genes are show by connecting lines.

Zooming and display manipulation works the same for the mouse as for the human maps.

II. Ensembl (European Bioinformatics Institute)

www.ensembl.org

The Ensembl genome browser is created and maintained jointly by the European Bioinformatics Institute (<http://www.ebi.ac.uk/>) and the Sanger Institute (<http://www.sanger.ac.uk/>). Access to the Ensembl website, data, and software is free.



Select element from dropdown menu to do a basic keyword search across all included organisms.

Choose a specific organism. Click on the yellow button to access a search page for that genome.

Click on an assembly to access information about the assembly.

Click on a date to access statistics about the particular build of that genome assembly.

Human Genome Page

Click on a chromosome to view and browse the maps.

Search Options:

- Select element from dropdown menu
- Select chromosome and region
- Keyword search
- BLAST option for searching with sequence
- Advanced Search: EnsMart
- Allows use of filters and can generate multiple data formats for results

Results

The results show two types of entries.

- Gene records from the Vertebrate Genome Annotation (VEGA) database
- Ensembl gene records

Click on the Gene Identifier number to view the full Gene Report. This provides a summary of information available on the gene with links to other databases. Click on the links in “Genomic Location” to access the genome maps.

Ensembl Contig View

- Overview

Provides a graphical view of the position of the gene on the chromosome displayed top-to-bottom, rather than left-to-right like the MapViewer. Includes: synteny maps to other species, contig and chromosomal band location, markers, and genes.

- **Detailed View**

Provides a graphical view of cDNAs, mRNAs, ESTs, proteins, and computer-predicted transcripts; somewhat similar to NCBI's evidence viewer. Can be scrolled or zoomed as desired.

- **BasePair View**

Provides the genomic sequence at the basepair level and includes all the translated amino acid sequence in all 6 possible reading frames. Each base and aa has a specific color. Red * signal a stop codon. Can be scrolled or zoomed as desired.

III. UCSC Genome Browser

<http://genome.ucsc.edu/>

Search Options

- **Genome Browser:** search by organism and chromosomal coordinates
- **Gene Sorter:** keyword search to obtain table of genes within a selected genome that are related to one another
- **Blat (“BLAST-Like Alignment Tool”):** fast sequence alignment tool similar to BLAST
- **Table Browser:** text-based access to the genome assemblies and annotation data stored in the Genome Browser database

Genome Browser

- Select from multiple organisms: human, chimp, mouse, rat, chicken, fruit fly and others
- Choose assembly date and image width for display
- Enter position within genome (examples of positional queries given on bottom of genome browser search webpage)

Gene Sorter

- Select an organism and assembly date from the pull-down menus.
- Type keyword(s) into the *search* text box
- Choose the gene relationship to sort the list using the *sort by* pull-down menu.
- Press the Go! button to display your results.
- The results are a table containing the specified gene, highlighted in green, and its relatives, each on a separate line. Use the *display* pull-down menu to adjust the number of genes shown.

The Gene Sorter will provide a list of genes based on your keywords and the chosen type of sort.

Type	Description
Expression (GNF Atlas2)	Difference in Expression with Selected Gene According to GNF Gene Expression Atlas2
Expression (GNF Atlas1)	Difference in Expression with Selected Gene
Protein Homology	Blastp E-value with Selected Gene
Gene Distance	Distance in Base Pairs from Selected Gene
Chromosome	Ordered by Chromosome Position
Name Similarity	Number of Leading Characters in Name that Match Selected Gene Name
Alphabetical	Alphabetical Order of Name
GO Similarity	Number of Shared Gene Ontology Terms

Results can be resorted by changing the selection in the drop down menu, re-entering the search keywords, or clicking on one of the genes in the results display.

Click on the chromosomal location of a gene in the results display to view the gene in its genomic context in a graphical map display.

The map display is vertical like the Ensembl display. Several sets of dropdown menus below the display allow manipulation of the maps.

One nice feature of the UCSC Genome Browser is the ability to display synteny blocks from multiple organisms very easily.

The map can be scrolled and zoomed at will.