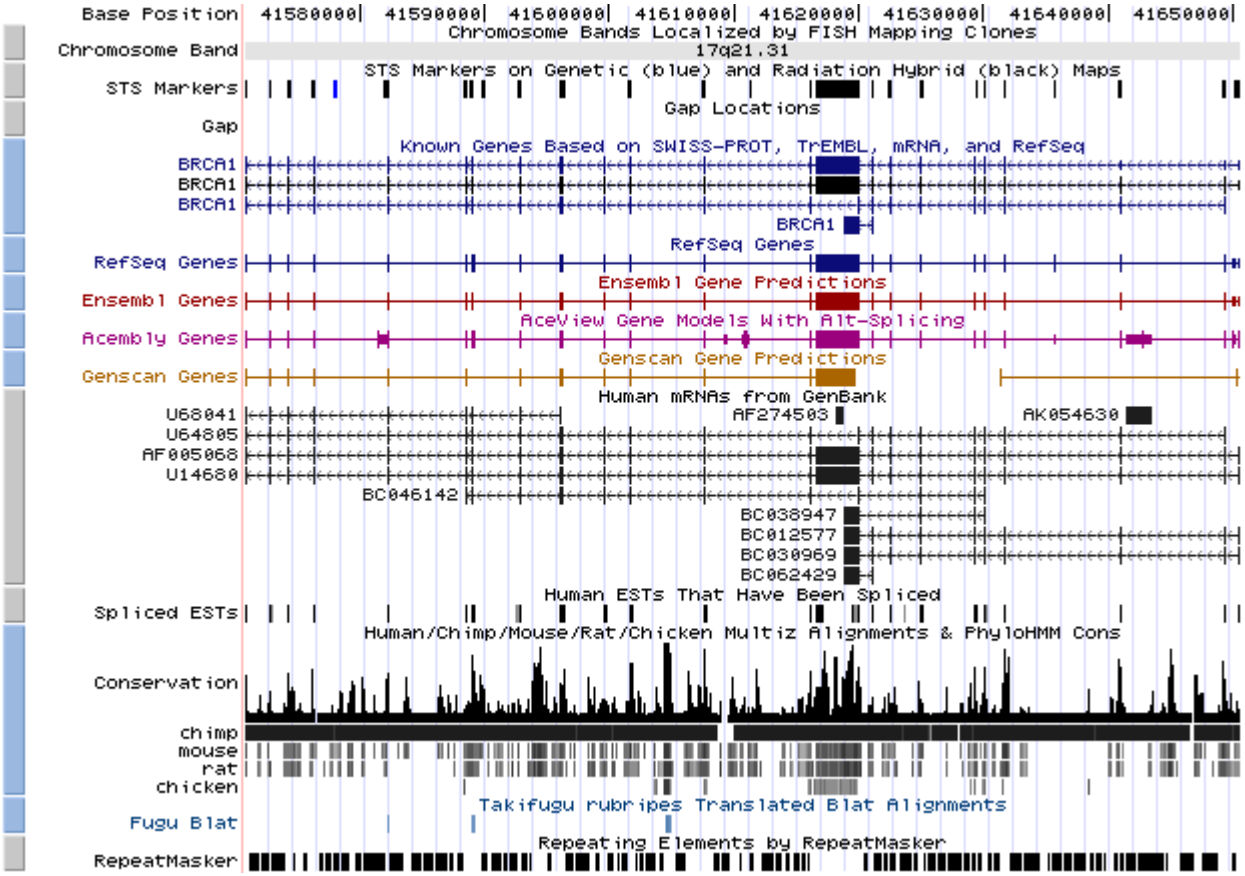
THE UCSC GENOME BROWSER

The UCSC Genome Browser is a web-based tool that allows researchers to view all 23 chromosomes of the human genome at any scale from a full chromosome down to an individual nucleotide. The browser integrates the work of countless scientists in laboratories worldwide, including work generated at UCSC, in an interactive, graphical display. The browser presents both experimentally validated and computer-predicted genes along with dozens of lines of evidence that help scientists recognize the key features of genes and predict their function. The databases for the genome browser are updated nightly with new information generated by researchers throughout the world.

Extremely fast search software on a remarkable computer system known as the KiloKluster allows researchers to match any DNA sequence to the human genome in seconds, thereby mapping experimental data to the reference sequence.

When directed to focus on a particular segment of the genome, the browser displays a range of data that is stacked vertically. At the top, it shows the chromosome number and the position on the chromosome. Underneath, it shows several rows of data about genes that have been found experimentally or have been predicted by a number of different methods. Below those are lines of information about gene expression and regulation, followed by comparisons with the genomes of other species and other information, such as single-nucleotide polymorphisms (SNPs).

Here's a fragment of a typical genome browser page:



The UCSC group continues to add functions to the genome browser, such as the UCSC Gene Sorter, which allows researchers to sort and filter data online, receiving detailed information about each known gene. The gene sorter groups genes according to several types of relationships, such as protein-level homology, similarity of gene expression profiles, or genomic proximity. By merging experimental results from multiple sources, this powerful tool allows researchers to better understand how genes function.

A next step beyond viewing a genome is gaining an understanding of the instructions encoded in it. Toward this end, clicking on a gene in the Known Genes track (see figure) leads to the UCSC proteome browser. The proteome browser elucidates gene expression through the encoded protein structures. It presents a rich set of useful protein properties and links to other valuable genomic and proteomic data sources available on the UCSC Genome Browser and on the web.

Together this information represents an extremely comprehensive view of the genome, helping scientists recognize important features of the sequence and providing strong evidence of function. For instance, the browser helps unravel the varied splicing patterns whereby one gene can make many different proteins. This process of alternative splicing is thought to explain how a human can be so complex, yet have only about twice as many genes as a roundworm.

Once the human genome sequence became available, other genome browsers also came on line, most notably those at the National Center for Biotechnology Information (NCBI) and at the European Bioinformatics Institute (EBI). Reciprocal links provided on each of the three browsers allow researchers to jump from any place in the human genome to the same region on either of the other two browsers.

Visit and use the UCSC Genome Browser at **www.genome.ucsc.edu**.

UCSC Proteome Browser

The UCSC Proteome Browser presents a rich set of useful protein properties, and it links to other valuable protein and genomic data sources available on the Web. For the first time, a user can access both the genome world and proteome world at the same time.

Integrated into the UCSC Genome Browser, the proteome browser can be accessed through the “known genes track” on genomes that include it (such as human, mouse, and rat). Clicking on a gene that has records associated with it makes the proteome browser available (under “Quick Links to Tools and Databases”).

For each protein, the browser displays its corresponding genomic exon structure (showing the coding portion of the genome) along with its amino acid sequence. A click on a particular exon will bring up the genome browser, which displays the genomic region of this exon together with other relevant data. The protein browser contains various data tracks that help visualize regions that may be of interest. It also contains three-dimensional structures and histograms showing a variety of properties, such as molecular weight, hydrophobicity, and frequencies of particular amino acids.