

DOI:
10.1038/nrg2243

MILESTONE 19

The clickable genome

“These heavily used genome browsers are now so much a part of the fabric of genome-based biological research that their contribution to progress would be difficult to overestimate.”

Consider a thought experiment in which the sequences of the human genome and the genomes of several model organisms were finished to a high standard, but the repositories that were needed to house and make sense of these data in an accessible manner did not exist. The sequence databases could be filled, but the usefulness of the data to biologists would be greatly diminished. In place of a powerful resource, we would have an ‘alphabet soup’.

The earliest repositories for DNA sequences established in the early-to-mid 1980s were the European Molecular Biology Laboratory Data Library, GenBank at the National Center for Biotechnology Information (NCBI) and the [DNA Data Bank of Japan](#). Although these databases have served the community admirably, as the pace of sequencing increased, it became clear that new graphical user interfaces would have to be developed in order to facilitate the viewing and manipulation of both the sequences and the subsequent annotations that would make them meaningful. An early and influential effort in this regard was ACeDB, a genomics database that was originally developed for the *Caenorhabditis elegans* Genome Project that could display genetic, cosmid and sequence maps in a flexible manner.



When the assembled draft sequence of the human genome was published in February 2001, it was made available through three public portals: [Ensembl](#), the [University of California, Santa Cruz \(UCSC\) Genome Browser](#) and the [NCBI Map Viewer](#). Mainly funded by the Wellcome Trust, Ensembl is a joint project of the European Bioinformatics Institute and the Wellcome Trust Sanger Institute. It presents a range of views of the human genome and the genomes of other organisms. Importantly, these are ‘clickable’ genomes, so the user can home in on small regions of a genome of interest to see protein-coding genes, RNA-coding genes, single-nucleotide polymorphisms, nucleotide composition, pseudogenes, contigs, expressed sequence tags, comparative alignments to other genomes and links to a suite of other databases that constitute the ongoing effort to produce a deep functional annotation of sequenced genomes. The UCSC Genome Browser, which was produced in its initial form by the Santa Cruz group that carried out the first genome assembly for the public Human Genome Project, can also present

a view of the genome at any scale, and offers annotations in a series of ‘tracks’ that can be added or eliminated depending on the interests of the user. These heavily used genome browsers are now so much a part of the fabric of genome-based biological research that their contribution to progress would be difficult to overestimate.

*Alan Packer, Senior Editor,
Nature Genetics*

ORIGINAL RESEARCH PAPERS Kent, W. J. *et al.* The human genome browser at UCSC. *Genome Res.* **12**, 996–1006 (2002) | Birney, E. *et al.* An overview of Ensembl. *Genome Res.* **14**, 925–928 (2004)

FURTHER READING Reed, J. L., Famili, I. & Thiele, I. & Palsson, B.O. Towards multidimensional genome annotation. *Nature Rev. Genet.* **7**, 130–141 (2006) | Eichler, E. E., Clark, R. A. & She, X. An assessment of the sequence gaps: unfinished business in a finished human genome. *Nature Rev. Genet.* **5**, 345–354 (2004) | Chintapalli, V.R., Wang, J. & Dow, J. A. Using FlyAtlas to identify better *Drosophila melanogaster* models of human disease. *Nature Genet.* **39**, 715–720 (2007) | Boffelli, D., Nobrega, M. A. & Rubin, E. M. Comparative genomics at the vertebrate extremes. *Nature Rev. Genet.* **5**, 456–465 (2004) | Ureta-Vidal, A., Ettwiller, L. & Birney, E. Comparative genomics: genome-wide analysis in metazoan eukaryotes. *Nature Rev. Genet.* **4**, 251–262 (2003)

WEB SITES

DNA Data Bank of Japan:

<http://www.ddbj.nig.ac.jp>

Ensembl: <http://www.ensembl.org>

NCBI Map Viewer: <http://www.ncbi.nlm.nih.gov/mapview>

UCSC Genome Browser:

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