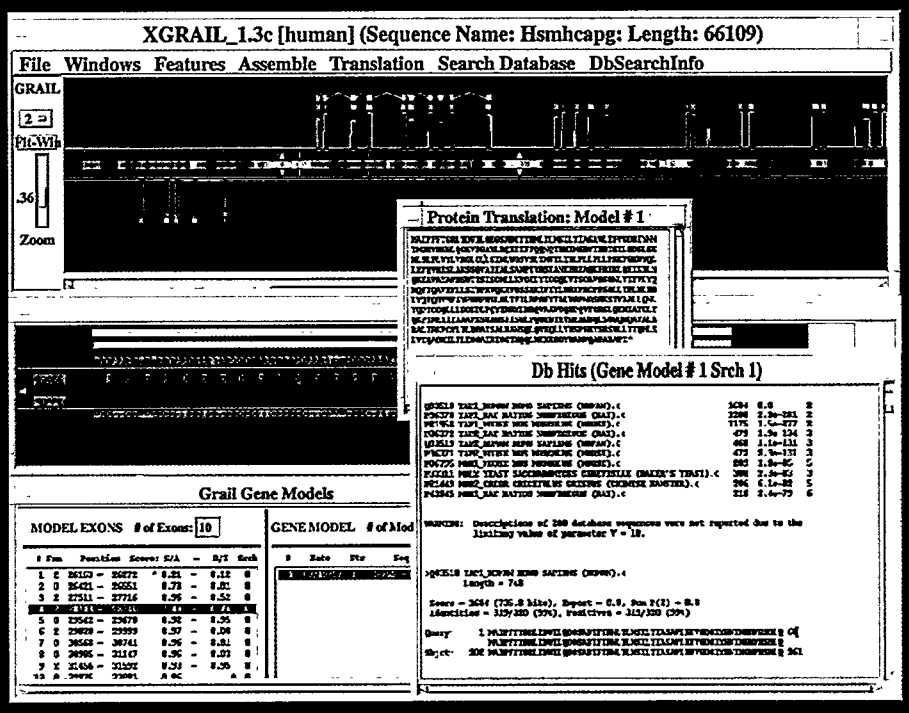
GRAIL and GenQuest

In 1996 the Gene Recognition and Analysis Internet Link (GRAIL) processed nearly 40 million bases of sequence per month, making it the most widely used "gene-finding" system available. Developed at Oak Ridge National Laboratory (ORNL) by a team led by Ed Uberbacher, GRAIL uses artificial intelligence and machine learning to discover complex relationships in sequence data. The genQuest server, also at ORNL, compares information generated by GRAIL with data in protein, DNA, and motif databases to add further value to annotation of DNA sequences.

GRAIL's latest version (1.3) combines a Motif Graphical Client with improved sensitivity and splice-site recognition, better performance in AT-rich regions, new analysis systems for model organisms, and frameshift detection. This system can be used on a wide variety of UNIX platforms, including Sun, DEC, and SGI. The many ways to access GRAIL include a command line sockets client that permits remote program calls to all basic GRAIL-genQuest analysis services, thus allowing convenient integration of GRAIL results into automated analysis pipelines.

Contact GRAIL staff through the Web site at *http://compbio.ornl.gov* or at *GRAILMAIL@ornl.gov* for e-mail and ftp access.



The figure above shows the GRAIL analysis of part of the human major histocompatibility locus, which carries genes responsible for cellular immunity. Included in this analysis are potential exons (gene-coding regions), gene models, CpG islands (areas rich in bases C and G found in most mammalian genes), and repetitive DNA elements. [Source: Richard Mural, ORNL]

and information: *http://beagle.colorado.edu/~eesnyder/GeneParser.html]*

The Baylor College of Medicine (BCM) Search Launcher improves user access to the wide variety of database-search tools available on the Web. Search Launcher features a single point of entry for related searches, the addition of hypertext links to results returned by remote servers, and a batch client. *[http://gc.bcm.tmc.edu:8088/search-launcher/launcher.html]*

FASTA-SWAP, also from the BCM group, is a new pattern-search tool for databases that improves sensitivity and specificity to help detect related sequences. BEAUTY, an enhanced version of the BLAST database-search program, improves access to information about the functions of matched sequences and incorporates additional hypertext links. Graphical displays allow correlation of hit positions with annotated domain positions. Future plans include providing access to information from and direct links to other databases, including organism-specific databases.

PROCRUSTES uses comparisons of the same gene of different species to delimit gene structure much more accurately. The product of a collaboration between Pavel Pevzner (University of Southern California) and two Russian researchers, PROCRUSTES is based on the spliced-alignment algorithm, which explores all possible exon assemblies and finds the multiexon structure that best fits a related protein. *[http://www-hto.usc.edu/software/procrustes]*