A DNA/protein sequence retrieval system: an Xtpanel application

Akihiro Oda [1] Jun-ichi Yano [1] Minoru Kanehisa [2]

[1] Information System Department
Molecular Biology Research Department
Nippon Shinyaku Co., LTD.
Nishiohji-hachijyo, Minami-ku Kyoto 601, Japan

[2] Institute for Chemical Research
Kyoto University
Uji Kyoto 611, Japan

Abstract

Computational methods are indispensable tools for researchers who determine a new nucleic acid or protein sequence. Such methods are used to obtain informaion to help identify and annotate the new sequence. Database searching is currrently the most popular way to get this information. Using homology- search programs such as FASTA or BLAST, one can search a database for sequences similar to a given query sequence. In addition, the alignment of multiple nucleotide or amino acid sequences is important for drawing phylogenetic trees and for predicting the structure of proteins. The multiple sequence alignment of proteins is also used to discover motifs and biologically important patterns.

We have developed a retrieval system for nucleic acid and/or protein sequences. The system is an Xtpanel application that runs on Sun Workstations with the X-Window system. Xtpanel, which is written by Steve Cole and Dave Nichols of Stanford University, allows one to build interactive X-Window system interfaces without having to do conventional X programming.

It supports FASTA and BLAST for homology searches, bfind and bget for database entry retrievals, and CLUSTAL V for multiple sequence alignments.

---

[1] 尾田 明博, 矢野 純一：日本新薬株式会社 情報システム部 分子生物学研究部, 〒601 京都市南区西大路通八条下ル

[2] 金久　實：京都大学化学研究所, 〒611 京都府宇治市五ケ庄

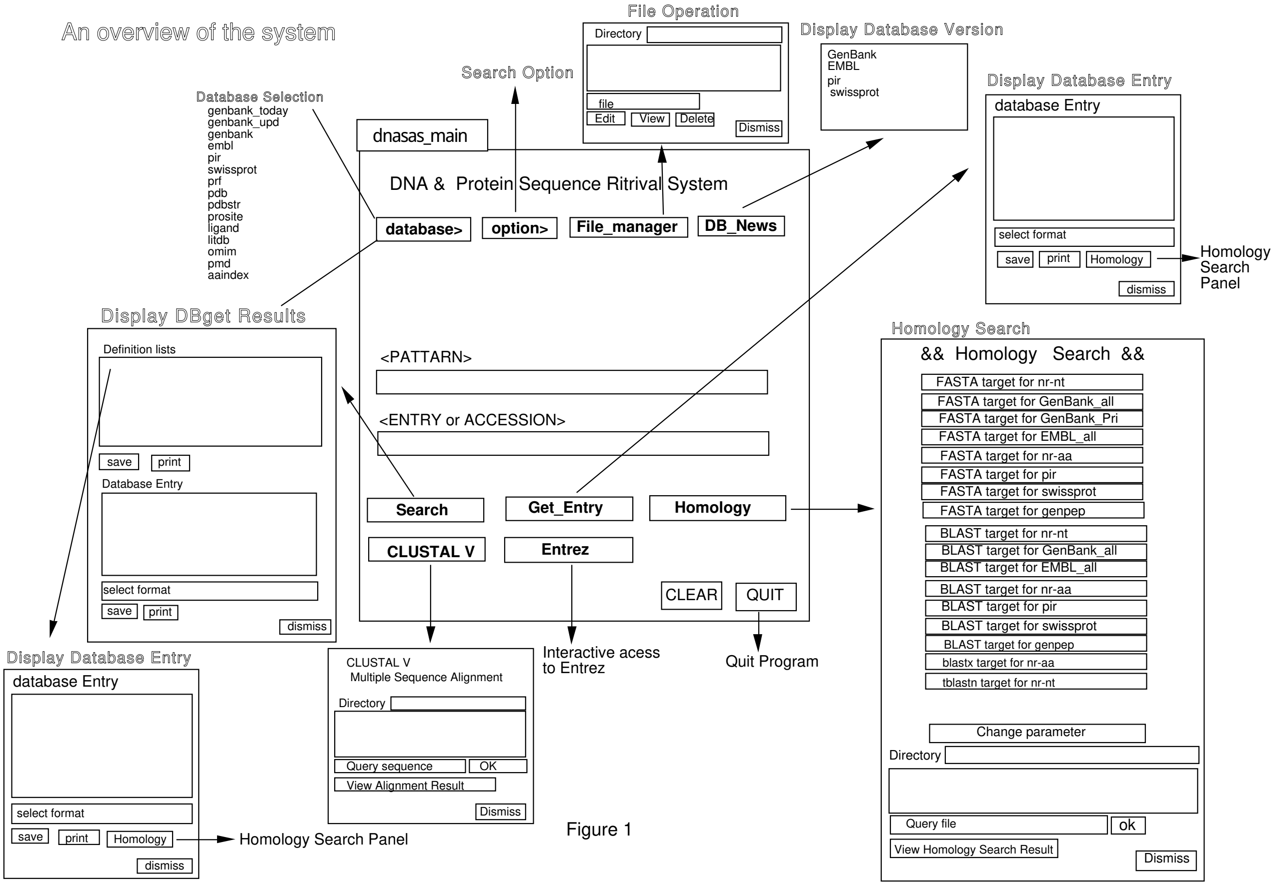

Figure 1

1 Function and graphical interface of the system

The system is easy to use because its functions are tightly coupled with its graphical user interface (see Figure 1 for an overview). The following functions are provided:

1. Keyword Searching
   The user can get the contents of entries by specifying an entry name or an accession number. Alternatively, he can obtain the list of entry names whose contents involve a given pattern.

2. Homology Searching
   The FASTA and BLAST programs are available.

3. Multiple Sequence Alignment
   The CLUSTAL V program is available.

4. Entrez
   The user can interactively access Entrez on the system.