

SEQANALREF: a sequence analysis bibliographic reference databank

Amos Bairoch

SEQANALREF is a bibliographic databank that stores the reference of articles from the expanding field of mathematical and computer analysis of biomolecular sequences. The majority of entries belong to one of the following categories:

Algorithms for protein and nucleic acid sequence analysis: primary, secondary and tertiary structure analysis; pattern matching; similarity searches; alignments, etc.

Algorithms for sequence-based phylogenetic analysis.

Description of biopolymer data banks: nucleic acid, protein, tertiary structure, carbohydrates, etc.

Description of software packages.

Description of on-line services for molecular biologists.

The format followed by this databank is a subject of that defined for the EMBL Nucleotide Sequence Data Library (1990) and SWISS-PROT protein sequence databanks (Bairoch, 1990). The line types currently used in the data bank are:

ID reference identifier
RA reference authors
RT reference title
RL reference location
KW keywords
CC comments
// separator line

Each ID line contains a unique identification code associated with the reference that it describes. The identification code is made of eight characters, in the following format:

AAAAYYNN

where 'AAAA' is an acronym generally made up from the first three letters and first initial of the name of the first author; 'YY' is the year of publication; and 'NN' is a serial number (starting with 01) which serves to differentiate otherwise identical identification codes.

An example of a SEQANALREF entry is given below:

ID JUNJ8401
RA Jungck J.R., Friedman R.M.;
RT 'Mathematical tools for molecular genetics data: an annotated bibliography';
RL Bull. Math. Biol. 46:699–744(1984).
KW NUCLEIC ACID; PROTEIN; BIBLIOGRAPHY.

The current release (release 14.0 of February, 1991) contains 1480 references.

SEQANALREF has been distributed since 1987 by various means and on various media. Currently, the most convenient sources for obtaining an up-to-date copy of this data bank are the following:

- (i) From the EMBL Data Library File Server (Stoehr and Omond, 1989). To obtain SEQANALREF you must send a message to the electronic mail address netserv@embl.bitnet. The message should contain the lines
GET REFLIST:SEQANALREF.DAT
GET REFLIST:SEQANALREF.DOC.
- (ii) On the GenBank On-line Service file FTP server (Benton, 1990). If your system is hooked to Internet you must FTP to the address genbank.bio.net (= 134.172.1.160). The SEQANALREF files are in the directory '/pub/db/seqanalref'.

References

- Bairoch A. (1990) *SWISS-PROT User Manual*, release 15.
EMBL Data Library Nucleotide Sequence Database User Manual, Release 23 of May 1990.
Stoehr, P.J. and Omond, R.A. (1989) *Nucleic Acids Res.*, **17**, 6763–6764
Benton, D. (1990) *Nucleic Acids Res.*, **18**, 1517–1520.

Circle No. 20 on Reader Enquiry Card