Cookies on EMBL-EBI website This website uses cookies to store a small amount of information on your computer, as part of the functioning of the site. Cookies used for the operation of the site have already been set.

To find out more about the cookies we use and how to delete them, see our <u>Cookie</u> and <u>Privacy</u> statements.

Sequence Similarity Searching

Tools > Sequence Similarity Searching

Sequence Similarity Searching is a method of searching sequence databases by using alignment to a query sequence. By statistically assessing how well database and query sequences match one can infer homology and transfer information to the query sequence.

The tools can be launched with different form pre-sets using the links - these can be changed on the tool page as well.

FASTA

FASTA

FASTA is another commonly used sequence similarity search tool which uses heuristics for fast **local** alignment searching.

▲ Protein ▲ Nucleotide ▲ Genomes
▲ Whole Genome Shotgun

SSEARCH

SSEARCH is an optimal (as opposed to heuristics-based) **local** alignment search tool using the Smith-Waterman algorithm. Optimal searches guarantee you find the best alignment score for your given parameters.

▲ Protein ▲ Nucleotide ▲ Genomes
▲ Whole Genome Shotgun

PSI-Search

PSI-Search combines the sensitivity of the Smith-Waterman search algorithm (SSEARCH) with the PSI-BLAST profile

BLAST

NCBI BLAST @

NCBI BLAST is the most commonly used sequence similarity search tool. It uses heuristics to perform fast **local** alignment searches.

🔦 Protein 🔌 Nucleotide 🔌 Vectors

WU-BLAST

WU-BLAST is similar to NCBI BLAST but combines multiple parameter options into a simpler 'sensitivity' setting.

🔦 Protein 🔌 Nucleotide

PSI-BLAST

PSI-BLAST allows users to construct and perform a BLAST search with a custom, position-specific, scoring matrix which can help find distant evolutionary relationships. PHI-BLAST functionality is also available to restrict results using patterns. construction strategy to find distantly related protein sequences.

🔦 Protein

GGSEARCH

GGSEARCH performs optimal **global-global** alignment searches using the Needleman-Wunsch algorithm.

🔦 Protein 🔌 Nucleotide

GLSEARCH

GLSEARCH performs an optimal sequence search using alignments that are **global** in the query but **local** in the database sequence. This can be useful when you want to match all of a short query sequence to part of a larger database sequence.

℁ Protein ℁ Nucleotide

FASTM/S/F ⁽²⁾

These specialist programs allow searches of databases using a group of short peptides as the query.

🔌 Protein 🔌 Nucleotide

The tools described on this page are provided using our **new bioinformatics analysis tools framework**. If you have any feedback or encounter any issues please let us know via <u>EMBL-EBI support</u>.

ENA Sequence Search

EMBL-EBI has a new nucleotide search tool which is far faster than BLAST for large datasets, with only a marginal loss in search sensitivity.

Try it out at ENA Sequence Search.