NCBI Cheat Sheet

by disha2020 via cheatography.com/147458/cs/32065/

Restricted searches

formula: term[Field] sometimes also: "term"[Field]

horse[Organism]

neoplasms[MeSH Terms]

* *MeSH (Medical Subject Headings)* see here

prolactin[Protein Name]

srcdb_refseq[Properties]

2010/06[Publication Date]

110:500[Sequence Length]

2015/3/1:2016/4/30[Publication Date]

NM_002299.4[Accession]

equidae[All Fields]

(Entrez) Searching options

Boolean operators:

AND (default operator) - finds documents that include both terms (example: Homo sapience[Organism] AND BRCA1[Gene Name])

OR - finds related documentation to either of terms (example: hemoglobin[protein] OR haemoglobin[protein])

NOT - finds information related to the left terms, but not to the right (example: promoters OR response elements NOT human AND mammals)

Truncation searching

NC_0000*[Accession] AND Human[Organism]

*" * " allows to search for unfinished term*

SPLIGN

Splign- is a utility for computing cDNA-to-Genomic

- identifies possible gene duplications

- recognizing introns, exones and splice signals

** intrones - non-coding regions of an RNA transcript*

** exone - coding regions of DNA* click here

Try SPLIGN here

BLAST (Basic Local Alignment Search Tool)

BLAST finds regions of similarity between biological sequences. The program compares nucleotide or protein sequences to sequence databases and calculates the statistical significance.

BLAST can be used to infer functional and evolutionary relationships between sequences as well as help identify members of gene families.

Try BLAST here

Conserved Domain Search

CD-Search is interface that searches through CDD (Conserved Domain Database) and returns an annotation of protein domains on the user query sequence

To compare two protein or nucleotide query, we can put them together in CD-Search. When results given we can check whether their domains have the same accession code.

Try CD-Search here

ORF finder

ORF findersearches for open reading frames (ORFs) in the DNA sequence you enter.

** open reading frames (ORFs) are defined as spans of DNA sequence between the start and stop codons (they include both introns and exons).*

The program returns the range of each ORF, along with its protein translation. Use ORF finder to search newly sequenced DNA for potential protein encoding segments, verify predicted protein using newly developed SMART BLAST or regular BLASTP (there are buttons to head to these interfaces).

The program returns the range of each ORF, along with its protein translation. Use ORF finder to search newly sequenced DNA for potential protein encoding segments, verify predicted protein using newly developed SMART BLAST or regular BLASTP (there are buttons to head to these interfaces).