



Study Guide to

Database Resources of the National Center for Biotechnology Information

Nucleic Acids Research, 2007, Vol.35, Database issue D5-D12

The NCBI home page web site is: _____.

The Entrez database includes ~ _____ DNA and protein sequences derived from several sources. It also includes over _____ gene-oriented sequence clusters in UniGene, _____ genetic variations in dbSNP and over _____ protein structures from the Molecular Modeling Database. How many online scientific textbooks does the database contain? _____

Results retrieved from an Entrez search may be sent to the Entrez _____ where they may be recalled later.

PubMed and PubMed Central

The PubMed database includes over _____ citations from greater than _____ life science journals for biomedical articles back to the 1950s.

PubMed Central is a digital archive of peer-reviewed journals in the life sciences providing access to greater than _____ - _____ articles! Participation in PMC requires a commitment to _____ to full text.

The NCBI taxonomy database is growing at the rate of _____ new taxa a month.

BLAST means:

(find this answer by looking at the 9th reference entry)

BLAST programs perform _____ similarity searches returning to the researcher a set of gapped alignments. Each returned alignment is scored and assigned a measure of statistical significance called the Expectation Value (_____). BLAST even takes into account the _____ composition of the query sequence in its estimation of statistical significance. A new _____ option

creates a dendrogram that clusters sequences according to their distances from the query sequence.

The database of single nucleotide polymorphisms (_____) is a repository for single base nucleotide _____ and short _____ and _____ polymorphisms. This database contains over _____ human SNPs and another _____ from a variety of other organisms.

Within the tools for gene-level analysis, the _____ (ORF) performs a six-frame translation of a nucleotide sequences.

An electronic PCR can be performed to estimate the genomic binding site, _____ and specificity for sets of primer pairs in specific organisms. This site is www.ncbi.nlm.nih.gov/sutils/e-pcr.

The Gene expression omnibus (_____) is a database for _____ and other forms of molecular data generated by the scientific community. The repository contains ~_____ individual measurements, derived from ~3000 array definitions, and spanning _____ organisms.

The Online Mendelian Inheritance in Man (OMIM) provides a catalog of human genes and _____. It contains information on disease _____ and genes. The database contains ~_____ entries.