

Database List, 04/05/12

NCBI GenBank: Leading database for peer-reviewed research for many conventional types of biological data (nt, AA, EST). Running BLAST against this database gives you a fair chance of identifying orthologous genes from a previously studied system.

<http://www.ncbi.nlm.nih.gov/genbank/>

Paleobiology Database (PBDB): Leading database for fossil record data. Molecular analyses characterizing branches in units of time should be consistent with the fossil record whenever possible.

<http://www.pbdb.org>

International Biogeography Society database list: A list of biogeographical, environmental, and phylogenetic databases for a wide variety of taxa.

<http://www.biogeography.org/html/Resources/databases.html>

American Society of Plant Taxonomists (ASPT) database list: A list of categorize plant-related databases, including genetics, morphologies, distributions, museums, etc.

<http://www.aspt.net/databases/>

DataDryad, TreeBASE: Data repositories for data from peer-reviewed studies – phylogenies, data alignments, etc. These data are useful for replicating results, or to incorporate into your study as evidence (i.e. to compare to your findings).

<http://www.datadryad.org>

<http://www.treebase.org>

2012 Nucleic Acid Research Summary Paper: An annual paper compiling molecular biology databases. The list contains 1380 databases (92 were added since 2011), which may be sorted categorically, alphabetically, or by paper abstract.

<http://www.oxfordjournals.org/nar/database/c/>

<http://nar.oxfordjournals.org/content/40/D1/D1.abstract>

Protein Databank (PDB): Database for experimentally determined protein structure. Say you identify some codon in an alignment is under positive selection. You find the ortholog of the gene is cataloged in PDB for some model organism. PDB may give you clues as to which amino acids are solvent accessible or functional, which may be used to corroborate with your phylogenetic evidence that your codon is under positive selection. PDB-101 provides a walkthrough for how to use resources.

<http://www.pdb.org>

Gene Ontology (GO): Database catalogs functions to genes. Evidence for function is associated with each entry (i.e. experimental, predicted, etc.). Say you sequence a gene of unknown function, and BLAST indicates your organism's gene is orthologous with some model organism's gene, GO may help you infer what the function of your organism's gene is.

http://amigo.geneontology.org/cgi-bin/amigo/browse.cgi?session_id=

modMine: Database for modENCODE – a project attempting to do a full characterization of all functional elements in model organisms' genomes (C. elegans, D. melanogaster). Accompanied by preliminary studies (general characterization of data). Good source of inspiration for studies in your system.

<http://intermine.modencode.org/>