

# Bioinformatics glossary based Database of Biological Databases: DBD

Siva Kiran RR, Setty MVN, Hanumatha Rao G\*

Received: 16 May 2009 / Received in revised form: 17 May 2009, Accepted: 18 May 2009, Published online: 6 June 2009  
© Sevas Educational Society 2008

## Abstract

Database of Biological/Bioinformatics Databases (DBD) is a collection of 1669 databases and online resources collected from NAR Database Summary Papers (<http://www.oxfordjournals.org/nar/database/a/>) & Internet search engines. The database has been developed based on 437 keywords (Glossary) available in <http://falcon.roswellpark.org/labweb/glossary.html>. Keywords with their relevant databases are arranged in alphabetic order which enables quick accession of databases by researchers. Database description provides brief information about the database with a link to main web page. DBD is available online and can be accessed at <http://www.biodbs.info>.

**Keywords:** Databases, biological databases, bioinformatics databases, bioinformatics glossary

## Introduction

Molecular biology has witnessed an information explosion in the last two decades as a consequence of development of rapid DNA sequencing techniques. The concomitant progress in computer based technologies has enabled molecular biologists in general and bioinformaticists in particular to cope with this information deluge. Information from various sources explicitly presented in the form of biological databases is generally maintained by many institutions. They vary widely in their content, format and access method [Kihler 2004]. Biological data are being deposited in a database and the need for data analysis has made molecular biology databases vital tools for research. Although several databases are reported in literature, many researchers are not aware of their availability and these databases were not effectively utilized (Ajay Babu et al. 2005).

**R.R. Siva Kiran, MVN Setty**

Department of Biotechnology, M S Ramaiah Institute of Technology, Bangalore – 560054, India

**G. Hanumatha Rao\***

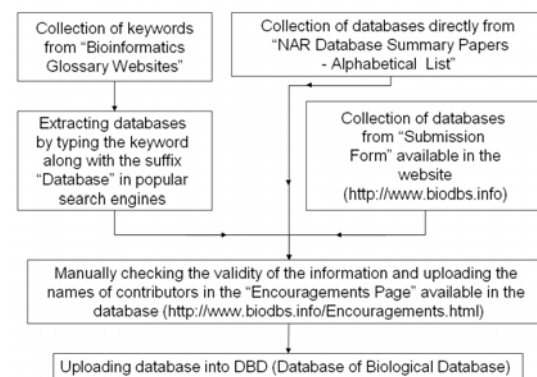
Center for Biotechnology, Department of Chemical Engineering, University College of Engineering, Andhra University, Visakhapatnam-530 003, India.

\*Tel: 0091 891 284 4878; Fax: 0091 891 274 7969  
Email: [ghr777au@gmail.com](mailto:ghr777au@gmail.com)

Databases of Biological Databases such as **DOD** – Database of Databases (<http://www.progenebio.in/DoD/DoD.htm>), **MetaBase** - The Database of Biological Database (<http://biodatabase.org>) and **The Molecular Biology Database Collection – Updates** (Baxevanis 2000), published annually by journal entitled “Nucleic Acids Research” help researchers to identify and correlate important queries beside providing a common platform for various molecular biology databases. Databases in DOD, Metabase & others have been grouped into categories as conceived by the respective authors. For instance, DOD has grouped 719 databases into 14 major categories (Galperin 2005) while Metabase has grouped 1119 databases into 21 categories.

The uniqueness of the present DBD is the alphabetic listing of all databases based on technical terms (keywords) viz. <http://falcon.roswellpark.org/labweb/glossary.html>. This facilitates very quick access to the desired databases.

**Figure 1:** Structure of DBD (Database of Biological Databases)



## Steps involved in construction of DBD (Figure 1)

- Extraction of suitable keywords from bioinformatics glossary website <http://falcon.roswellpark.org/labweb/glossary.html>.
- Typing the keyword along with the suffix “Database” in popular search engines like <http://www.google.com>, <http://www.yahoo.com> etc.

**Table 1:** Alphabetical listing of the number of keywords with the total number of databases linked to DBD (Database of Biological Database)

Alphabets	Total keywords	Total number of databases	Some major keywords with total number of database linked to DBD
A	16	64	Alignment – 8, Alternative Splicing – 22, Annotation – 12, Agricultural – 4
B	20	68	Bacterial – 8, Bacillus – 6, Binding – 6, Biochemical Pathway – 9, Biological databases listing - 11
C	48	88	Caenorhabditis elegans – 5, Cancer – 14, Chromosome – 9, Cytogenetics – 3, Cluster - 4
D	33	76	Disease – 12, Domain – 9, Drosophila -15, DNA – 18
E	25	86	<i>E.coli</i> – 14, Enzyme – 16, Eukaryotic – 8, EST – 11, Expression databases – 5, EBI – 3, Evolution - 3
F	12	19	Fish – 3, Fly -3, Functional Genomics - 4
G	15	77	Genome – 38, Gene & Gene Expression – 16, Genetics – 7, Germ - 6
H	25	93	HIV – 21, Homologous & Homology – 17, Human – 18, <i>Haemophilus</i> – 3, Higher taxon - 5
I	23	56	Inhibitor - 10, Intron – 6, Isotope - 4
J	Nil	Nil	Nil
K	4	4	Kinases – 1, KEGG - 1
L	9	32	Ligand – 23, Liver proteome - 1
M	41	258	Motif - 10, Metabolic Pathway - 11, Microarray - 22, Microbial - 15, Mitochondrial - 25, Mouse - 30, Mutation - 40
N	11	34	NMR - 5, Nucleic - 4, Nucleotide - 6
O	11	38	Oligo – 38, Organelles - 7, Ortholog - 9
P	52	297	Parasite - 10, Pathway - 12, PCR - 8, Peptide - 34, Phenotype - 9, Plant - 5, Plasmid - 10, Primer - 10, Probe - 13, Promoter – 15, Protease - 12, Protein - 53, Proteome - 18
Q	3	11	QSAR - 6, Quantitative trait loci - 4
R	19	113	Radiation hybrid - 6, Rat - 20, Retrovirus - 5, RNA - 39, Ribosomal – 7, Rice - 17
S	34	132	Sanger - 4, Sequence - 44, Signal Transduction - 8, SNP - 16, Species - 5, Structure - 13, Synapse - 3
T	19	47	Taxonomy - 3, Transcript - 5, tRNA - 3, Transgenic - 3, Tumor - 9, Tree - 3
U	3	7	Unigene - 4, UniProt - 2
V	5	31	Variation - 9, Virus - 16, Vector - 4
W	3	4	Worm - 2
X	3	4	X chromosome - 2, <i>Xenopus laevis</i> - 1
Y	2	25	Yeast - 23
Z	1	5	<i>Zea mays</i> - 5

For example genome keyword can be typed as “Genome Database” in search engine.

- Collection of databases from “NAR Database Summary Papers-Alphabetical List” available in the website (<http://www.oxfordjournals.org/nar/database/a/>)
- Listing all databases in tables along with suitable keywords.

**Figure 2:** Screenshot of Database of Biological Databases (DBD)



Keywords	Database Description	
<b>Abstract</b>	Enza Via A Database of Highly Structured and Complex Contents	<a href="http://www.enzavial.com/">http://www.enzavial.com/</a>
<b>Acidic</b>	A hierarchical database system for displaying genomic data (Genomikahitzya)	<a href="http://www.acidic.org/">http://www.acidic.org/</a>
<b>Activity</b>	Functional DNA/RNA site activity	<a href="http://www.mga.hawaii.edu/ndg/ndg.html">http://www.mga.hawaii.edu/ndg/ndg.html</a>
<b>Aspergillus</b>	Aspergillus genome database for the per. sp. (Aspergillus genome)	<a href="http://www.aspergillus.com/">http://www.aspergillus.com/</a>
<b>Asparagine</b>	Asparagine Database	<a href="http://www.asparagine.com/">http://www.asparagine.com/</a>
<b>Alignment</b>	Homologous Structure Alignment Database (HOMSTRAC)	<a href="http://www.crysl.bior.cam.ac.uk/homstrac/">http://www.crysl.bior.cam.ac.uk/homstrac/</a>
	Protein Alignment organized as Structural Superfamilies	<a href="http://www.ncbi.nlm.nih.gov/blast/">http://www.ncbi.nlm.nih.gov/blast/</a>
	EMBL/CCO Information for Database	<a href="http://www.ebi.ac.uk/EMBL/CCO/">http://www.ebi.ac.uk/EMBL/CCO/</a>
	Standard Alignment Database Homepage	<a href="http://www.ebi.ac.uk/EMBL/CCO/">http://www.ebi.ac.uk/EMBL/CCO/</a>
	BLAST/BLAST 2.0: A benchmark alignment database	<a href="http://www.ncbi.nlm.nih.gov/blast/">http://www.ncbi.nlm.nih.gov/blast/</a>
	The NCBI/BLAST 2.0 protein sequence alignment package (User Guide)	<a href="http://www.ncbi.nlm.nih.gov/blast/">http://www.ncbi.nlm.nih.gov/blast/</a>
	SAD: A specialized structural alignment database (mapping sequence-structure alignments)	<a href="http://portal.nrc.ca/ncic/ncic/ncic.html">http://portal.nrc.ca/ncic/ncic/ncic.html</a>
	A Multiple Alignment Database	<a href="http://www.ncbi.nlm.nih.gov/blast/">http://www.ncbi.nlm.nih.gov/blast/</a>
	COO: Cloned alignments of protein domains from Pfam, SMART and COO databases	<a href="http://www.ncbi.nlm.nih.gov/blast/">http://www.ncbi.nlm.nih.gov/blast/</a>
	CE: Combinatorial Extension method to compute and refine 3D protein structure alignments	<a href="http://www.ncbi.nlm.nih.gov/blast/">http://www.ncbi.nlm.nih.gov/blast/</a>
<b>Alkyl</b>	Alkyl Degradation and DNA polymerase	<a href="http://www.ncbi.nlm.nih.gov/blast/">http://www.ncbi.nlm.nih.gov/blast/</a>
<b>Alternative Splicing</b>	ASDB: Alternative Splicing Database	<a href="http://www.ncbi.nlm.nih.gov/blast/">http://www.ncbi.nlm.nih.gov/blast/</a>
	EDAS: EST Derived Alternative Splicing Database	<a href="http://www.ncbi.nlm.nih.gov/blast/">http://www.ncbi.nlm.nih.gov/blast/</a>

## Features of DBD:

All keywords are arranged in alphabetic order as shown in Figure 2. Databases are listed along with the keywords (Table 1). The description of database has been reduced so as to accommodate more databases with keywords. Websites pertaining to bioinformatics keywords like Alignment, SQL, Base Pair, Blast, Biotechnology etc are also listed for helping researchers to reduce their time in finding suitable information. Tutorial websites are also

listed along with databases. The listed databases or tutorial websites are based solely on the display of the relevant websites present in search engines.

## Conclusion

The primary aim of DBD is to provide an easy access to a number of databases, tutorial websites etc. based on keywords. Brief description about the databases and keyword helps the users to know the contents of the database, their respective keywords and database links. The DBD shall be updated periodically for the benefit of researchers and more bioinformatics/biology keywords will also be added to the database. We are confident, that the DBD will be highly useful tool in the hands of molecular biologists and bioinformaticians.

## Acknowledgement

The authors thank Rajesh K of Department of Pharmacy, Andhra University, India, Rohit MD, Mubashir HM, Yomeen RM, of Dept. of Biotechnology, MS Ramaiah Institute of Technology, Bangalore, India and Sangeeta Uranakar, Naganandini G and Usha D of University Visweswaraiiah College of Engineering, K R Circle, Bangalore, India for their contributions in the development of this database.

## References

Attwood TK, Parry-Smith DJ (2000) Introduction to Bioinformatics, 1<sup>st</sup> edn. Prentice Hall, UK

Kihler J (2004) Integration of life science databases. DDT: BIOSILICO 2:61-69

Ajay BP, Radha B, Lakshmi VV, Rao NG (2005) DoD: Database of Databases - updated molecular biology databases. In Silico Biology 5:0054 </isb/2005/05/0054>

Baxevanis AD (2000) The Molecular Biology Database Collection: an online compilation of relevant database resources. Nuc Acids Res 28(1):1-7

Galperin MY (2005) The Molecular Biology Database Collection: 2005 update

Nuc Acids Res. doi:10.1093/nar/gki139