

La bioinformatica include:

- Database che raccolgono dati sperimentali prodotti nei laboratori di ricerca
- Software per la navigazione in banche dati

Un database biologico è un insieme organizzato di informazioni e dati che provengono da studi in laboratori di ricerca (analisi sia in *vitro* che in *vivo*), da centri di bioinformatica (analisi in *silico*) e da pubblicazioni scientifiche.

I Database sono costituiti da “entries”.

I dati sono strutturati e organizzati in modo da essere consultabili e utilizzabili nel miglior modo possibile.

I Database di primo livello raccolgono sequenze nucleotidiche (DNA, RNA) o proteiche e rendono disponibili informazioni per identificare le specie da cui hanno origine le sequenze e le loro funzioni.

I Database specializzati raccolgono informazioni più specifiche sulla tassonomia, le funzioni, le pubblicazioni scientifiche, le malattie correlate alle mutazioni, delle sequenze nucleotidiche.

Ecco una lista dei Database utili per i laboratori di bioinformatica (puoi copiare gli indirizzi tra i preferiti).

<http://www.ensembl.org>

*Ensembl* (il nome ricorda la parola francese “*ensemble*” e al contempo “EMBL” *European Molecular Biology Laboratory*) è un database nato dalla collaborazione dell’ EMBL - European Bioinformatics Institute (EBI) e il Wellcome Trust Sanger Institute (WTSI) per sviluppare un sistema di software che produce e gestisce in modo automatico le annotazioni su alcuni genomi eucariotici.

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

NCBI (National Biotechnology Institute, NIH) ha creato un database pubblico e ha sviluppato software per analizzare i dati del genoma.

<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=OMIM>

OMIM, Online Mendelian Inheritance in Man, è una raccolta dei geni umani e delle malattie ad essi correlate. Costruito e pubblicato da Dr. Victor A. McKusick e dal suo collega Johns Hopkins, e sviluppato per internet dall’NCBI, the National Center for Biotechnology Information. Il database contiene testi e pubblicazioni scientifiche sulle malattie genetiche e molti link a MEDLINE e al database di sequenze Entrez, oltre che ad altre risorse presenti nell’NCBI.

<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?CMD=search&DB=pubmed>

PubMed, raggiungibile dal sistema dell’ NCBI Entrez è stato sviluppato dal National Center for Biotechnology Information (NCBI) e dal National Library of Medicine (NLM), che si trova nel U.S. National Institutes of Health (NIH). PubMed permette l’accesso a informazioni bibliografiche scientifiche e mediche.

<http://www.expasy.org/sprot/sprot-top.html>

<http://www.uniprot.org/>

Swiss Prot è un database di sequenze proteiche che contiene un gran numero di annotazioni (come la descrizione della funzione di una proteina, i suoi domini, le modificazioni post-trascrizionali, le varianti,...), un livello molto basso di ridondanza e un alto livello di integrazione con gli altri database biologici.