Introduction to Bioinformatics and relations to other fields

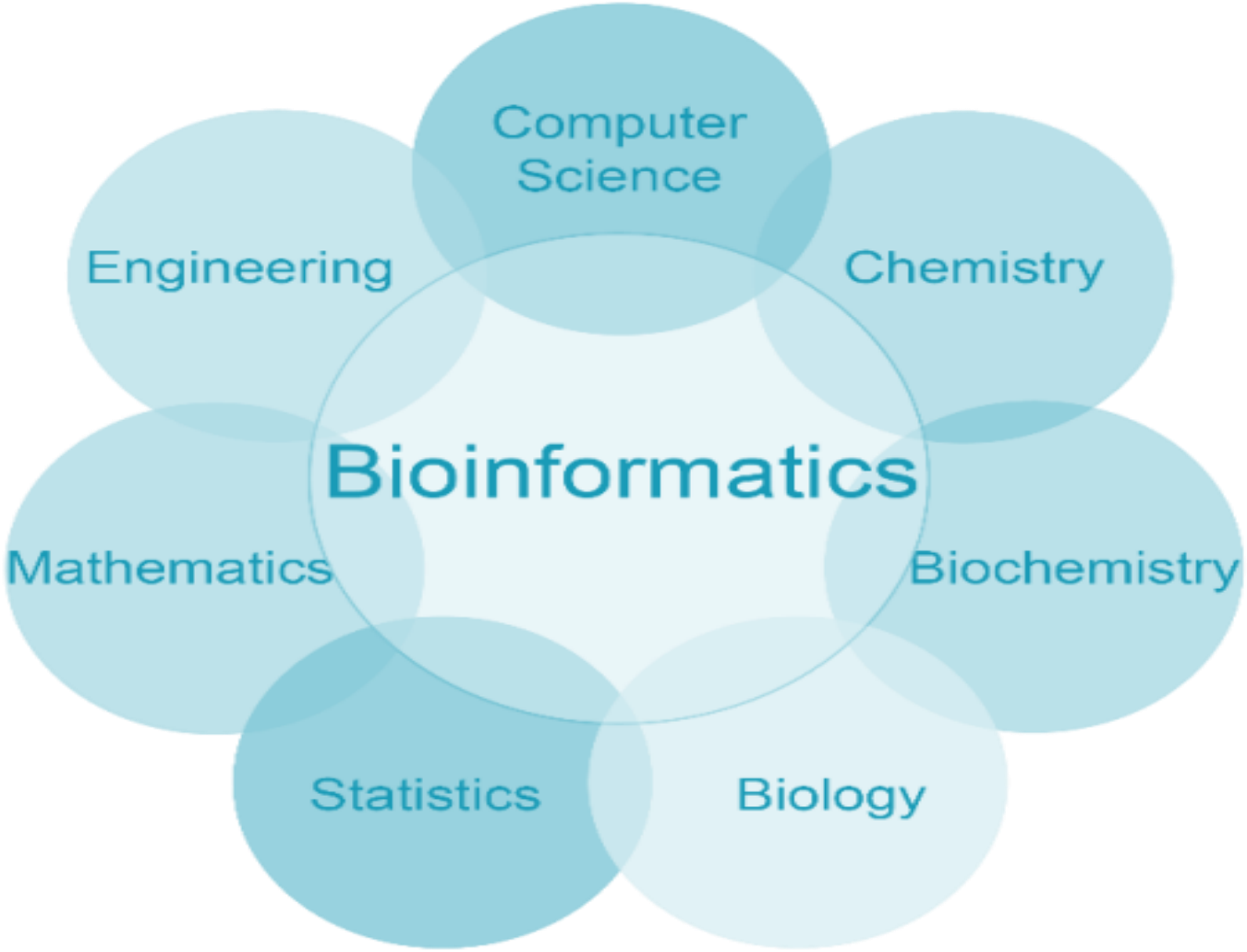
م.م. شيرين نزار قاسم
كلية المعلوماتية الطبية الحيوية
قسم الانظمة الطبية الذكية

DEFINITION OF BIOINFORMATICS

A field of science that uses computers, databases, math, and statistics to collect, store, organize, and analyse large amounts of biological, medical, and health information.

Its deals with DNA microarrays, comparative, functional and structural genomics, proteomics & medical information's

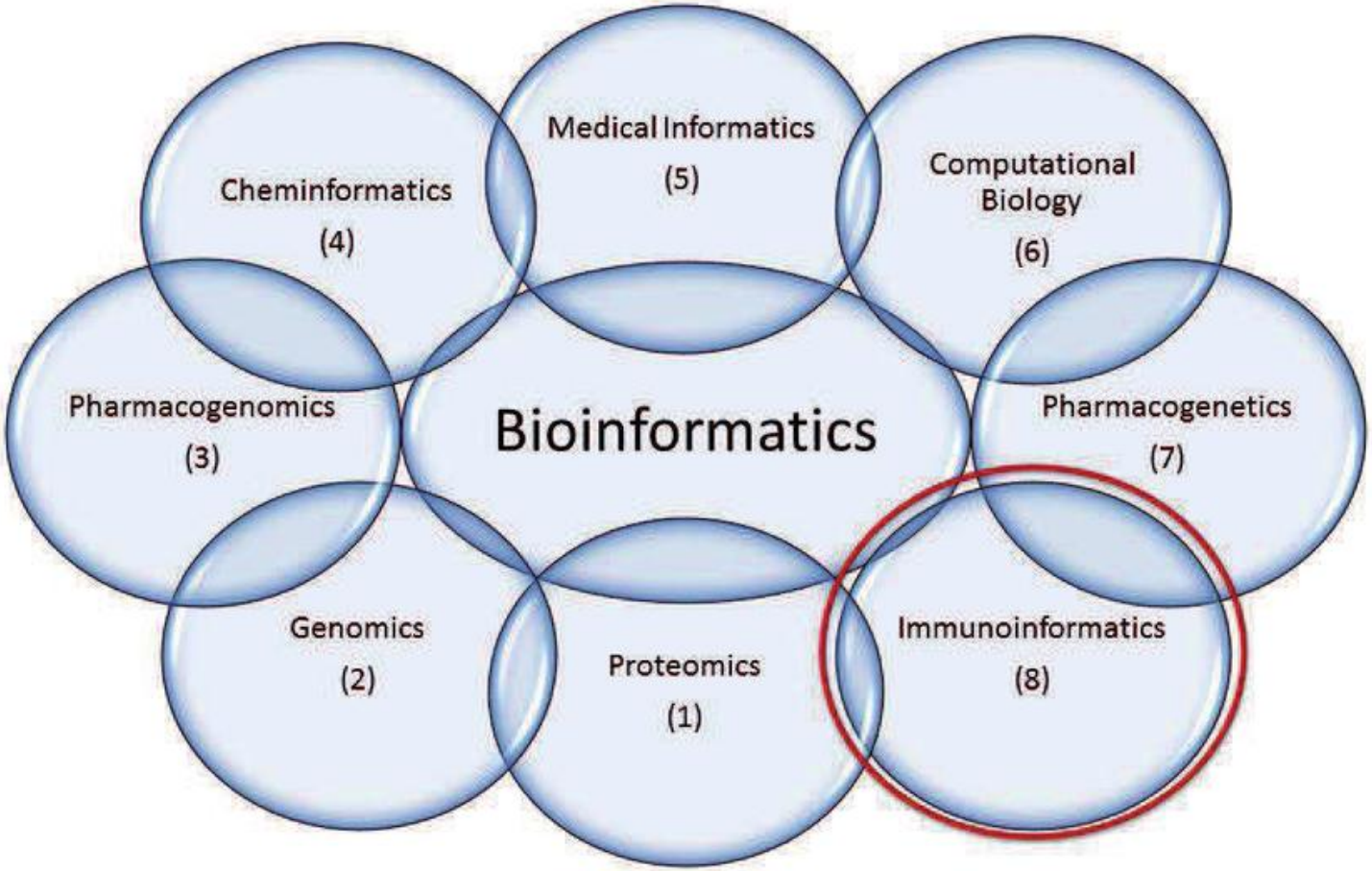
Interaction with other fields



METHODS OF PERFORMING EXPERIMENTS:

- before the era of bioinformatics, only two ways of performing biological experiments were available: within a living organism (**so- called in vivo**) or in an artificial environment (**so-called in vitro**) from the Latin in glass.
- Taking the analogy further, we can say that bioinformatics is in fact **in silico** biology, from the silicon chips on which microprocessors are built.

FIELDS RELATED TO BIOINFORMATICS



FIELDS RELATED TO BIOINFORMATICS

- **What is Biophysics?**
- An interdisciplinary field which applies techniques from the physical sciences to understanding biological structure and function”
- **What is Cheminformatics?**
- "the combination of chemical synthesis, biological screening, and data-mining approaches used to guide drug discovery and development"
- **What is Computational Biology?**
- Computational biology is not a "field", but an "approach" involving the use of computers to study biological processes and hence it is an area as diverse as biology itself.“

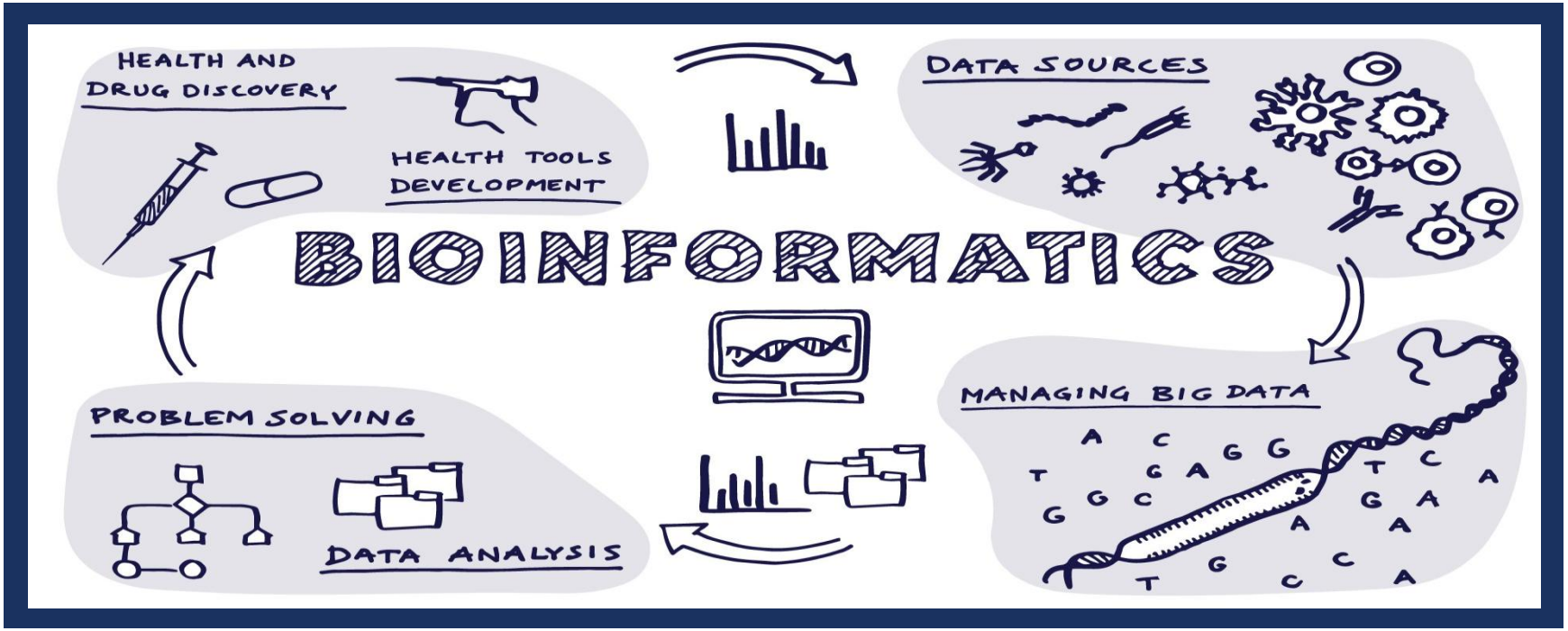
FIELDS RELATED TO BIOINFORMATICS

- **What is Genomics?**
- "Genomics is any attempt to analyze or compare the entire genetic complement of a species or species (plural)".
- **What is Medical informatics/ Med informatics?**
- "Study, invention, and implementation of structures and algorithms to improve communication, understanding and management of medical information.“
- **What is Pharmacogenomics?**
- "Pharmacogenomics is the application of genomic approaches and technologies to the identification of drug targets".

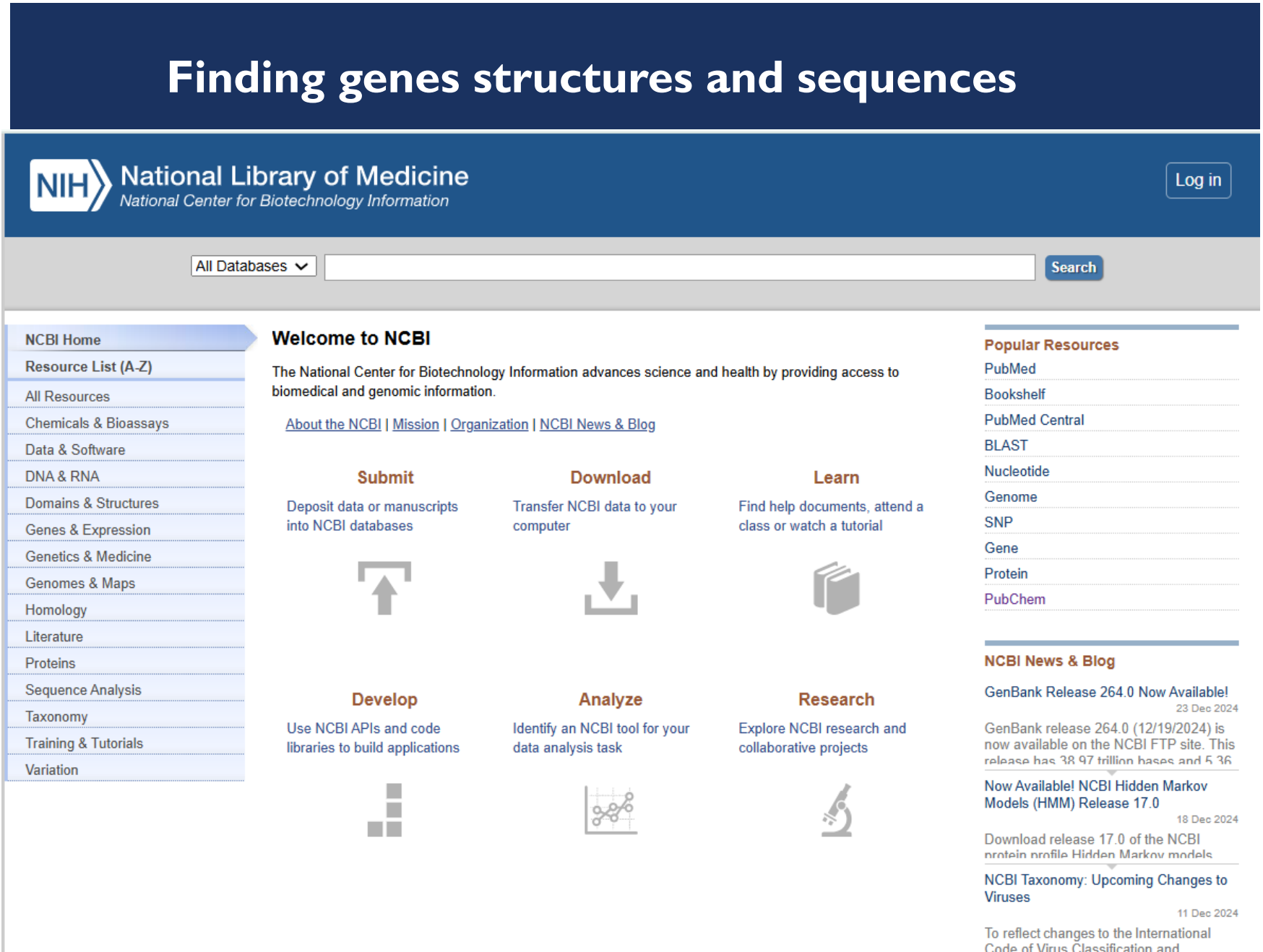
BIOINFORMATICS HISTORY

- The father and mother of Bioinformatics, Dr. Margaret Belle Dayhoff (March 11, 1925 – February 5, 1983) was an American biophysicist and a pioneer in the field of bioinformatics. She engineered the application of mathematics and computational methods to the field of biochemistry.
- Her work of collection the protein sequence **(Dayhoff's Atlas of Protein Sequence and Structure, 1954-1965)**
- he EMBL established their data library in 1980 to collect, organize and distribute nucleotide sequence data and related information.
- • NCBI was established in U.S.A. NCBI serves as primary information databank and provider of information.
- • The National Biomedical Research Foundation established the PIR in 1984.

Some Applications for Bioinformatics



Finding genes structures and sequences



Search NCBI databases

Help

homo sapiens tpo Search

Results found in 28 databases for "homo sapiens tpo"

Literature

Books	68	books and reports
MeSH	0	ontology used for PubMed indexing
NLM Catalog	1	books, journals and more in the NLM Collections
PubMed	3,478	scientific & medical abstracts/citations
PubMed Central	3,790	full-text journal articles

Health

ClinVar	0	human variations of clinical significance
dbGaP	0	genotype/phenotype interaction studies
GTR	0	genetic testing registry
MedGen	0	medical genetics literature and links
OMIM	0	online mendelian inheritance in man
PubMed Health	7	clinical effectiveness, disease and drug reports

Genomes

Genes

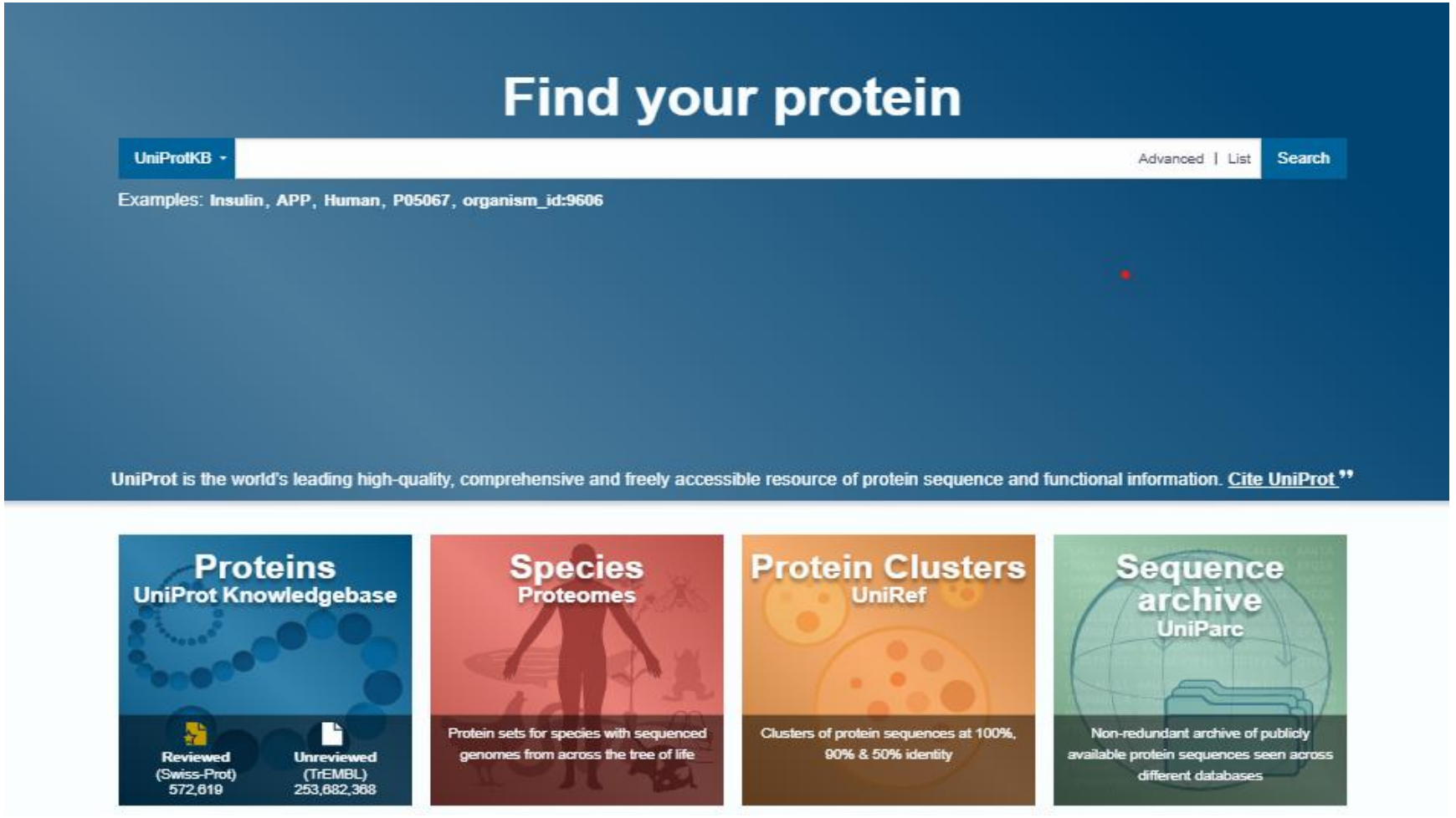
EST	2	expressed sequence tag sequences
Gene	49	collected information about gene loci
GEO DataSets	1,230	functional genomics studies
GEO Profiles	50,081	gene expression and molecular abundance profiles
HomoloGene	3	homologous gene sets for selected organisms
PopSet	1	sequence sets from phylogenetic and population studies
UniGene	4	clusters of expressed transcripts

Proteins

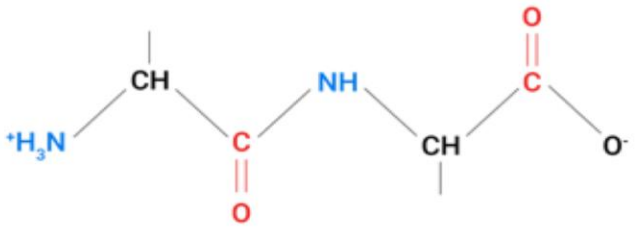
Conserved Domains	0	conserved protein domains
Protein	752	protein sequences
Protein Clusters	14	sequence similarity-based protein clusters
Structure	21	experimentally-determined biomolecular structures

Finding protein structures and sequences

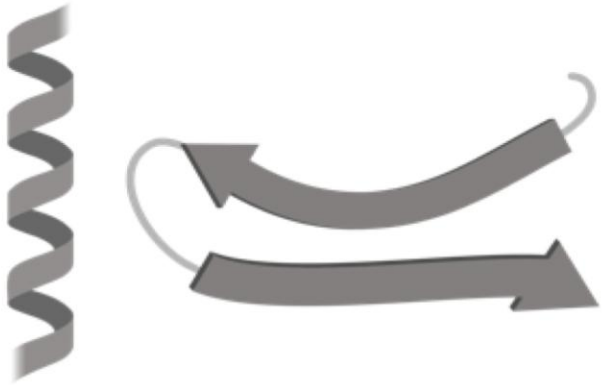
UniProt is the world's leading high-quality, comprehensive and freely accessible resource of protein sequence and functional information.



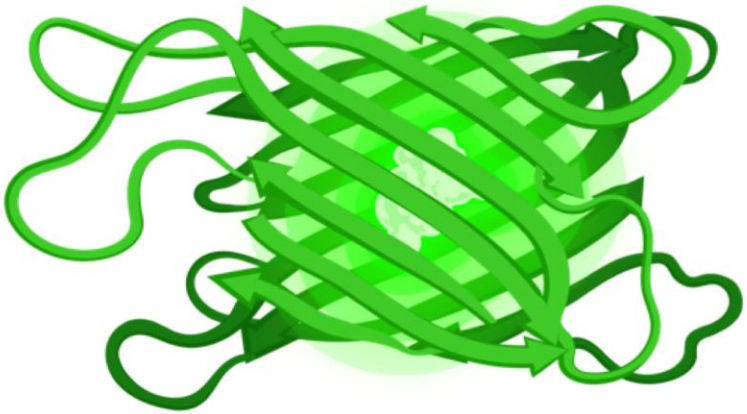
Primary



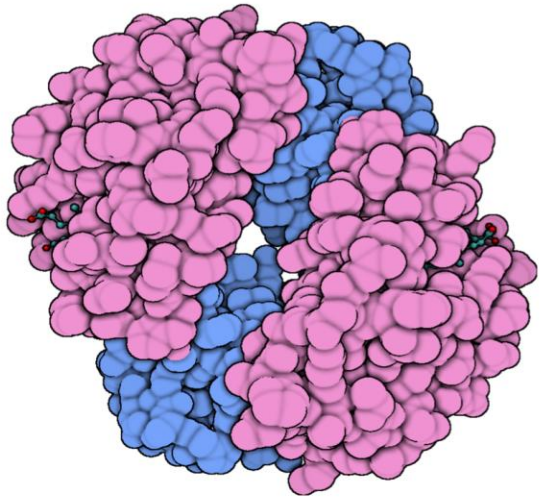
Secondary

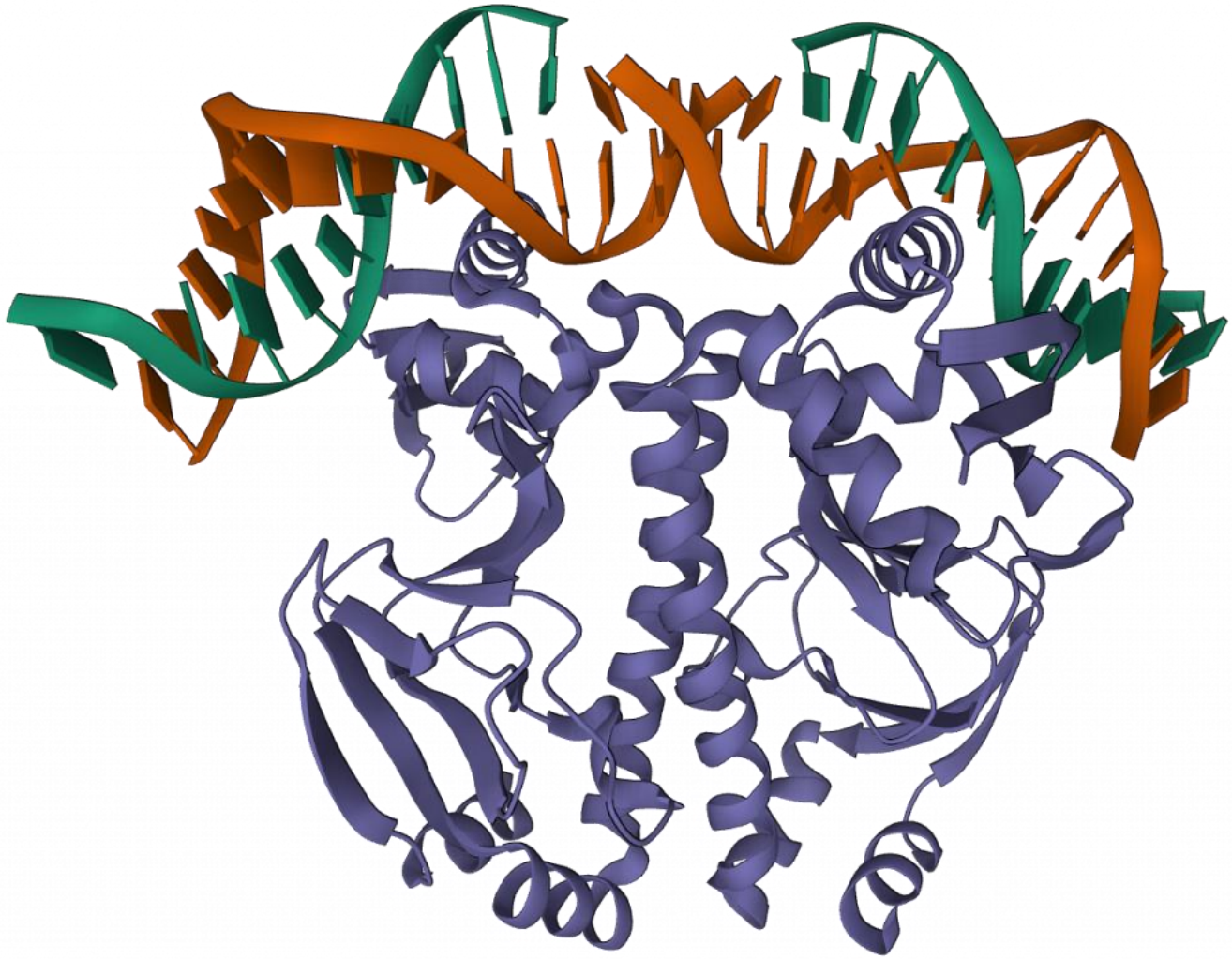


Tertiary



Quaternary





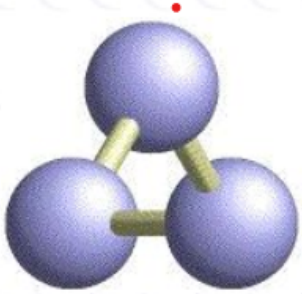
RasMol is a important scientific tool for molecular graphics visualisation

Copying and Distribution	Contents	Software Distributions	Latest Windows Installer	External Packages
RasMol Manual	RasMol Blog	Frequently Asked Questions	RasMol 2.7 Series History	RasMol and OpenRasMol
SourceForge OpenRasMol Site	Click Here to Make a Donation	RasMol SourceForge Site

Home Page for RasMol and OpenRasMol

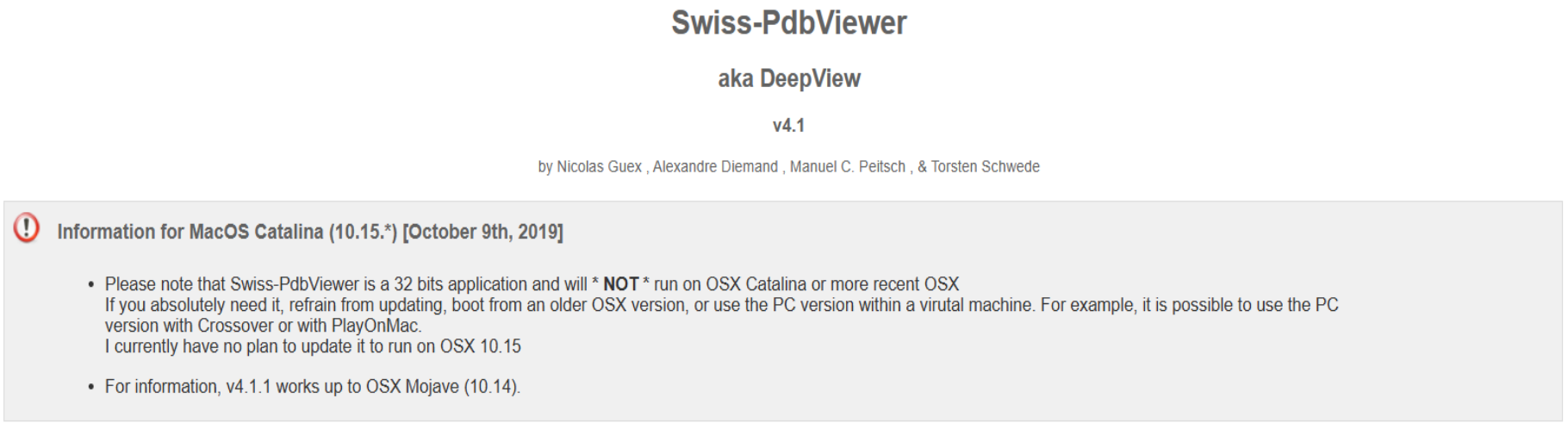
Molecular Graphics Visualisation Tool

- RasMol Latest Windows Installer
- RasMol Latest Source Tarball
- RasMol Latest Manual
- Donate to Support RasMol
- Register your RasMol

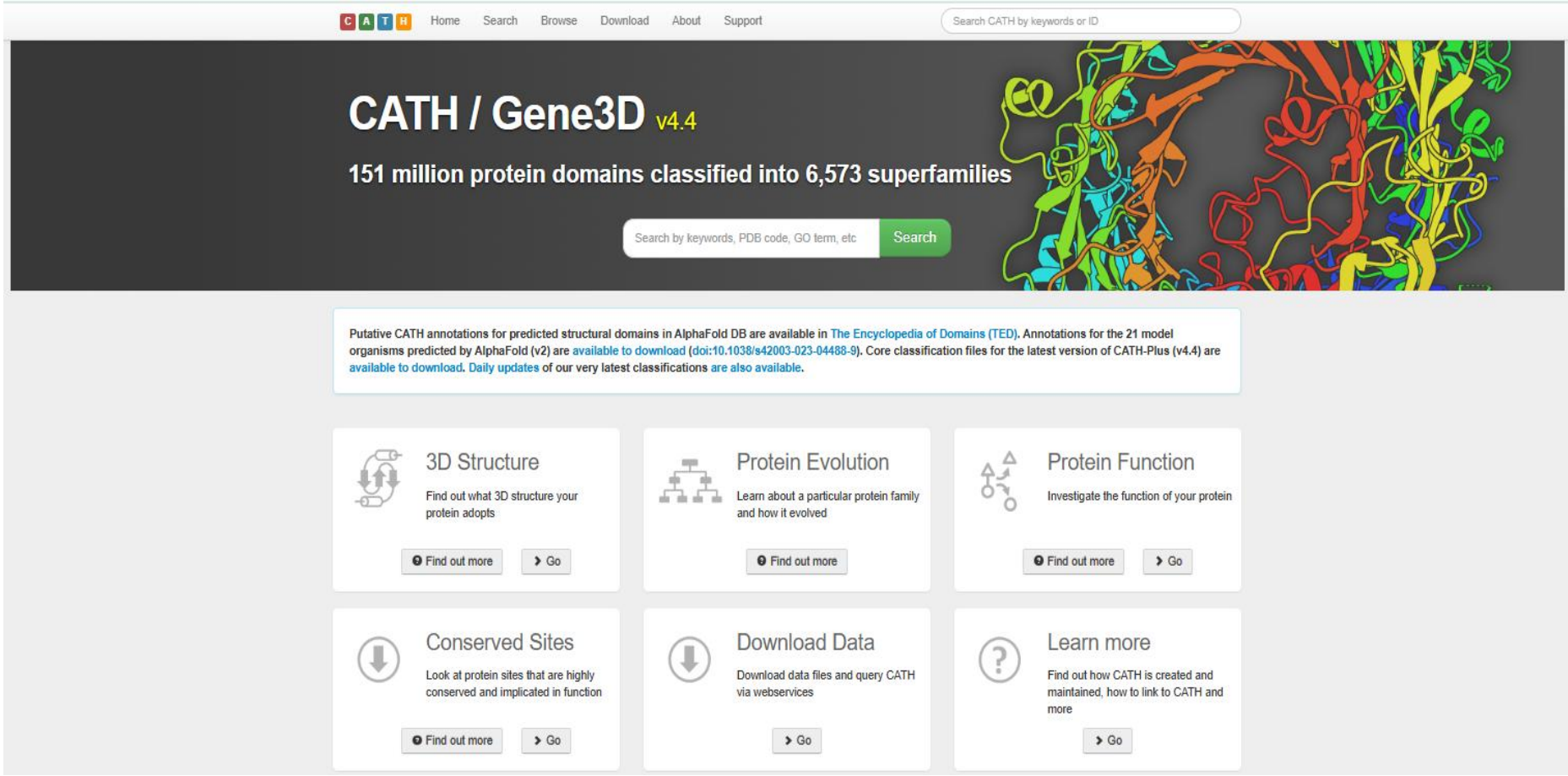


- RasMol 2.7.5 Windows Installer
- RasMol 2.7.5 Source Tarball
- RasMol 2.7.5 Manual
- Donate to Support RasMol
- Register your RasMol

Swiss-PdbViewer (aka DeepView) is an application that provides a user friendly interface allowing to analyse several proteins at the same time. The proteins can be superimposed in order to deduce structural alignments and compare their active sites or any other relevant parts. Amino acid mutations, H-bonds, angles and distances between atoms are easy to obtain thanks to the intuitive graphic and menu interface.



CATH is a classification of protein structures downloaded from the Protein Data Bank.



The Dali server is a network service for comparing protein structures in 3D

The Dali server is a network service for comparing protein structures in 3D. You submit the coordinates of a query protein structure and Dali compares them against those in the Protein Data Bank (PDB). In favourable cases, comparing 3D structures may reveal biologically interesting similarities that are not detectable by comparing sequences.

Check queue status here. Megausers please consider downloading the standalone program.

You can perform three types of database searches:

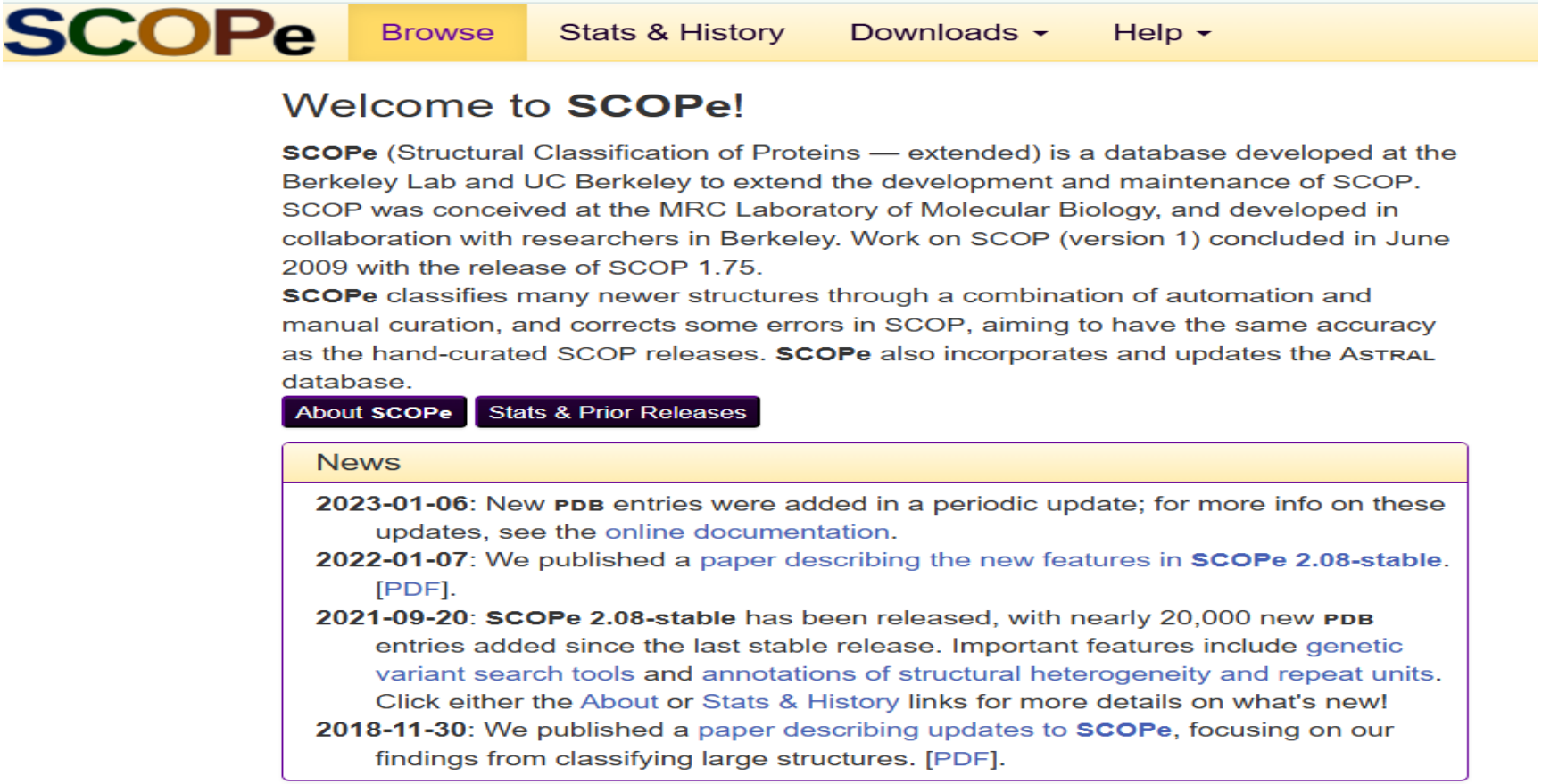
- Heuristic **PDB search** - compares one query structure against those in the Protein Data Bank
- Exhaustive **PDB25** search - compares one query structure against a representative subset of the Protein Data Bank
- Hierarchical **AF-DB** search - compares one query structure against a species subset of the AlphaFold Database v2

- beta testing:AF-DB heuristic - search AlphaFold Database v2 in minutes

and two types of structure comparisons of user selected structures:

- **Pairwise** structure comparison - compares one query structure against those specified by the user
- **All against all** structure comparison - returns a structural similarity dendrogram for a set of structures specified by the user

SCOPe is a database for (Structural Classification of Proteins — extended)



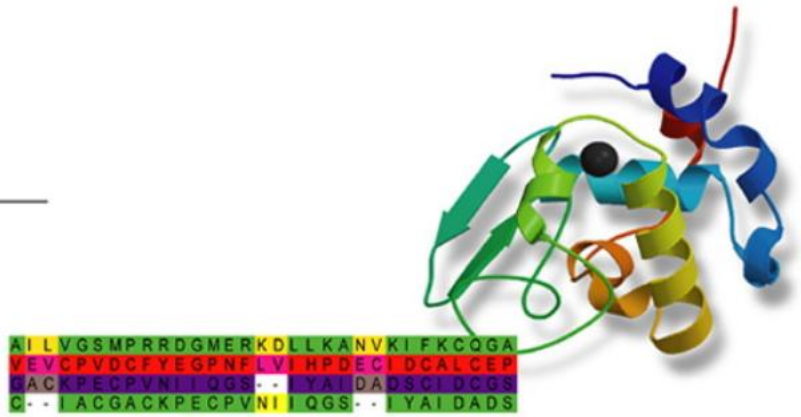
MODELLER is used for homology or comparative modelling of protein three-dimensional structures

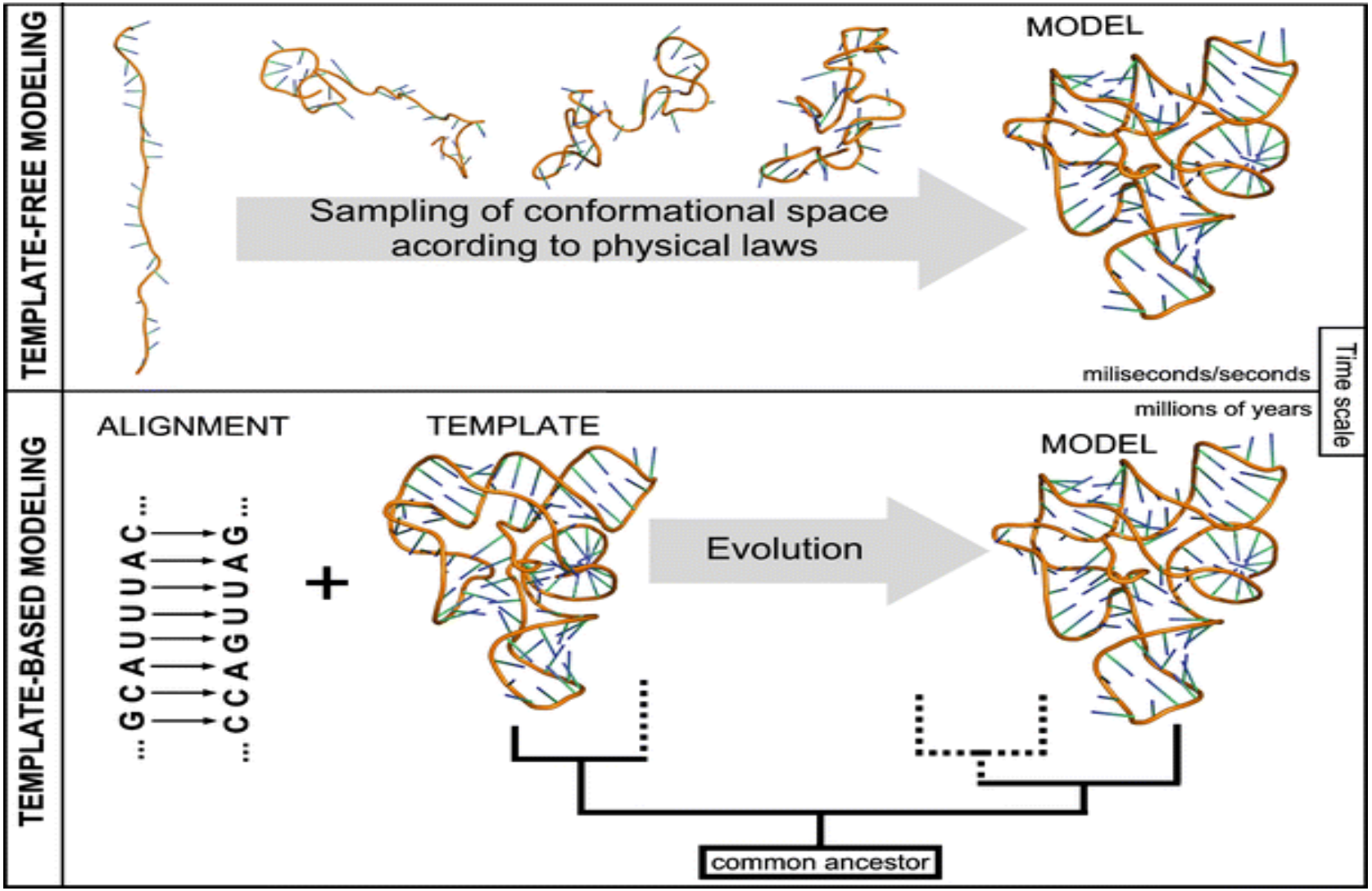
The user provides an alignment of a sequence to be modelled with known related structures and MODELLER automatically calculates a model containing all non-hydrogen atoms.

MODELLER implements comparative protein structure modelling by satisfaction of spatial restraints and can perform many additional tasks, including de novo modelling of loops in protein structures, optimization of various models of protein structure with respect to a flexibly defined objective function, multiple alignment of protein sequences and/or structures, clustering, searching of sequence databases, comparison of protein structures, etc

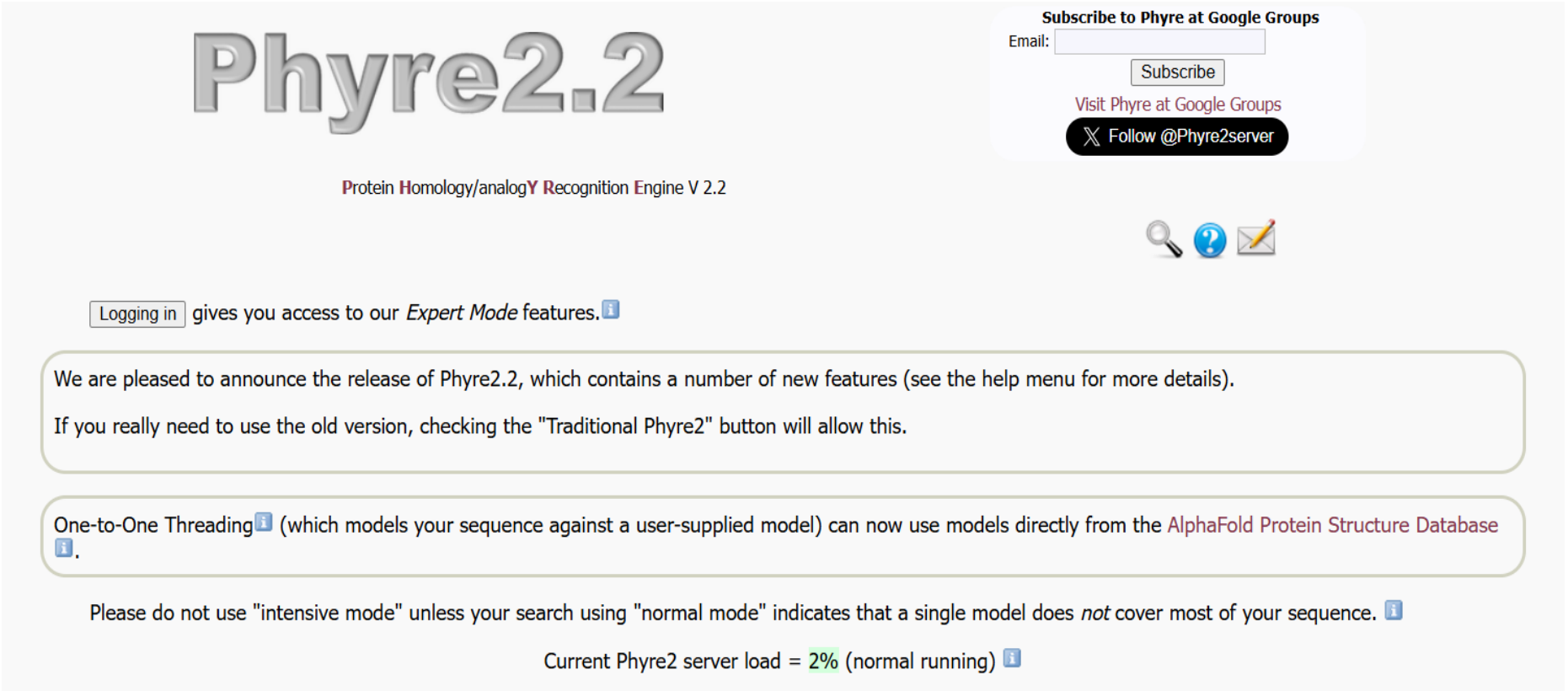
Modeller

Program for Comparative Protein Structure Modelling by Satisfaction of Spatial Restraints

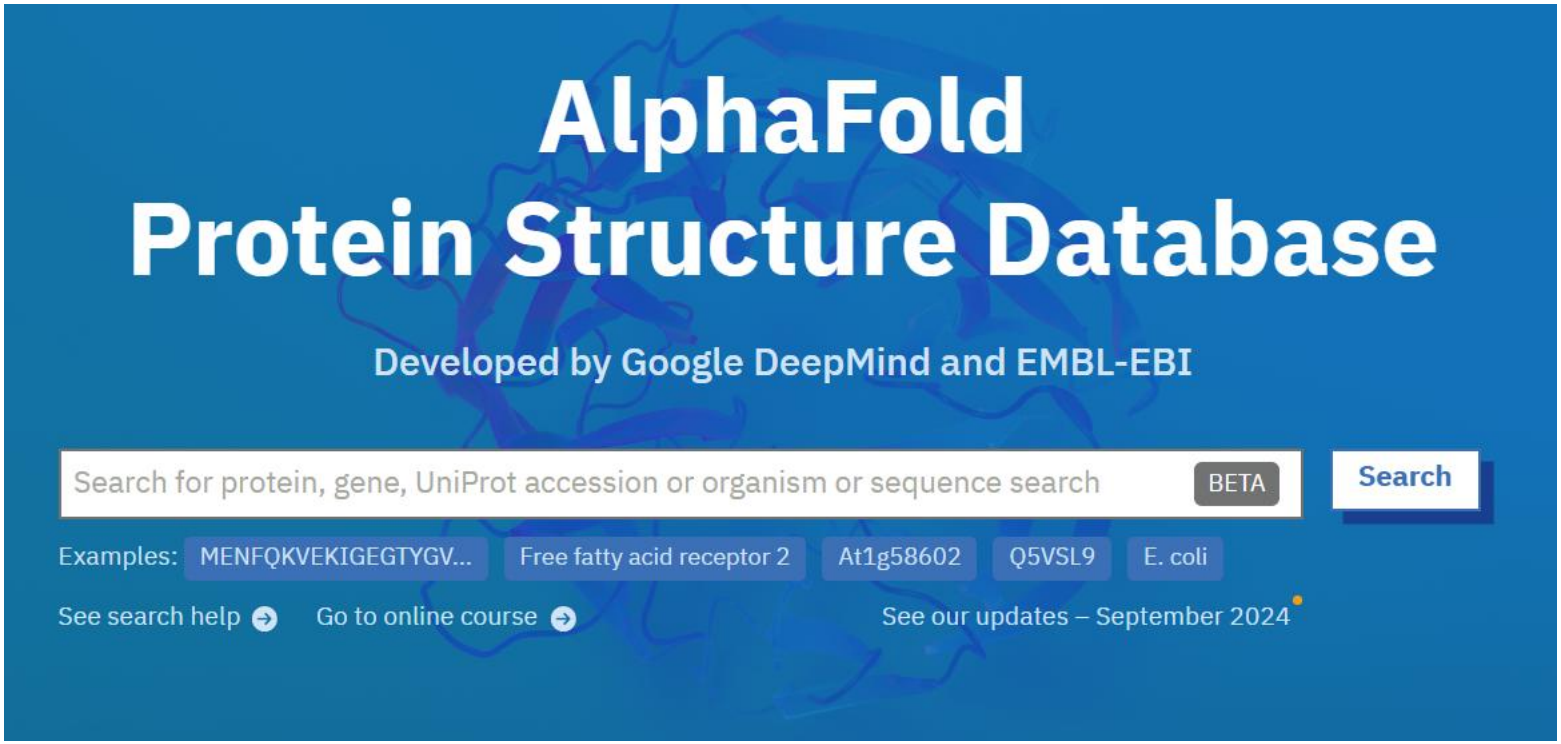



TEMPLATE-FREE MODELING
MODEL
Sampling of conformational space
acording to physical laws
miliseconds/seconds
Time scale
millions of years
TEMPLATE-BASED MODELING
ALIGNMENT
...GCAUUUAC...
...CCAGUUAG...
+
TEMPLATE
Evolution
MODEL
common ancestor

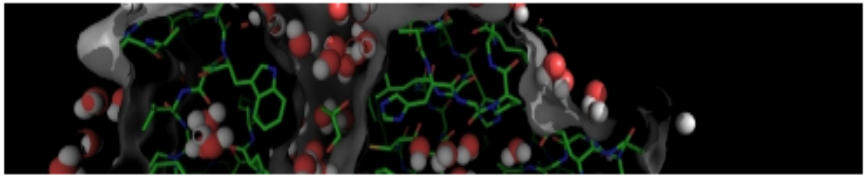
Protein Homology/analogy Recognition Engine



Alphafold Is An AI System Developed By Google Deepmind That Predicts A Protein's 3D Structure From Its Amino Acid Sequence. It Regularly Achieves Accuracy Competitive With Experiment.



ELNÉMO Allows You To View 3-D Animations Of The Protein Movement



home | start a new run | job status | references&downloads | examples | help

Should you encounter any unexpected behaviour,
please let us know.

Welcome to *elNémo* !

elNémo is the Web-interface to the *Elastic Network Model* (ENM), a fast and simple way for computing the low frequency normal modes of a macromolecule (Tirion, 1996). Note that, thanks to the RTB approximation (Durand et al., 1994; Tama et al., 2000), this server can perform calculations for all-atom systems.

Predicting interactions (molecular docking)

Two different protein molecules (or a protein and a small molecule) might interact with each other. This is technically referred to as molecular docking.

Prediction of Protein 3D structure

0101#01001010#10111010#
01010001#10010#1001#101
10010#1001001001011#0

Residue	ϕ	ψ	ω
THR	0.0	147.7	172.9
THR	107.2	-125.3	187.4
CYS	123.4	63.6	103.7
PRO	60.3	83.9	-116.7

DNA → **Algorithm** → **Protein Model**

Natural Evolution →

page discussion view source history

Biopython

(Redirected from Main Page)

Introduction

Biopython is a set of freely available tools for biological computation written in Python by an international team of developers.

It is a distributed collaborative effort to develop Python libraries and applications which address the needs of current and future work in bioinformatics. The source code is made available under the Biopython License, which is extremely liberal and compatible with almost every license in the world. We work along with the Open Bioinformatics Foundation, who generously host our website, bug tracker, and mailing lists.

This wiki will help you download and install Biopython, and start using the libraries and tools.

Get Started

- Download Biopython
- Installation help (PDF)

Get help

- Tutorial (PDF)
- Documentation on this wiki
- Cookbook (working examples)
- Discuss and ask questions

Contribute

- What's being worked on
- Developing on Github
- Google Summer of Code

The latest release is Biopython 1.55 beta, released on 10 August 2010.

OBF News

- Biopython 1.55 beta released
- Biopython 1.54 released
- OBF Google Summer of Code Accepted Students
- Illumina FASTQ files 1 Read Segment Quality Control Indicator
- Partial sequence files with Biopython
- Making Biopython SeqIO and AlignIO easier
- Biopython 1.54 beta released
- OBF in Google Summer of Code
- Sanger FASTQ format and the Solexa/Illumina variants
- Biopython 1.53 released

See also our news page, and twitter.

navigation

- Main Page
- Downloads
- Mailing lists
- Documentation
- Cookbook
- News
- Bugs
- Source Code
- Participants
- Script Central
- Recent changes
- Random page

search

Go Search

HOMEWORK

As seen in Slide -3- the interaction of bioinformatics with other fields, write a short assay about one of them