



Bioinformatics

Applications

Amandeep Singh

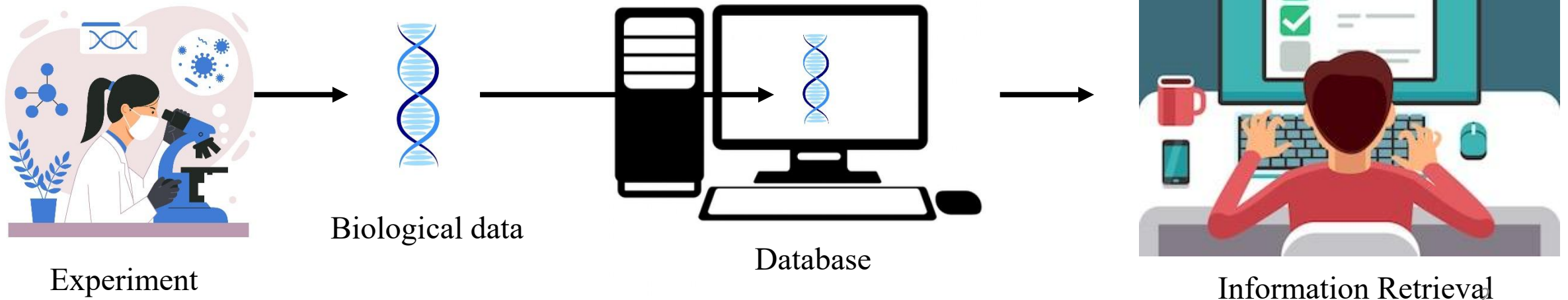
“Exploring Science, Inspiring Learning”

Applications of Bioinformatics

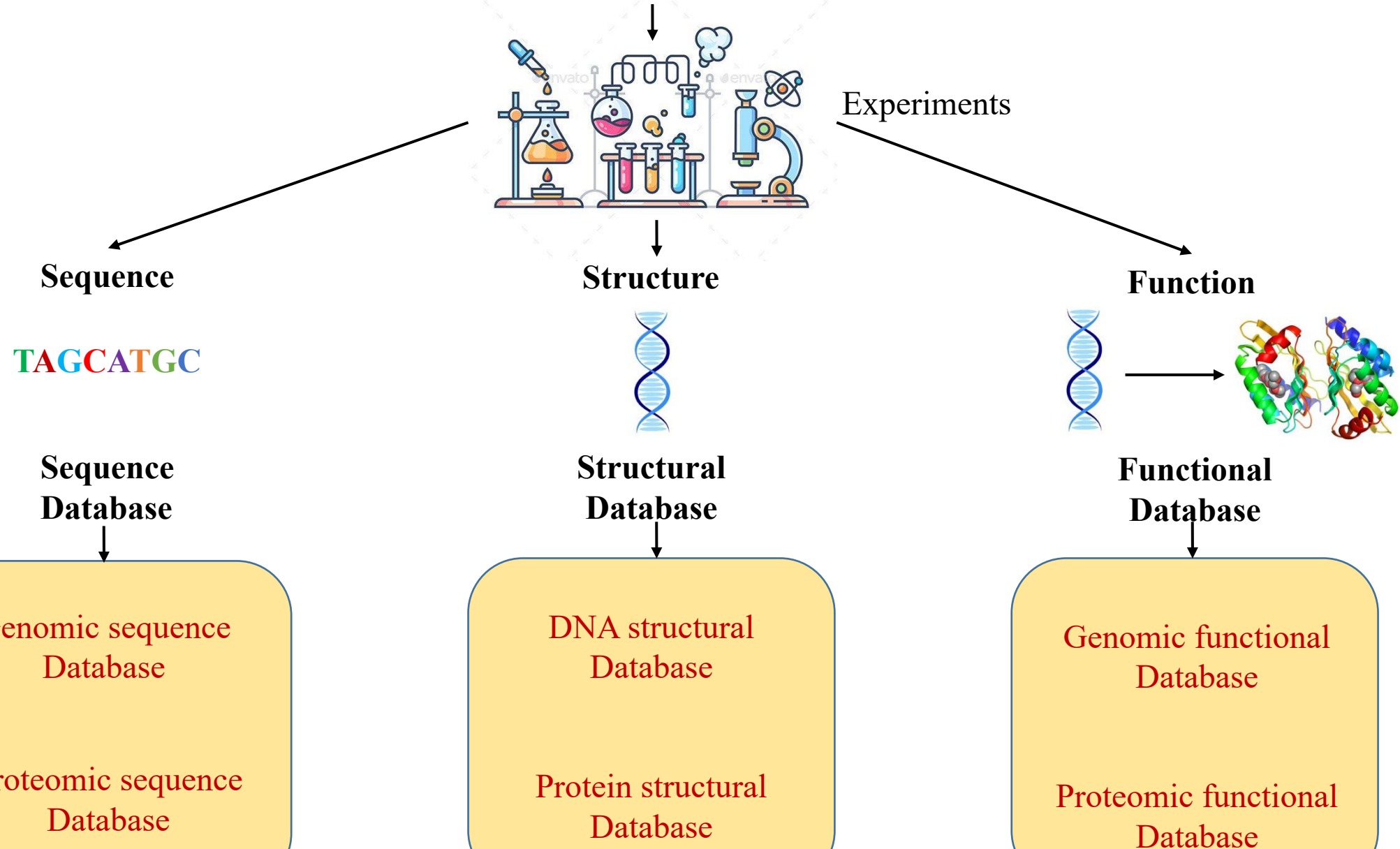
Applications of Bioinformatics operates at 3 levels:

1. Basic level

Used to **organize biological data** (in the form of **databases**) to help the researchers to access information or retrieval.

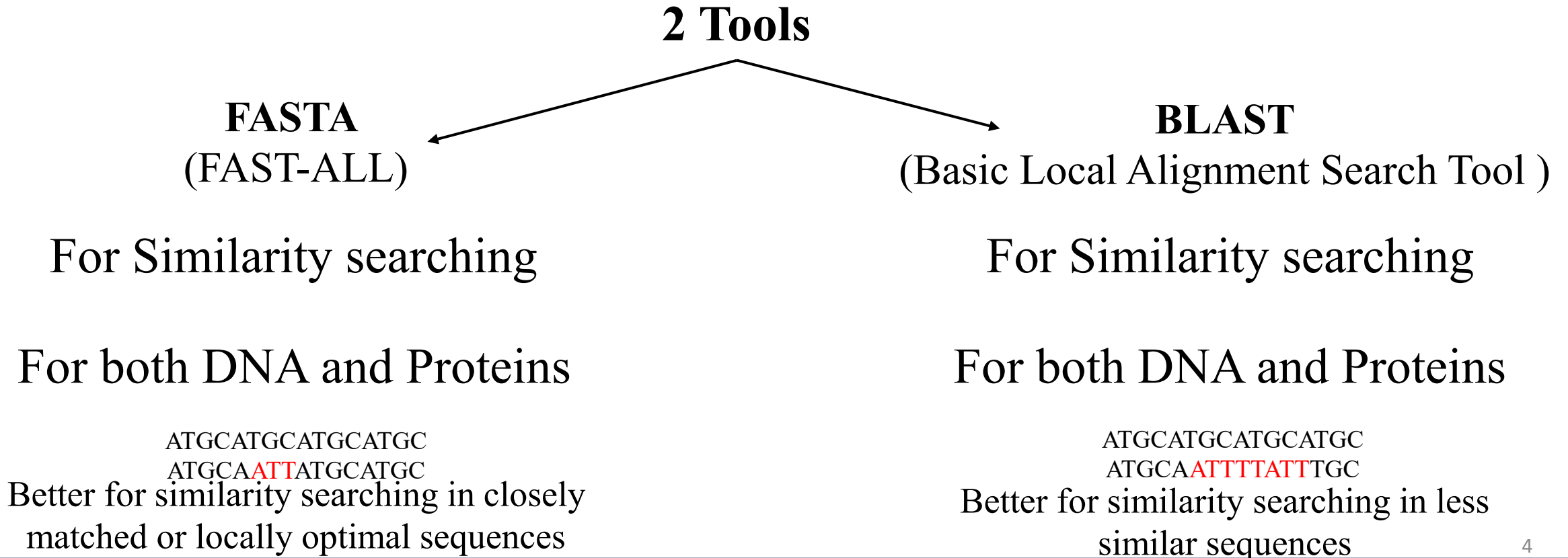


Retrieval of Biological data



2. Second level

To develop **tools and resources** that helps to analyze data i.e. for **data analysis**.



3. Third level

To use these tools to **analyze the data** and **interpret the results** in a biological meaningful manner, for global analysis.



Biological data

Tools → ANALYSIS →



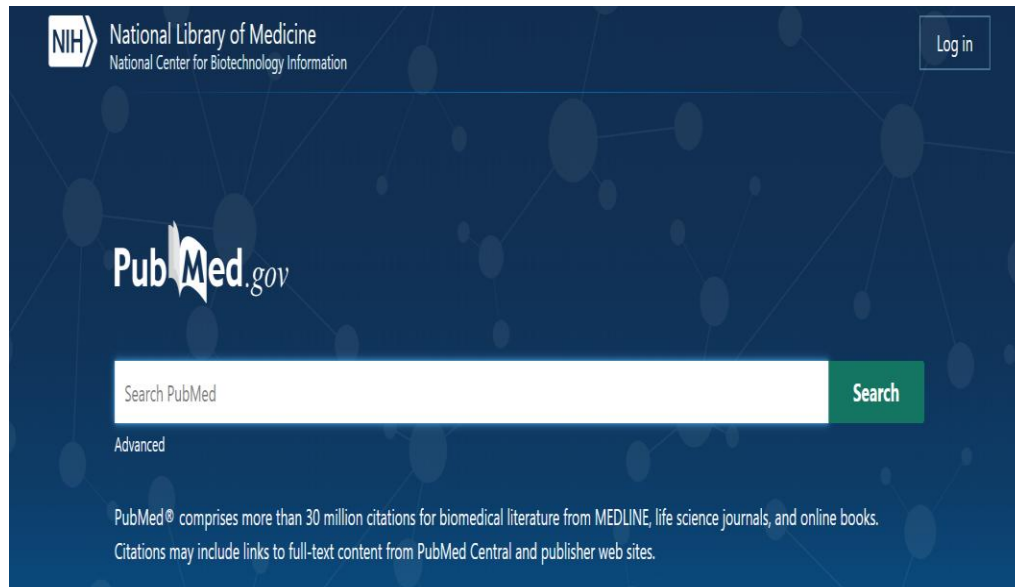
Global Analysis

Information Search and Retrieval

Bioinformatics provides resources for various **search engines** e.g. PubMed (for information search) and **applications** e.g. eTBLAST (for retrieval) of biological data.

PubMed

PubMed is a free search engine accessing primarily the MEDLINE database of references and abstracts on life sciences and biomedical topics.



eTBLAST

Application that is used to compare a query set of sentences with a database of other text to identify the text in the database that is most similar to the query.



Paper 1

Abstract



Paper 2

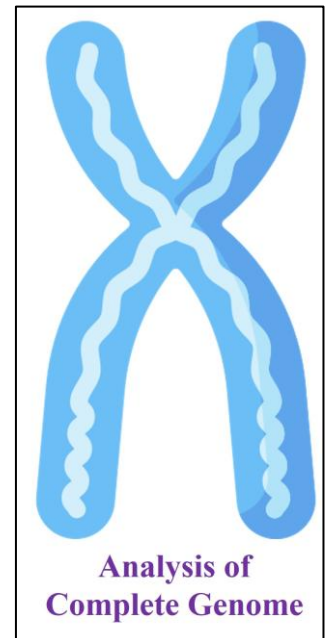
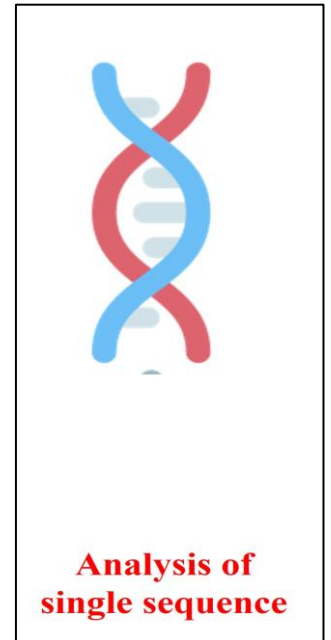
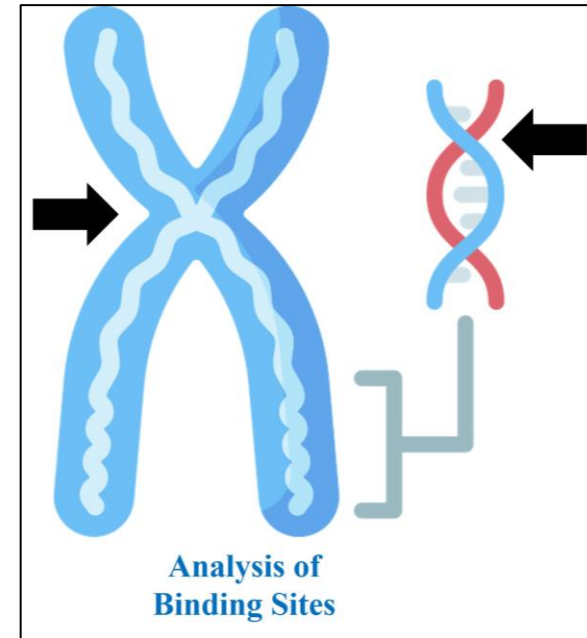
Abstract

————→
Compare

Genetics Related Applications

There are three types of computational problems in genetics:

1. **Analysis of single sequence** to access similarities with known genes.
2. Identification of typical features such as **binding sites** or derive **evolutionary relationships** through phylogenetic trees.
3. **Analysis of Complete genome** to identify members of gene families, determination of the chromosomal location of the gene etc.

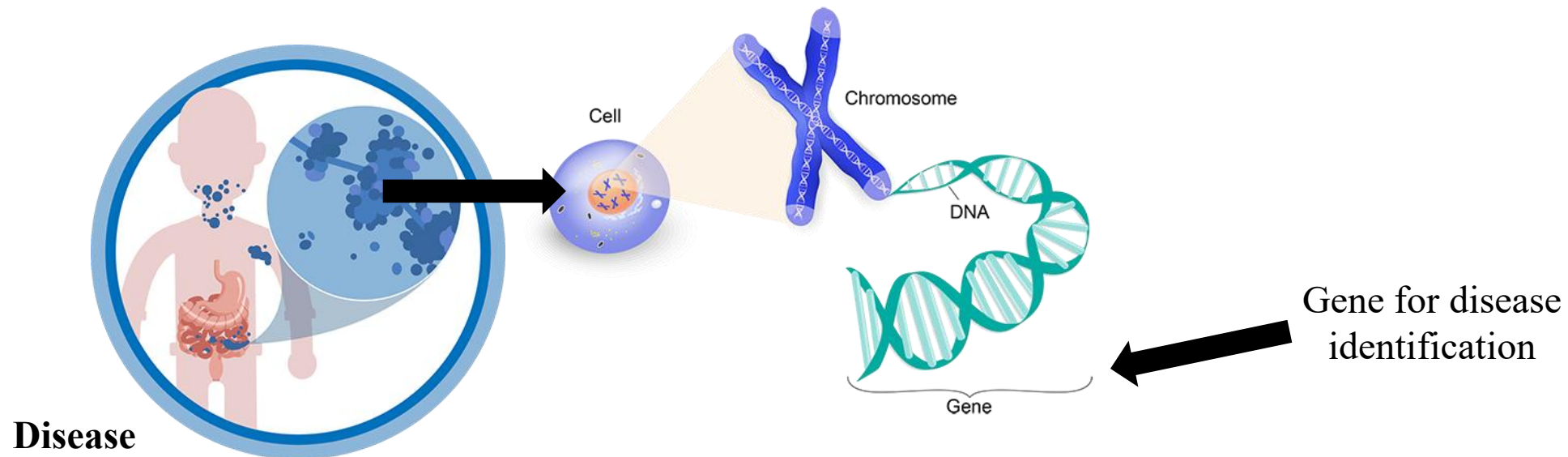


Sequence Comparison

Tools used for sequence comparison are: BLAST and FASTA

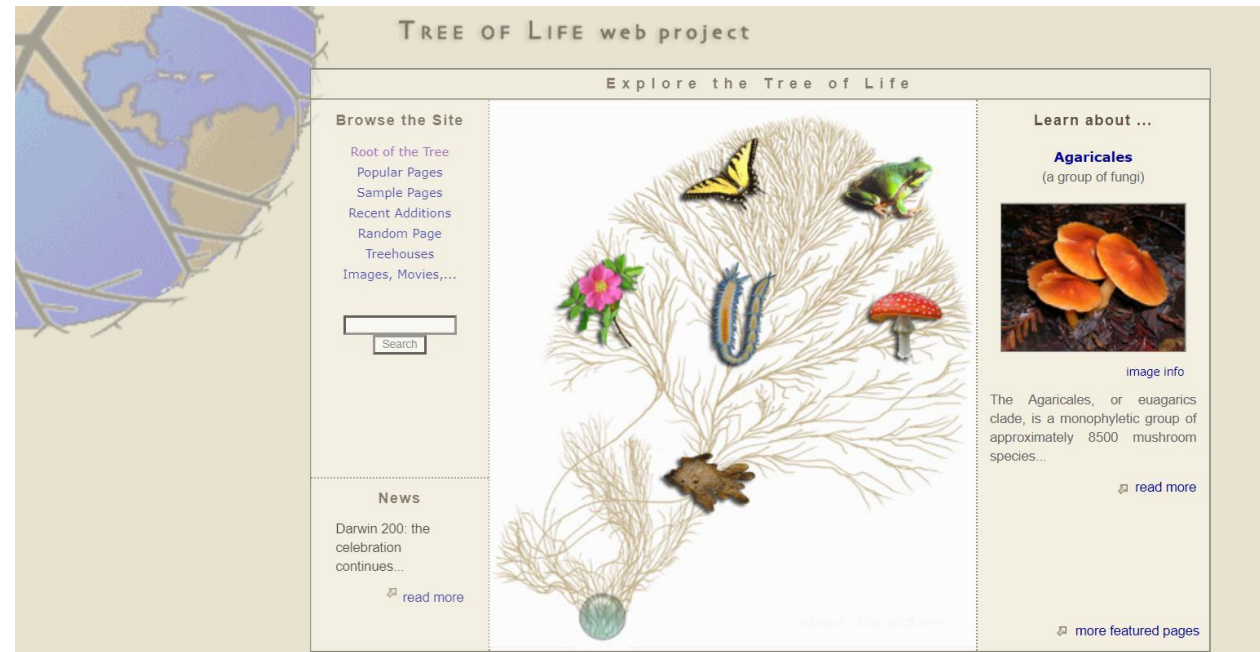
Linkage Analysis

- Involves analysis of large amount of data.
- It is used to **identify** the chromosomal location of **genes** that is used for **disease** identification.
- **Tool: PubMed, PedMiner**



Phylogenetic Analysis

- Phylogenetic Analysis is also known as molecular taxonomy.
- It uses the representation of evolutionary information in the form of phylogenetic trees.
- Tool: **PHYLIP**
- The **Tree of Life** (<http://tolweb.org/tree>) is collaborative Internet project containing information about phylogeny and biodiversity.



GENOMICS

It refers to *mapping*, *sequencing* and *analysis* of genomes.



STRUCTURAL GENOMICS



Linkage analysis



Molecular cytogenetics



Physical mapping



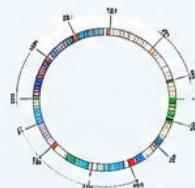
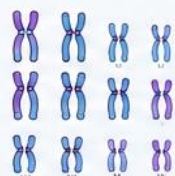
EST sequencing



Genome sequencing



Genome organization



FUNCTIONAL GENOMICS



Gene expression



Forward genetics
(phenotype → genotype)



Reverse genetics
(genotype → phenotype)



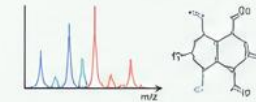
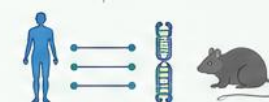
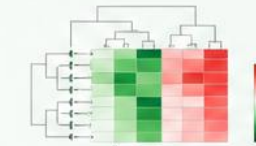
Comparative genomics



Proteomics



Metabolomics



From **structure** to **function**,
genomics drives discovery and innovation in life sciences.



Microarrays

The **transcriptional profile of most genes** within a genome can be analyzed using microarrays.

Transcriptional Profile

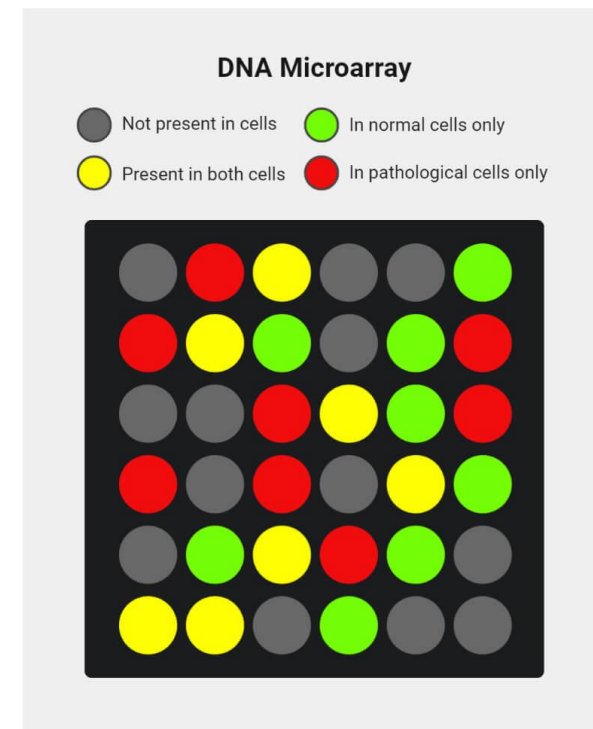
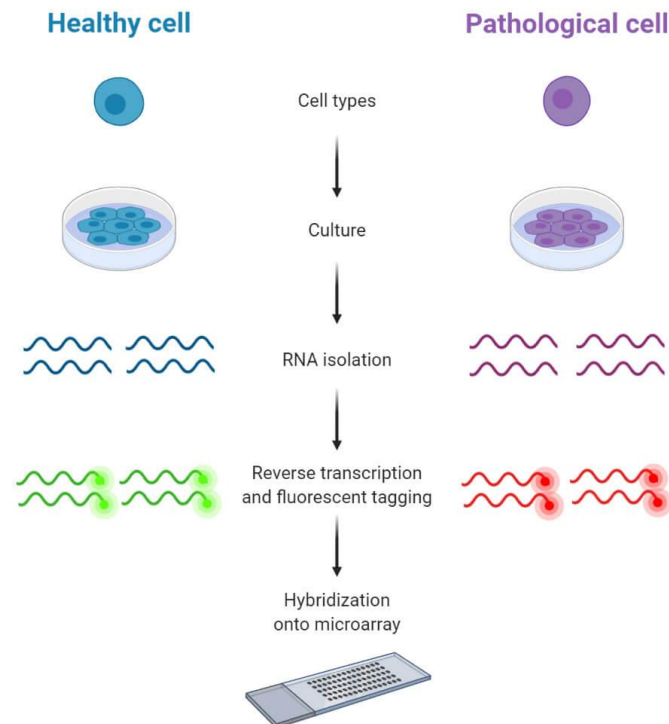


Gene expression



Cell Type

ArrayExpress is a public repository of microarray data.

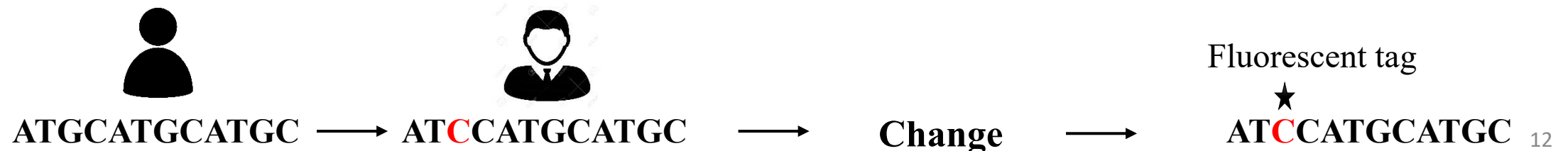


Sequence Assembly

Used to assemble large number of individual sequence reads.

Tools:

1. **Phred/Phrap/Consed Suite**: used to do the assembly and finishing of shotgun sequencing information.
2. **PolyPhred**: used for comparison with fluorescent based sequences across traces obtained from different individuals to identify heterozygous sites for single nucleotide substitution.





Genome Annotation

Genome annotation is the process of **identifying functional elements** along the sequence of a genome.



Tools:



Genome Annotation Consortium (GAC):

It is a multi-institution collaboration to assist in the annotation and analysis of genome sequences by bioinformatics.



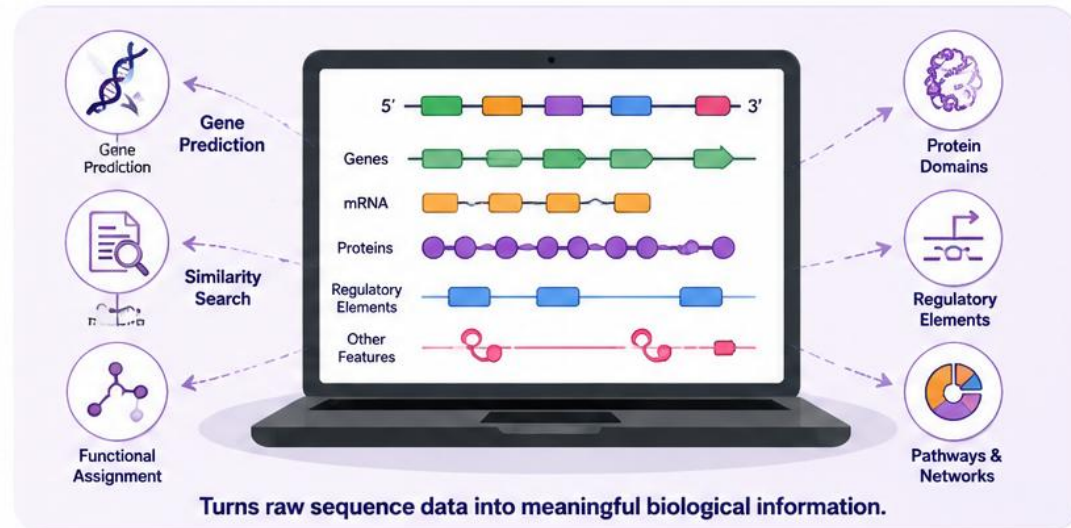
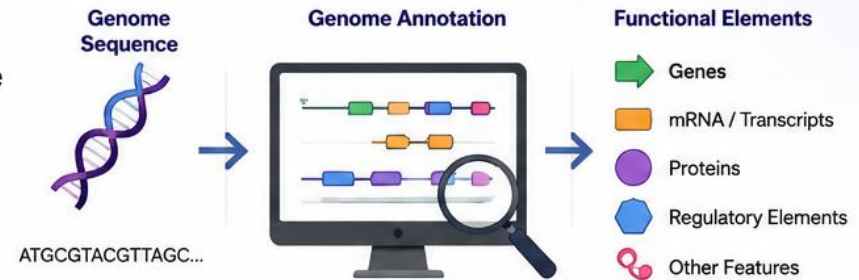
Other tools:

GRAIL, GRAIExp and Pipeline III

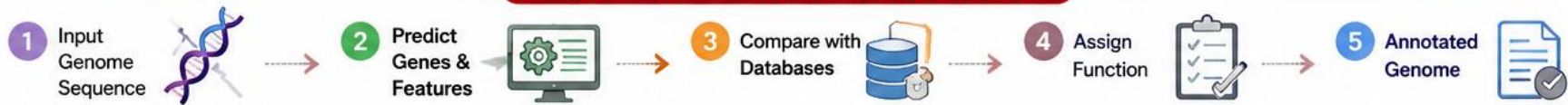


Genie:

It is a tool to locate genes and it uses Hidden Markov Models (HMMs).



HOW GENOME ANNOTATION WORKS



More tools & resources:



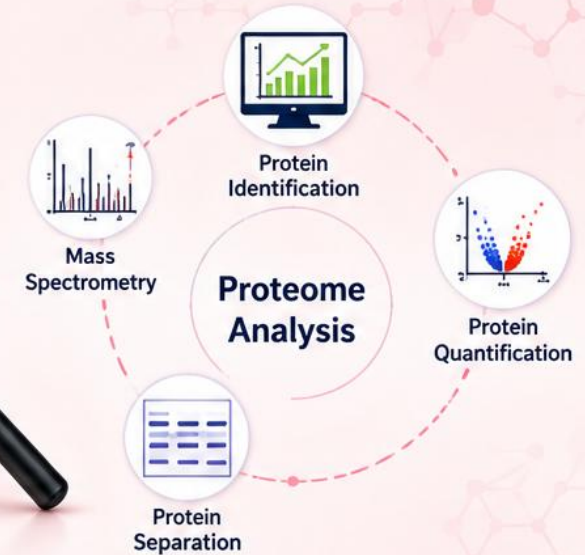
There are several tools available at

<http://www.softberry.com>



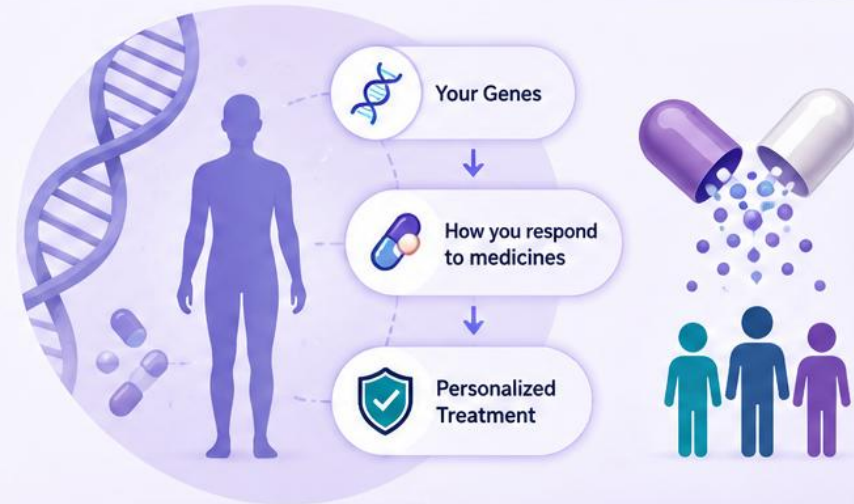
Proteomics

- ✓ It is the study of **proteomes** (proteins expressed by genome).
- ✓ It refers to all proteins produced by a species and it **varies with time**.



Pharmacogenomics

- ✓ It is the study of how **genes** affect a person's **response to drugs**.
- ✓ It combines **pharmacology** (the science of drugs) and **genomics** (the study of genes and their functions) to develop **effective, safe medications** and doses that will be **tailored** to a person's genetic makeup.

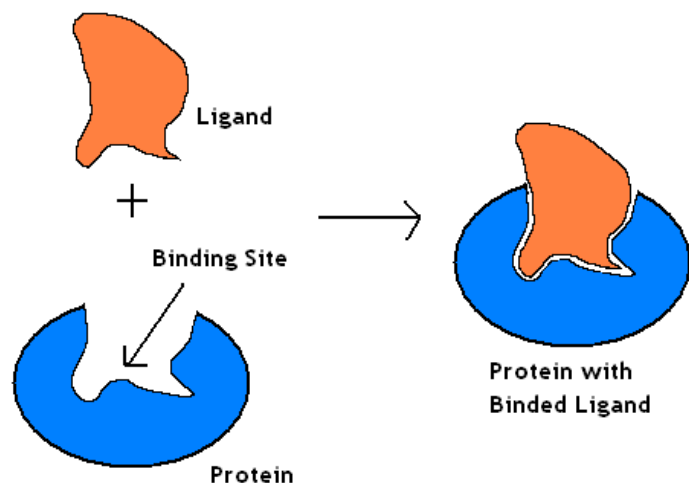


From proteins to personalized medicine,
integrating biology for better health outcomes.

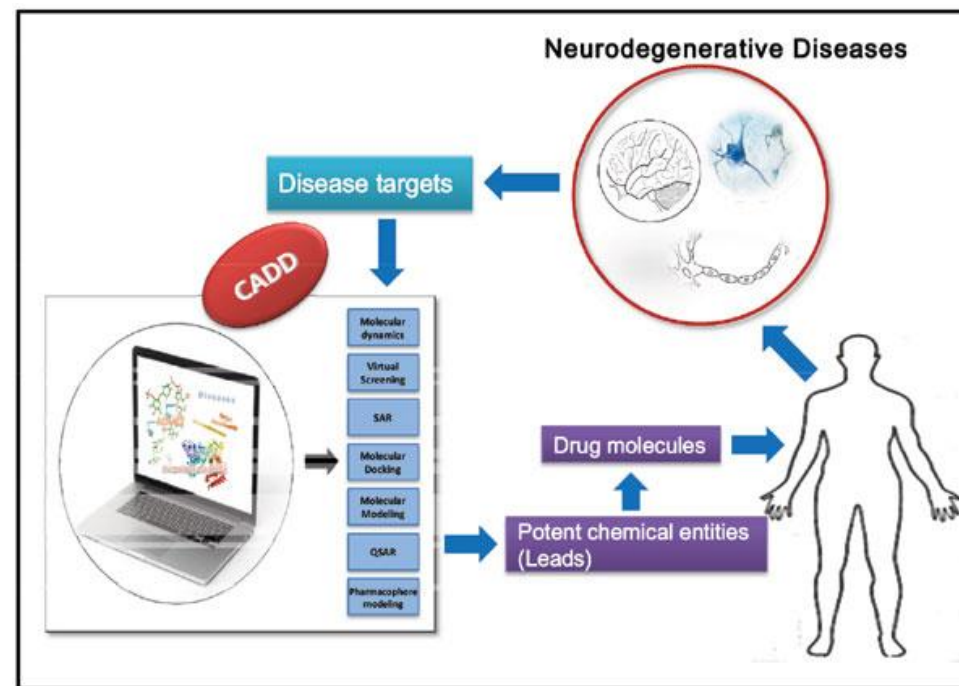


Drug discovery and Computer-aided Drug Design

Drug Discovery: Analysis of binding sites in target proteins or identification of structural features common to active compounds.



Computer-aided Drug Design
Uses bioinformatics tools for drug discovery.





System Biology



It is an integrated multi-disciplinary approach in which you study **pathways** and **networks**.



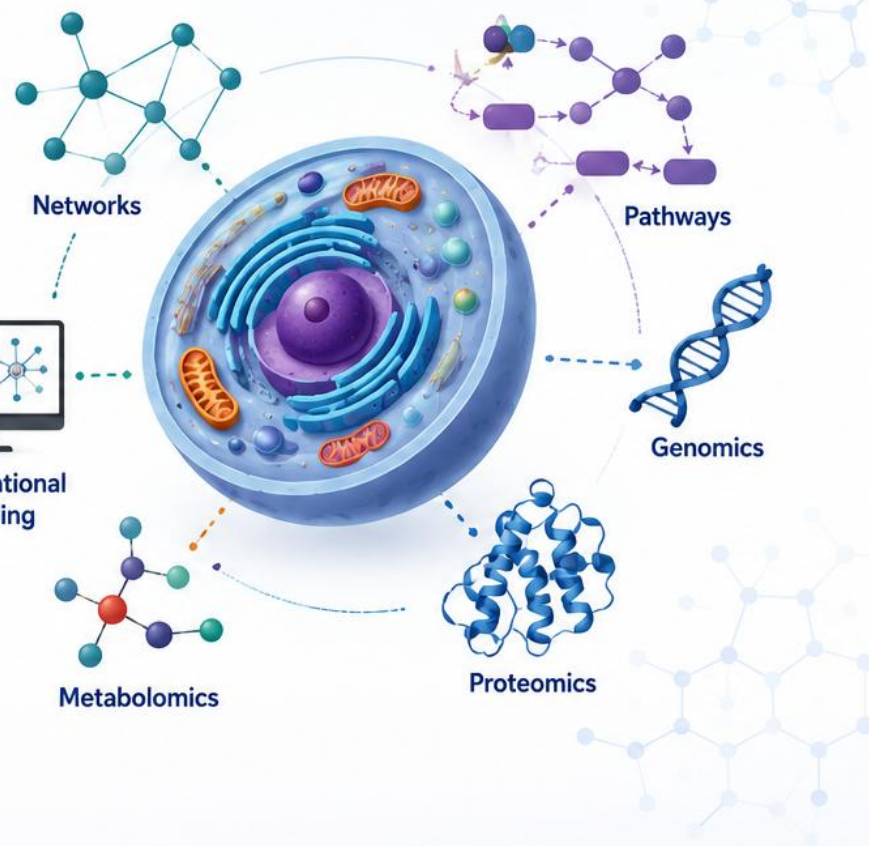
Tool:



NRCAM:
General Computational tool,
the virtual cell for modelling
cellular processes.



Computational
Modeling



WHAT SYSTEMS BIOLOGY DOES



Collects data
(genomics,
proteomics,
metabolomics)



Builds computational
models



Analyzes pathways
and networks



Generates insights



Guides experiments
and decisions



Explore NRCAM:



<https://www.genecards.org/cgi-bin/carddisp.pl?gene=NRCAM>

