# BHARATHIDASAN UNIVERSITY

Tiruchirappalli- 620024,
Tamil Nadu, India.

## Programme : M.Sc., Biomedical Science

**Course Title : Bioinformatics**

**Course Code : BM35S1BI**

## Unit – I

## TOPIC: BASICS OF BIOINFORMATICS

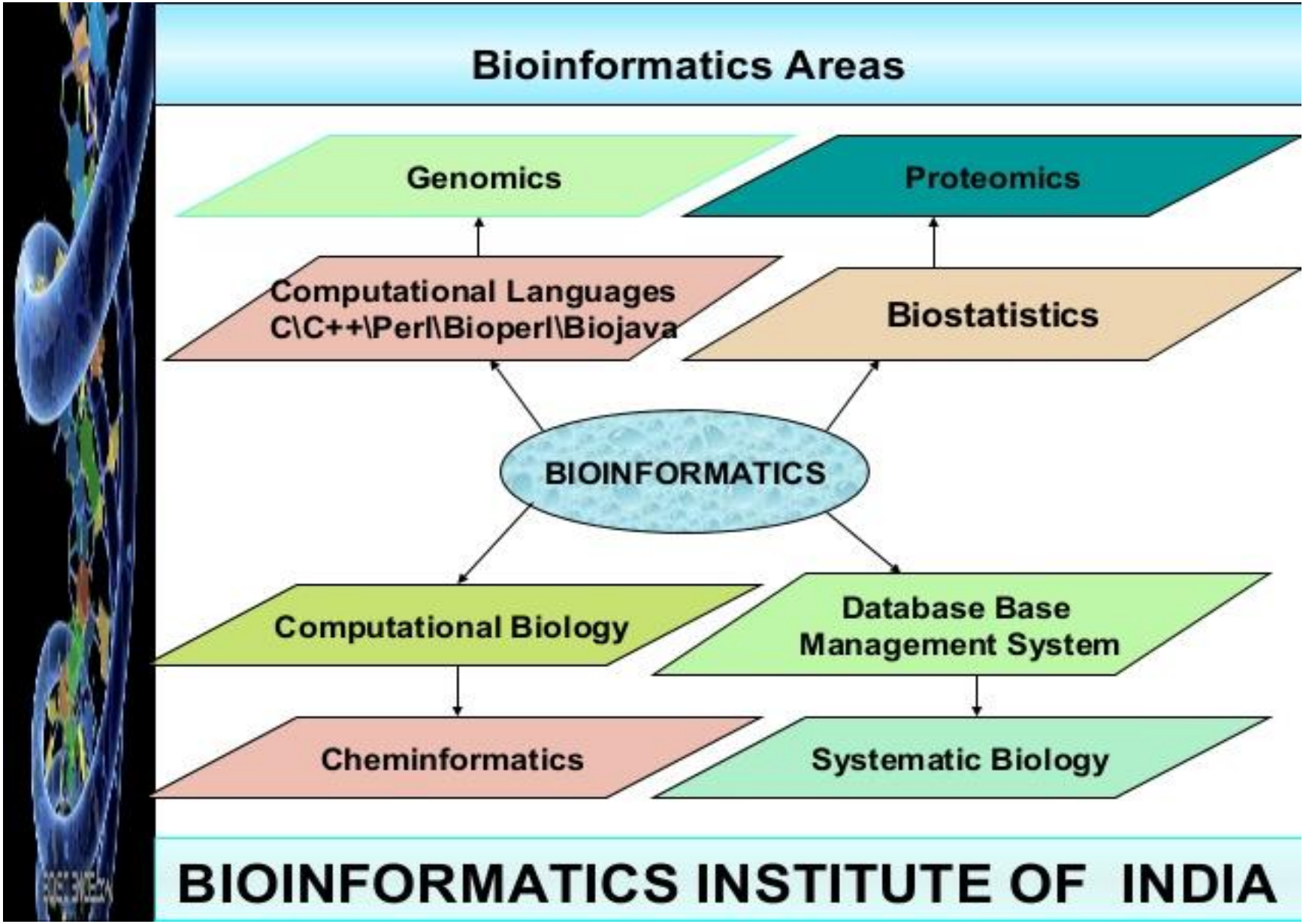
**Dr. P. JEGANATHAN**
Guest Lecturer
Department of Biomedical Science

# BASICS OF BIOINFORMATICS

## Definition of Bioinformatics

General Definition: A computational approach ,Solves the biological problem.

- Bioinformatics is emerging and advance branch of biological science , contain Biology mathematics and Computer Science.
- Bioinformatics developed a new thought , to maintain the concepts and store .The huge amount of Biological data.
- Bioinformatics concepts and Method are different than the Biological concepts and method.
- Bioinformatics, A logical and technical means by which not only solve the Biological problems but also can predicts the new aspects.


Bioinformatics Areas
Genomics
Proteomics
Computational Languages
C\C++\Perl\Bioperl\Biojava
Biostatistics
BIOINFORMATICS
Computational Biology
Database Base
Management System
Cheminformatics
Systematic Biology
BIOINFORMATICS INSTITUTE OF INDIA

## *Insilico* Areas of Bioinformatics

Computational Biology

- Comparative Homology Modeling
- Phylogenetic Analysis
- Protein structure prediction
- Protein Folding Problem
- Micro array analysis
- Docking Approaches& New Drug Discovery

**BIOINFORMATICS INSTITUTE OF INDIA**

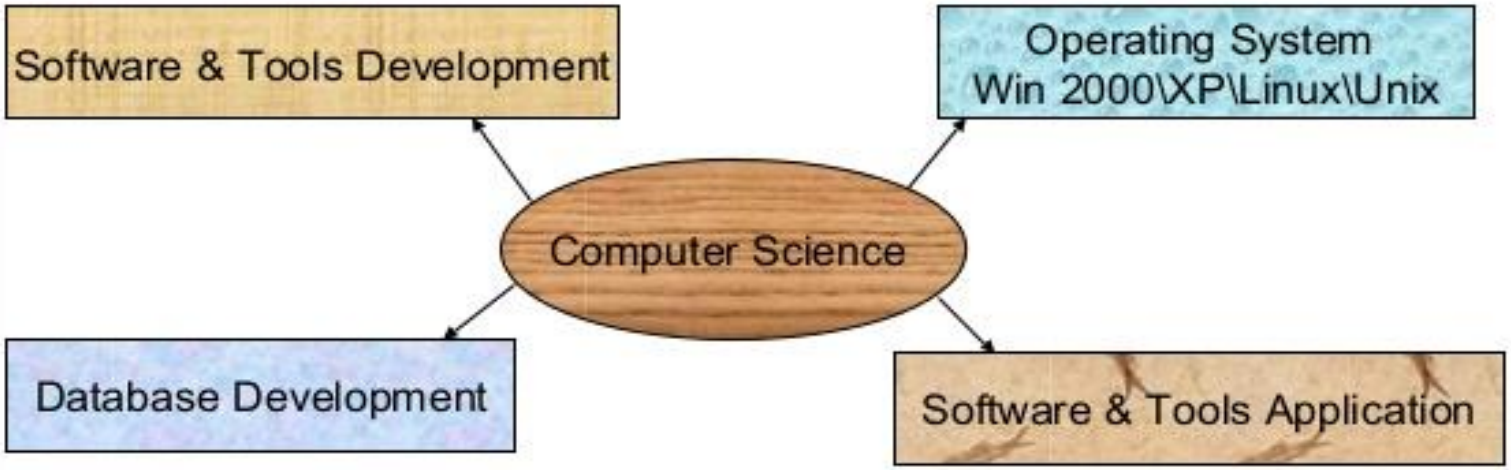
# Internet and Bioinformatics

✓Internet plays an important role to retrieve the biological information.

✓Bioinformatics emerging new dimension of Biological science, include The computer science ,mathematics and life science.

✓The Computational part of bioinformatics use to optimize the biological problems like (metabolic disorder, genetic disorders).

✓Computational part contains:



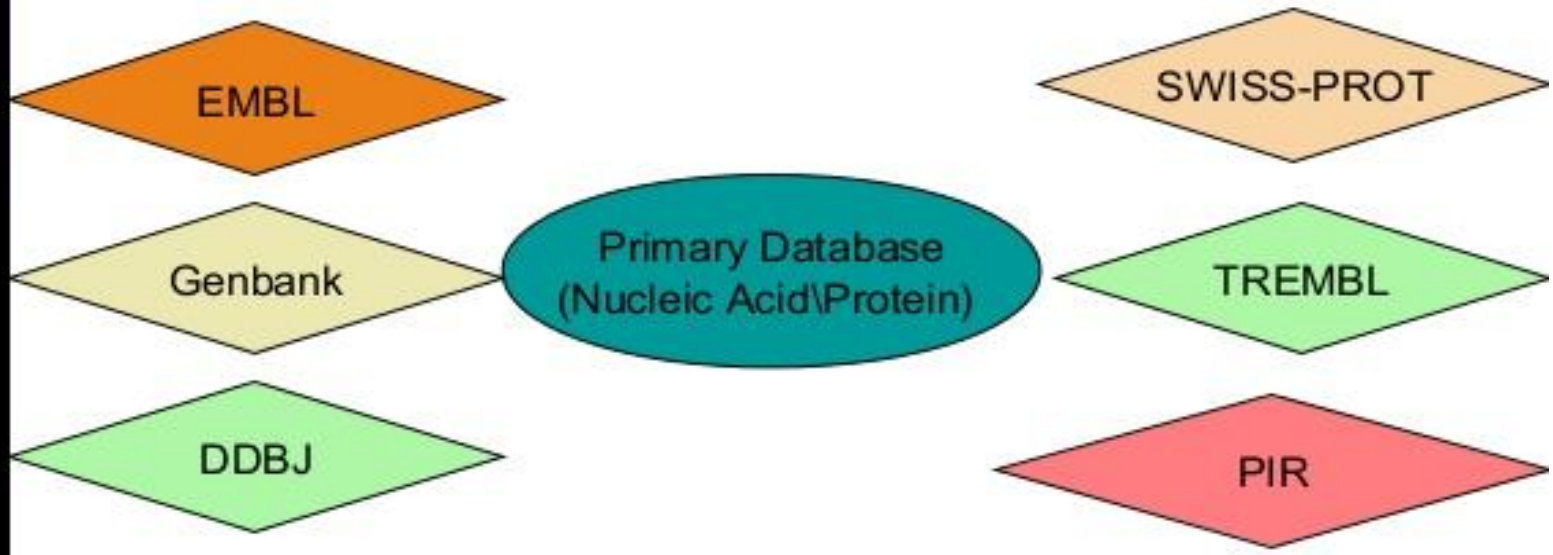
## Biological Databases

| Type of databases | Information Contain |
|---|---|
| Bibliographic databases | Literature |
| Taxonomic databases | Classification |
| Nucleic acid databases | DNA information |
| Genomic databases | Gene level information |
| Protein databases | Protein information |
| Protein families, domains and functional sites | Classification of proteins and identifying domains |
| Enzymes/ metabolic pathways | Metabolic pathways |

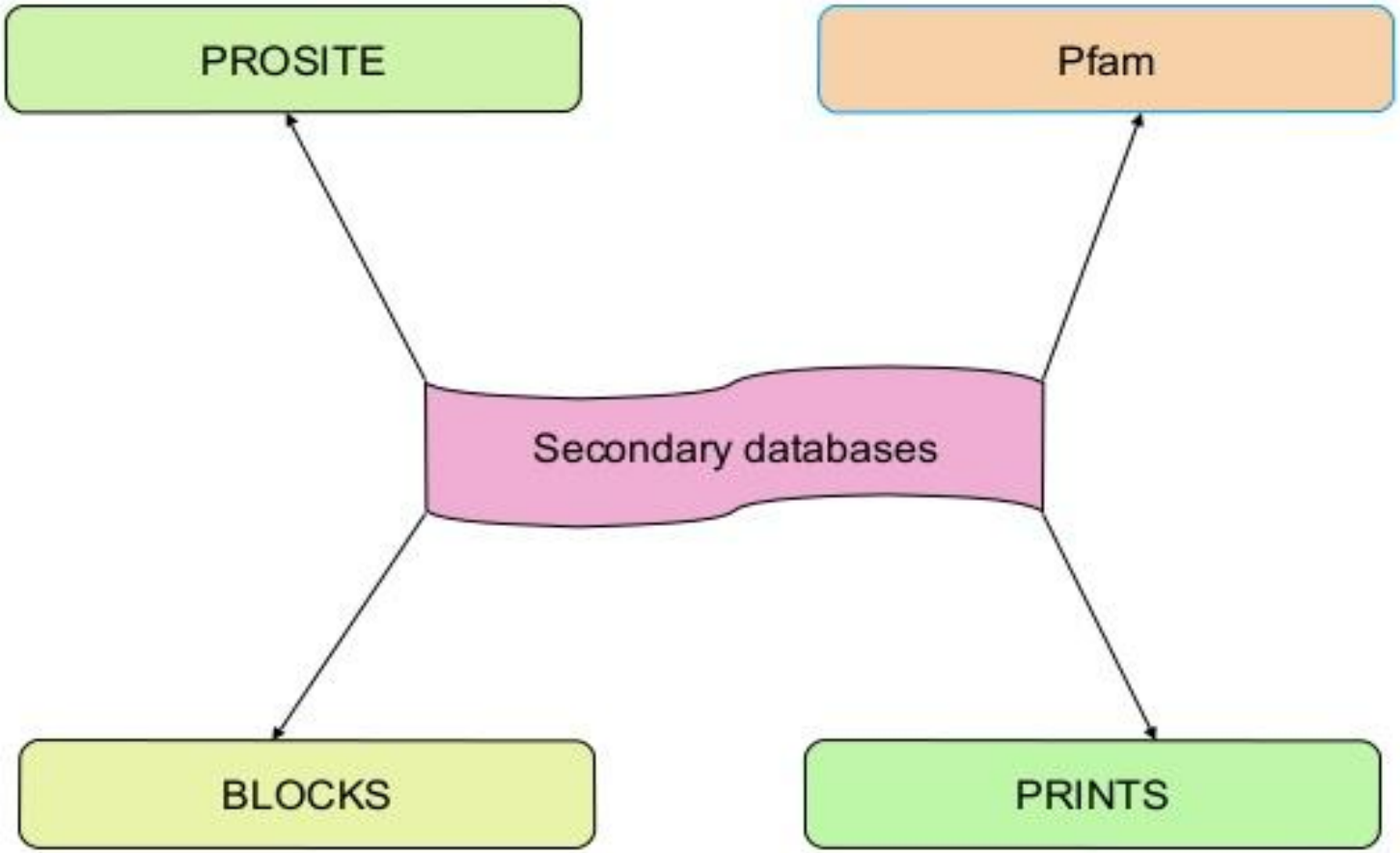
## Types Of Biological Databases Accessible

There are many different types of database but for routine sequence analysis, the following are initially the most important.

- Primary databases
- Secondary databases
- Composite databases

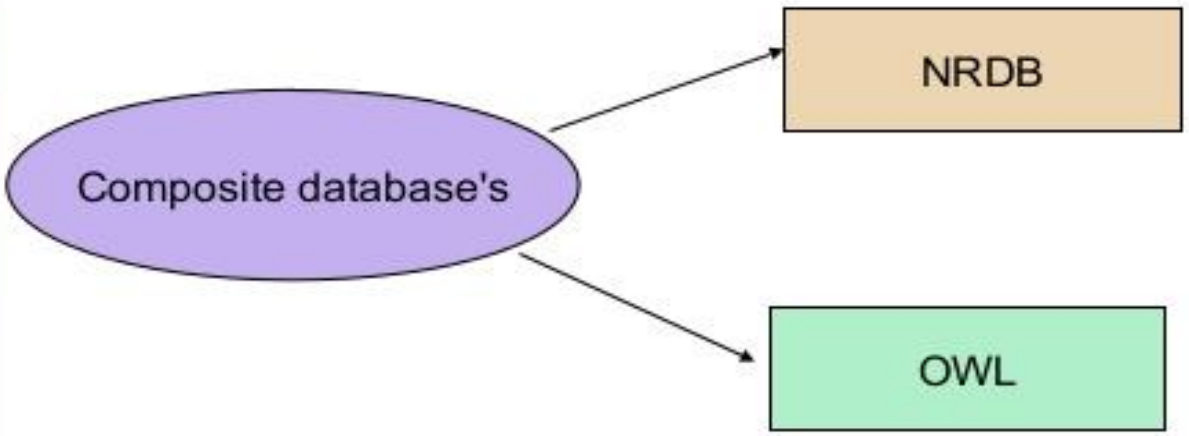


## Secondary databases



## Composite databases

✓Combine different sources of primary databases.



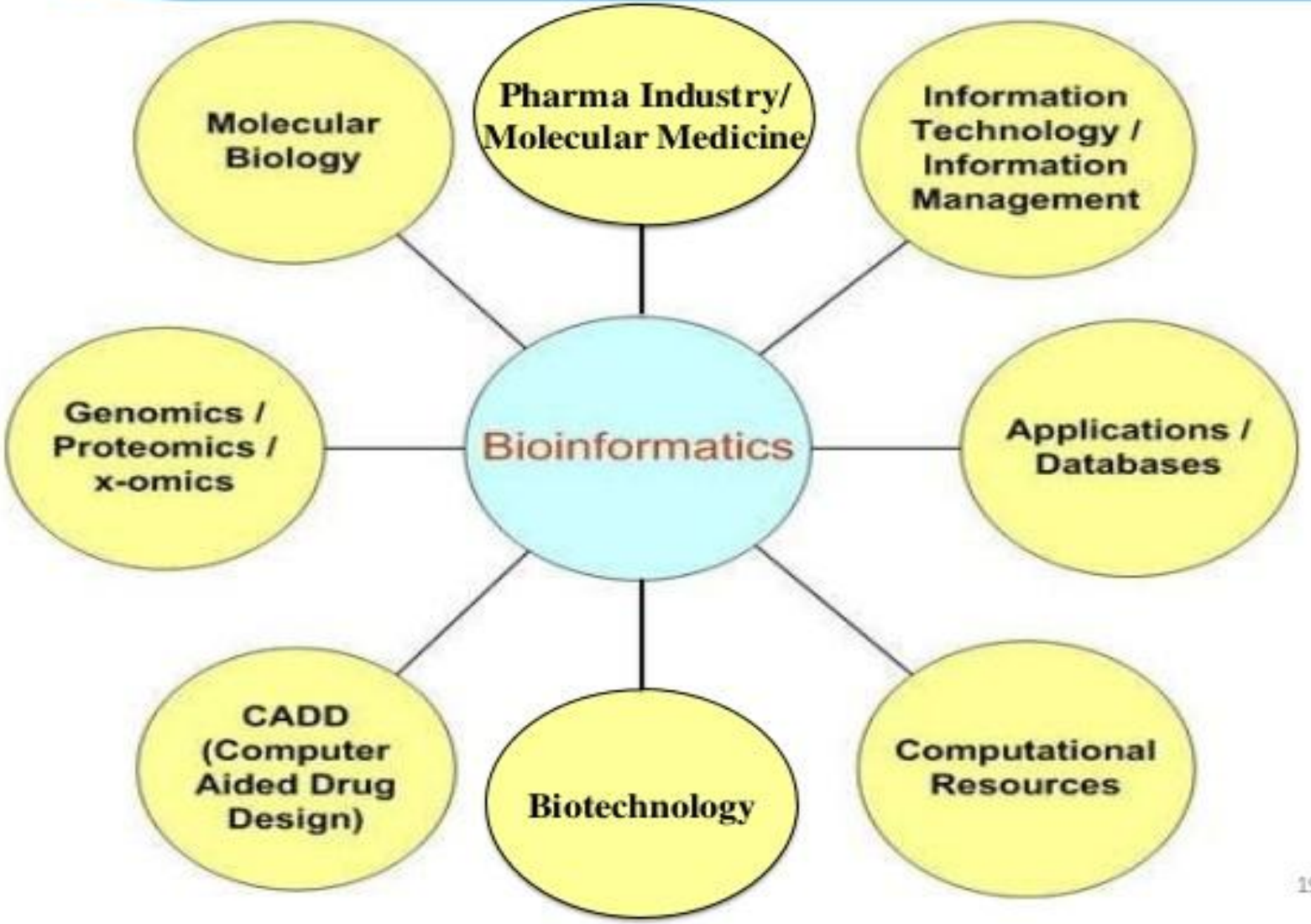
# ROLES OF COMPUTERS IN BIOLOGY

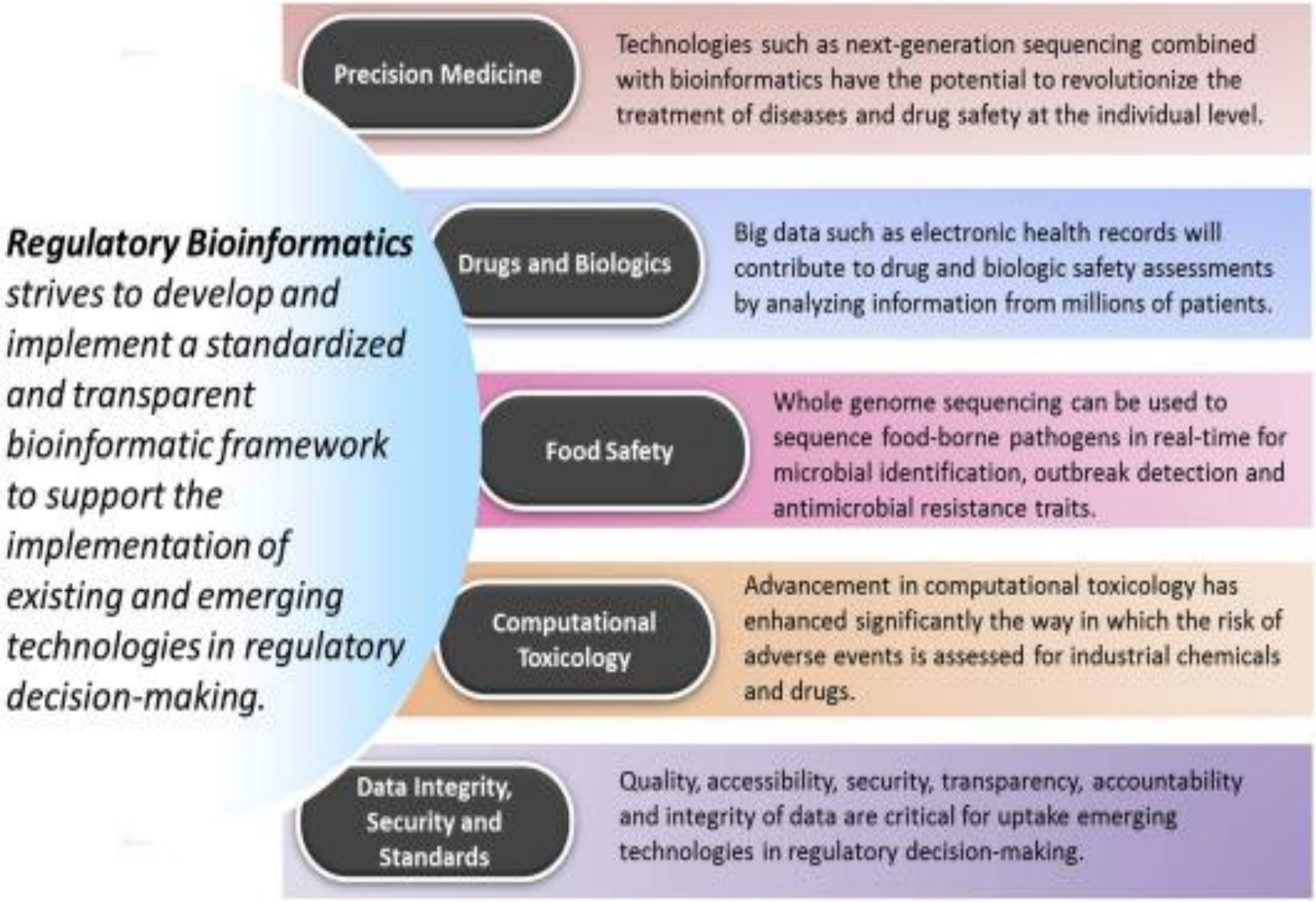
- The **biological computer** is an implantable device that is mainly used for tasks like monitoring the body's activities or inducing therapeutic effects, all at the molecular or cellular level. This is made up of RNA, DNA and proteins and can also perform simple mathematical calculations.
- **Computers** are probably the biggest advance in **scientific** technology. They allow us to analyze huge data sets and run statistical analyses far quicker than in the past. **Scientists** also use them to run simulations and build electronic models.

# BIOCOMPUTING

- **Biocomputing** is defined as the process of building computers that use biological materials, mimic biological organisms or are used to study biological organisms.
- **Bio computers** use systems of biologically derived molecules—such as DNA and proteins—to perform computational calculations involving storing, retrieving and processing data.
- **Biocomputers** are computers made of proteins, genes, and cells, and capable of performing mathematical operations.

# Applications of Bioinformatics




Regulatory Bioinformatics strives to develop and implement a standardized and transparent bioinformatic framework to support the implementation of existing and emerging technologies in regulatory decision-making.
Precision Medicine
Technologies such as next-generation sequencing combined with bioinformatics have the potential to revolutionize the treatment of diseases and drug safety at the individual level.
Drugs and Biologics
Big data such as electronic health records will contribute to drug and biologic safety assessments by analyzing information from millions of patients.
Food Safety
Whole genome sequencing can be used to sequence food-borne pathogens in real-time for microbial identification, outbreak detection and antimicrobial resistance traits.
Computational Toxicology
Advancement in computational toxicology has enhanced significantly the way in which the risk of adverse events is assessed for industrial chemicals and drugs.
Data Integrity, Security and Standards
Quality, accessibility, security, transparency, accountability and integrity of data are critical for uptake emerging technologies in regulatory decision-making.

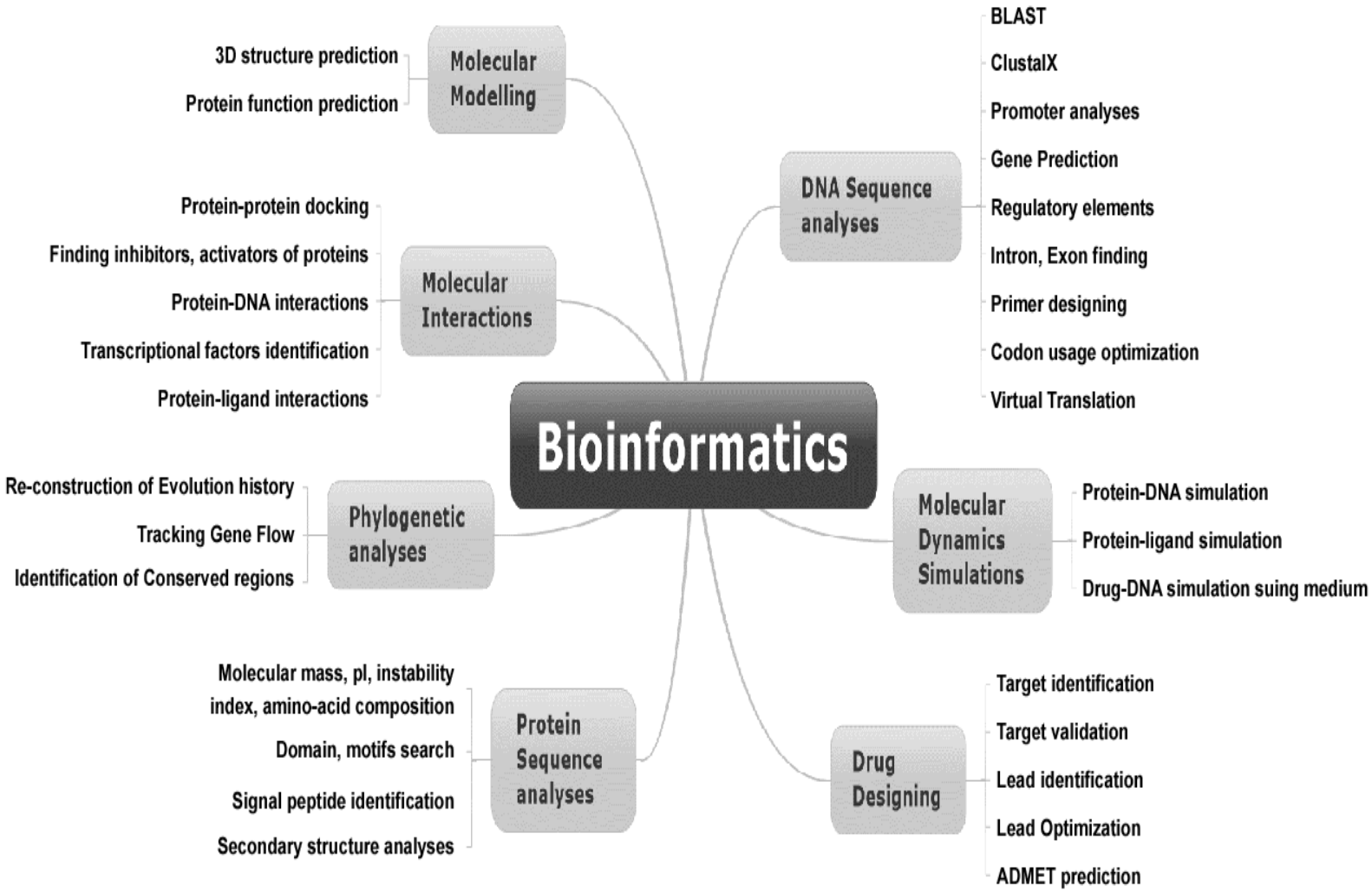
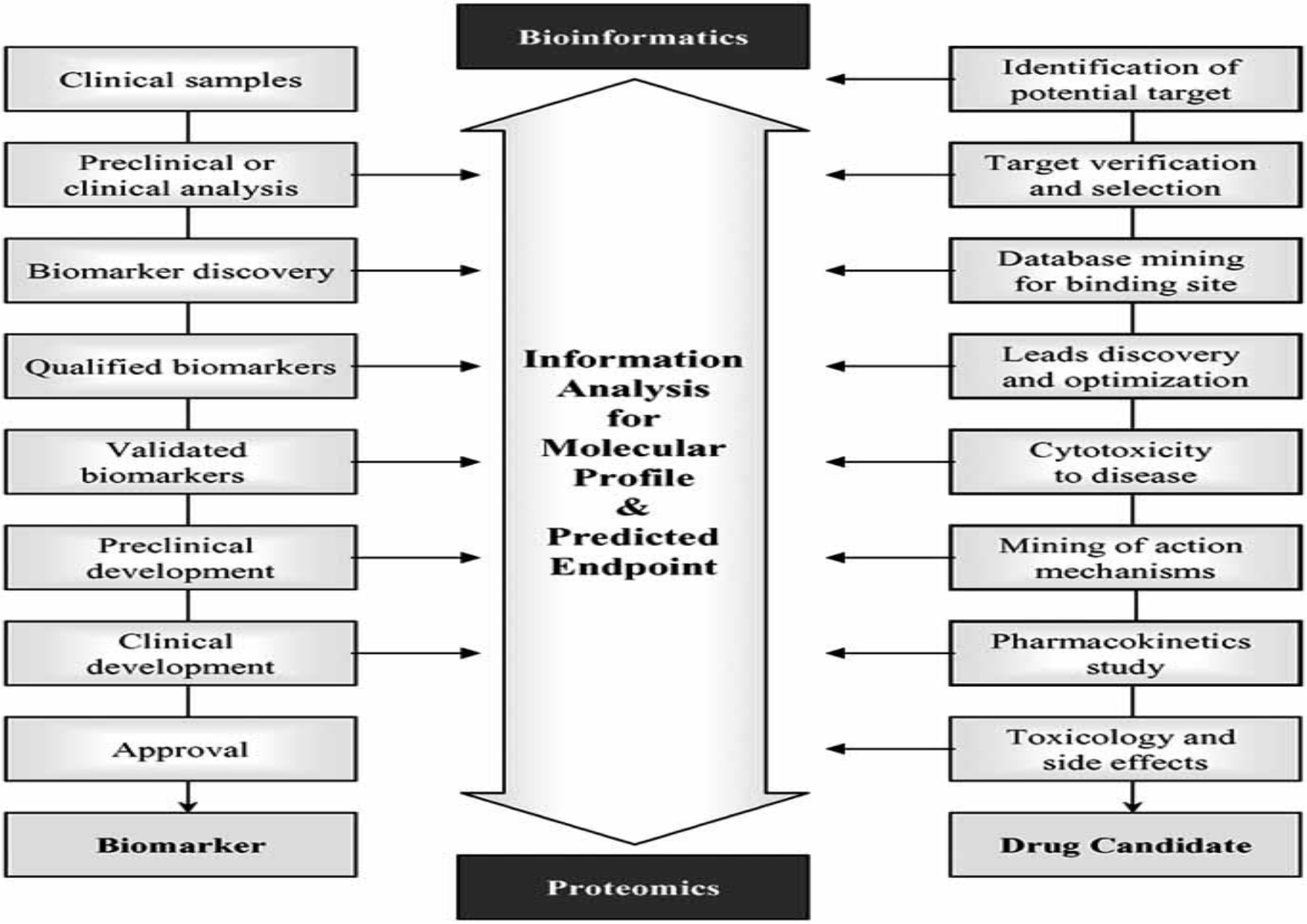
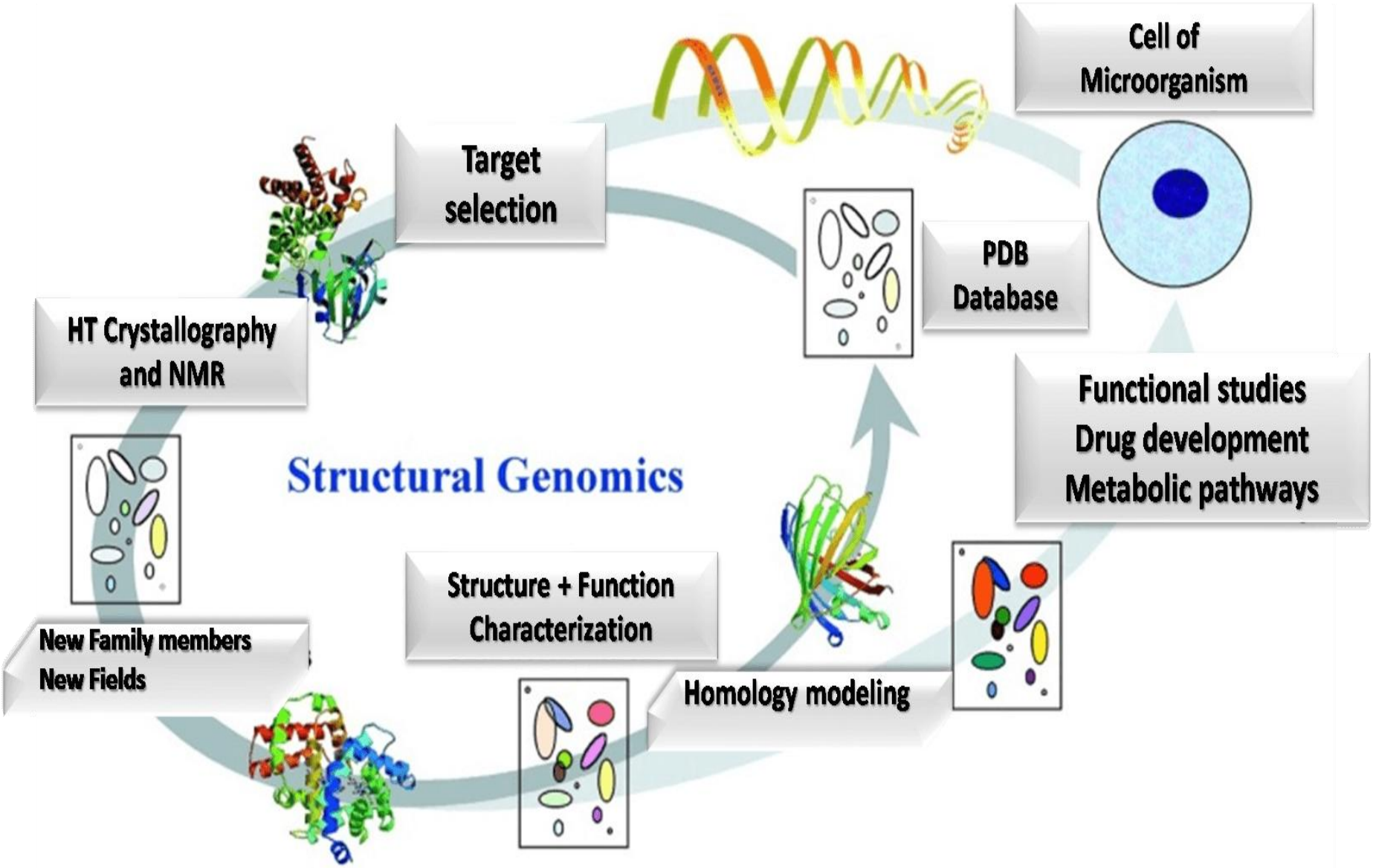



**Figure 1:** Application of bioinformatics tools in various areas of biological sciences.


Bioinformatics
Clinical samples
Preclinical or clinical analysis
Biomarker discovery
Qualified biomarkers
Validated biomarkers
Preclinical development
Clinical development
Approval
Biomarker
Information Analysis for Molecular Profile & Predicted Endpoint
Proteomics
Identification of potential target
Target verification and selection
Database mining for binding site
Leads discovery and optimization
Cytotoxicity to disease
Mining of action mechanisms
Pharmacokinetics study
Toxicology and side effects
Drug Candidate


Cell of Microorganism
Target selection
PDB Database
HT Crystallography and NMR
Functional studies
Drug development
Metabolic pathways
Structural Genomics
Structure + Function Characterization
New Family members
New Fields
Homology modeling

# Scope of Bioinformatics

- Bioinformatics deals with methods for starting, retrieving and analysing biological data such as nucleic acid (DNA /RNA) and protein sequences , structure, functions pathways and genetic interactions.
- The computational methods in bioinformatics extend information for probing not only at genome level or protein level but up to whole organism level, or ecosystem level of organization.
- It provides genome level data for understanding normal biological processes and explains the malfunctioning of genes leading to diagnosing of diseases and designing of new drugs.

# BIOINFORMATICS INSTITUTES

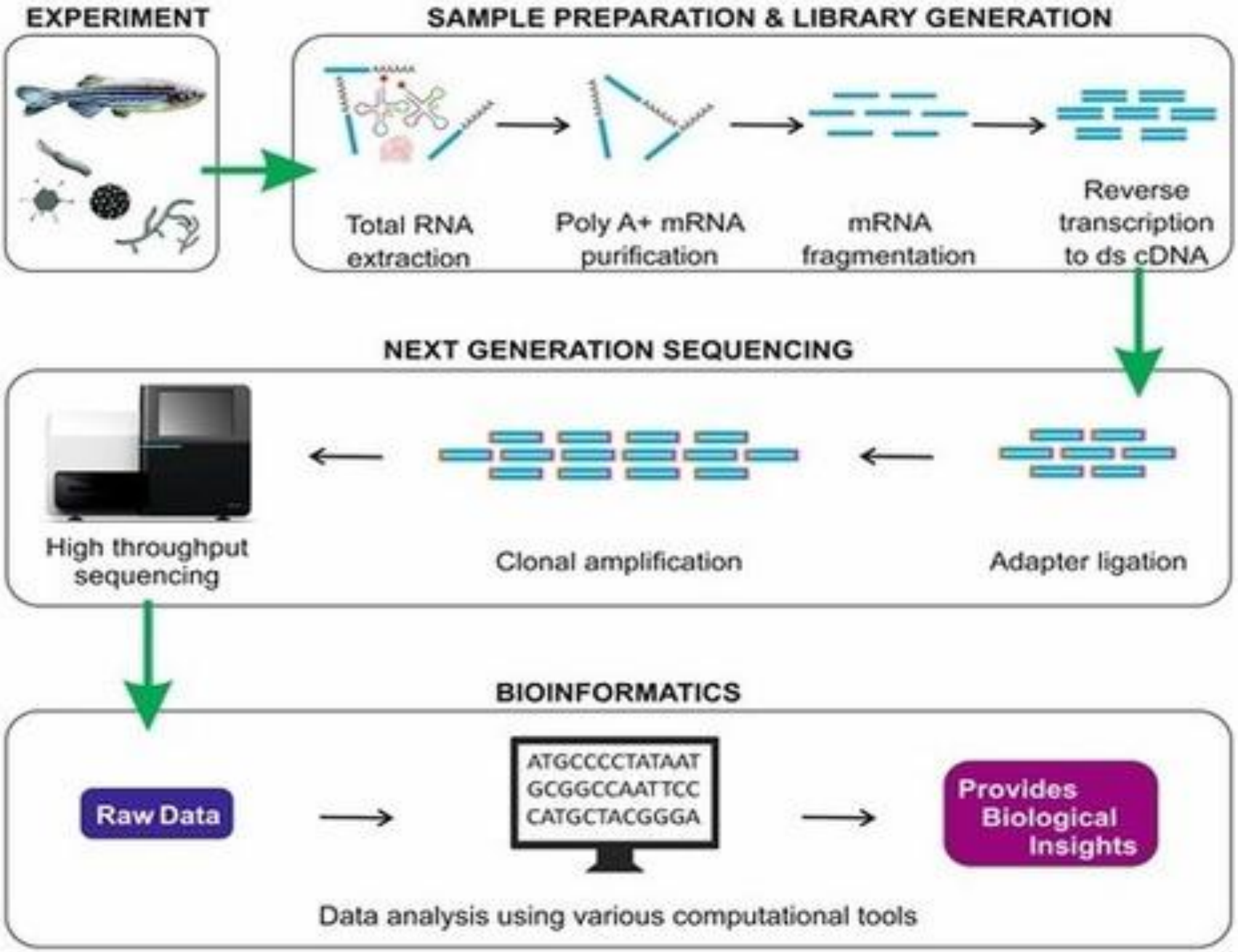
- National Center for Biotechnology Information (NCBI)
- European Bioinformatics Institute (EMBL-EBI)
- Australia Bioinformatics Resource (EMBL-ABR)
- Swiss Institute of Bioinformatics (SIB)
- Scripps Research Institute (TSRI)
- European Molecular Biology Laboratory (EMBL)
- Wellcome Trust Sanger Institute (WTSI)
- Computational Biology Department
- Broad Institute
- Whitehead Institute
- The Institute for Genomic Research
- Center for Biomolecular Science and Engineering
- Netherlands Bioinformatics Centre
- COSBI
- Max Planck Institute for Molecular Cell Biology and Genetics (MPI-CBG)
- Partner Institute for Computational Biology
- Flatiron Institute
- DDBJ Center (DDBJ)
- Database Center for Life Science (DBCLS)

# BIOINFORMATICS INDUSTRIES

- Applied Maths provides the software suite BioNumerics
- Astrid Research
- BIOBASE
- BioBam Bioinformatics creator of Blast2GO
- Biomax Informatics AG bioinformatics services.
- Biovia (formerly Accelrys).
- Chemical Computing Group MOE software for structural modelling
- CLC Bio Bioinformatics workbenches.
- DNASTAR provides DNA sequence assembly and analysis.
- Gene Codes Corporation
- Genedata software for data analysis and storage.
- GeneTalk web-based services.
- GenoCAD
- Genomatix
- Genostar provides streamlined bioinformatics.
- Inte:Ligand
- Integromics

# BIOINFORMATICS INDUSTRIES

- Invitrogen creator of Vector NTI
- Korea Computer Centre Sinhung Company
- Leidos Biomedical Research Inc. formerly SAIC. Services are aimed at the Federal Government market.
- MacVector
- QIAGEN Silicon Valley (formerly Ingenuity Systems)
- Qlucore
- Phalanx Biotech Group
- SimBioSys created the eHITS software
- SRA International services aimed at the Federal Government market.
- Strand Life Sciences
- TimeLogic offers DeCypher FPGA-accelerated BLAST, Smith-Waterman, HMMER and other sequence search tools.


EXPERIMENT
SAMPLE PREPARATION & LIBRARY GENERATION
Total RNA extraction
Poly A+ mRNA purification
mRNA fragmentation
Reverse transcription to ds cDNA
NEXT GENERATION SEQUENCING
High throughput sequencing
Clonal amplification
Adapter ligation
BIOINFORMATICS
Raw Data
ATGCCCCTATAAT
GCGGCCAATTCC
CATGCTACGGGA
Provides Biological Insights
Data analysis using various computational tools

# References

1. Statistics for Bioinformatics

-Julie Dawn Thompson

2. Multiple Sequence Alignment methods

-David J.Russell

3. Bioinformatics for DUMMIES (2nd Edition)

-Jean – Micha Claverie, Cedric Notredame

# THANK YOU