G P S Raghava, Head Bioinformatics Ce

Institute of Microbial Technology, Chandigarh, India

Human Genome   Drug Discovery   Vaccine Design
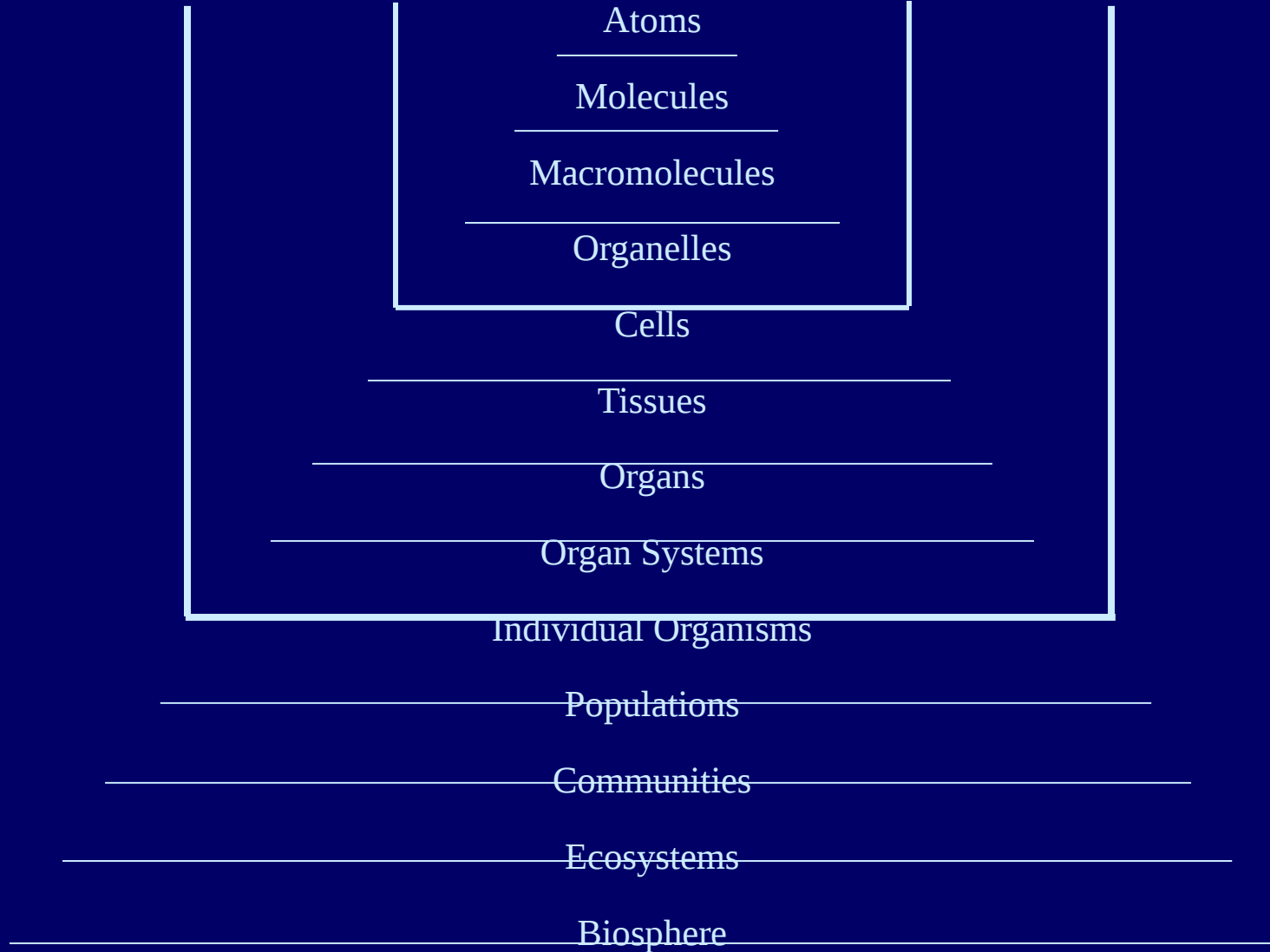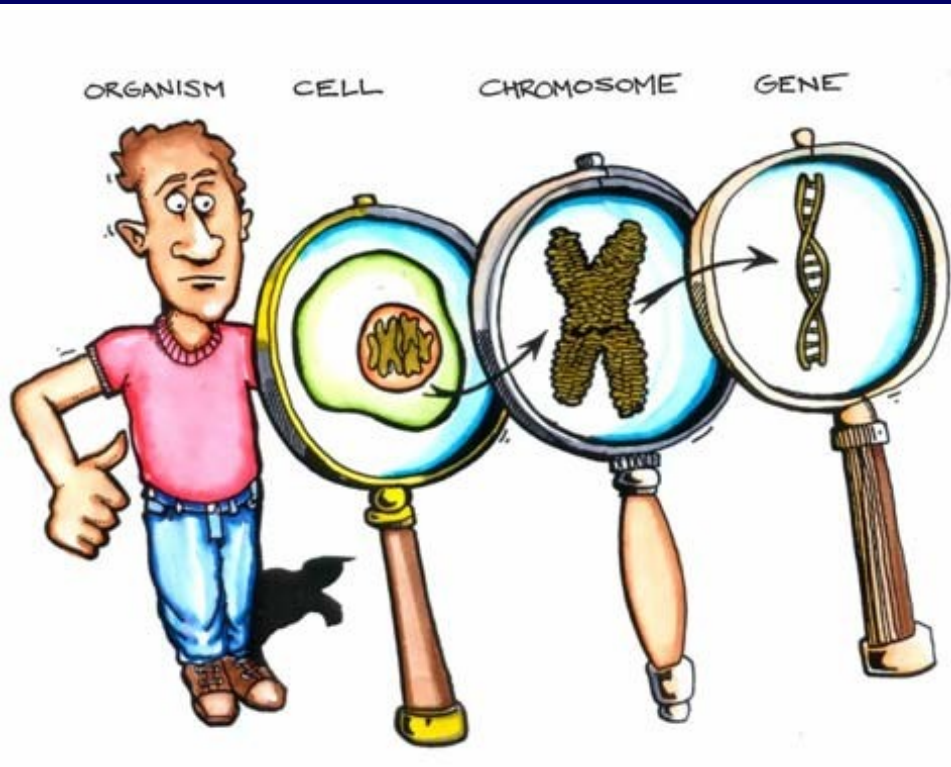
Email: raghava@imtech.res.in   **http://crdd.osdd.net/**

http://ww.imtech.res.in/raghava/

Hierarchy in Biology

Atoms

Molecules

Macromolecules

Organelles

Cells

Tissues

Organs

Organ Systems

Individual Organisms

Populations
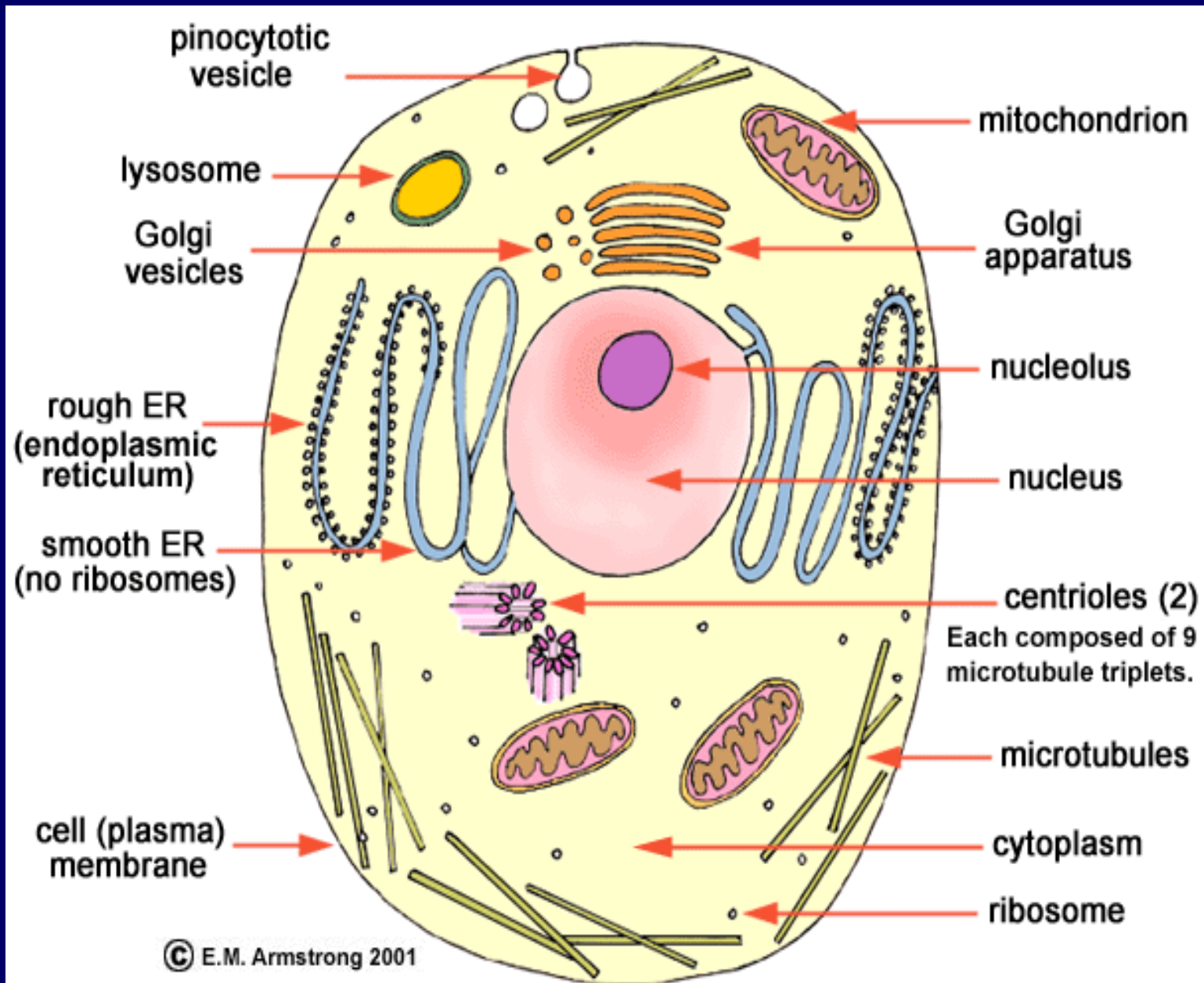
Communities

Ecosystems

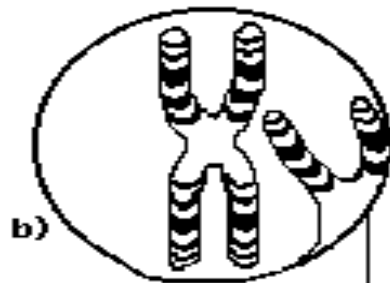Biosphere

Cells and DNA

- Human body contains ~100 trillion cells

- Each cell contains 23 pairs of chromosomes (= **genome**)

- Chromosomes contain DNA
  - DNA is made of 4 nucleotide bases (Adenine, Guanine, Cytosine & Thymine) = AGCT sequence

- Every cell (except a few) in an individual contains the same exact genome

Animal cell

pinocytotic vesicle

mitochondrion

lysosome

Golgi vesicles

Golgi apparatus

rough ER (endoplasmic reticulum)

nucleolus

nucleus

smooth ER (no ribosomes)

centrioles (2)
Each composed of 9 microtubule triplets.

cell (plasma) membrane

microtubules

cytoplasm

ribosome

© E.M. Armstrong 2001

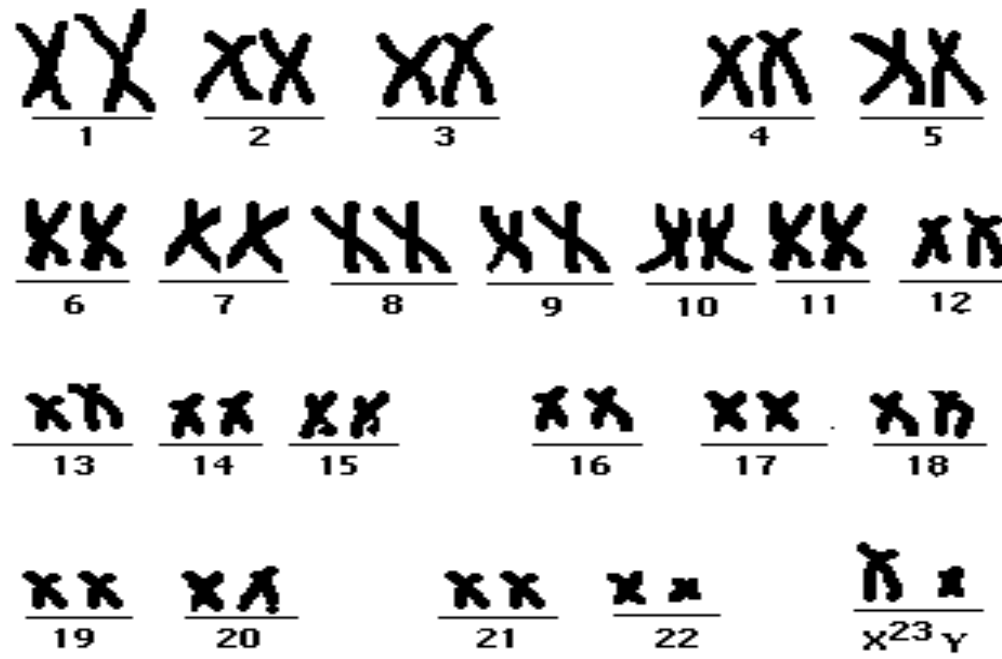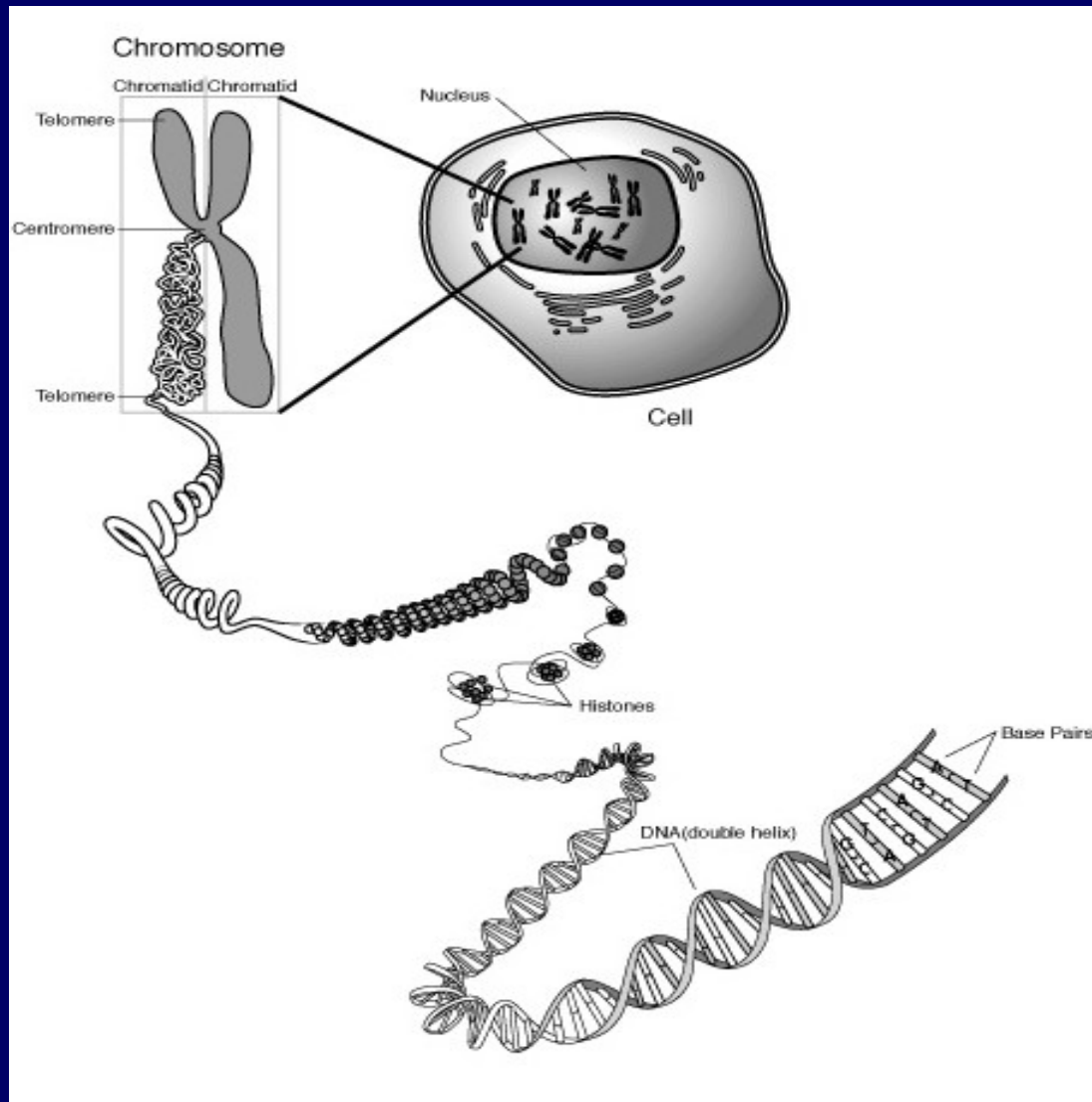# Human Chromosomes

HUMAN CHROMOSOMES
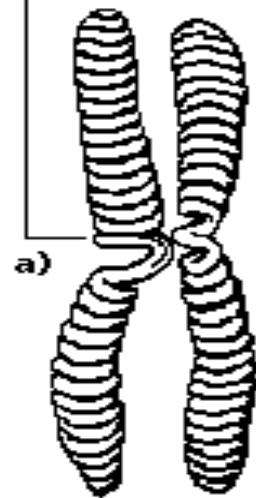
Chromosomes and DNA

Genome Annotation

The Process of Adding Biology Information and

Predictions to a Sequenced Genome
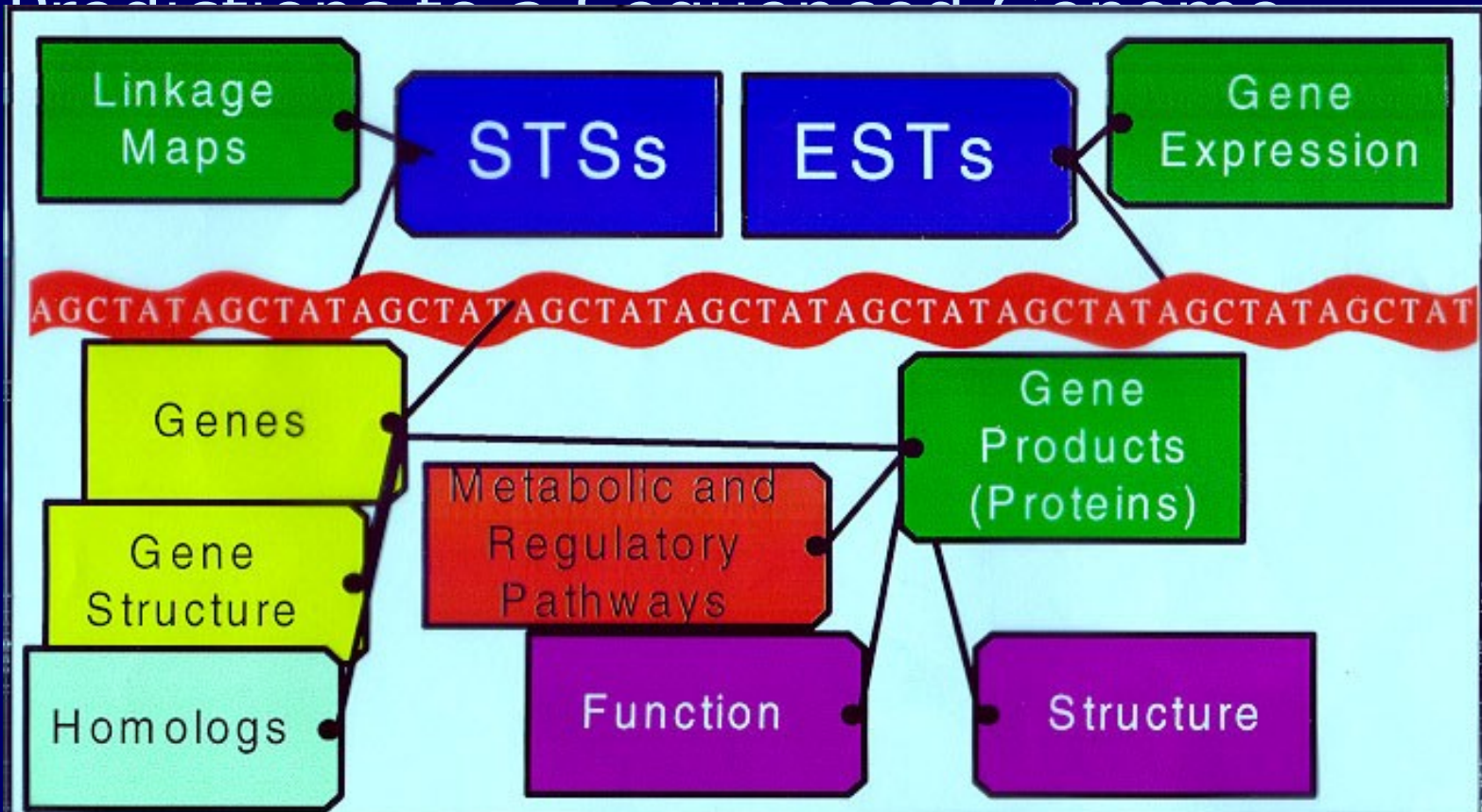
WHOLE GENOME SEQUENCING

Genome annotation tools at IMTECH

- **Protein Structure prediction servers**

- **Servers for predicting function of proteins**

- **Servers for designing epitope based vaccine**

- **Genome annotation**

- **Molecular Interactions & Modifications**

- **Designing of Therapeutic Molecules**

- **Computer Aided Drug Design**

http://www.imtech.res.in/raghava/

Hierarchical shotgun sequencing

Genomic DNA
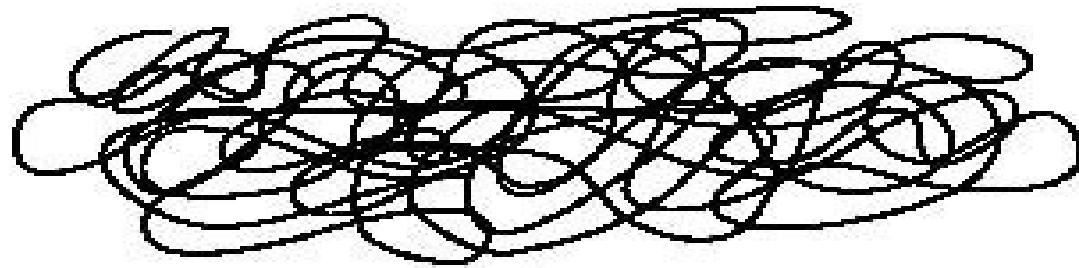
BAC library

Organized mapped large clone contigs
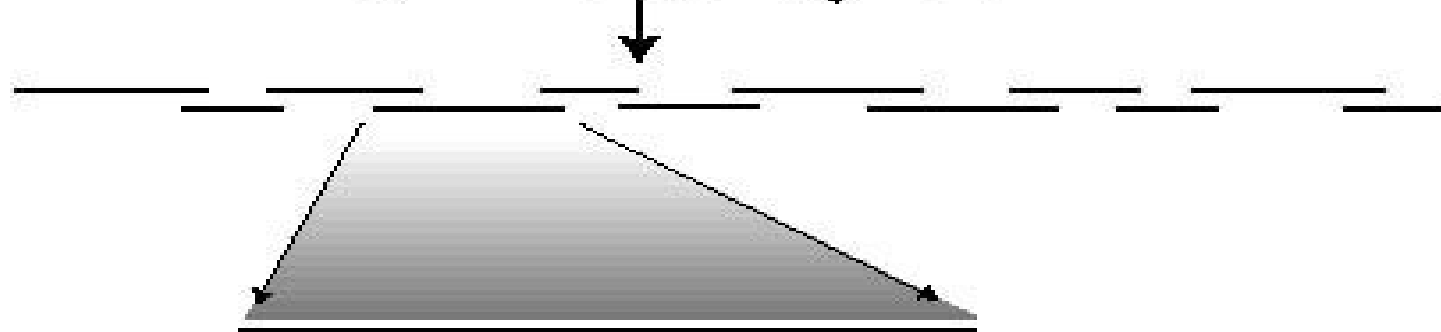
BAC to be sequenced

Shotgun clones

Shotgun sequence
```
...ACCGTAAATGGGCTGATCATGCTTAAA
          TGATCATGCTTAAACCCTGTGCATCCTACTG...
```

Assembly
```
...ACCGTAAATGGGCTGATCATGCTTAAACCCTGTGCATCCTACTG...
```
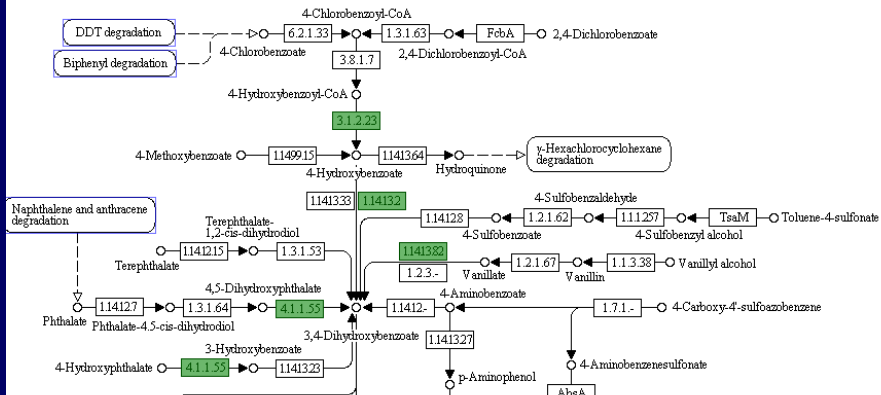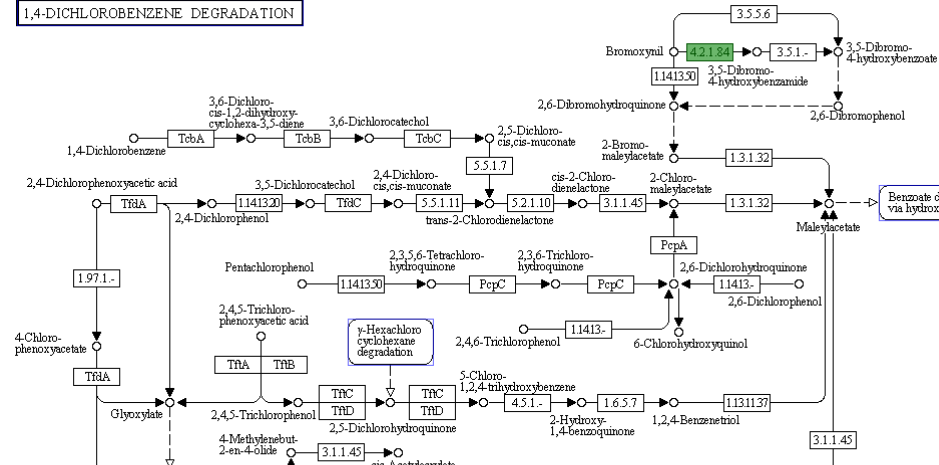
Journal of Bacteriology

Genome Sequence of the Nitroaromatic Compound-Degrading Bacterium Burkholderia sp. Strain SJ98

Shailesh Kumar, Surendra Vikram and Gajendra Pal Singh Raghava
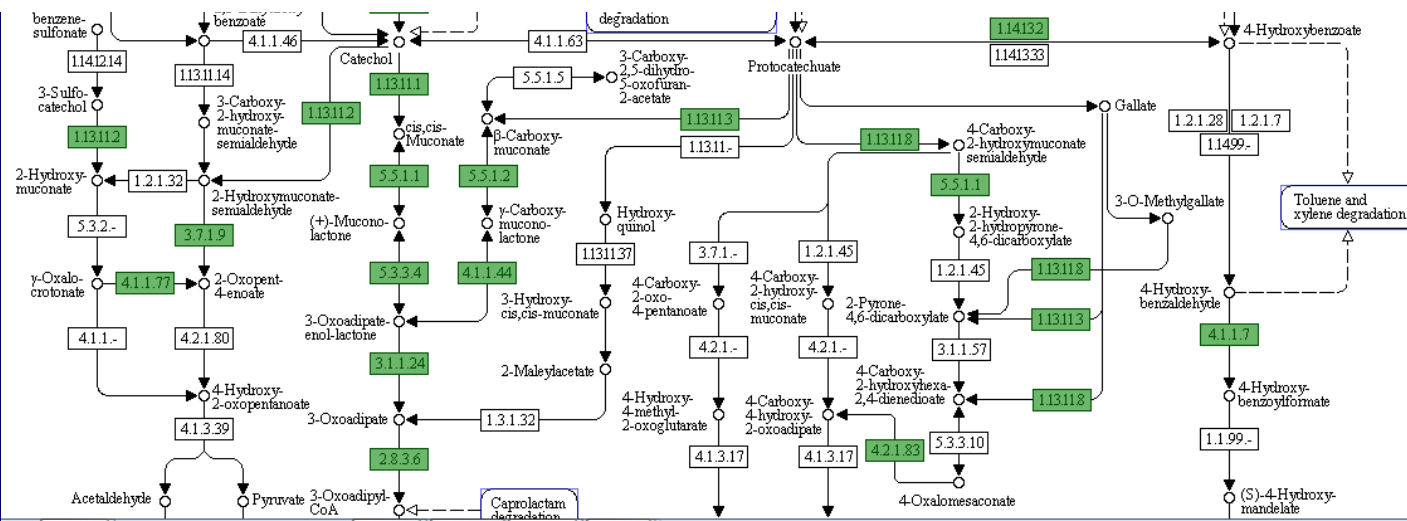
Genome assembly and annotation done at IMTECH

- *Burkholderia sp.* SJ98 (Kumar *et al.* 2012).
- *Debaryomyces hansenii* MTCC 234 (Kumar *et al.* 2012).
- *Imtechella halotolerans* K1$_T$ (Kumar *et al.* 2012).
- *Marinilabilia salmonicolor* JCM 21150$_T$ (Kumar *et al.* 2012).
- *Rhodococcus imtechensis sp.* RKJ300 (Vikram *et al.* 2012).
- *Rhodosporidium toruloides* MTCC 457 (Kumar *et al.* 2012).

Concept of Drug and Vaccine

- **Concept of Drug**
  - Kill invaders of foreign pathogens
  - Inhibit the growth of pathogens
- **Concept of Vaccine**
  - Generate memory cells
  - Trained immune system to face various existing disease agents

The Future of Genomics in Medicine
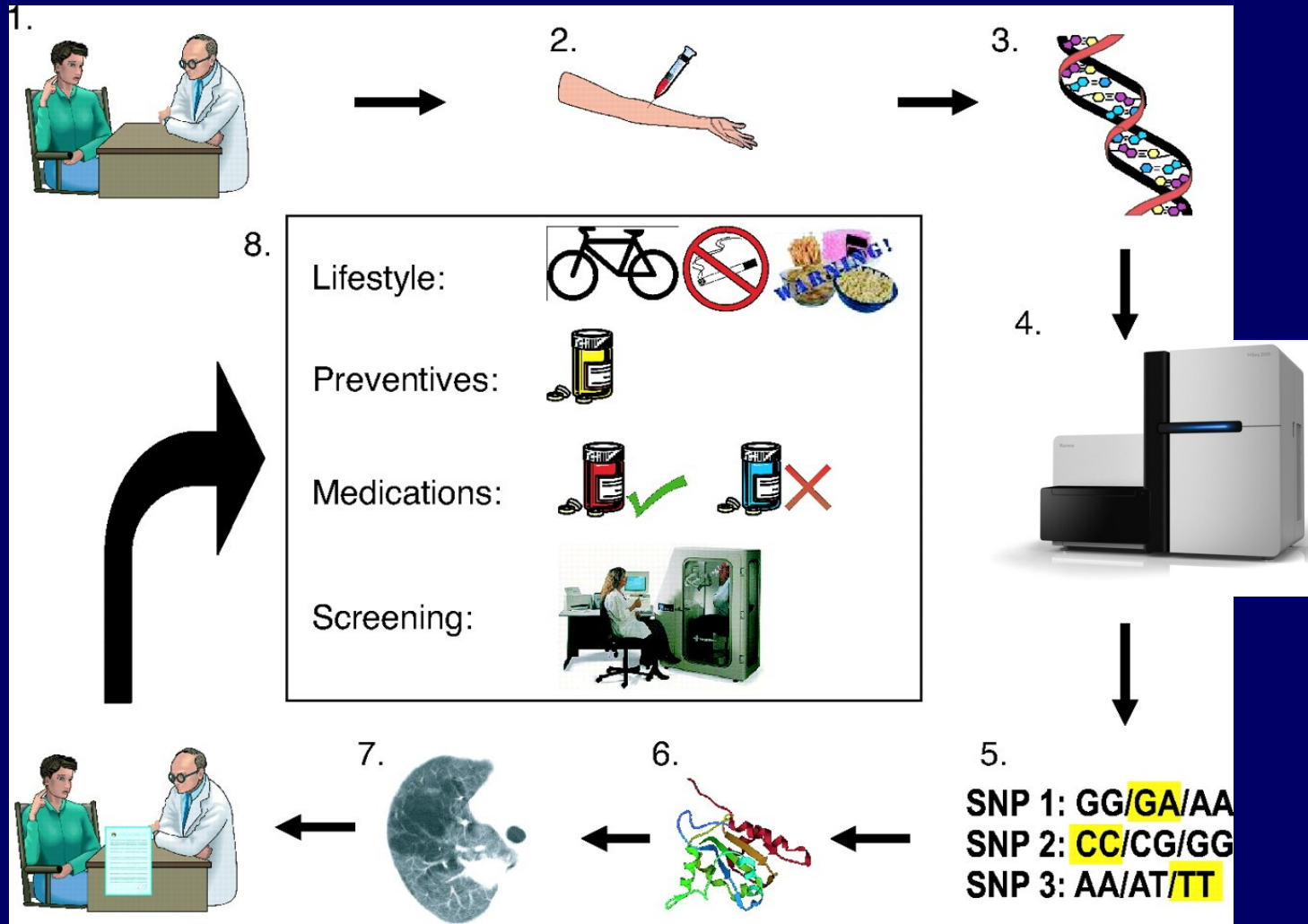
Antibiotics and drug resistance

Drug Discovery and Design

History of Drug/Vaccine development

- **Plants or Natural Product**
  - Plant and Natural products were source for medical substance
  - Example: foxglove used to treat congestive heart failure
  - Foxglove contain digitalis and cardiotonic glycoside
  - Identification of active component
- **Accidental Observations**
  - Penicillin is one good example
  - Alexander Fleming observed the effect of mold
  - Mold(Penicillium) produce substance penicillin
  - Discovery of penicillin lead to large scale screening
  - Soil micoorganism were grown and tested
  - Streptomycin, neomycin, gentamicin, tetracyclines etc.

Drug Discovery and Design

- **Chemical Modification of Known Drugs**

  - Drug improvement by chemical modification

  - Pencillin G -> Methicillin; morphine->nalorphine

- **Receptor Based drug design**

  - Receptor is the target (usually a protein)

  - Drug molecule binds to cause biological effects

  - It is also called lock and key system

  - Structure determination of receptor is important

- **Ligand-based drug design**

  - Search a lead ocompound or active ligand

  - Structure of ligand guide the drug design process

process

GENOMICS, PROTEOMICS & BIOPHARM.

Potentially producing many more targets and "personalized" targets

Identify disease
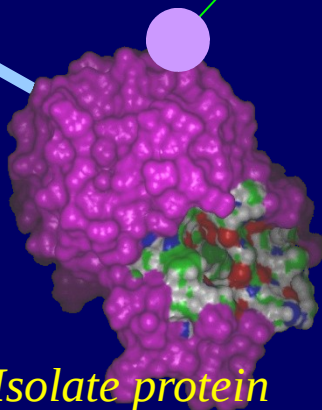
HIGH THROUGHPUT SCREENING

Screening up to 100,000 compounds a day for activity against a target protein

VIRTUAL SCREENING

Using a computer to predict activity

Isolate protein

COMBINATORIAL CHEMISTRY

Rapidly producing vast numbers of compounds
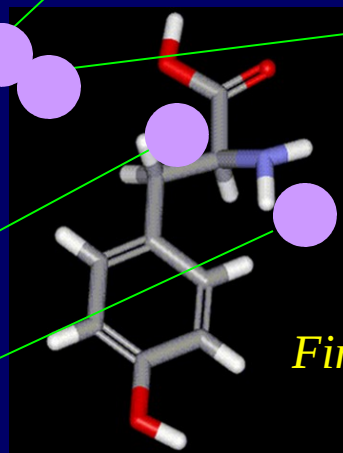
Find drug

MOLECULAR MODELING

Computer graphics & models help improve activity

IN VITRO & IN SILICO ADME MODELS

Tissue and computer models begin to replace animal testing

Preclinical testing

Drug Design based on Bioinformatics Tools

- **Detect the Molecular Bases for Disease**
  - Detection of drug binding site
  - Tailor drug to bind at that site
  - Protein modeling techniques
  - Traditional Method (brute force testing)

- **Rational drug design techniques**
  - Screen likely compounds built
  - Modeling large number of compounds (automated)
  - Application of Artificial intelligence
  - Limitation of known structures

Important Points in Drug Design based on Bioinformatics Tools

- **Application of Genome**
    - 3 billion bases pair
    - 30,000 unique genes
    - Any gene may be a potential drug target
    - ~500 unique target
    - Their may be 10 to 100 variants at each target gene
    - 1.4 million SNP
    - 10^{200} potential small molecules

The amount of fund required depends on the success rate at the clinical trial stage
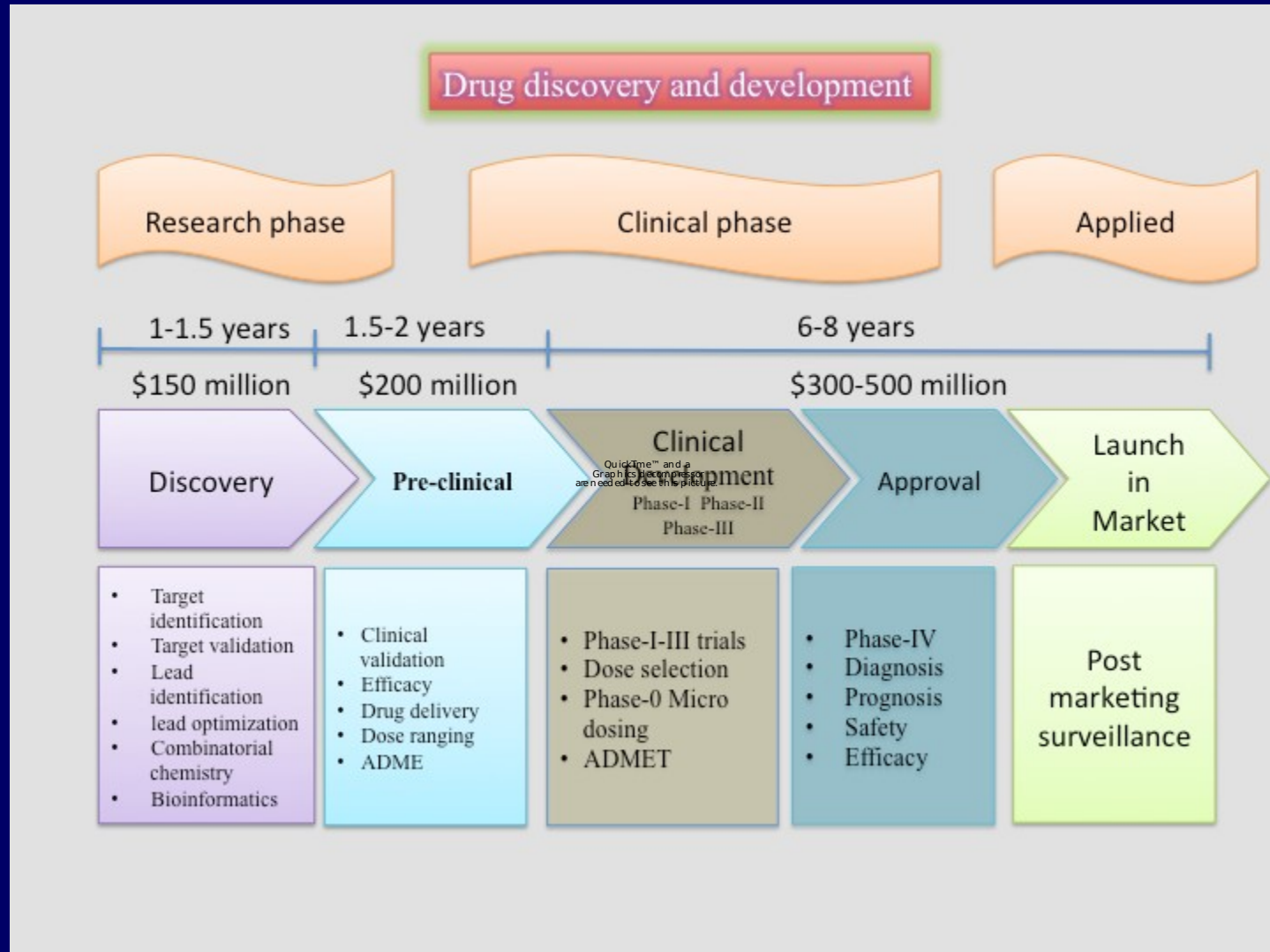
An overview of the workflow of *in silico* drug designing process

Identification of drug target protein

Chemical database/Literature → Inhibitors with known bioactivity → Select best inhibitors

Drawing Chemical Structure

Descriptors Calculation, Analysis and Optimization

Building QSAR models

Generate Analogues

Calculate descriptors

Chemical Database/Libraries → Database Mining using best models

Virtual screening

Active compounds/Analogues

Docking of compounds at the binding site

Prediction of drug-like and ADMET properties

Select Best-Fit

Drug-like inhibitors against Potential target

Software Development for Drug Discovery

Importance of Open Source for Drug Discovery

- Discovery of Drug by Public for Public

- Drugs for Disease Specific to Developing Countries (like India)

- Development of Drugs for diseases of poor persons

- Process of Discovery will be fast (few to many contributors)

- Academic institutes/universities and small industry may afford

Examples of open source software

- Operating Systems
  - Linux
  - FreeBSD, OpenBSD, and NetBSD
- Internet
  - Apache (> 50% of the world's web servers)
  - BIND: DNS for the entire Internet.
  - Sendmail (Most email servers)
  - OpenSSL (standard for secure communication)

- Programming Tools
  - Languages (Perl, Python, PHP)
  - GNU compilers and tools (GCC, Make)

Challenges in Drug Discovery

Computational Resources for Drug Discovery

CSIR

Home | OSDD | Raghava | News | Rewards | Challenges | OSDDpub | Forum | Indipedia | Drugpedia | FAQ | License | Register ...

Search [   ]

Target Identification

Genome Annotation
Proteome Annotation
Potential Targets
Protein Structure

Virtual Screening

QSAR Techniques
Docking & QSAR
Chemoinformatics
siRNA/miRNA

Drug Design

Lead Optimization
Pharmainformatics
ADMET
Clinical Informatics

How to Contribute?

Expermentalists
Virtual Trainees/Jobs
Software Developers

Computational Resources

Library Interfaces
Meta Servers
Publishing Document
Data on *M.tb.*

**** Prediction of cytochrome P450 isoform responsible for metabolizing a drug molecule ****

Computational Resources for Drug Discovery

OSDD Forum is an initiative with a vision to provide affordable healthcare to the developing world. The OSDD concept aims to synergize the power of genomics, computational technologies and facilitate the participation of young and brilliant talent from Universities and industry. It seeks to provide a global platform where the best brains can collaborate and collectively endeavor to solve the complex problems associated with discovering novel therapies for neglected diseases like Tuberculosis.

CRDD (Computational Resources for Drug Discovery) is an important module of the *in silico* module of OSDD. The CRDD web portal provides computer resources related to drug discovery on a single platform. Following are major features of CRDD:

- CRDD provides computational resources for researchers in the field of computer-aided drug design.
- CRDD allows users to discuss their problem with other members.
- CRDD gives equal opportunity to those willing to solve these problems.
- CRDD Wiki maintain wikipedia related to drug discovery.
- Contributors may host their database or web server on CRDD portal.

Thus, CRDD provides a platform for researchers having limited resources.

MetaPred:A webserver for the Prediction of Cytochrome P450 Isoform responsible for Metabolizing a Drug Molecule

Toxipred |KiDoQ | GDoQ | NPTOPE |KetoDrug |CRDD |OSDD |IMTECH | Raghava

please cite:: **Prediction of cytochrome P450 isoform responsible for metabolizing a drug molecule BMC Pharmacology**

» Home
» Submit
» Algorithm
» Developers
» Contact Us
» Help
» Dataset

Cytochrome P450 enzymes (CYPs) are a multi gene family of heme-containing isoenzymes that are involved in oxidative metabolism of drug, steroids and carcinogens. About sixty CYPs are reported in human genome, but more than 90% of all therapeutic drugs are metabolized by five isoforms i.e. CYP1A2, CYP2C9, CYP2C19, CYP2D6 and CYP3A4.

MetaPred Server predict metabolizing CYP isoform of a drug molecule/substrate, based on SVM models developed using CDK descriptors.This server will be helpful for researcher working in the field of drug discovery.This study demonstrates that it is possible to develop free web servers in the field of chemoinformatics. This will encourage other researchers to develop web server for public use, which may lead to decrease the cost of discovering new drug molecules. In the following flow digaram we have given the example of CYP3A4, how this study will be helpful in drug design.
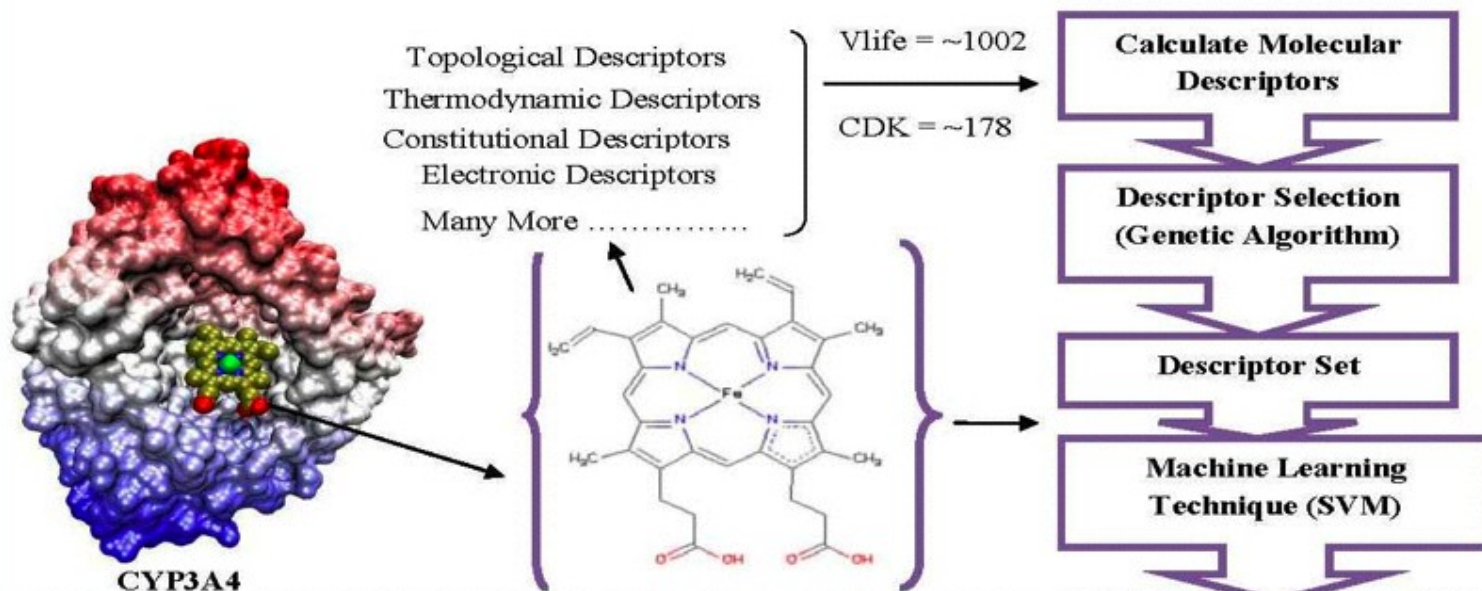
Home Page of HIVcoPRED

Server for prediction of HIV coreceptor usage

Home Page of HIVcoPRED

***** **Reference:** Kumar, R. and Raghava, GPS (2013) Hybrid Approach for Predicting Coreceptor Used by HIV-1 from Its V3 Loop Amino Acid Sequence. PLoS ONE 8(4): e61437 ******

CPPsite: a webSite for Cell Penetrating Peptides

Cytoplasm

Nucleus

CPP MEDIATED CARGO DELIVERY
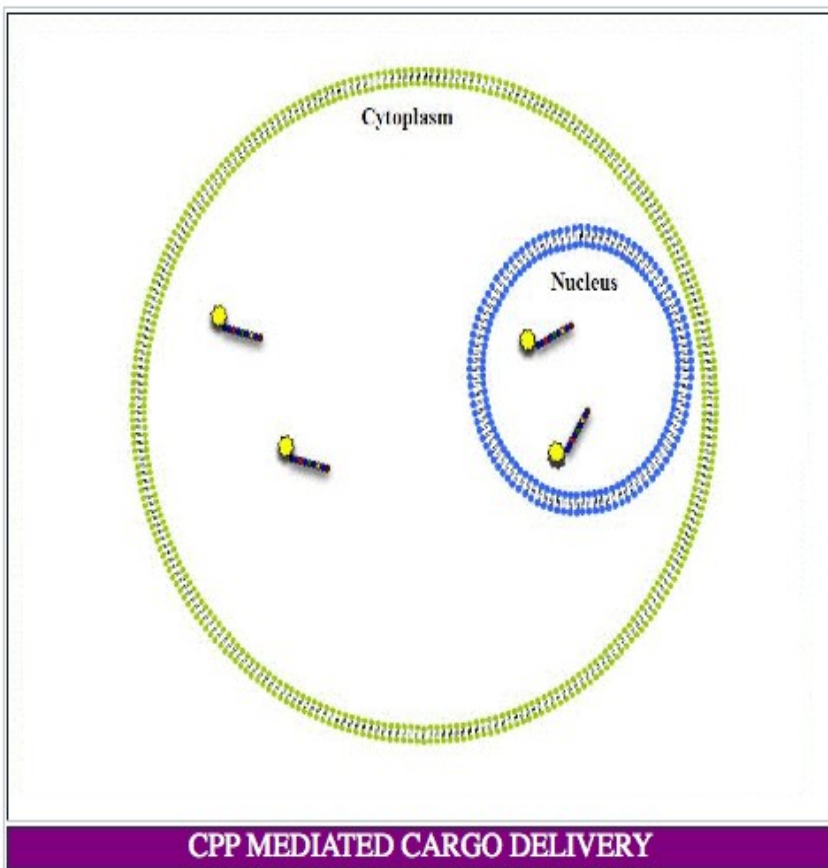
CPPsite: CPPsite is a database of experimentally validated Cell Penetrating Peptides (10-30 amino acids).

Importance of CPPsite: CPPs have tremendous therapeutic applications. These are widely used to promote intracellular uptake of conjugated cargos (nucleic acids, peptide nucleic acids, proteins, drugs, liposomes etc.) and thus play role to overcome the problem of poor delivery and low bioavailability of therapeutic molecules. CPP conjugated drugs when delivered *in vivo* have s hown promising results with high efficacy. Many CPP-conjugated compounds are under clinical trials. CPPsite database provides comprehensive information on CPPs, which may be helpful to scientific community working in the area of peptide based drug discovery.

What type of information it has: CPPsite database's current version contains comprehensive information of 843 CPPs with multiple entries in terms of peptide sequence, source/origin, localization, uptake efficiency, uptake mechanism, hydrophobicity, charge etc.

Is it a manually curated database: Yes, we have collected and compiled all the information from published literature. In addition, we have also generated structural information of CPPs. We predicted tertiary and secondary structure of these peptides using PepStr and DSSP.

Work in Progress

1. Prediction of CPP   2. Designing CPP 3. Scanning in proteins

Computer-Aided Drug Discovery
Searching Drug Targets: Bioinformatics

Genome Annotation

FTGpred: Prediction of Prokaryotic genes
EGpred: Prediction of eukaryotic genes
GeneBench: Benchmarking of gene finders
SRF: Spectral Repeat finder

Comparative genomics

GWFASTA: Genome-Wide FASTA Search
GWBLAST: Genome wide BLAST search
COPID: Composition based similarity search
LGEpred: Gene from protein sequence

Subcellular Localization Methods

PSLpred: localization of prokaryotic proteins
ESLpred: localization of Eukaryotic proteins
HSLpred: localization of Human proteins
MITpred: Prediction of Mitochndrial proteins
TBpred: Localization of mycobacterial proteins

Prediction of drugable proteins

Nrpred: Classification of nuclear receptors
GPCRpred: Prediction of G-protein-coupled receptors
GPCRsclass: Amine type of GPCR
VGIchan: Voltage gated ion channel
Pprint: RNA interacting residues in proteins
GSTpred: Glutathione S-transferases proteins

Protein Structure Prediction

APSSP2: protein secondary structure prediction
Betatpred: Consensus method for β-turns prediction
Bteval: Benchmarking of β-turns prediction
BetaTurns: Prediction of -turn types in proteins
Turn Predictions: Prediction of α/ β/γ -turns in proteins
GammaPred: Prediction of-turns in proteins
BhairPred: Prediction of Beta Hairpins
TBBpred: Prediction of trans membrane beta barrel proteins
SARpred: Prediction of surface accessibility (real accessibility)
PepStr: Prediction of tertiary structure of Bioactive peptides
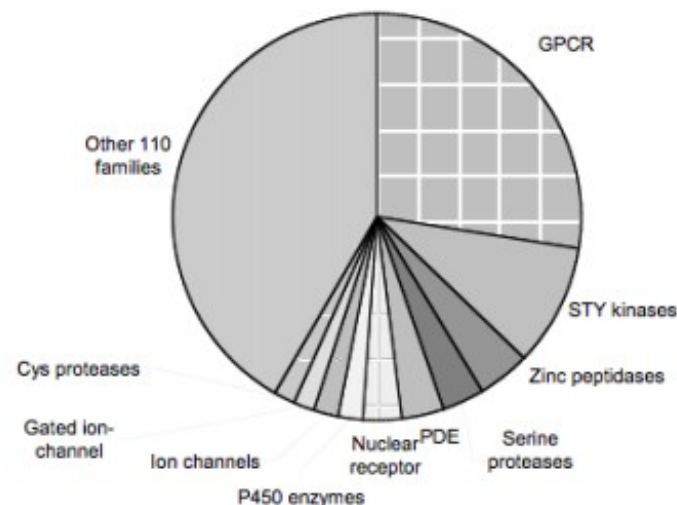
Understanding Immune System

Major impact on public health and incidence of infectious diseases
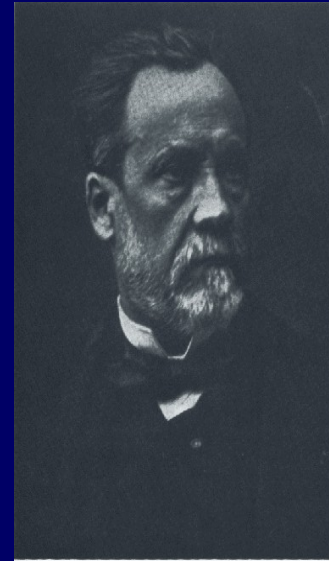
➤ **E. Jenner, the pioneer of vaccination in the Western world,**
➤ **Lady M. Montagu, an early advocate of smallpox inoculation**
➤ **L. Pasteur, who discovered attenutation**

Human Vaccines against pathogens

Organism	Type	Vaccine Type	Year
Variola virus	Virus	Live	1798
Rabies virus	Virus	Inactivated	1885
Salmonella typhi	Bacteria	Live	1896
Vibrio cholerae	Bacteria	Inactivated	1896
Yersinia pestis	Bacteria	Inactivated	1897
Corynebacterium diphtheriae	Bacteria	Toxoid	1923
Bordetella pertussis	Bacteria	Acellular	1926
Clostridium tetani	Bacteria	Toxoid	1927
Mycobacterium tuberculosis	Bacteria	Live	1927
Yellow fever virus	Virus	Live	1935
Influenza virus type A	Virus	Inactivated	1936
Influenza virus type B	Virus	Inactivated	1936
Coxiella burnetii	Bacteria	Inactivated	1938
Rickettsia prowazekii	Bacteria	Inactivated	1938
Rickettsia rickettsii	Bacteria	Inactivated	1938
Central European encephalitis virus	Virus	Inactivated	1939
Poliovirus types 1, 2, and 3	Virus	Inactivated/Live	1962
Measles virus	Virus	Live	1963
Mumps virus	Virus	Live	1967
Rubivirus	Virus	Live	1969
Staphylococcus aureus	Bacteria	Staphage lysate	1976
Streptococcus pneumoniae	Bacteria	Polysaccharide	1977
Human adenovirus types 4 and 7	Virus	Live	1980
Neisseria meningitidis	Bacteria	Polysaccharide	1981
Hepatitis B	Virus	Recombinant	1986
Haemophilus influenzae	Bacteria	Conjugate	1987
Hantaan virus	Virus	Inactivated	1989
Japanese encephalitis virus	Virus	Inactivated	1992
Varicella-zoster virus	Virus	Live	1994
Hepatitis A	Virus	Inactivated	1995
Escherichia coli	Bacteria	Inactivated	1995
Junin virus	Virus	Live	1996
Bacillus anthracis	Bacteria	Adsorbed	1998
Borrelia burgdorferi	Bacteria	Recombinant	1998

History of Immunization

- Children protected who recovered from smallpox
- Immunity induce, a process known as variolation
- Variolation spread to England and America
- Stopped due to the risk of death
- Edward Jenner found that protection against smallpox
- Inoculation with material from an individual infected with cowpox
- This process was called vaccination (cowpox is vaccina)
- Inoculum was termed a vaccine
- Protective antibodies was developed

Traditional Vaccine to Epitope Based Vaccines in Genomics Era

- **Concept of vaccine and Drug**
  - **Drug:** Kill invaders/pathogens and/or Inhibit the growth of pathogens
  - **Vaccine:** Trained immune system to face various existing disease agents

- **Type of Vaccines**
  - Whole Organism of Pathogen (MTb, 4000 proteins)
  - Target proteins/antigens which can activate immune system
  - Subunit Vaccine, Epitope Based Vaccines (T & B cell epitopes)

- **Limitations of present methods of subunit vaccine design**
  - Developed for one or two MHC alleles (not suitable for large population)
  - Do not consider pathways of antigen processing
  - Difficult to detect experimentally known epitopes

- **Initiatives taken by Bioinformatics Centre at IMT, Chandigarh India**

Different vaccine design strategies
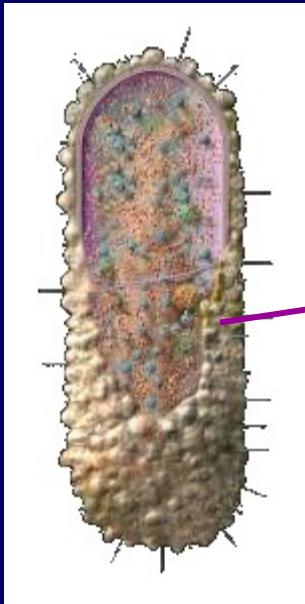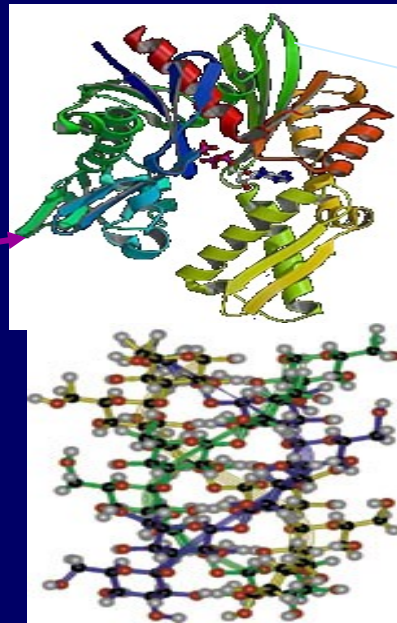
WHOLE ORGANISM → **Purified Antigen** → **Epitopes (Subunit Vaccine)**

attenuated

T cell epitope

Multiple layers of the immune system

Pathogens

Skin

Biochemical barriers

Phagocyte

Innate immune response

Lymphocytes

Adaptive immune response
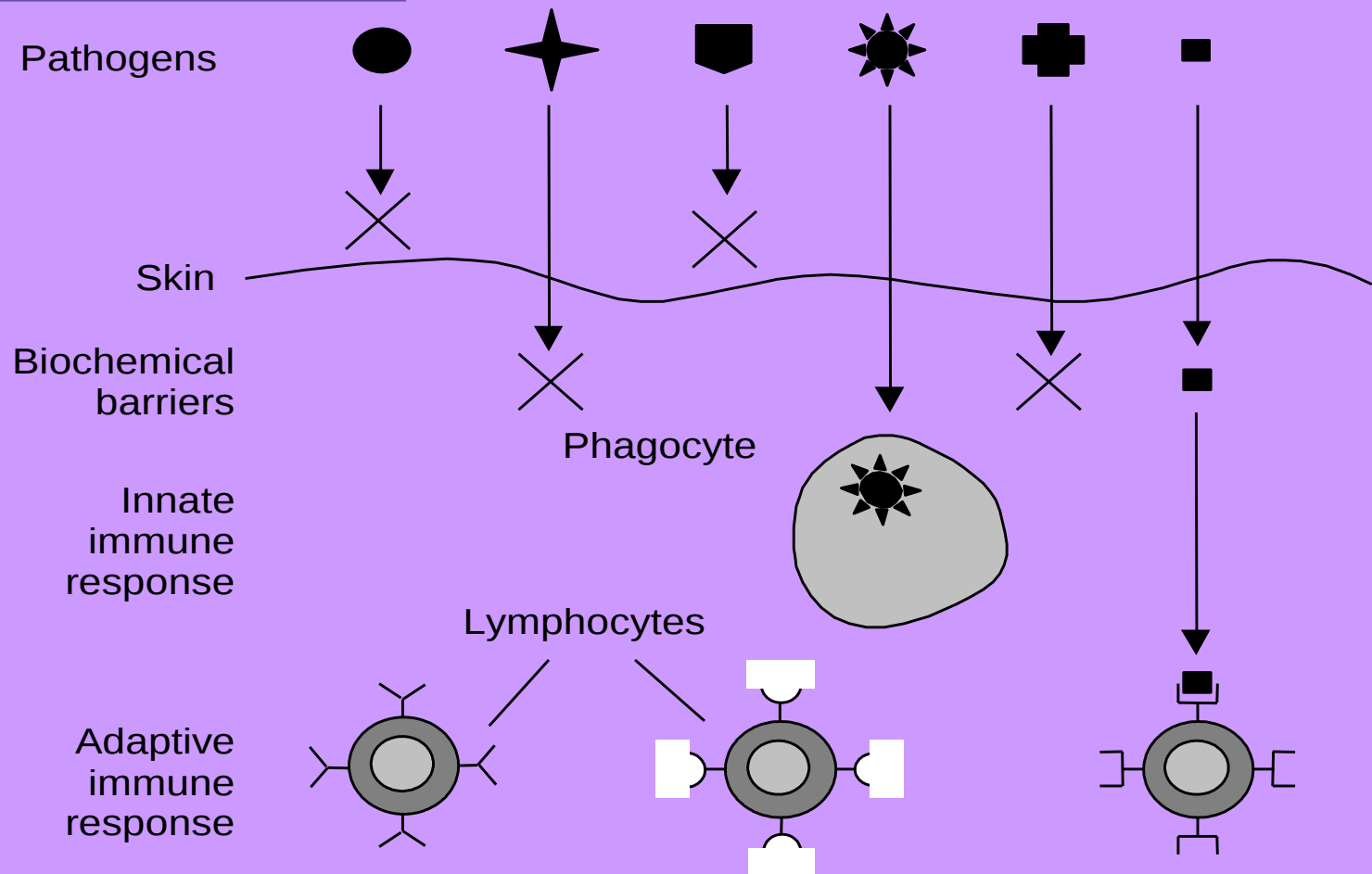
Immune Defense and Long Term Protection

Disease Causing Agents
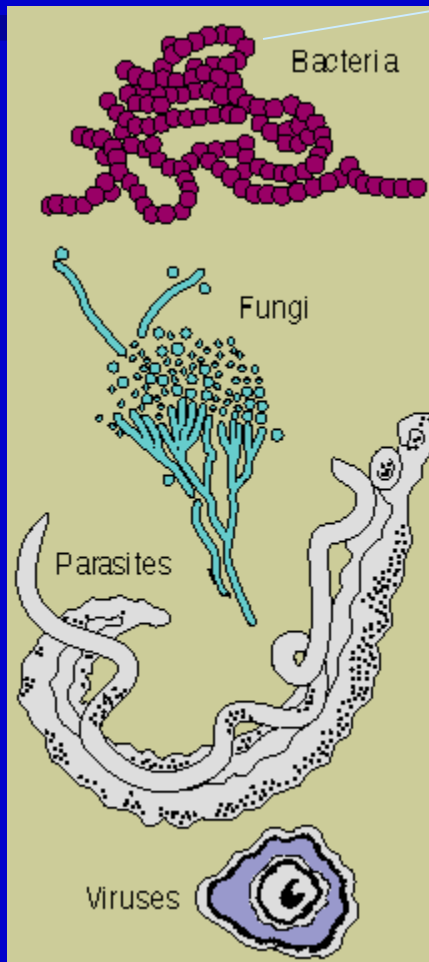
Pathogens/Invaders
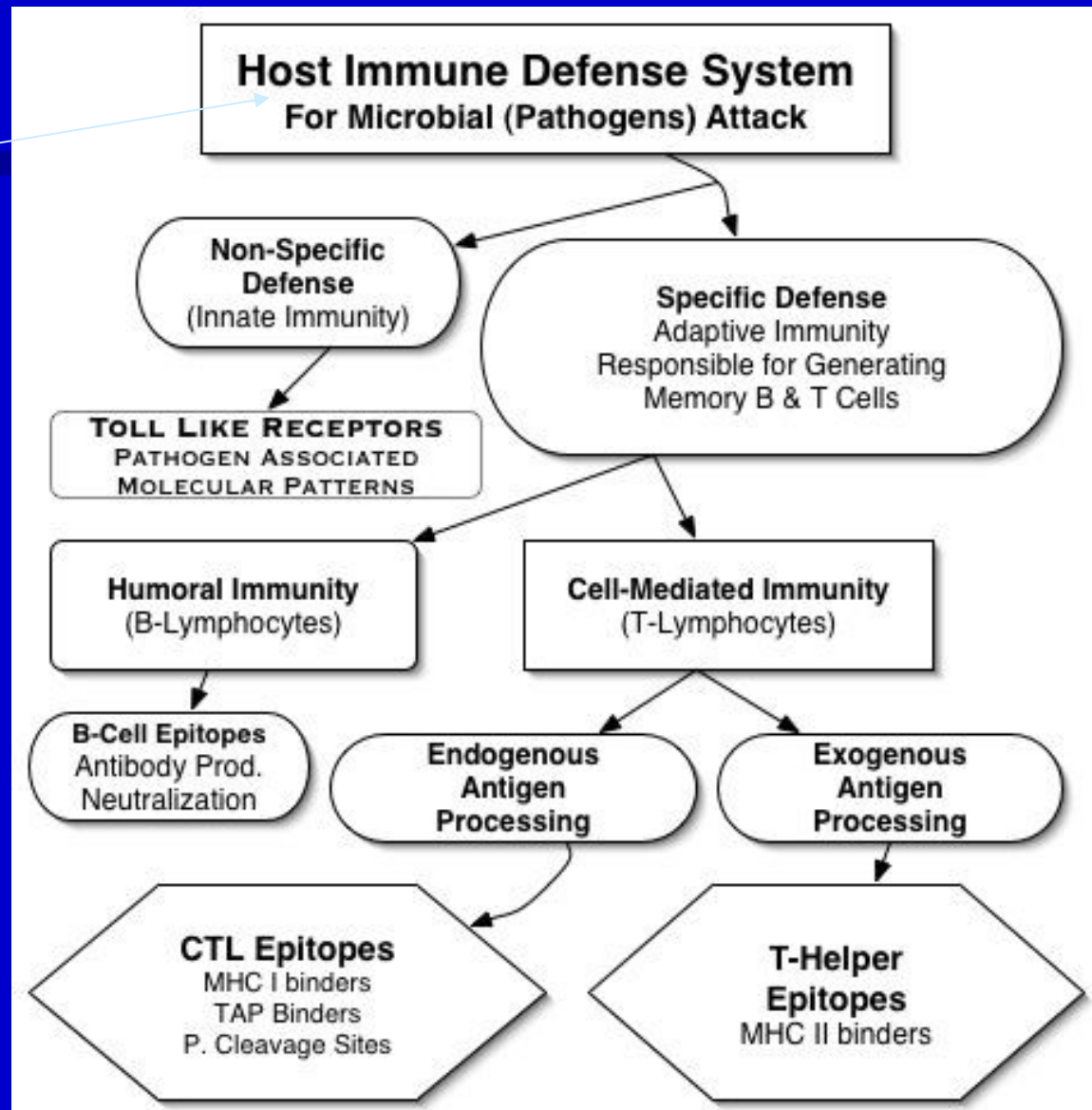
Challenge in Vaccine Design
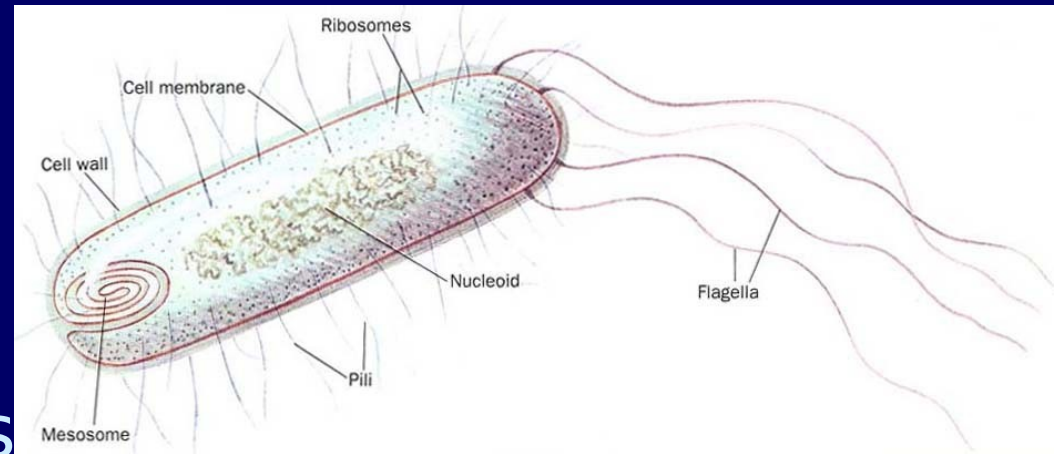
- 10^{12} types of B-cells
- Lack of effective vaccine against HIV, Cancer, M.Tb., Malaria
- Universal vaccines (MHC alleles)
- Bacterial genomes
  - Around 10 million bases
  - 4000 genes, 4000 proteins/antigens & cell
  - Selection of antigen/vaccine target
  - Identification of antigenic regions (epitopes)

OSDD ⬡ LINUX

CONNECT Next

A Custor
fo

- General Inoformation ⊕
- Software Packages ⊕
- Install/Download ⊕
- Service to Community ⊕
- OSDDlinux Online ⊕
- Important Resources ⊕

Install/Download ⊖

- Live DVD/USB
- Full Installation
- Virtual Box
- On existing machine
- Package Repository
- Upgrade/New Packages

General Inoformation ⊖

- Major Features
- Installation Guide
- Users Guide
- Drug Discovery manual
- GPSR manual
- Required Package
- List of Packages

Service to Community ⊖

- Command Mode
- Web Services
- Galaxy Portal
- GUI-based Software

Software Packages ⊖

- Bioinformatics
- Vaccine Informatics
- Drug Informatics
- Biotherapeutics
- Analysis of NGS data
- Education & Research
- Basic Scripts

crdd.osdd.net/raghava/web.php

Apps Google State Bank of India Home – PubMed – N Yahoo! Google Maps W Wikipedia YouTube News Popular

Raghava's group

| Protein Structure | Protein Function | Vaccine Design | Genome Annotation | Biological Database | Therapeutic Peptides | Molecular Interaction |

Home Page

Web Services

Research

Services Offered

Group Members

Miscellaneous

Contact Us

Software name	Description
AntiCP	Prediction and design of anticancer peptides.
Toxinpred	Prediction and designing of toxic/non-toxic peptides.
AntiBP	Mapping of antibacterial peptides in a protein sequence.
AntiBP2	Advanced server for predicting antibacterial peptides with high precision.
CPPsite	CPPsite is a database of experimentally validated cell penetrating Peptides
CellPPD	Computer-aided Designing of efficient cell penetrating peptides.
TumorHoPe	A comprehensive database experimentally characterized tumor homing peptides.
TumorHPD	Server dedicated for designing tumor homing peptides.
Destamp	Designing of stable antibacterial peptides
HEMOLYTIK	A repository of experimentally validated hemolytic and non-hemolytic peptides.
ParaPep	Compilation of experimentally validated anti-parasitic peptides and their structure

Chemoinformatics and Pharmacoinformatics

Web Server	Description
DrugMint	A Server for Identification of Drug-like Molecules
ABMPred	Prediction of AntiBacterial Compounds against MurA Enzyme
MDRIpred	Prediction of Inhibitor against Drug Resistant M.Tuberculosis
DMKpred	Prediction of Drug molecules for kinase protein
KiDoQ	Prediction of inhibition constant of a molecule against Dihydrodipicolinate synthase enzyme
TOXIpred	Prediction of aqueous toxicity of small chemical molecules in T. pyriformis.
MetaPred	Prediction of cytochrome P450 isoform responsible for metabolizing a drug molecule.
GDoQ	Model for prediction of GLMU inhibitors using QSAR and docking apprach.
KetoDrug	Binding affinity prediction of ketoxazole derivatives against fatty acid amide hydrolase.
WebCDK	Web Interface for CDK libraries
TLR4HI	SVM based model for computing inhbitors against human TLR4 (Toll like receptor).
DMKPred	A webserver for the prediction of binding of chemical molecules with specific kinases.
ntEGFR	Predicting and designing imidazothiazoles/pyrazolopyrimidines based inhibitors against wild/mutant EGFR.
CancerIn	Classification and designing of anti-cancer inhibitors.
EGFRpred	Prediction of inhibitor of anti-EGFR molecules of diverse class.
DiPCell	Designing of inhibitors against pancreatic cancer cell lines.
HIVfin	Prediction of fusion protein inhibitors against HIV.

Molecular Interactions

Software name	Description
ADPint	Prediction of ADP interacting residues in a protein.
ATPint	Identification of ATP binding sites in ATP-binding proteins.
DOMprint	SVM based model for predicting domain-domain interaction (DDI).
GlycoEP	Prediction of C-, N- and O-glycosylation site in eukaryotic proteins.
GlycoPP	Prediction of potential N-and O-glycosites in prokaryotic proteins.
GTPbinder	Identification of GTP binding residue in protein sequences.
MYCOprint	A tool fort exploration of the interactome of Mycobacterium tuberculosis.
NADbinder	Prediction of NAD binding proteins and their interacting residues.
Pprint	ANN based method for identification of RNA-interacting residues in a protein.
PreMieR	Identification of mannose interacting residues (MIRs) in protein sequences.
PROprint	Prediction of physical/functional interaction between two protein molecules.
RNApin	A server for the prediction of protein interacting nucleotides in RNA sequences.
tRNAmod	Prediction of post transcriptional modifications in transfer-RNA (tRNA) sequence.
VitaPred	Identification of different class of vitamin interacting residues in a protein.

Biological Databases

Database name	Description
MHCBN	A curated database of MHC-binding, Non-binding peptides and T-cell epitopes.
Bcipep	A database of B-cell epitopes.
HaptenDB	A database of hapten molecules that can not activate immune system.
PolysacDB	Compilation of antigenic polysaccharides found on surface of microbial organism.
TumorHope	A database of experimentally characterized tumor homing peptides.
AntigenDB	Information about a wide range of experimentally-validated antigens.
OXDBase	Compilation of oxygenases involved in the biodegradation on xenobiotic compounds.
HMRBase	A manually curated database of hormones and their Receptors.
CPPsite	Compilation of experimentally validated Cell Penetrating Peptides (10-30 amino acids).
BIAdb	Information about Benzylisoquinoline Alkaloid molecules
HIVsir	A manually curated database of anti-HIV siRNAs.
CCDB	Catalog of genes involved in the different stages of cervical carcinogenesis.
ProGlycProt	Repository of experimentally characterized eubacterial and archaeal glycoproteins.
NPACT	A database of plant derived natural compounds that exhibit anti-cancerous activity.
CancerDR	Compilation of anticancer drugs and their effectiveness against various cancer cell lines.
ccPDB	Compilation and creation of datasets from PDB for structural/functional annoation of proteins.
ParaPep	HIPdb is a manually curated database of experimentally validated antiparasite peptides.
EGFRindb	Collection of EGFR inhibitors from literature.
CancerPPD	Collection and compilation of experimentally validated anticancer peptides
PCMdb	Pancreatic cancer methylation database provides large scale collection of metylated genes.
HerceptinR	Information about assays performend to test sensitivity/resistance of Herceptin Antibody.
HemolytiK	A resource of experimentally tested hemolytic peptides.
CancerTope	A database of epitopes found in protein involved in cancer.
AHTPDB	AHTPDB is an ideal platform for complete & relevant information for large number of antihypertensive peptides.

Genome Annotations

Server Name	Description
FTG	Locating probable protein coding region in nucleotide sequence using FFT based algorithm.
GWBLAST	Genome wide similarity search using BLAST
GWFASTA	Genome Wise Sequence Similarity Search using FASTA.
EGPred	Prediction of gene (protein coding regions) in eukaryote genomes that includes introns/exons.
SVMgene	SVM based approach to identify the protein coding regions in human genomic DNA.
SRF	Find repeats through an analysis of the power spectrum of a given DNA sequence.
MyPattern	A program for detection of a 'motif' in DNA sequence using an exact search method.
GeneBench	A suite of datasets and tools for evaluating gene prediction methods.
FTGPred	A web server for predicting genes in a DNAsequence.
PHDcleav	Prediction of Human Dicer cleavage sites.
PolyApred	Prediction of polyadenylation signal (PAS) in human DNA sequence.
siRNAPred	Predicting actual efficacy of both 21mer and 19mer siRNAs with high accuracy.
ECGPred	Analsis of expresion data and correlation between gene expression and nucleotides composition of genes.
desiRam	Designing of highly efficient siRNA with minimum mutation approach
MARSpred	Discriminating between Mitochondrial and Cytosolic Aminoacyl tRNA Synthetases
Icaars	Identification & Classification of Aminoacyl tRNA Synthetases.
LGEpred	Prediction of correlation between amino acid residue and gene expression level.

Immunoinformatics or Vaccine Informatics

Software name	Description
T-Helper Epitopes or MHC/HLA Class II binders **(Adaptive Immunity, Exogenous Antigen)**	
MHCBN	A database of MHC-Binding, Non-binding peptides and T-cell epitopes.
ProPred	Identification of promiscuous MHC Class-II binding regions in an antigen sequence
HLA-DR4Pred	Identification of HLA-DRB1*0401(MHC class II alleles) binding peptides.
MHC	Matrix Optimization Technique for identification of binding core in MHC II binding peptides
MHC2pred	The MHC2Pred is an SVM based method for prediction of promiscuous MHC class II binding peptides.
MHCBENCH	Benchmarking of MHC binding peptide prediction algorithms.
FDR4	Prediction of binding affinity of HLA-DRB*0401 binders in an antigenic sequence.
IL4pred	In silico platform for designing and disovering of interleukin-4 inducing peptides.
IFnepitope	Designing of interferon-gamma inducing epitopes.
CTL Epitopes or MHC/HLA Class I binders **(Adaptive Immunity, Endogenous Antigens)**	
PROPRED1	Prediction of promiscuous binders for 47 MHC/HLA class I alleles using quantitative matrices;
Pcleavage	Identification of protesosomal cleavage sites in a protein sequence.
TPPred	Prediction of TAP binding peptides for understanding of peptide internalization to endoplasmic reticulum
CTLPred	A direct method for prediction of CTL epitopes.
nHLApred	This is a comprehensive method for prediction of MHC binding peptides or CTL epitopes of 67 MHC class alleles.
MMBPred	Prediction of mutated MHC class I binders in an antigen, having high affinity and promiscuousity.
HLAPRED	The method can identify and predict HLA (both class I & II) binding regions in an antigen sequence.
Linear & Conformational B-cell Epitopes	
BCIPEP	Collection & compilation of B-cell epitopes from literature
BCEPRED	Prediction of linear B-cell epitopes, using Physico-chemical properties
ABCPred	Mapping of B-cell epitope(s) in an antigen sequence, using artificial neural network.

Functional Annotation of Proteins

Server name	Description
NRpred	Prediction and classification of nuclear receptors, SVM models based on composition.
GPCRpred	Prediction of families and superfamilies of G-protein coupled receptors (GPCR)
ESLpred	Subcellular localization of the eukaryotic proteins using dipeptide compostion and PSI-BLAST.
PSLPred	Prediction of subcellular localization of bacterial proteins
BTXPred	It predicts bacterial toxins and their function from primary amino acid sequence.
GPCRsclass	This webserver predicts amine type of G-protein coupled receptors
Mitpred	Specifically trained to predict mitochondrial proteins with high accuracy
Oxypred	Classification and prediction of oxygen binding proteins.
VGIchan	Classification and prediction of proteins involved in voltage gated ion channels.
HSLpred	Subcellular localization of human proteins with high accuracy
DNAsize	Compute length of DNA or protein fragments from gel using a graphical method.
GSTpred	SVM-based method for predicting Glutathione S-transferase protein.
Mango	A server for predicting functional class of a protein.
LGEpred	Calculate correlation coefficient between amino acid residue and gene expression level.
NTXpred	Identification of neurotoxins their source and function from primary amino acid sequence.
VICMpred	Classification of bacterila proteins particularly virulent proteins
ALGpred	Prediction of allergenic proteins and mapping of IgE epitopes in antigens.
PseaPred	Prediction of proteins secreted by Malarial Parasite P. falciparum into infected-erythrocyte.
RSLPred	SVM based method for subcellular localizaton of rice proteins.
COPid	Composition based identification and classification of proteins.
ESLPred2	Advanced method for subcellular localization of eukaryotic proteins.
ISSpred	Identification of Inteins hiding in their protein sequences.
CyclinPred	CyclinPred is a SVM based prediction method to identify novel cyclins.

Proteins Structure Prediction

Web Server	Description
AlphaPred	A neural network based method for predicting alpha-turn in a protein.
APSSP2	Prediction of secondary structure of proteins from their amino acid sequence.
AR_NHPred	Identification of aromatic-backbone NH interaction in protein residues.
BetatPred	Statistical-based method for predicting Beta Turns in a protein.
Betatpred2	Prediction of Beta-turns with high accuracy using multiple sequence alignment.
BetaTurns	It predict different types of beta-turns (e.g., Types I/II/IV/VIII) in a protein.
BhairPred	Prediction of beta hairpins in proteins using ANN and SVM techniques.
CHpredicts	Prediction of CH-O, CH-PI interactions in backbone residues of a protein
GammaPred	Identification of gamma-turn containing residues in a given protein sequence.
PEPstr	Prediction of tertiary structure of small peptides (7 to 25 residues).
Proclass	Classification of proteins based on secondary structure contents.
PSA	Analyze the amino acid sequence and multiple sequence alignment of proteins.
RPFOLD	A fold recognition server for searching protein fold in PDB.
SARpred	ANN-model for redicting real-value of surface acessibility of protein residues.
TBBpred	This server predict Transmembrane Beta Barrel regions in a protein.
PEP2D	This server allows you to predict secondary structure of peptides.

Thanks