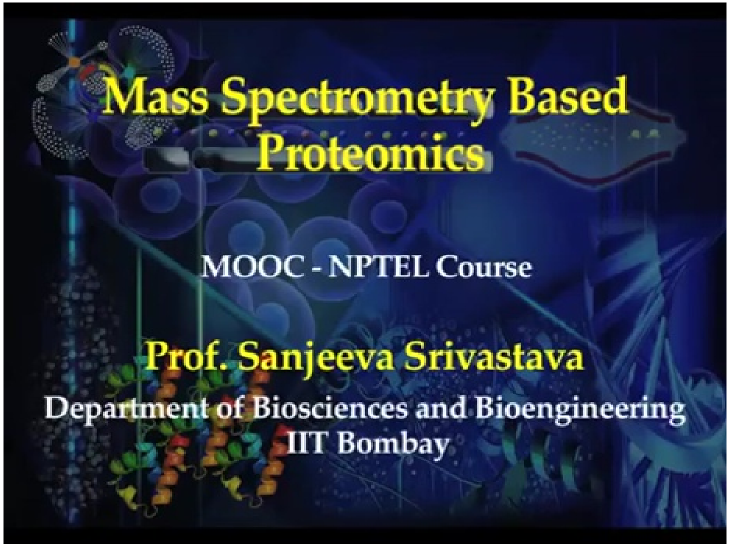
Mass Spectrometry Based Proteomics
Professor Sanjeeva Srivastava
Department of Biosciences and Bioengineering
Indian Institute of Technology, Bombay
Mod 04 Lecture Number 17

(Refer Slide Time 00:10)



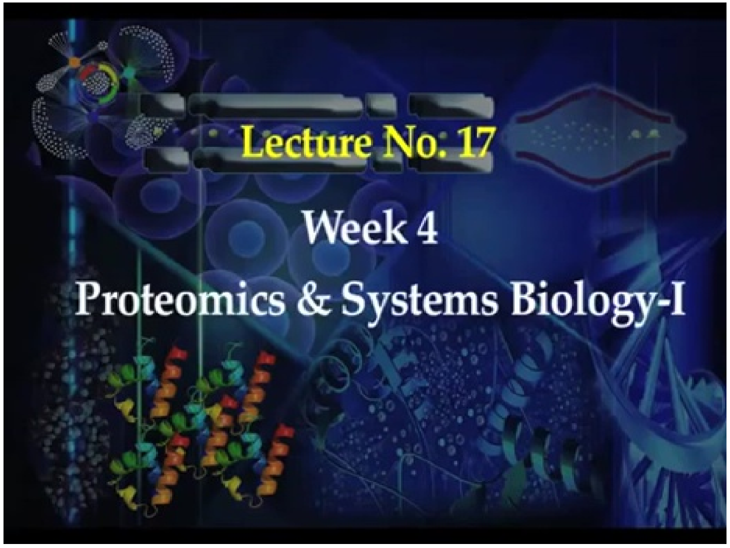
Mass Spectrometry Based Proteomics

MOOC - NPTEL Course

Prof. Sanjeeva Srivastava

Department of Biosciences and Bioengineering
IIT Bombay

(Refer Slide Time 00:15)

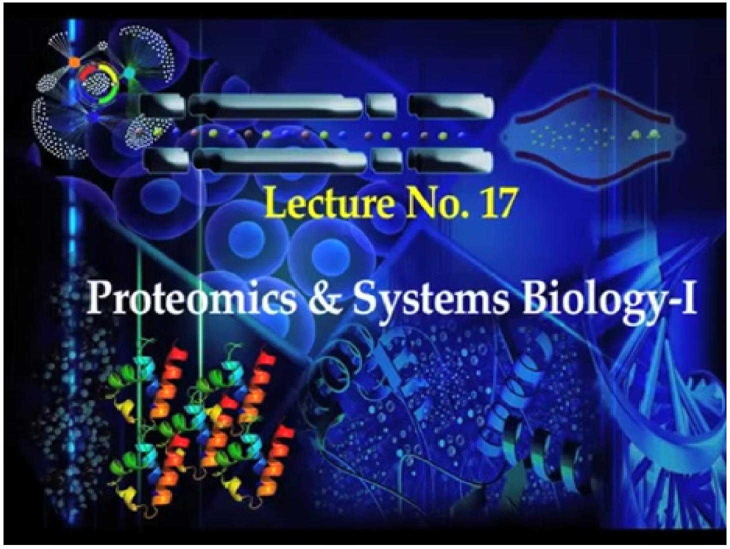


Lecture No. 17

Week 4

Proteomics & Systems Biology-I

(Refer Slide Time 00:19)

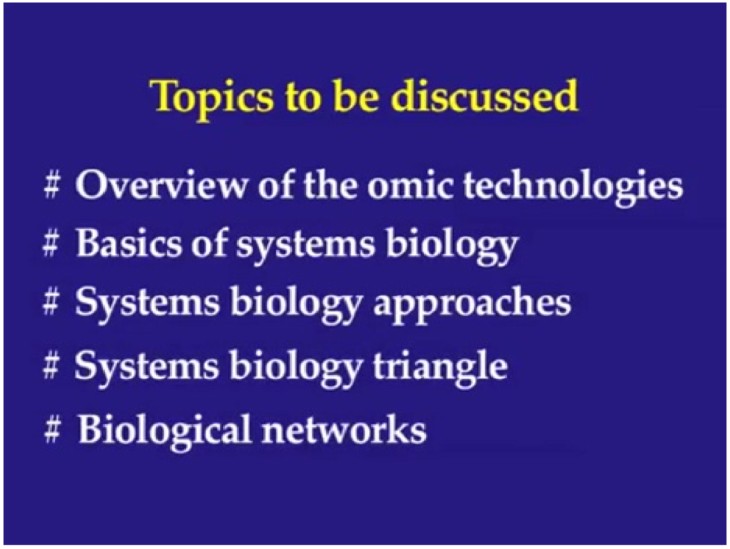


Lecture No. 17

Proteomics & Systems Biology-I

Today's lecture is on Proteomics and Systems Biology.

(Refer Slide Time 00:26)

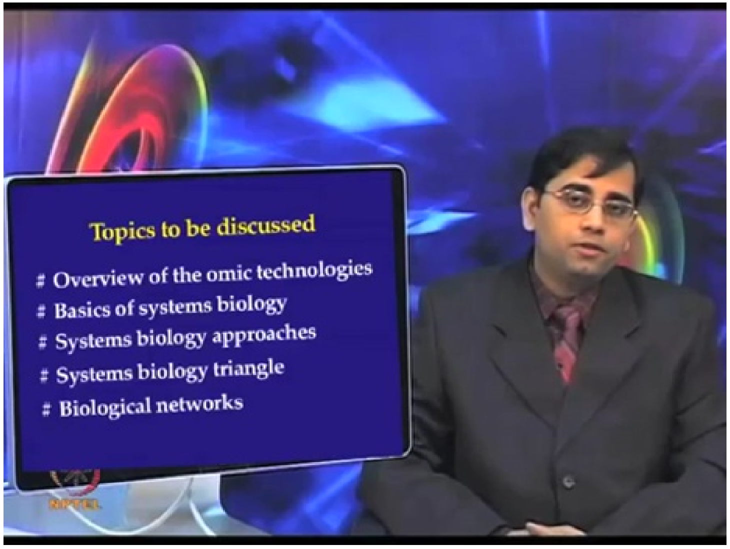


Topics to be discussed

- \# Overview of the omic technologies
- \# Basics of systems biology
- \# Systems biology approaches
- \# Systems biology triangle
- \# Biological networks

The lecture outline: First I will start on proteomics and then we will move on to the Systems Biology.

(Refer Slide Time 00:37)



Topics to be discussed

- Overview of the omic technologies
- Basics of systems biology
- Systems biology approaches
- Systems biology triangle
- Biological networks

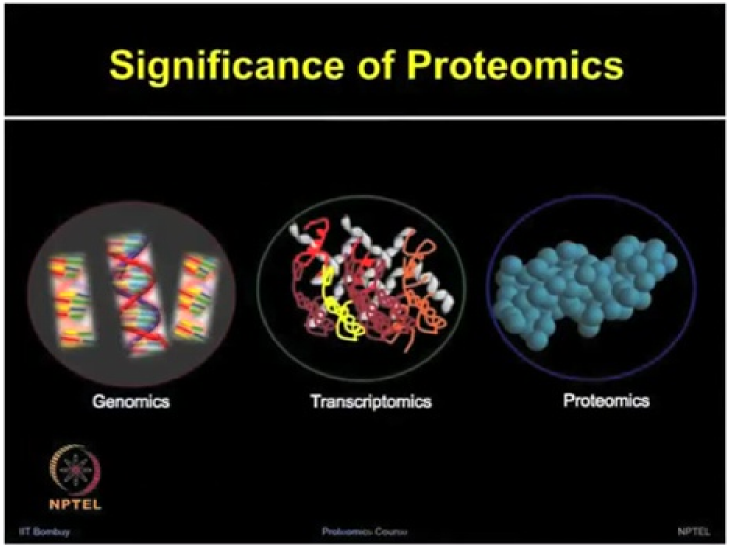
Today we will continue on first, Proteomics and then how different types of Omic data can be applied for Systems Approach Analysis, that will be discussed.

(Refer Slide Time 00:48)

Section I

Overview of the omics technologies

(Refer Slide Time 00:53)



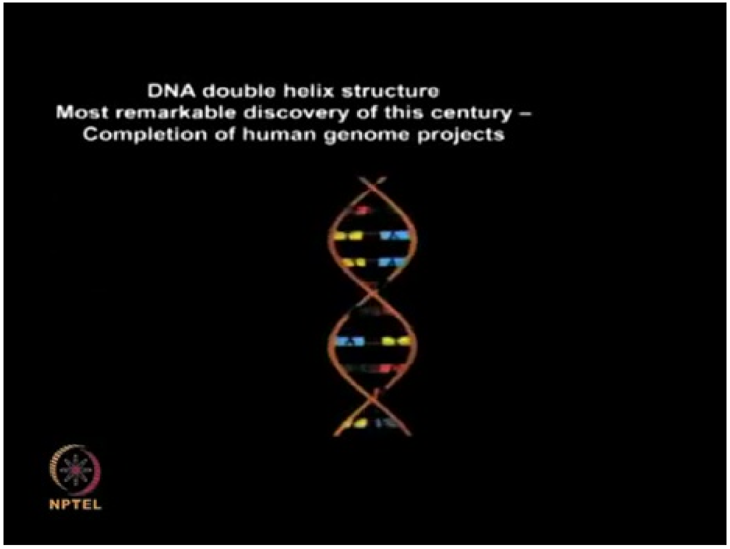
Significance of Proteomics

Central dogma concept which is an orderly and uni-directional flow of information encoded in base sequences of cells passed on from DNA to RNA and then to the proteins. This is the simplest definition of central dogma.

The Genome sequencing projects, they have provided researchers with unprecedented information of genome sequences. However there are numerous proteins which can be encoded by the genome. Therefore the analysis of the static genome by doing sequencing alone is not sufficient.

A gene can code for several types of proteins because of its alternative splicing and post-transcriptional and post-translational modifications. Therefore, it suggests that studying proteins is more challenging than genome or transcriptome. Therefore proteomics has great significance to understand the biological systems.

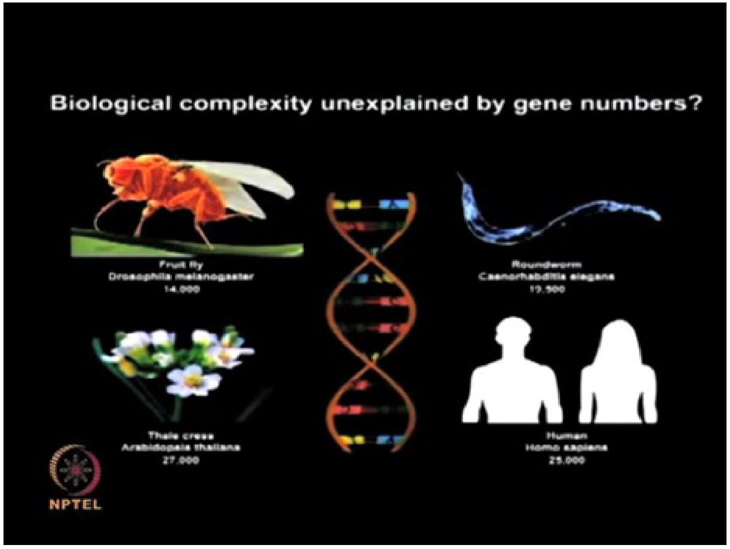
(Refer Slide Time 02:08)



DNA double helix structure
Most remarkable discovery of this century –
Completion of human genome projects

The completion of genome sequencing projects of several organism including human

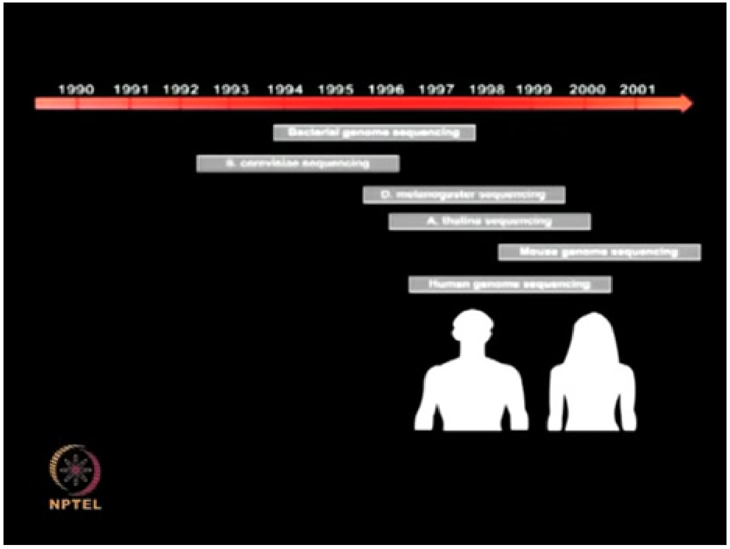
(Refer Slide Time 02:15)



Biological complexity unexplained by gene numbers?

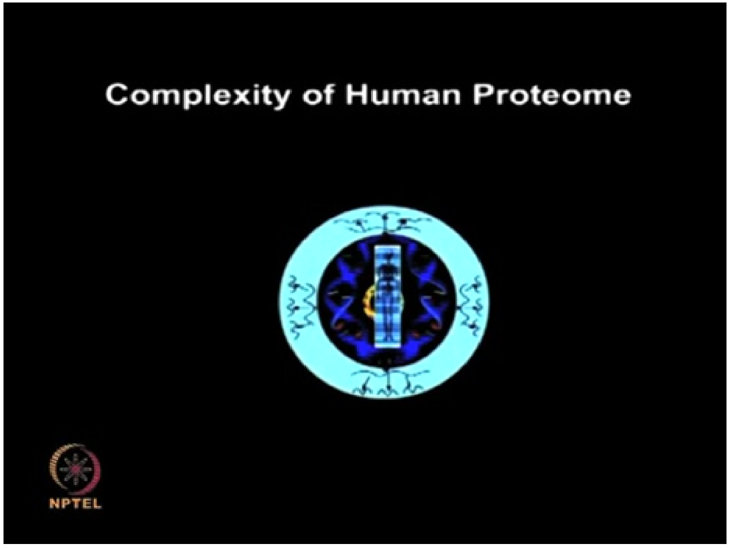
has been one of the most remarkable achievements of this century. However, these have not been sufficient to unravel the mystery of complex biological processes. The similar gene numbers of many diverse group of organisms has failed to explain their varying biological complexity.

(Refer Slide Time 02:43)



A more meaningful understanding of biological function can be obtained through the characterization of products of gene expression, the protein which serve as ultimate effective molecule of biological systems.

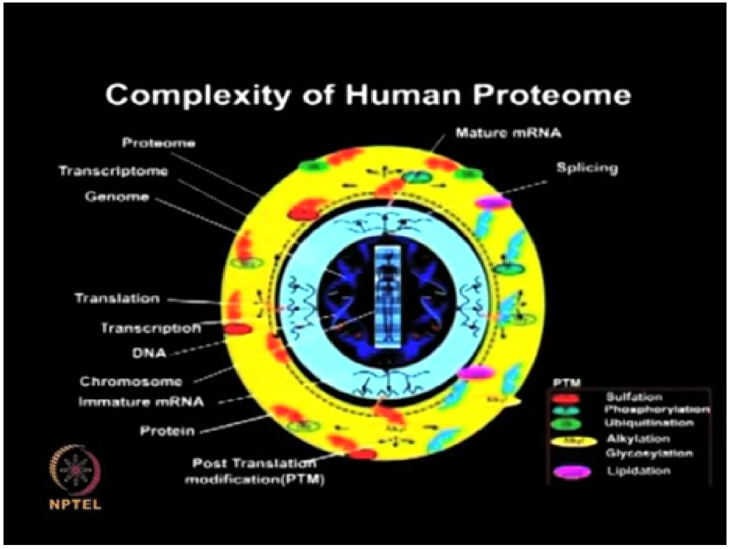
(Refer Slide Time 03:02)



Complexity of Human Proteome

The proteomics refers to the study of entire protein complement expressed by an organism at any given time.

(Refer Slide Time 03:14)

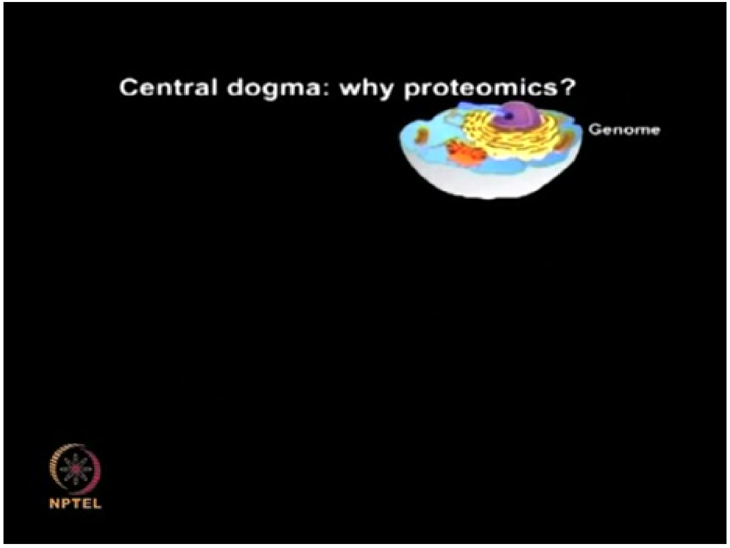


Complexity of Human Proteome

While the genome of an organism is mostly static, the proteome is dynamic and it changes with environment and time, thereby elevating its complexity level.

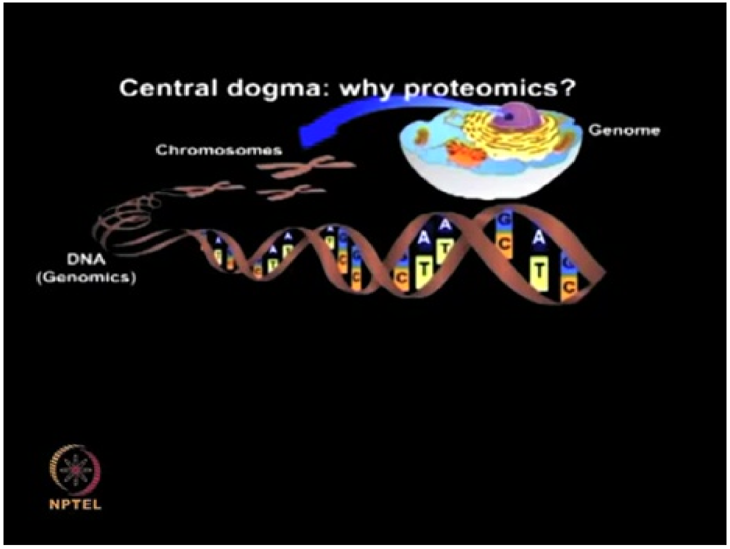
The gene regulation is regulated by several post-transcriptional and post-translational modifications due to which the number of proteins expressed in a cell is much greater than its genomic counterpart. The proteomics aims to

(Refer Slide Time 03:46)



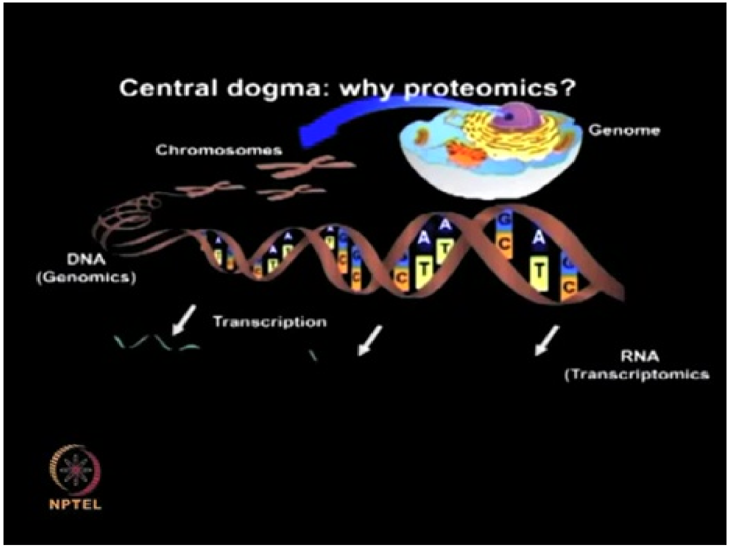
Central dogma: why proteomics?

(Refer Slide Time 03:51)


Central dogma: why proteomics?

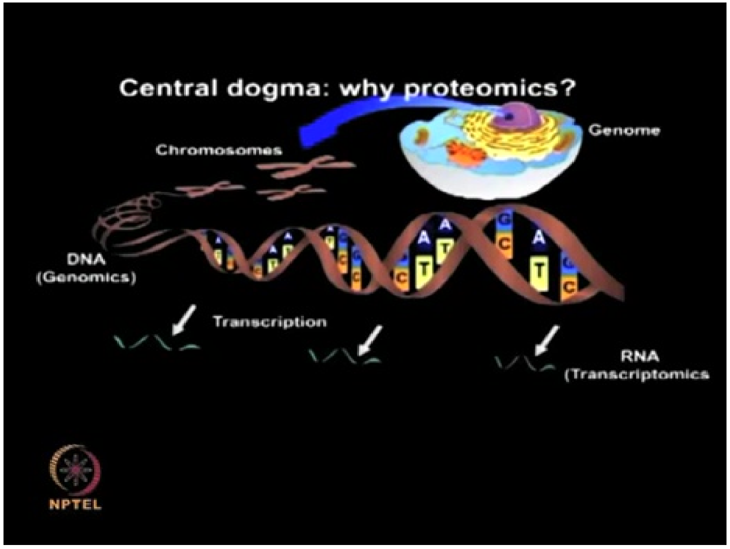
decipher the structure and function of all proteins

(Refer Slide Time 03:57)


Central dogma: why proteomics?

in a given cell under

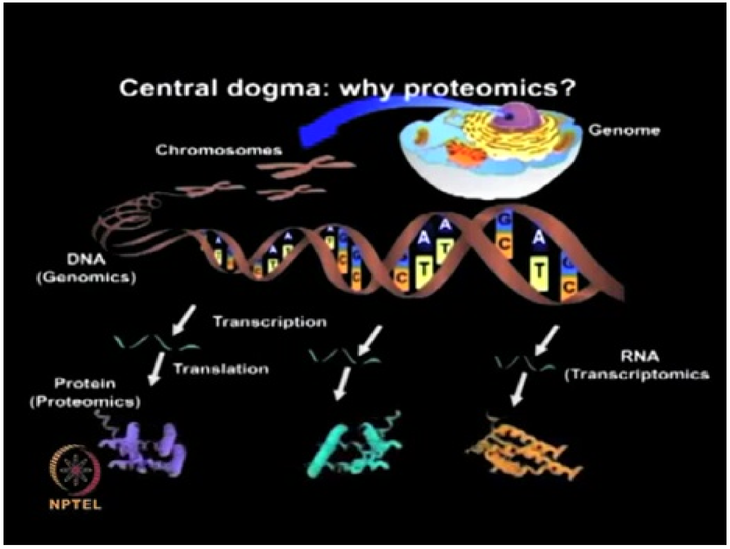
(Refer Slide Time 03:59)



Central dogma: why proteomics?

specific conditions and to obtain a global view of cellular processes at the protein level.

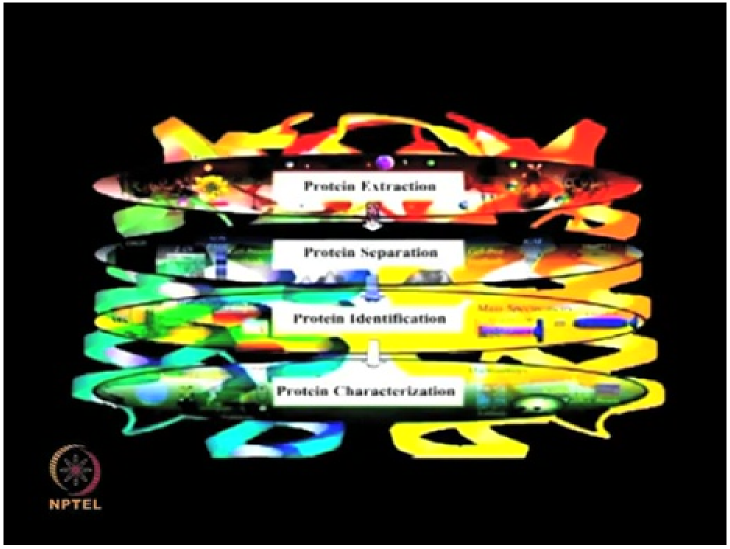
(Refer Slide Time 04:10)



Central dogma: why proteomics?

Study at the DNA level is known as genomics, RNA level transcriptomics and protein level

(Refer Slide Time 04:18)



Protein Extraction

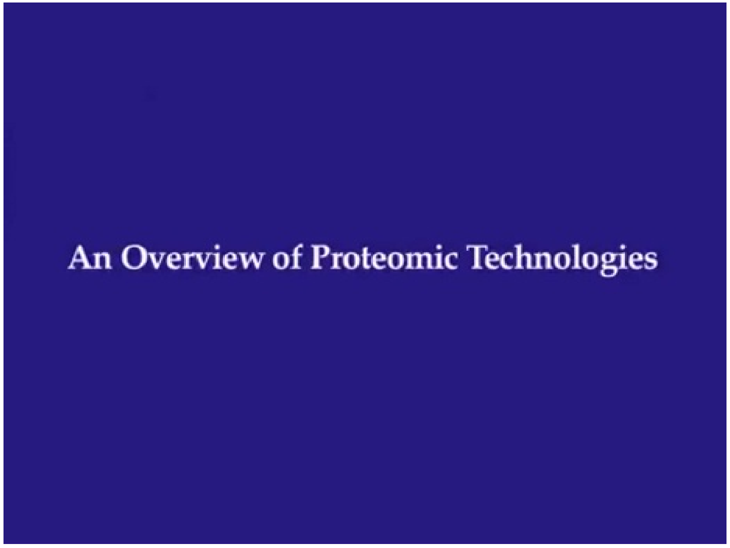
Protein Separation

Protein Identification

Protein Characterization

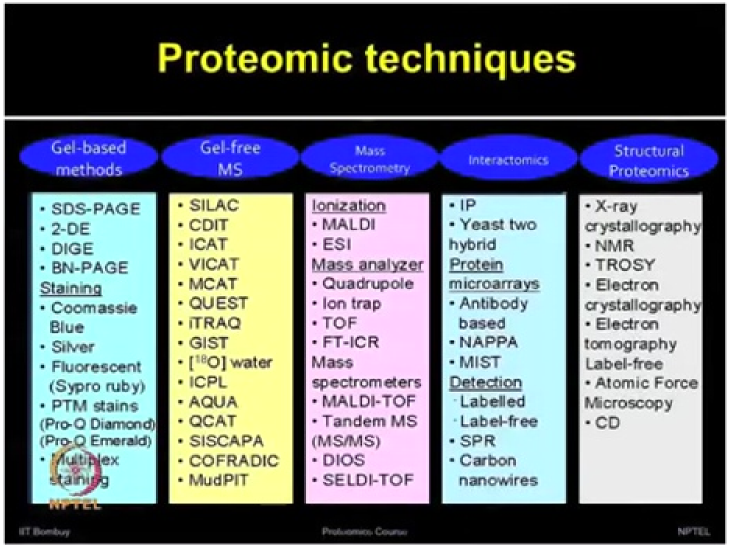
proteomics. Analysis of the proteome involves protein extraction, separation, identification and finally characterization of various proteins.

(Refer Slide Time 04:34)



An Overview of Proteomic Technologies

(Refer Slide Time 04:38)



Proteomic techniques

Gel-based methods
- SDS-PAGE
- 2-DE
- DIGE
- BN-PAGE

Staining
- Coomassie Blue
- Silver
- Fluorescent (Sypro ruby)
- PTM stains (Pro-Q Diamond) (Pro-Q Emerald)
- Multiplex staining

Gel-free MS
- SILAC
- CDIT
- ICAT
- VICAT
- MCAT
- QUEST
- iTRAQ
- GIST
- [^{18}O] water
- ICPL
- AQUA
- QCAT
- SISCAPA
- COFRADIC
- MudPIT

Mass Spectrometry

Ionization
- MALDI
- ESI

Mass analyzer
- Quadrupole
- Ion trap
- TOF
- FT-ICR

Mass spectrometers
- MALDI-TOF
- Tandem MS (MS/MS)
- DIOS
- SELDI-TOF

Interactomics
- IP
- Yeast two hybrid

Protein microarrays
- Antibody based
- NAPPA
- MIST

Detection
- Labelled
- Label-free
- SPR
- Carbon nanowires

Structural Proteomics
- X-ray crystallography
- NMR
- TROSY
- Electron crystallography
- Electron tomography Label-free
- Atomic Force Microscopy
- CD

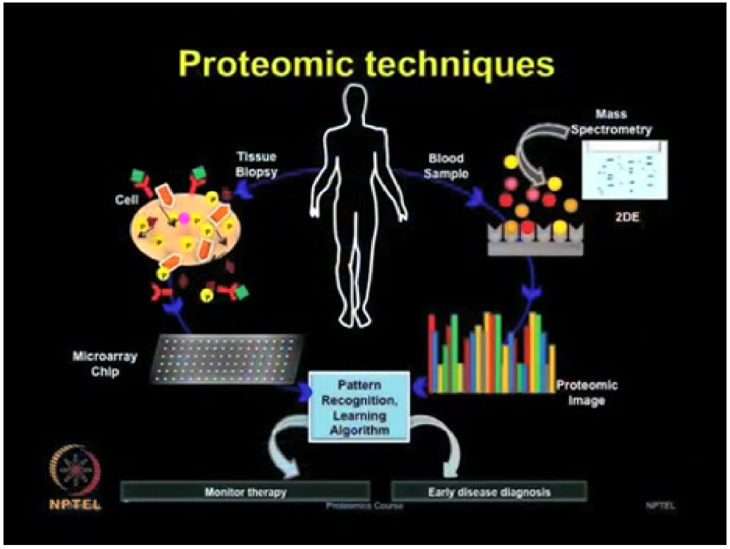
IIT Bombay Proteomics Course NPTEL

Very broadly, you can group these technologies as gel-based methods, gel-free mass spectrometry based methods, mass spectrometry techniques, techniques to study protein interactions and structure proteomics

So as we have seen, there are large number of proteomic technologies which are currently available for various applications. Many times, to address one biological question different types of methodologies come together and then provide solutions to that problem.

(Refer Slide Time 05:16)

Proteomic techniques

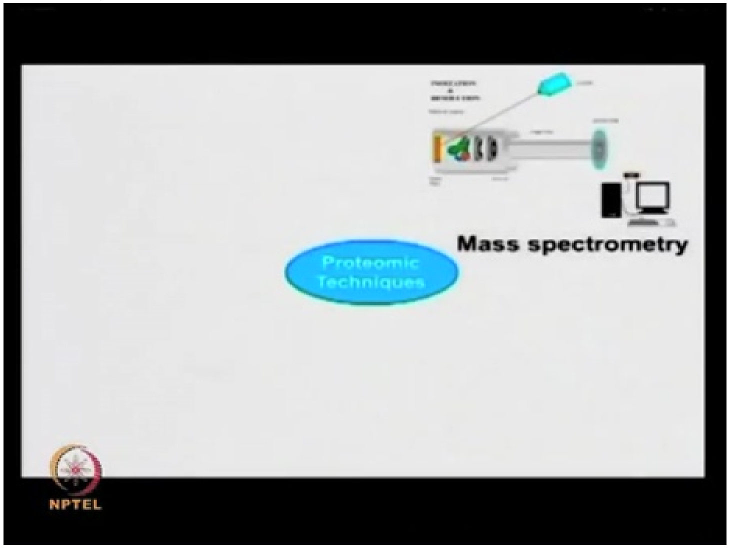


For example, looking at some clinical sample for identification of biomarkers of a disease, one can employ the samples such as tissue or blood or different type of body fluids.

And then either directly extract the protein and subject it to the mass spectrometry or first resolve on two dimensional electrophoresis, followed by identification on mass spectrometry or take these samples, directly apply on the microarray based platforms and then detect using label-based or label-free methodologies.

Eventually these types of results will enhance knowledge for the monitoring the therapy response as well as identification of early disease diagnosis. This is just one example. Similarly multiple type of proteomic technologies can be used for different applications.

(Refer Slide Time 06:18)

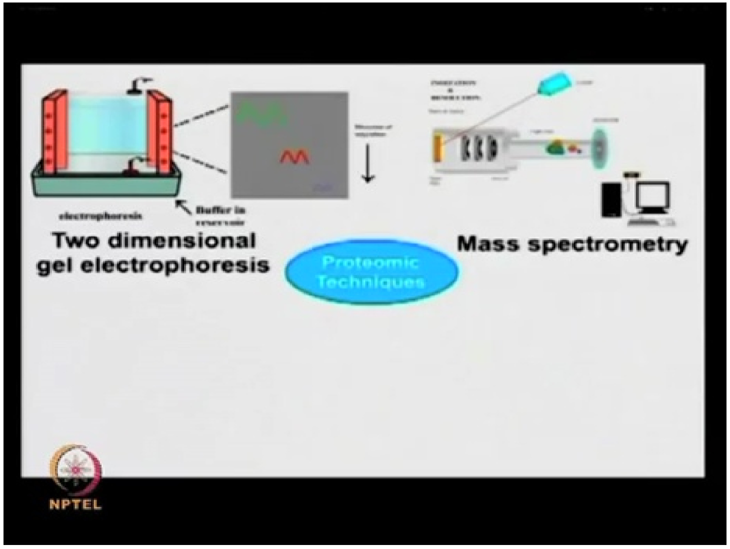


Mass spectrometry

Proteomic Techniques

There are several proteomic techniques which are employed for studying these proteins such as

(Refer Slide Time 06:27)



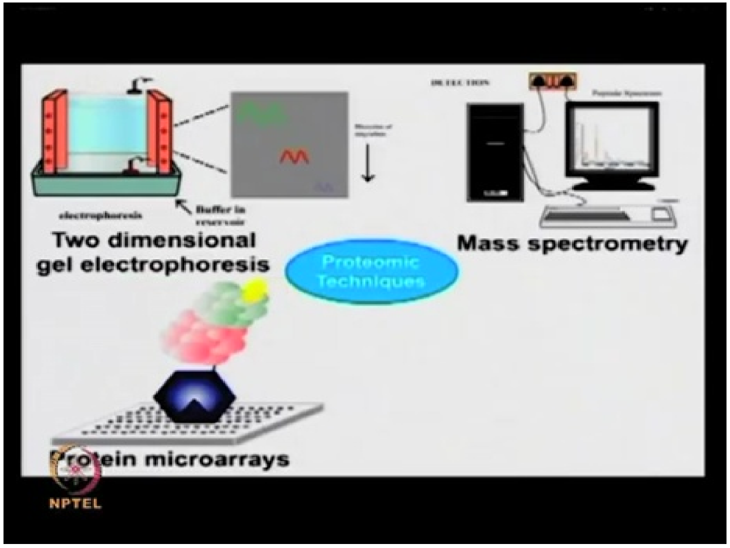
Two dimensional gel electrophoresis

Mass spectrometry

Proteomic Techniques

two dimensional gel electrophoresis, mass spectrometry, protein microarrays, as well as some label-free detection techniques such as Surface Plasma Resonance

(Refer Slide Time 06:39)



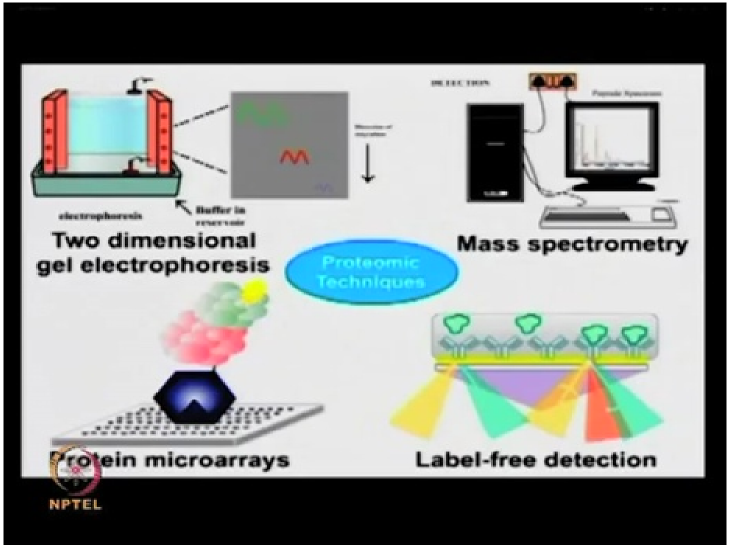
Two dimensional gel electrophoresis

Mass spectrometry

Proteomic Techniques

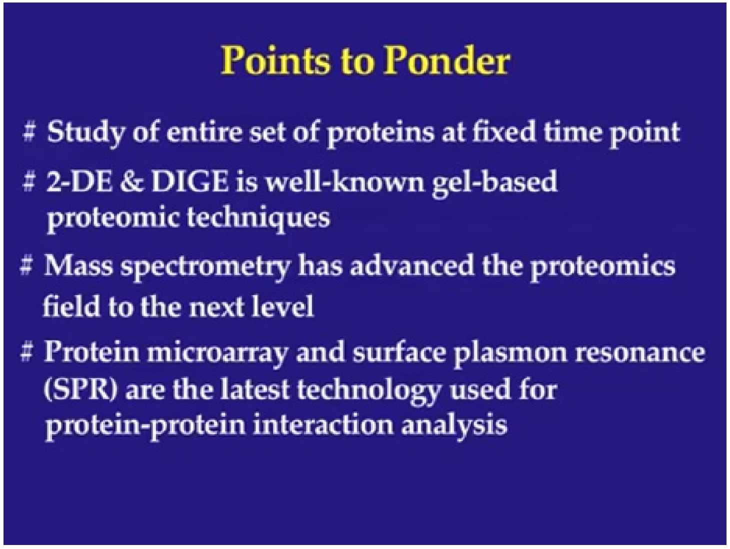
Protein microarrays

(Refer Slide Time 06:47)



SPR.

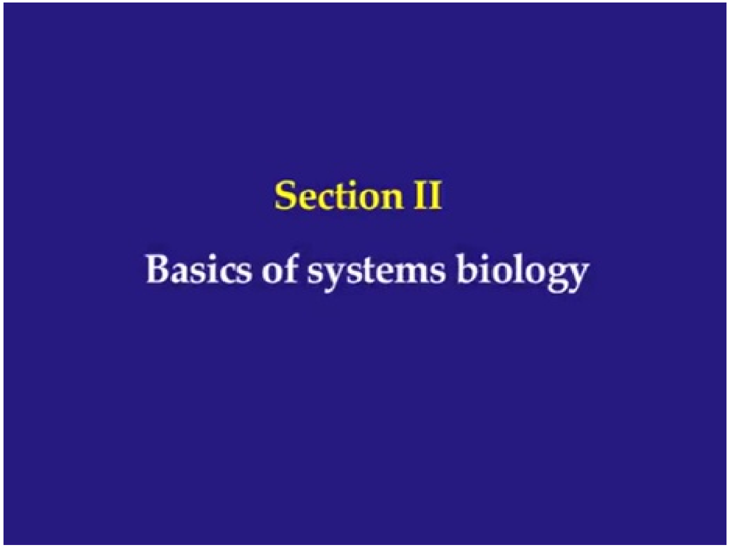
(Refer Slide Time 07:00)



Points to Ponder

- # Study of entire set of proteins at fixed time point
- # 2-DE & DIGE is well-known gel-based proteomic techniques
- # Mass spectrometry has advanced the proteomics field to the next level
- # Protein microarray and surface plasmon resonance (SPR) are the latest technology used for protein-protein interaction analysis

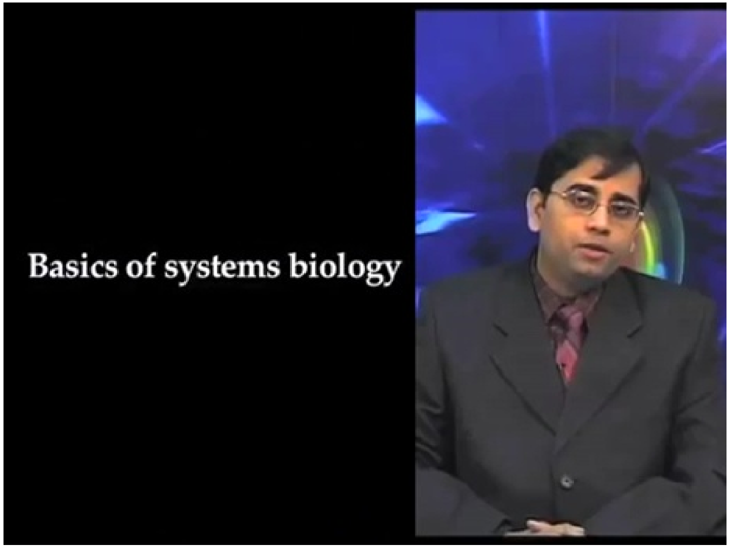
(Refer Slide Time 07:07)



Section II

Basics of systems biology

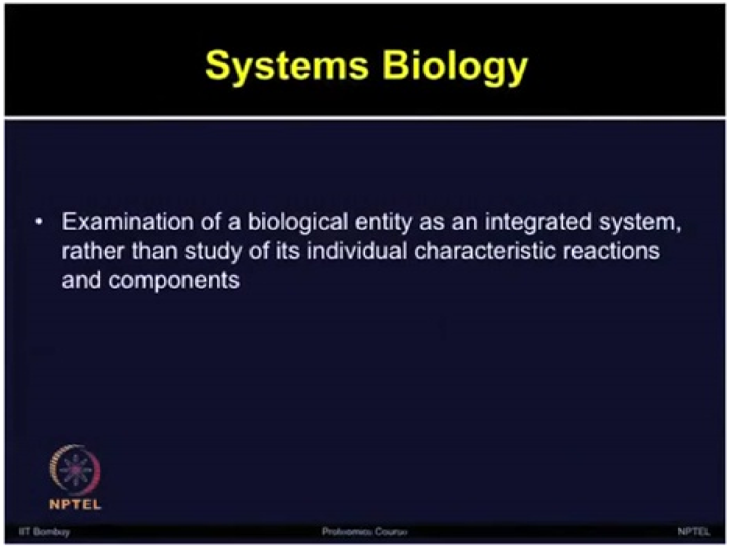
(Refer Slide Time 07:16)



Basics of systems biology

After discussing proteomics, now let’s talk about Systems biology. So what is Systems Biology? Is that an historic knowledge, a method to understand biological systems or a tool to solve the practical problems?

(Refer Slide Time 07:30)



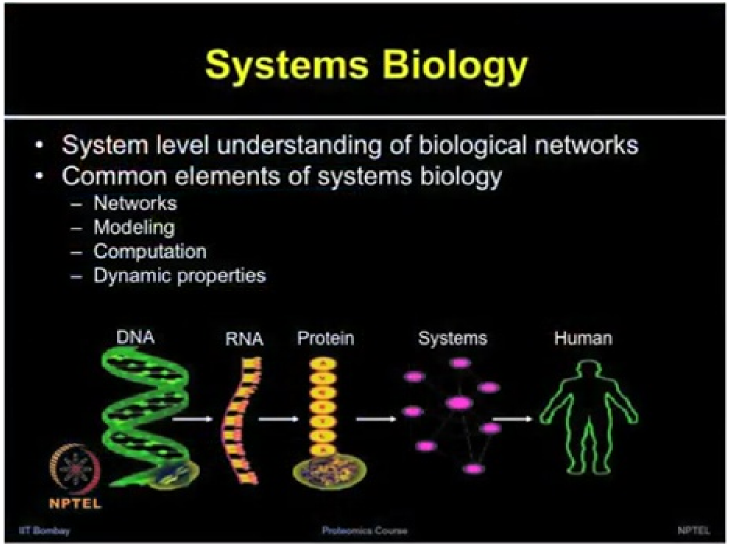
Systems Biology

- Examination of a biological entity as an integrated system, rather than study of its individual characteristic reactions and components

IIT Bombay Proteomics Course NPTEL

Systems Biology is the examination of a biological entity as an integrated system rather than the study of its individual characteristic reactions and components which is termed as Systems Biology. The study of all the mechanisms underlying the complex biological processes in the form of integrated system of many interacting components is studied under Systems Biology.

(Refer Slide Time 08:03)



Systems Biology

- System level understanding of biological networks
- Common elements of systems biology
  - Networks
  - Modeling
  - Computation
  - Dynamic properties

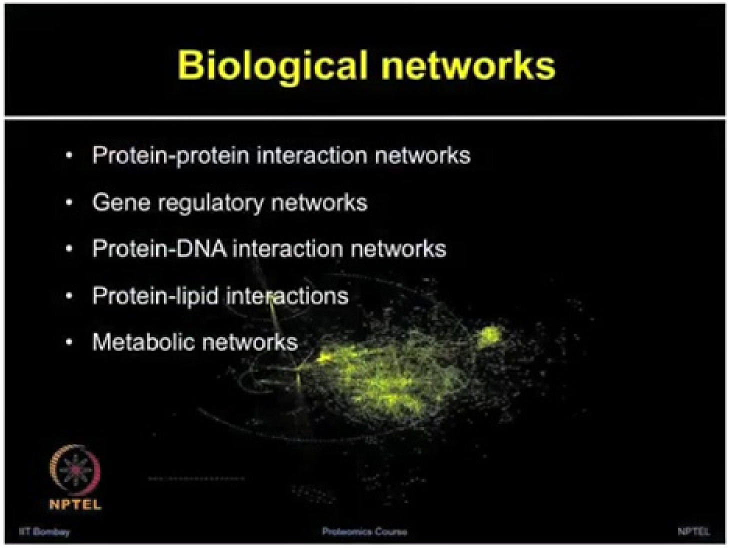
IIT Bombay Proteomics Course NPTEL

The systems level understanding of biological networks require information from different levels. As you can see from DNA to RNA to proteins forming systems and then that information can be applied to understand a complex system for different organism. The biological information is represented by the networks of interacting elements and dynamic responses to the perturbations.

These networks provide insights which cannot be analyzed from the isolated components of the system. The common elements of the Systems Biology include networks, modeling, computation and dynamic properties.

There are different types of biological networks such as

(Refer Slide Time 09:04)

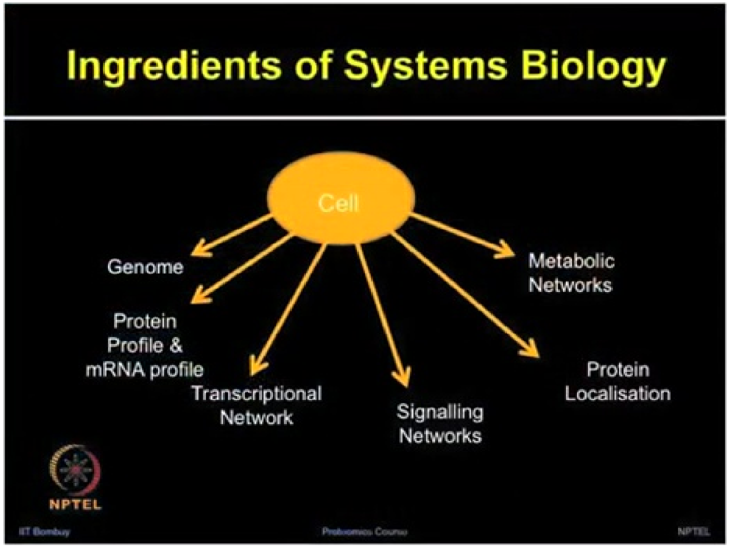


Biological networks

- Protein-protein interaction networks
- Gene regulatory networks
- Protein-DNA interaction networks
- Protein-lipid interactions
- Metabolic networks

protein-protein interaction networks, gene regulatory networks, protein and DNA interaction networks, protein-lipid, protein-other biomolecule networks and metabolic networks.

(Refer Slide Time 09:18)



Ingredients of Systems Biology

Cell

Genome

Protein Profile & mRNA profile

Transcriptional Network

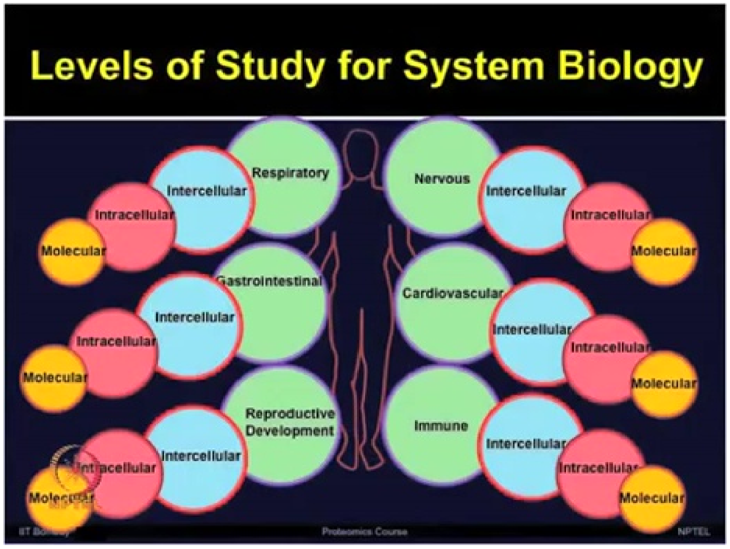
Signalling Networks

Protein Localisation

Metabolic Networks

The various ingredients of systems biology, for example if you are studying about the cell and the system's behavior; one needs to look at the genome, its transcriptome profile, proteome profile, how protein DNA and different types of transcription networks are altered, protein-protein signaling networks, multimeric complexes how they are formed how protein is localized in the intracellular dynamics and metabolic networks. All of these are ingredients of study about a system.

(Refer Slide Time 09:57)

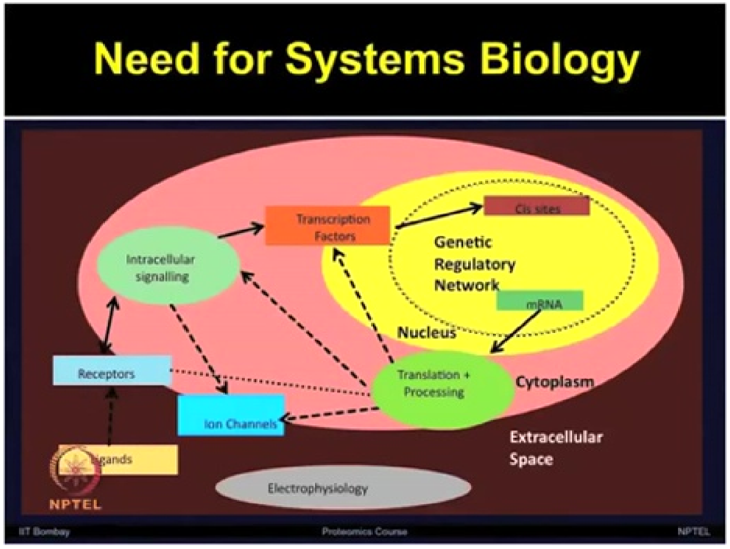


Levels of Study for System Biology

IIT Bombay Proteomics Course NPTEL

System Biology study can be done at different levels. For example, to study the complex physiology of humans one can look at individual system such as respiratory system, nervous system or other physiological systems. Studies can be done at the intercellular or intracellular level and finally at the molecular level involving genomics, transcriptomics and proteomics.

(Refer Slide Time 10:28)



Need for Systems Biology

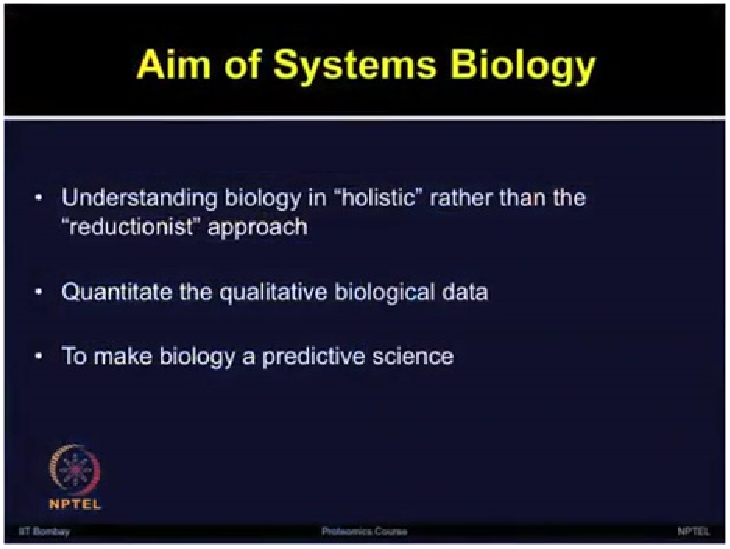
IIT Bombay Proteomics Course NPTEL

So why there is need for Systems Biology? The study of biology at the system and subsystem level for understanding the biological processes and networks is very much required. As you can see, to understand even a simple system of a cell; how it is regulated with the extracellular space, the cytoplast and different other components.

Examination of the structure and dynamics of cellular and organisable function is very much required for understanding the systems rather than the characteristics of isolated parts of the cell or the organism.

So what is the aim of Systems Biology? …To understand the biology in holistic approach rather than the reductionist approach..

(Refer Slide Time 11:26)



Aim of Systems Biology

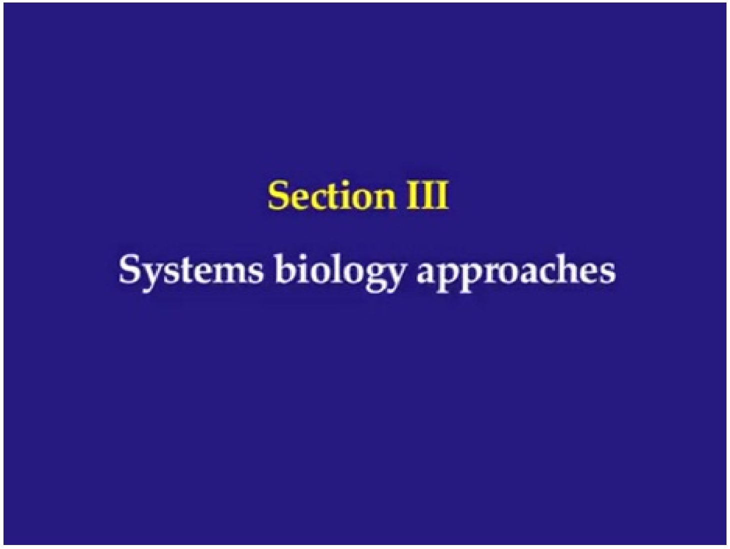
- Understanding biology in “holistic” rather than the “reductionist” approach
- Quantitate the qualitative biological data
- To make biology a predictive science

NPTEL

IIT Bombay Proteomics Course NPTEL

The Systems Biology aims to quantitate the qualitative biological data and provide some level of predictions by applying different types of computational modeling

(Refer Slide Time 11:40)

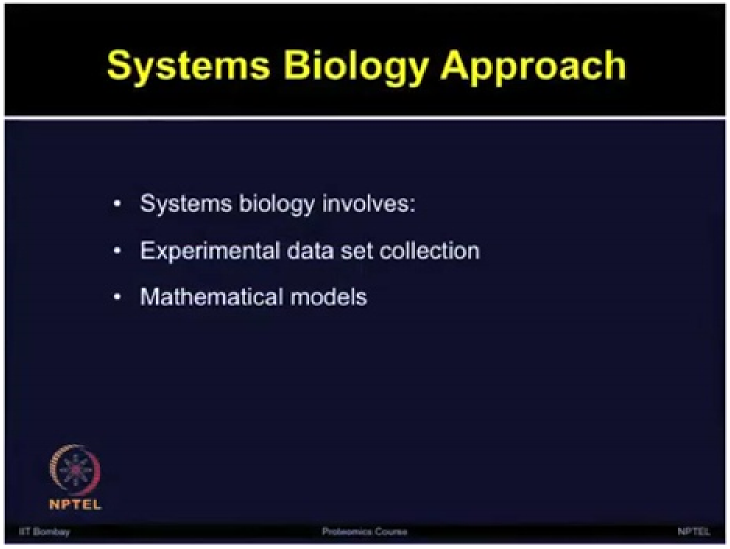


Section III

Systems biology approaches

The systems biology approach involves,

(Refer Slide Time 11:49)



Systems Biology Approach

- Systems biology involves:
- Experimental data set collection
- Mathematical models

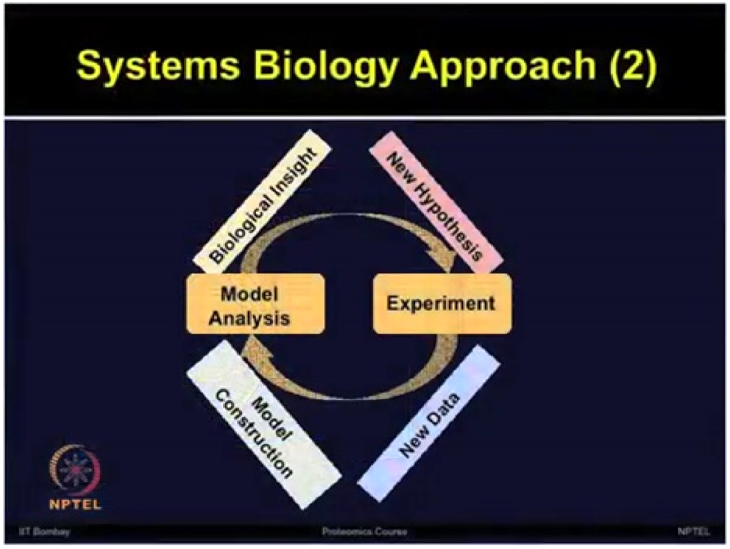
NPTEL

IIT Bombay Proteomics Course NPTEL

first of all collection of large experimental datasets and then mathematical models to provide insight of some significant aspects of the datasets.

The simple systems biology approach would involve experiments by adding new datasets

(Refer Slide Time 12:12)



Systems Biology Approach (2)

Biological Insight

New Hypothesis

Model Analysis

Experiment

Model Construction

New Data

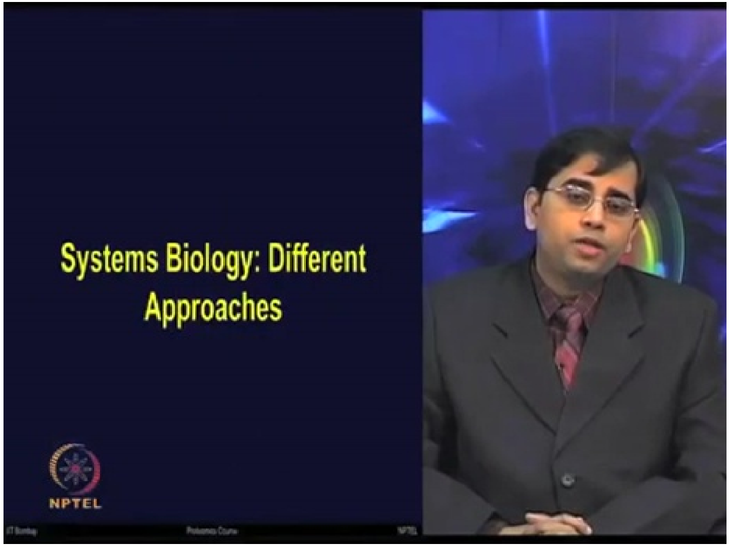
NPTEL

IIT Bombay Proteomics Course NPTEL

which will be used for model construction and model analysis and the biological insight derived from these models can be used to propose new hypothesis.

So the properties of a system are probably more than just the sum of all its individual properties or its components. Therefore it is possible that system may have its own property by applying all the components.

(Refer Slide Time 12:46)



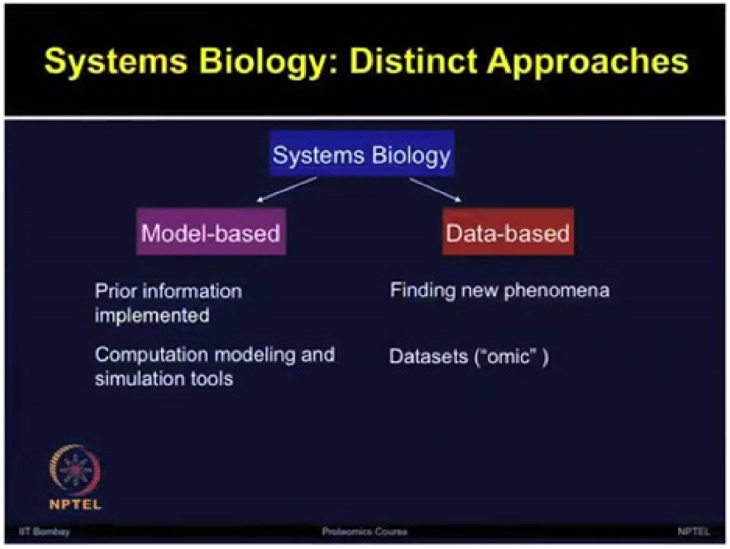
Systems Biology: Different Approaches

NPTEL

IIT Bombay Proteomics Course NPTEL

So what all different approaches have been taken to study the systems biology?

(Refer Slide Time 12:52)



Systems Biology: Distinct Approaches

Systems Biology

Model-based

Prior information implemented

Computation modeling and simulation tools

Data-based

Finding new phenomena

Datasets ("omic")

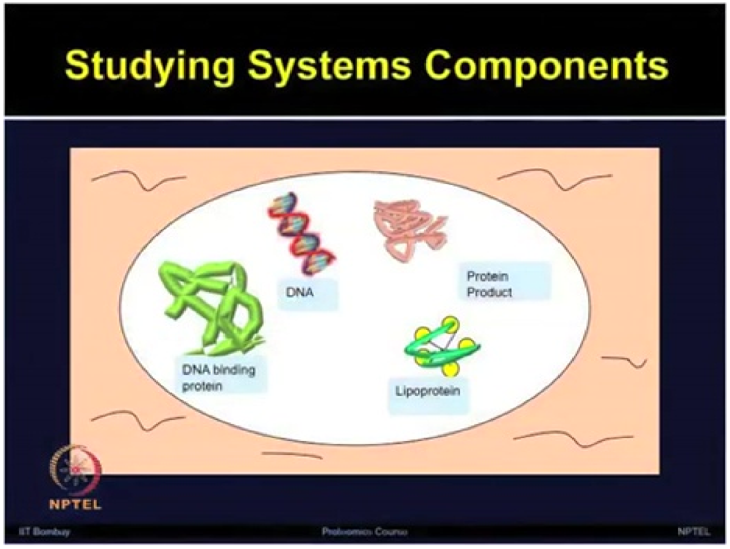
The distinct approaches of systems biology include the model-based and data-based methods.

The model-based approach involves some prior information which can be implemented in the models where as in the data-based, the objective is to find the new phenomenon.

The model-based relies on computational modeling and simulation tools whereas the database method relies on the Omics datasets.

So in model-based systems biology approach, it is difficult to build the detailed kinetic models but in data-based system the complex relationship among various types of Omics information and metabolic pathways and networks can be created.

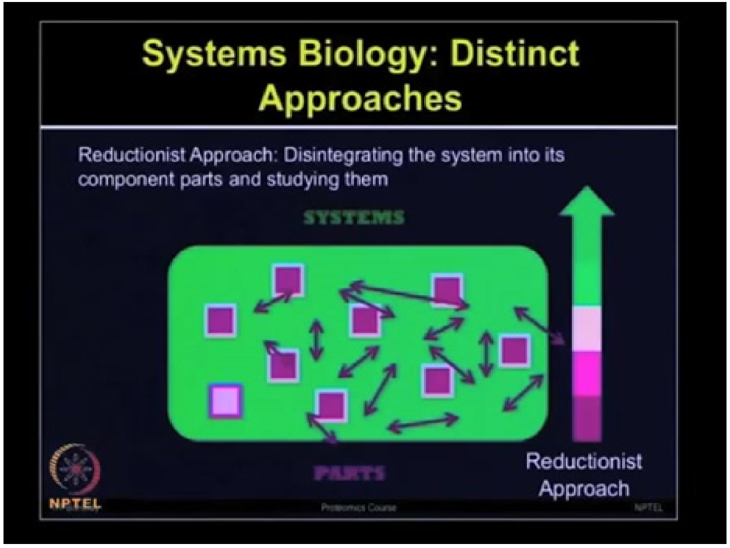
(Refer Slide Time 13:47)



Studying Systems Components

Studying systems component is very challenging. Systems biology and biological network modeling aims to understand the system's structure and function for better understanding of system's properties like its robustness as well as use for the prediction of systems behavior in response to the perturbations.

(Refer Slide Time 14:15)

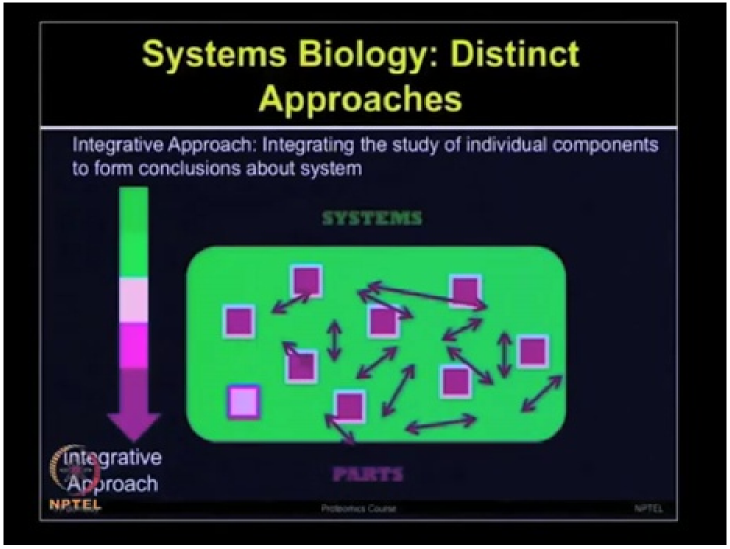


Systems Biology: Distinct Approaches

Reductionist Approach: Disintegrating the system into its component parts and studying them

The reductionist approach involves disintegrating the system into its component parts and studying them where as the integrative approach involves integrating the study of individual components to form conclusions about the system.

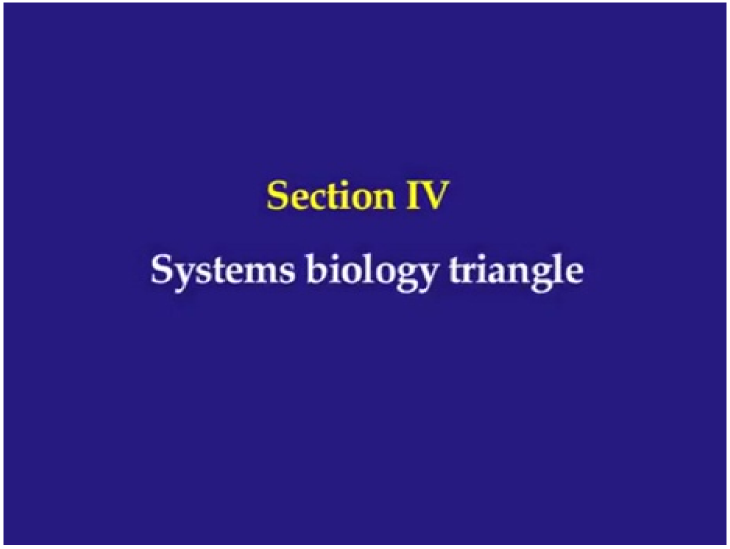
(Refer Slide Time 14:33)



Systems Biology: Distinct Approaches

Integrative Approach: Integrating the study of individual components to form conclusions about system

(Refer Slide Time 14:36)



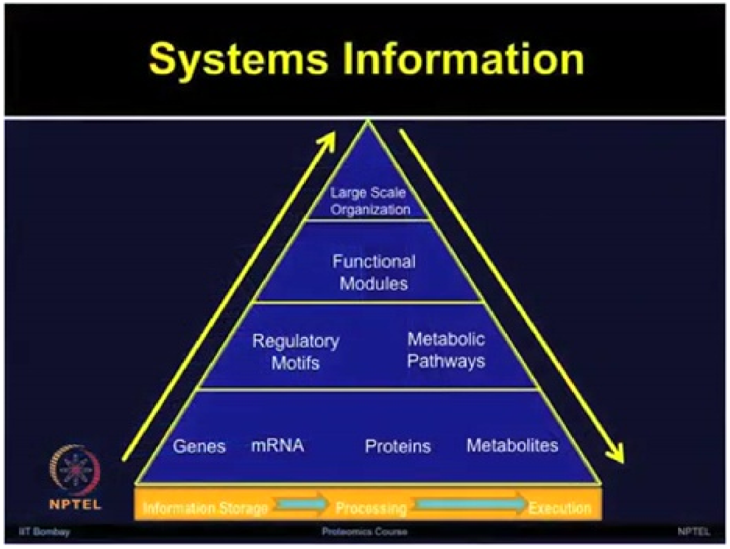
Section IV

Systems biology triangle

What is Systems Biology triangle?

So first of all, the systems information is generated at various levels. As we have discussed starting from genes to mRNA to proteins to metabolite;

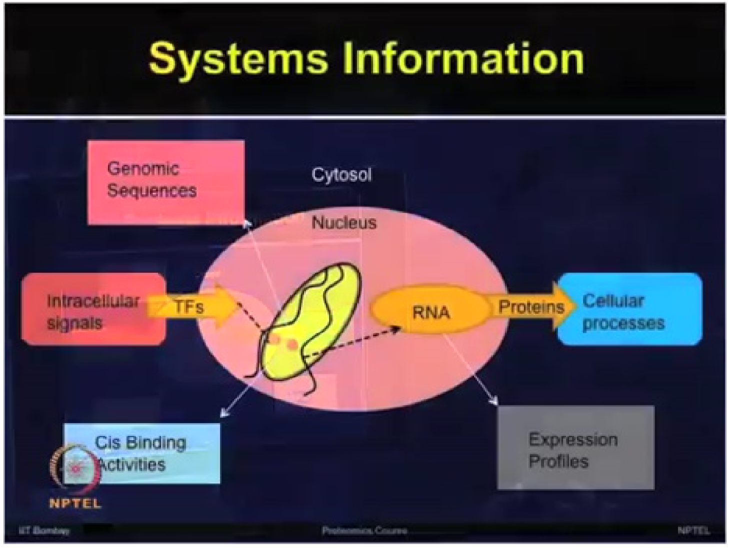
(Refer Slide Time 15:01)


Systems Information

or identifying regulatory motors, metabolic pathways, functional modules and different large scale organizations. This information has to be stored, processed and further executed to identify the system level information.

Even simplest system such as cell can be linked with various properties.

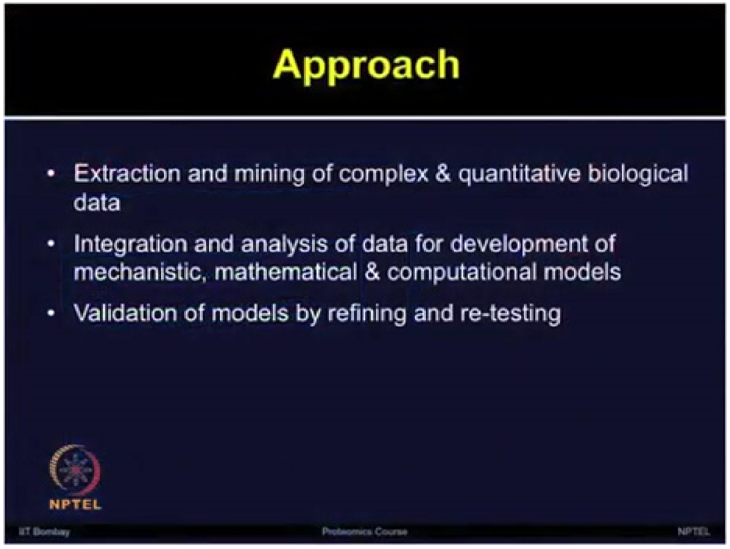
(Refer Slide Time 15:29)


Systems Information

Its genome sequences of different molecules, intracellular signals, transcription factors, different type of Cys binding activities, the expression profiling of RNA and proteins and different types of cellular processes.

So what approach one can take to study about the system?

(Refer Slide Time 15:52)



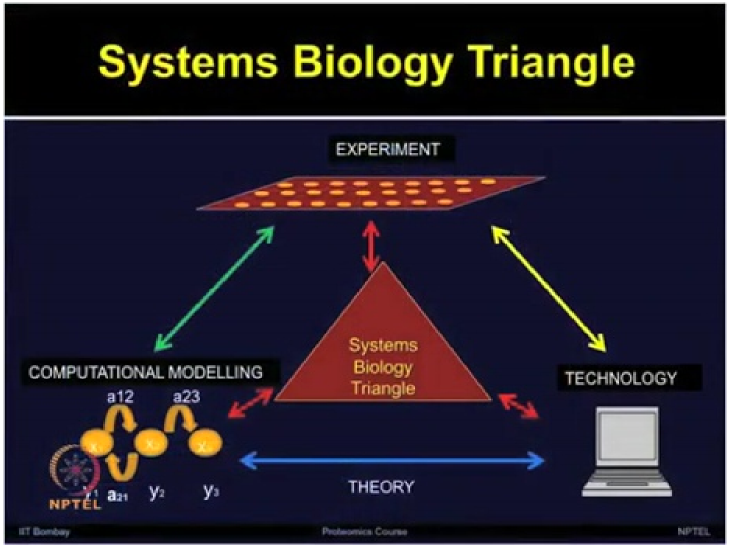
Approach

- Extraction and mining of complex & quantitative biological data
- Integration and analysis of data for development of mechanistic, mathematical & computational models
- Validation of models by refining and re-testing

Extraction and mining of complex and quantitative biological data, integration and analysis of these datasets for development of mechanistic, mathematical and computational models, validation of these models by retesting and refining after proposing some hypothesis

Different online databases and repositories are nowadays developed for sharing large datasets and various systems models. The systematic approach to study how molecules act together within the network of interaction that make up life is definitely going to be useful to understand the Systems biology.

(Refer Slide Time 16:48)

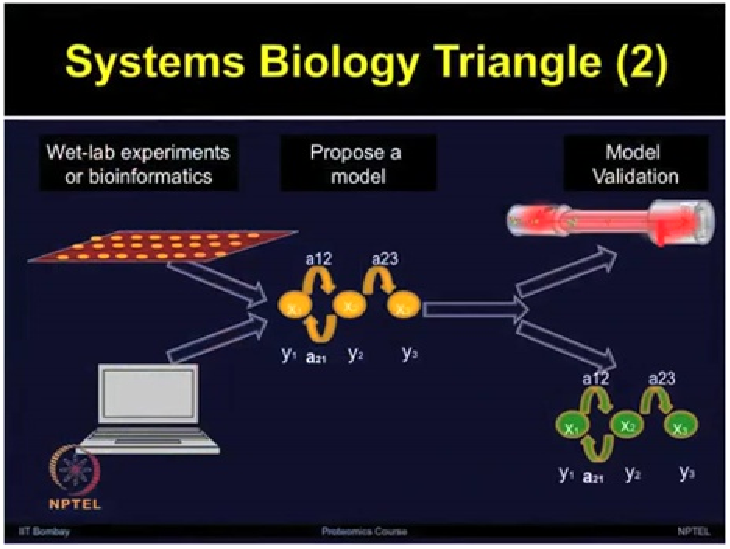


Systems Biology Triangle

The Systems Biology Triangle, as you can see here involves the experimental dataset could be derived from different types of Omics platforms; technologies, how the computational analysis can be performed; different types of bioinformatics software and tools and then computational modeling by obtaining some theoretical concepts.

The synergistic application of the experiment, theory and technology with modeling to enhance the understanding of biological processes as a whole system rather than the isolated part is termed as Systems Biology Triangle .

(Refer Slide Time 17:30)



Systems Biology Triangle (2)

Wet-lab experiments or bioinformatics

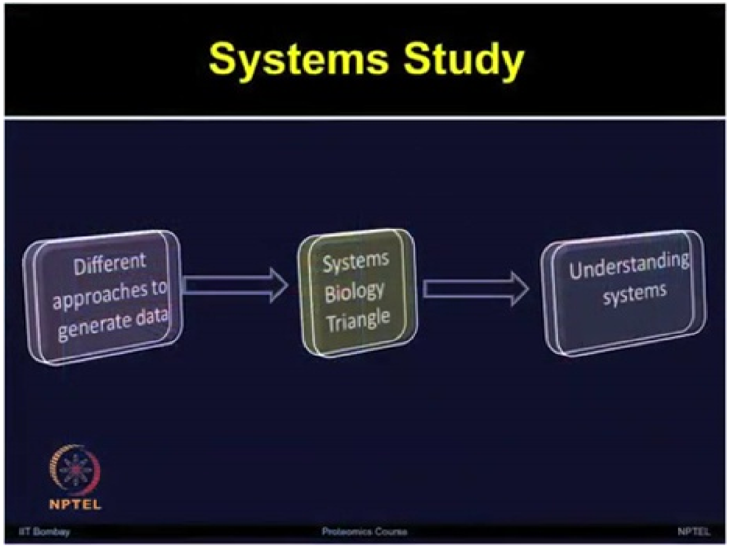
Propose a model

Model Validation

So, in the Systems Biology Triangle, the wet labs experiments or the bioinformatics based data analysis can be used to propose a model.

The model building as an aid to understand the complex system and some hypothesis can be generated, which could be used further to propose more quantitative models or predictive models and also it can be used for independent techniques for the model validation.

(Refer Slide Time 18:03)



Systems Study

Different approaches to generate data

Systems Biology Triangle

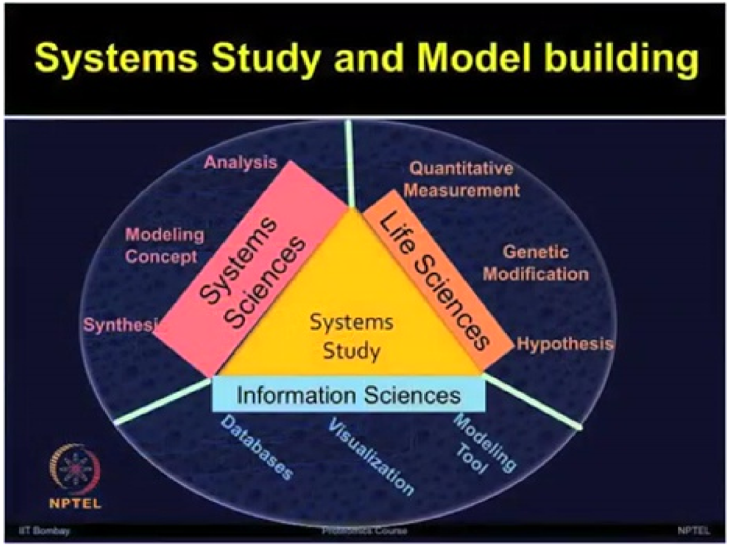
Understanding systems

NPTEL

So what is Systems Study? First of all the difference between systems study and components study one needs to understand and this is what we have tried to emphasize in the previous slides.

After generating the datasets and creating systems biology triangle, then this information can be used for understanding the systems in the more complex and mechanistic level.

(Refer Slide Time 18:32)



Systems Study and Model building

Analysis

Modeling Concept

Synthesi

Systems Sciences

Quantitative Measurement

Genetic Modification

Hypothesis

Life Sciences

Systems Study

Information Sciences

Databases

Visualization

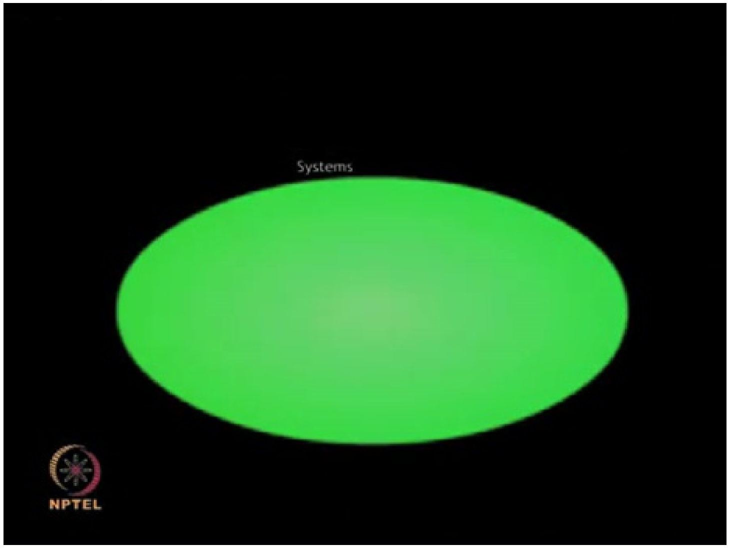
Modeling Tool

NPTEL

Systems study and Model Building: The systems science includes synthesis, modeling concepts, analysis. Life sciences discipline provides quantitative measurements, genetic modifications and deriving some hypothesis.

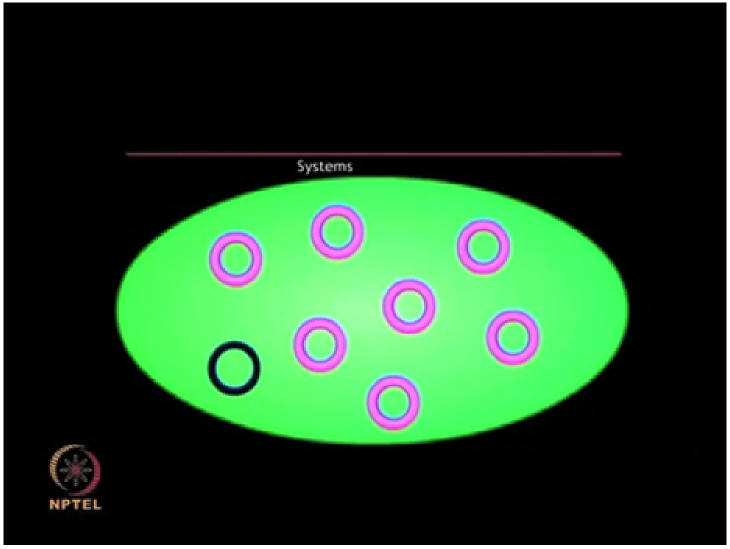
The information sciences enable the visualization, the modeling tools and different databases. So this model-building as an aid to understand the complex system is very useful for systems level investigation.

(Refer Slide Time 19:09)



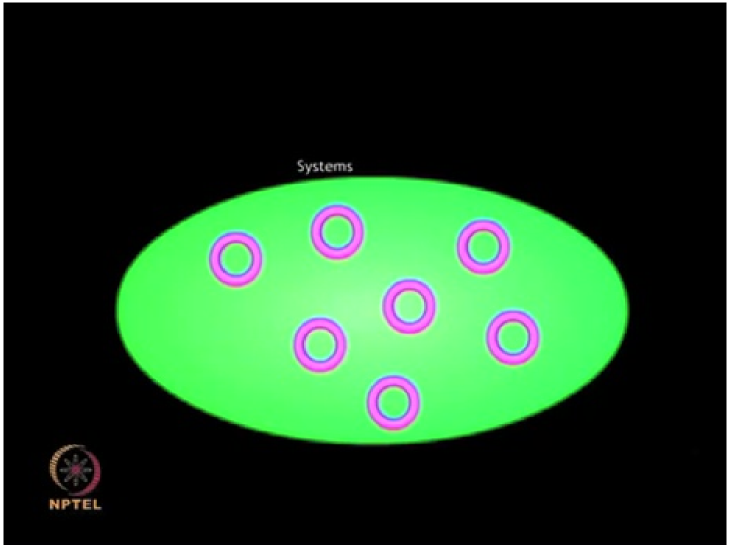
Systems is an entity which maintains its existence through mutual interaction of its constituent parts. The systems biology research consists of identification of the parts, characterization of the components,

(Refer Slide Time 19:28)



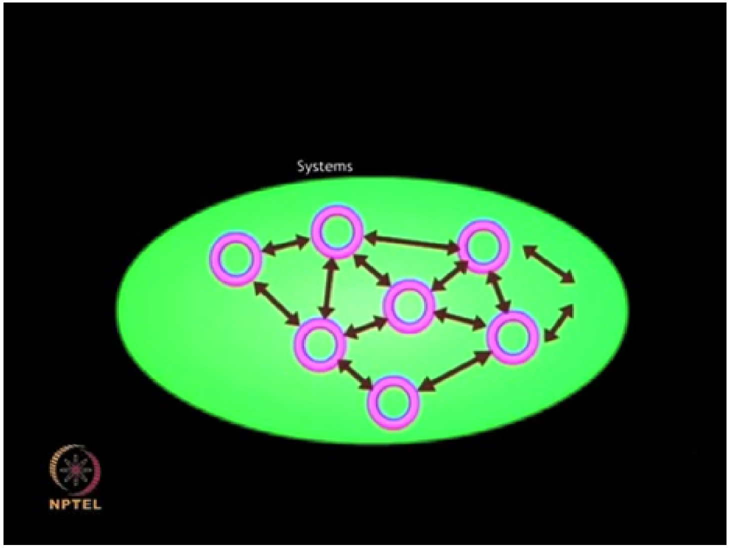
exclude the ones which are

(Refer Slide Time 19:32)



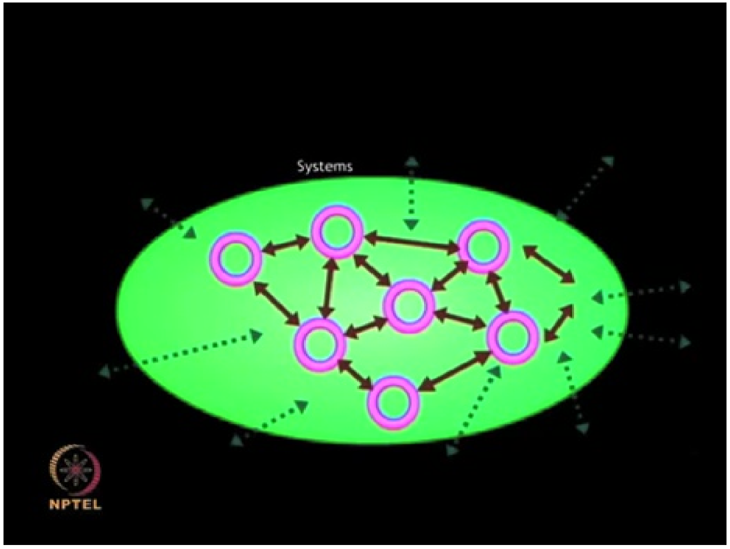
not a part of the system,

(Refer Slide Time 19:36)



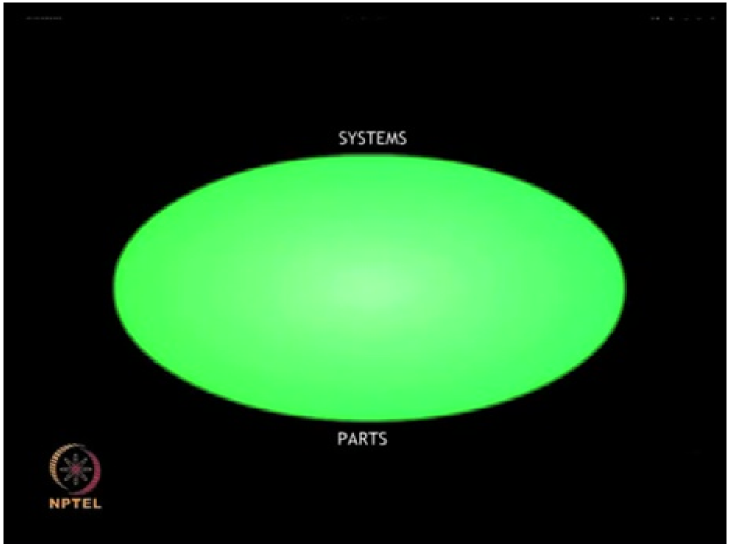
identify the interaction of the components with each other and identify

(Refer Slide Time 19:41)



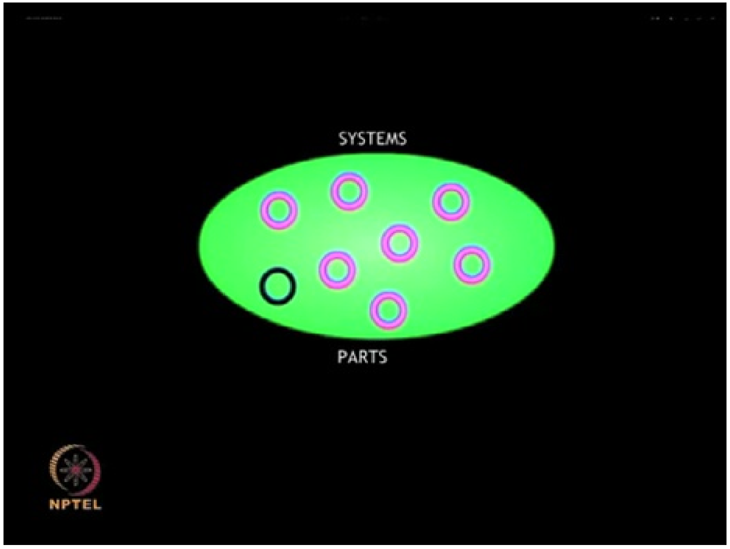
the interaction of the components with the environment which modulate the parts either directly or indirectly through modulation of internal interactions.

(Refer Slide Time 19:56)



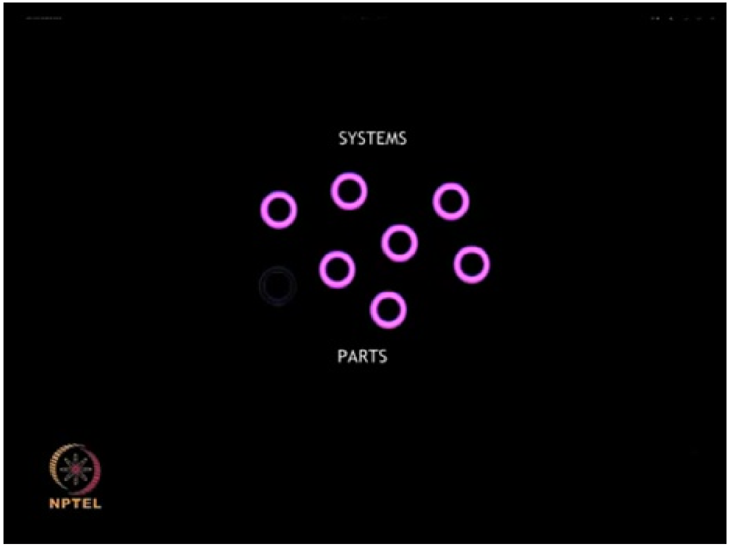
The systems biology concept can be understood with help of two approaches such as reductionist approach and integrative approach.

(Refer Slide Time 20:09)



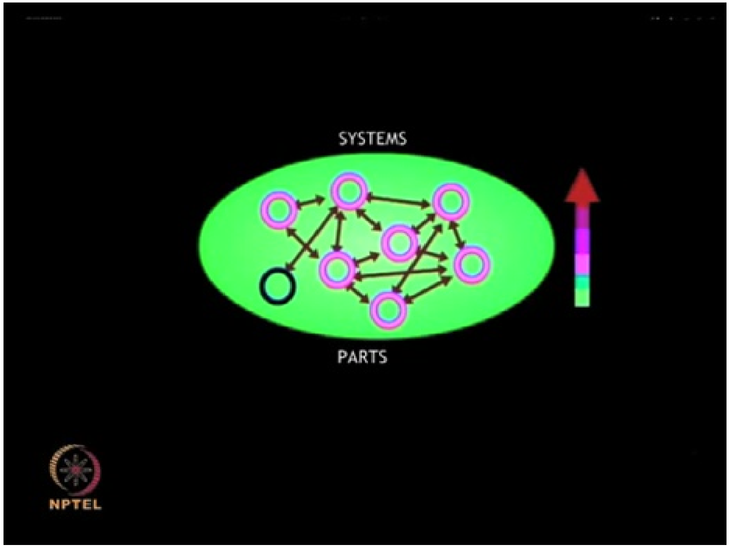
The reductionist approach focuses on disintegrating the system into its component parts and studying them

(Refer Slide Time 20:23)



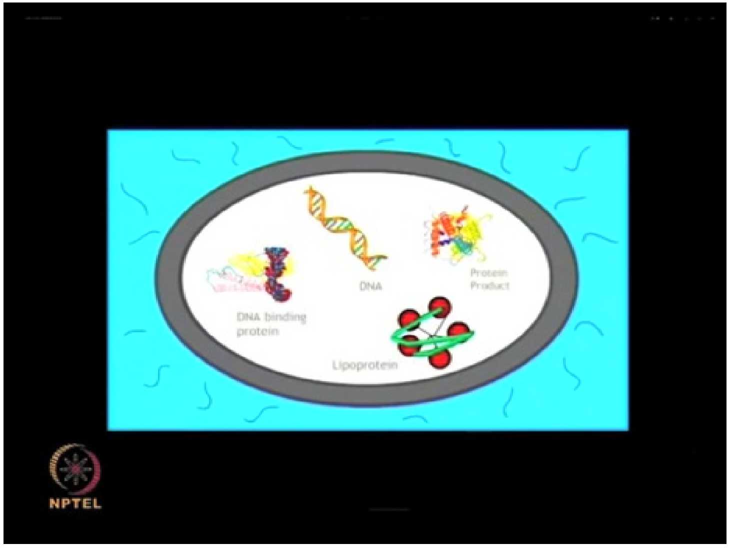
where as the

(Refer Slide Time 20:29)



integrative approach focuses on integrating the study of individual component to form conclusions about the system.

(Refer Slide Time 20:35)

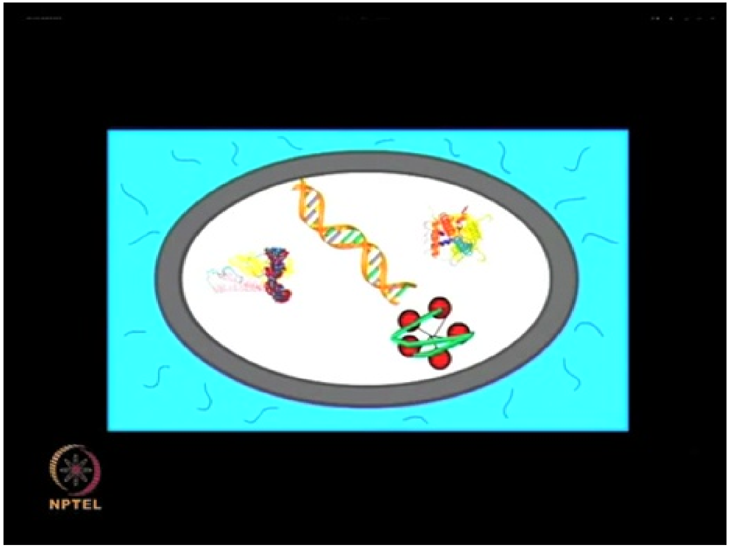


Consider a cell with its component molecules. Let's say we want to study the metabolic pathway as a biological system. When the environment of the cell is perturbed a little, the individual components undergo unique changes such as

(Refer Slide Time 20:56)

increase in the production rate

(Refer Slide Time 20:57)

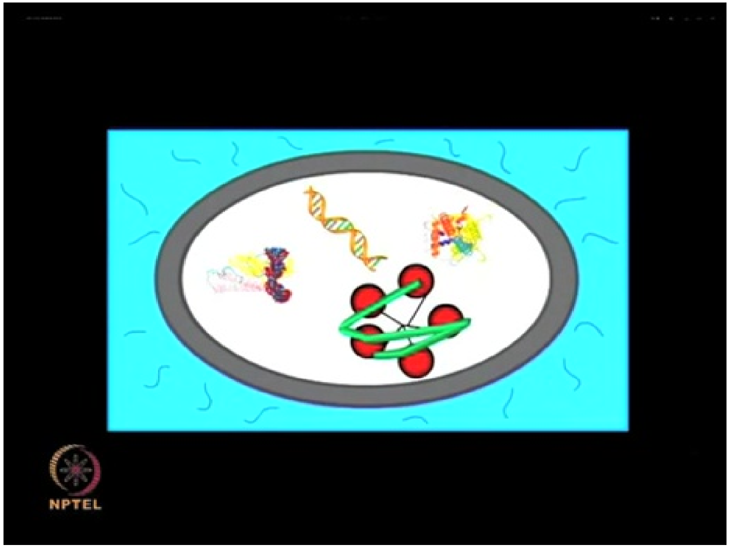


or decrease in their amount.

(Refer Slide Time 21:01)

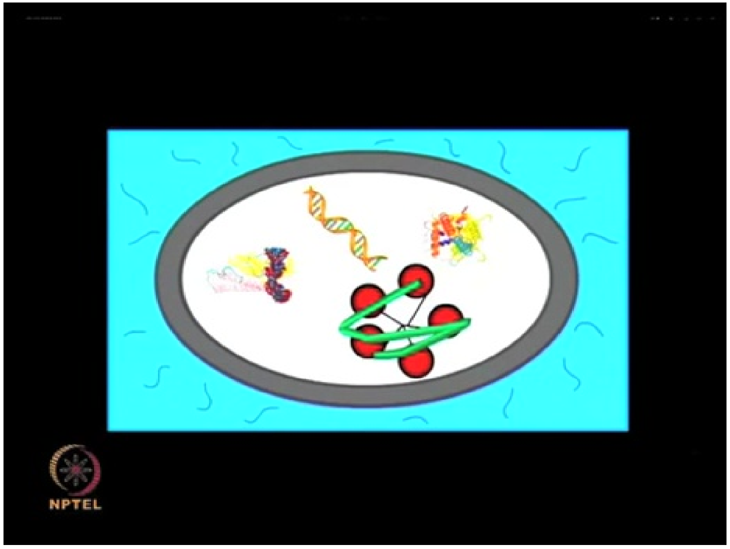
At this stage

(Refer Slide Time 21:02)



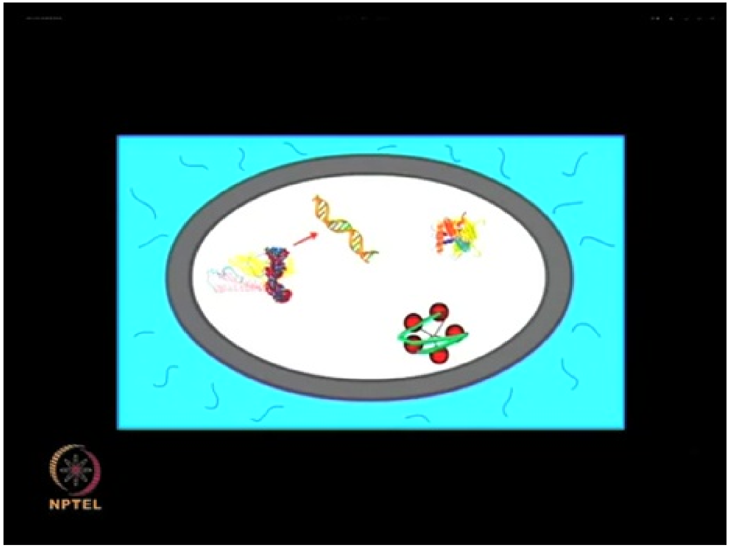
, due to lack of knowledge of the nature of interaction of the proteins, we cannot interpret how the system gets affected. But when study the interaction of one component with the other, we can conclude that the increase in rate; that if DNA binding protein leads to increase in the synthesized amount of DNA

(Refer Slide Time 21:29)



which further changes

(Refer Slide Time 21:35)



the final amount of lipoproteins produced

(Refer Slide Time 21:40)

Thus we can see,

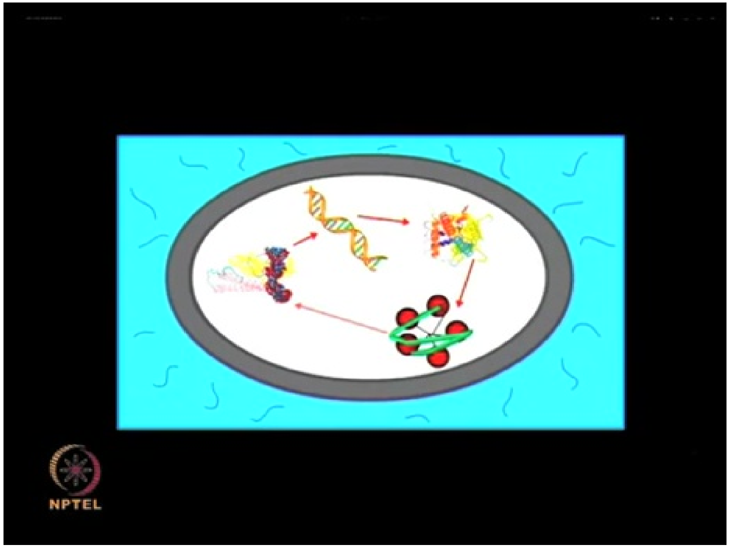
(Refer Slide Time 21:41)

that to study a system

(Refer Slide Time 21:42)

we need to analyze

(Refer Slide Time 21:43)



not just the components but their interactions. These biological systems can be protein-protein interaction networks, gene regulatory networks protein DNA networks, protein lipid networks and metabolic networks.

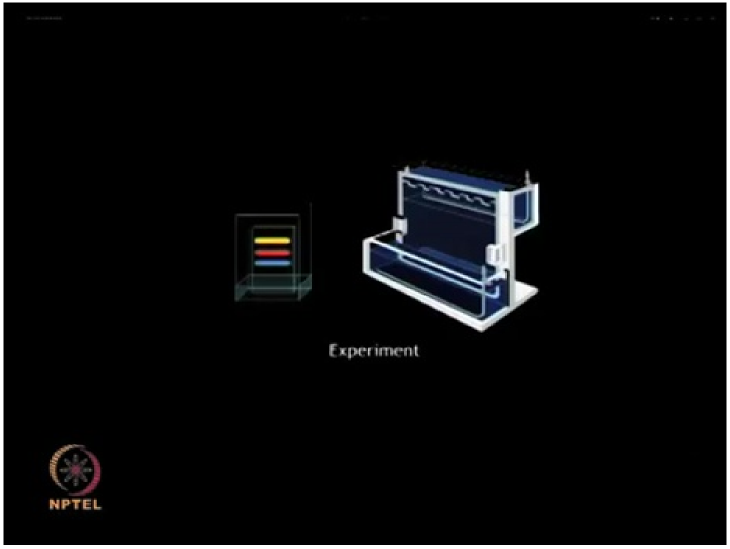
(Refer Slide Time 22:10)

To study the system we need to know about the components and its interactions. The data about the components comes from

(Refer Slide Time 22:20)

genomic and proteomic studies. The information about the molecular interactions comes from the interatomic studies.

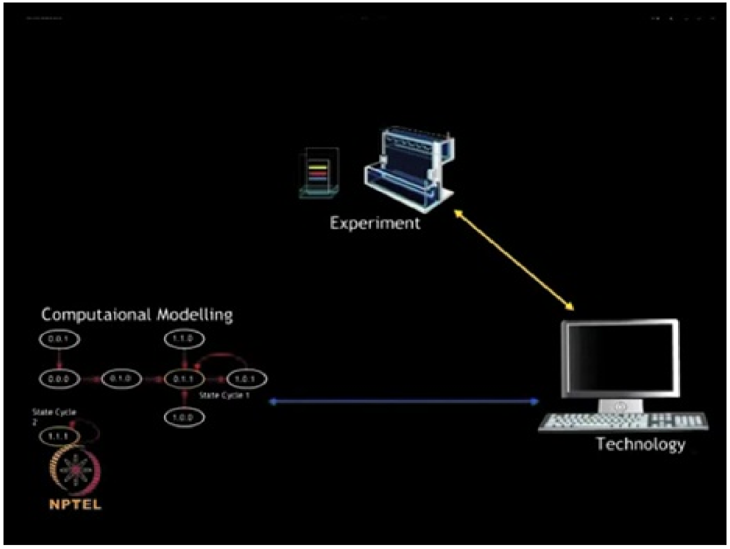
(Refer Slide Time 22:30)



Experiment

Here it is shown that in a systems approach, experiment,

(Refer Slide Time 22:40)



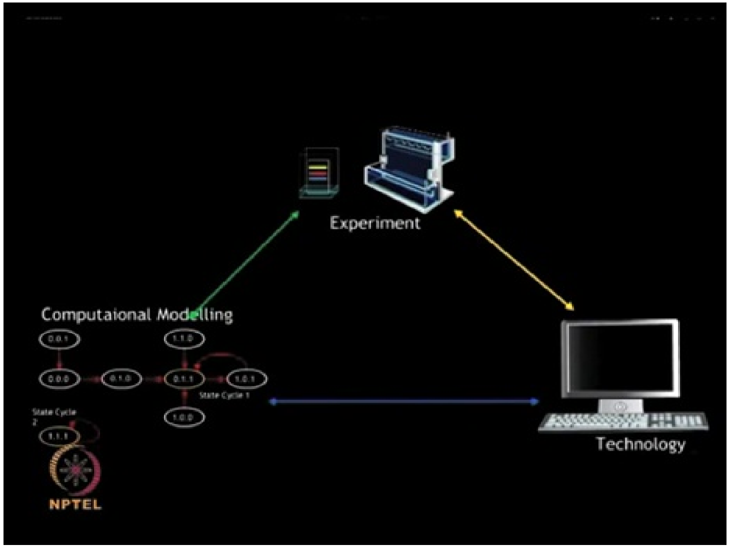
Experiment

Computaional Modelling

Technology

technology and computational modeling;

(Refer Slide Time 22:43)



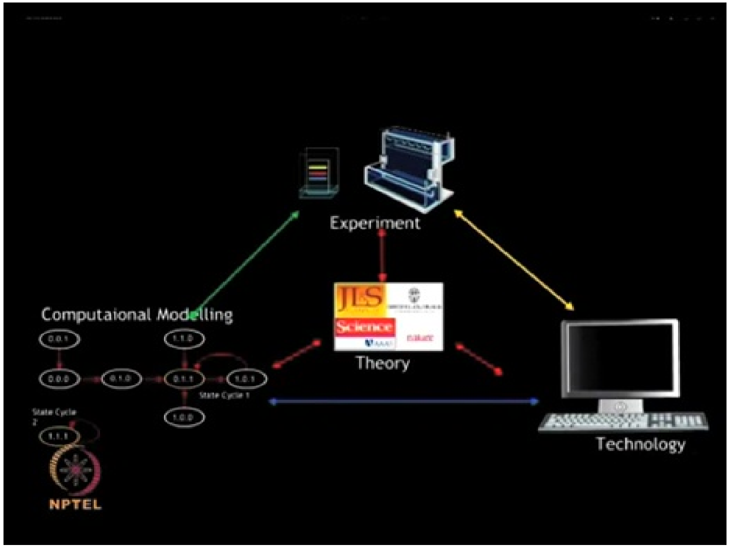
Experiment

Computaional Modelling

Technology

this triangle is very important which has to be linked with the

(Refer Slide Time 22:47)



Experiment

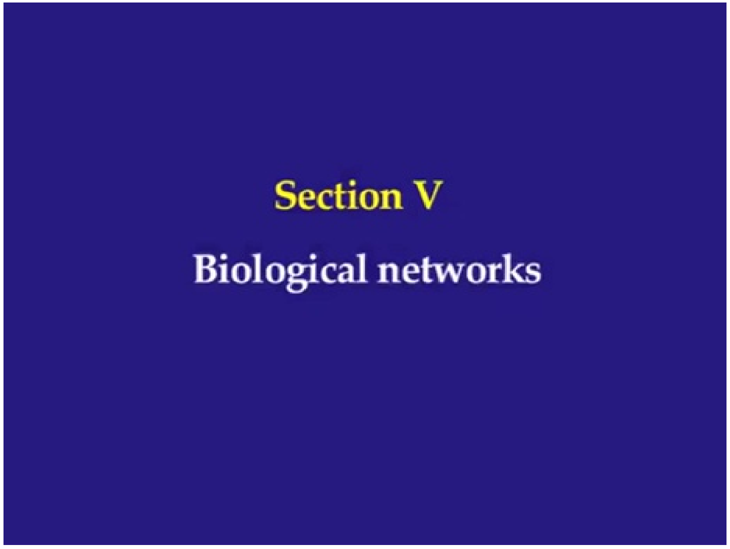
Theory

Computaional Modelling

Technology

theory to form a Systems Triangle

(Refer Slide Time 22:53)



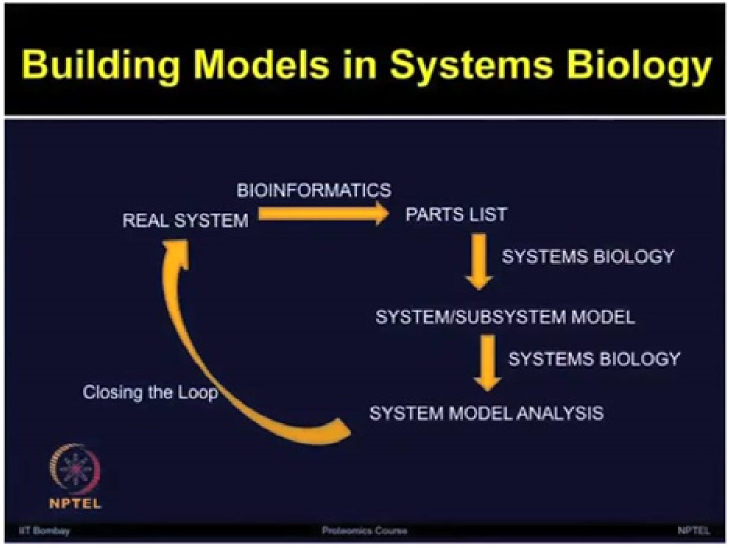
Section V

Biological networks

Lets now talk about how to model biological networks.

To build a model in Systems Biology, first of all the parts list can be generated by using data sets derived from the using Biology approaches. The Systems or subsystems models can be generated

(Refer Slide Time 23:21)



Building Models in Systems Biology

REAL SYSTEM → BIOINFORMATICS → PARTS LIST

PARTS LIST → SYSTEMS BIOLOGY → SYSTEM/SUBSYSTEM MODEL

SYSTEM/SUBSYSTEM MODEL → SYSTEMS BIOLOGY → SYSTEM MODEL ANALYSIS

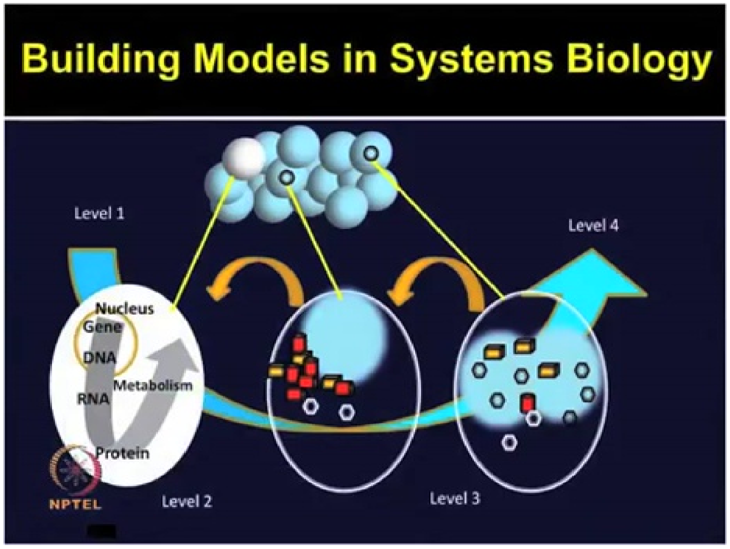
SYSTEM MODEL ANALYSIS → Closing the Loop → REAL SYSTEM

which can be used for Systems model analysis.

Now this could be applied for the real systems and by applying the knowledge, using bio-informatics tools, it could be again applied back to the original components which could be used to derive some hypothesis and validation of the datasets. So it wil work like a closed loop.

To build the models in Systems Biology, information is generated at different levels.

(Refer Slide Time 23:55)



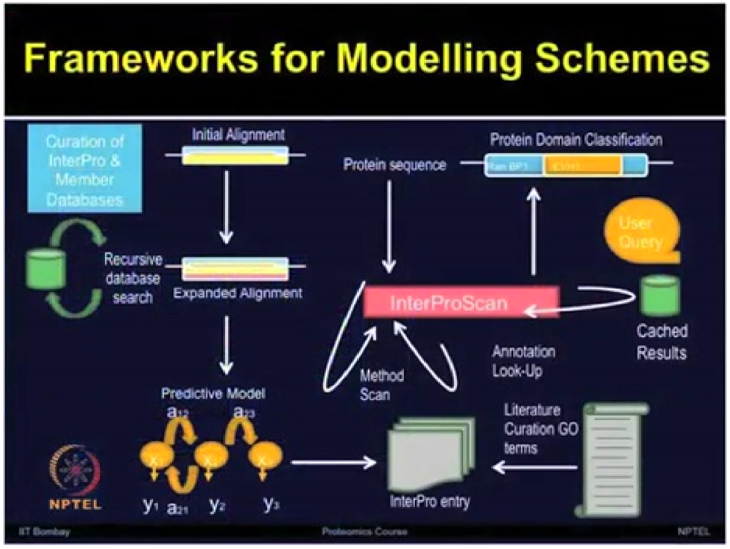
Building Models in Systems Biology

Level 1, such as DNA and gene expression, level 2 the intracellular networks, level 3 cell-cell and transmembrane signals, and level 4 integrated organ level information.

What are the frameworks required for modeling schemes?

Different types of deterministic or stochastic models have been proposed. The compartmental variables or individual or functional variables have been studied. The especially homogenous or specially explicit models are generated, which could be applied in the uniform time scale or separated time scale. This framework would involve single scale entities or cross scale entities.

(Refer Slide Time 24:48)



Frameworks for Modelling Schemes

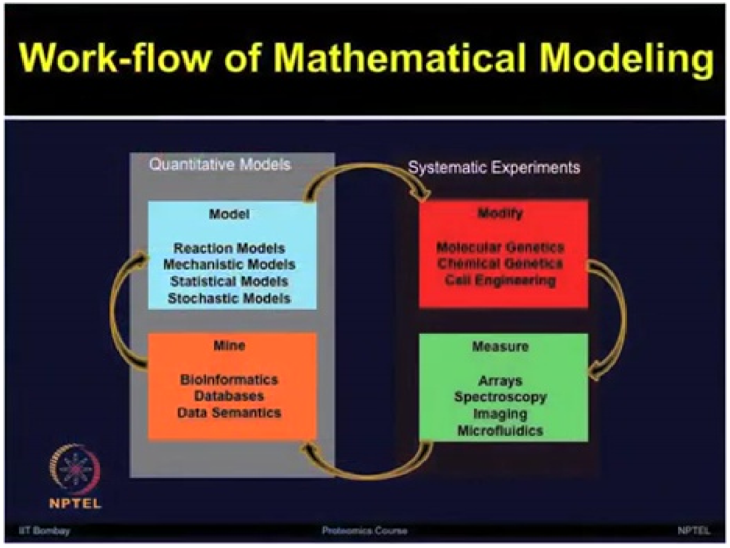
Curation of InterPro & Member Databases
Initial Alignment
Recursive database search
Expanded Alignment
Predictive Model
a_{12} a_{23}
y_1 a_{21} y_2 y_3
Protein sequence
Protein Domain Classification
User Query
InterProScan
Cached Results
Method Scan
Annotation Look-Up
Literature Curation GO terms
InterPro entry
NPTEL
IIT Bombay Proteomics Course NPTEL

As you can see here, this framework requires different level of information in very complex manner; whether it is curation of the databases, how to align these information using bioinformatic tools to generate the predictive models which could be also developed by using the literature curated data sets or experimental data sets and finally it could be used to study the systems level properties.

Let's discuss the workflow of mathematical modeling.

A paradigm can be proposed based on modify, model, measure and mine.

(Refer Slide Time 25:37)



Work-flow of Mathematical Modeling

Quantitative Models

Model
Reaction Models
Mechanistic Models
Statistical Models
Stochastic Models

Mine
Bioinformatics
Databases
Data Semantics

Systematic Experiments

Modify
Molecular Genetics
Chemical Genetics
Cell Engineering

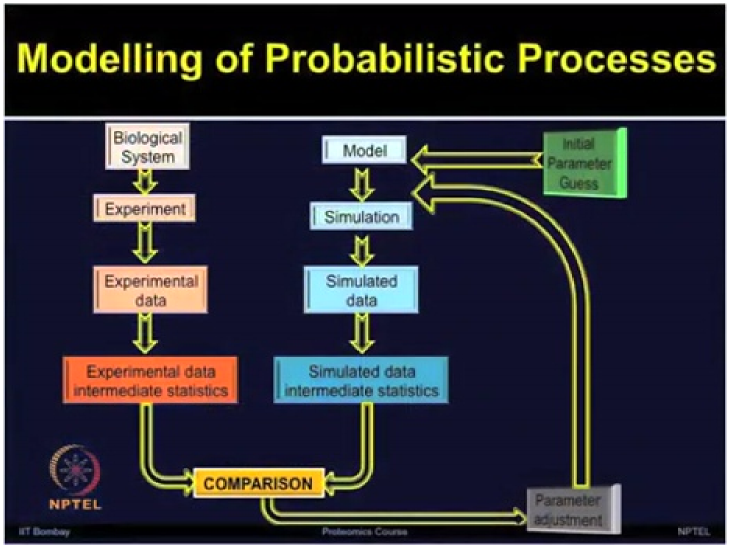
Measure
Arrays
Spectroscopy
Imaging
Microfluidics

NPTEL
IIT Bombay Proteomics Course NPTEL

Systematic experiments, different type of molecular genetics, chemical genetics and cell engineering approaches can be used for modifying and different level of measurements by applying microarrays, spectroscopy imaging and microfluidics based approaches from proteomics and genomics can be used further for mining which involves bioinformatics, databases and data semantics.

Now these datasets could be used to derive the model, which could be reaction, mechanistic, statistical or stochastic models. So,starting from systematic experiments to reaching and deriving the quantitative models, this workflow can be applied.

(Refer Slide Time 26:33)



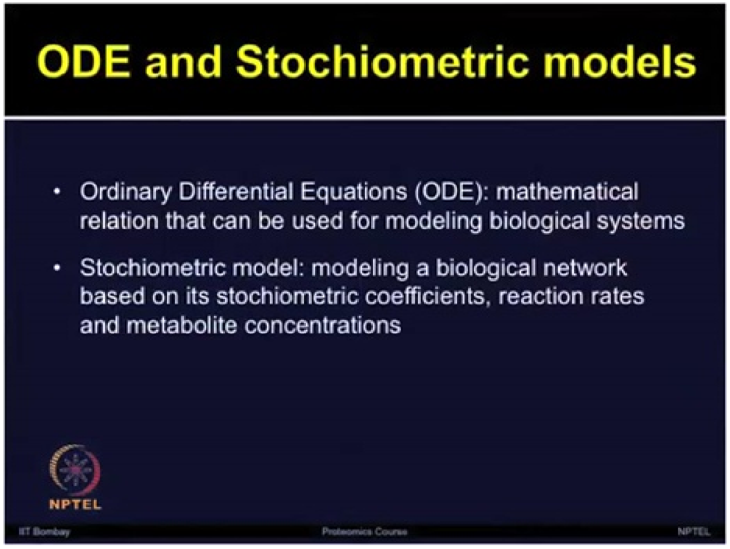
Modelling of Probabilistic Processes

The modeling of probabilistic processes involves, let’s say, you want to study a biological system. So some experiment has to be performed. The experimental datasets will be generated from which some statistics can be applied, which can be used for the comparison.

Now different types of models can be generated using simulations and simulations datasets, which can be used for intermediate statistics

And by comparing these two types of information and adjusting the parameters, one can study the systems and derive probabilistic processes.

(Refer Slide Time 27:21)



ODE and Stochiometric models

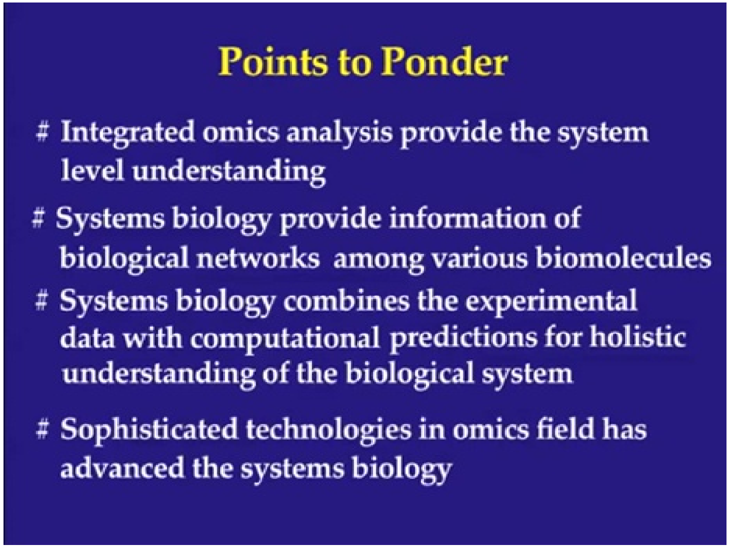
- Ordinary Differential Equations (ODE): mathematical relation that can be used for modeling biological systems
- Stochiometric model: modeling a biological network based on its stochiometric coefficients, reaction rates and metabolite concentrations

What is ordinary differential equations and stochiometric models? The quantitative analysis measures and aims to make models for precise kinetic parameters of a system's network component. It also uses the properties of network connectivity.

The ODE is a mathematical relation that can be used for modeling the biological systems. The quantitative models mostly use Ordinary Differential Equations or ODE to link the reactants and products concentrations through the reaction rate constants.

To develop the computationally efficient and reliable models of the underlying gene regulatory networks, these ODE models can be used. The stochiometric model, it is modeling a biological network based on its stochiometric coefficients, reaction rates and metabolite concentrations

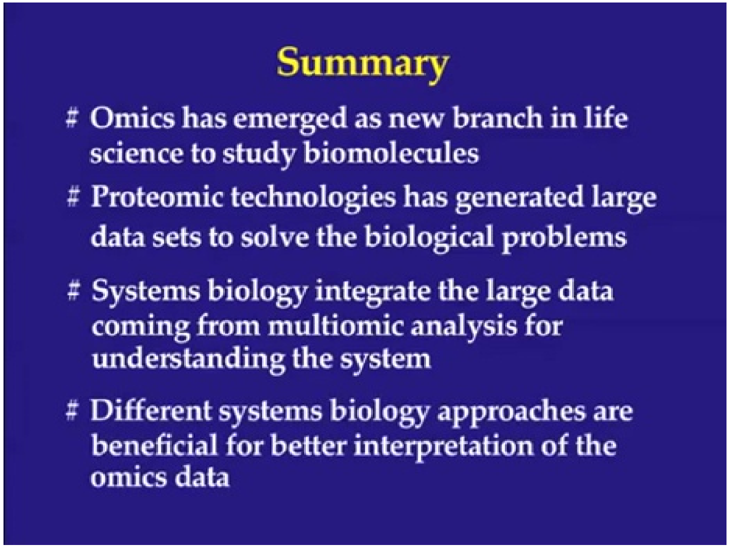
(Refer Slide Time 28:42)



Points to Ponder

- # Integrated omics analysis provide the system level understanding
- # Systems biology provide information of biological networks among various biomolecules
- # Systems biology combines the experimental data with computational predictions for holistic understanding of the biological system
- # Sophisticated technologies in omics field has advanced the systems biology

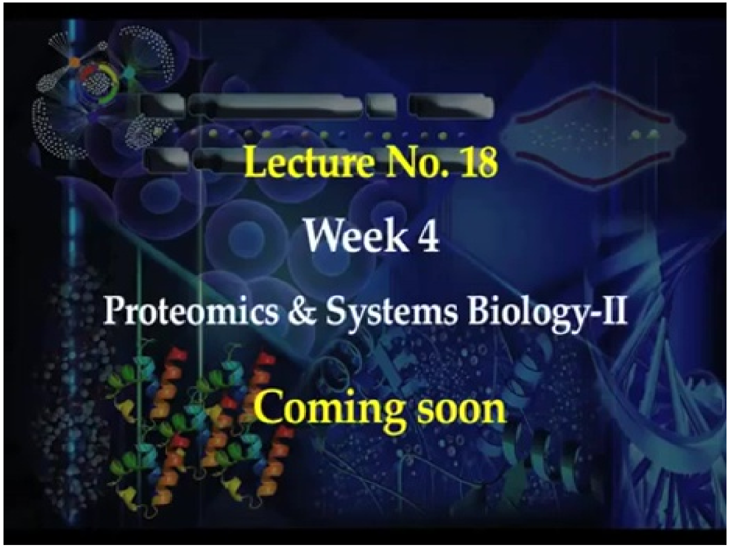
(Refer Slide Time 29:01)



Summary

- # Omics has emerged as new branch in life science to study biomolecules
- # Proteomic technologies has generated large data sets to solve the biological problems
- # Systems biology integrate the large data coming from multiomic analysis for understanding the system
- # Different systems biology approaches are beneficial for better interpretation of the omics data

(Refer Slide Time 29:08)



Lecture No. 18

Week 4

Proteomics & Systems Biology-II

Coming soon