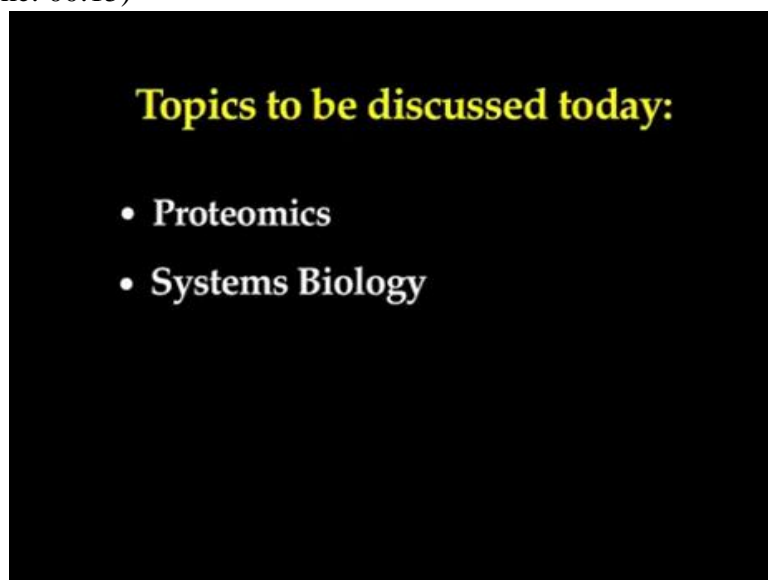


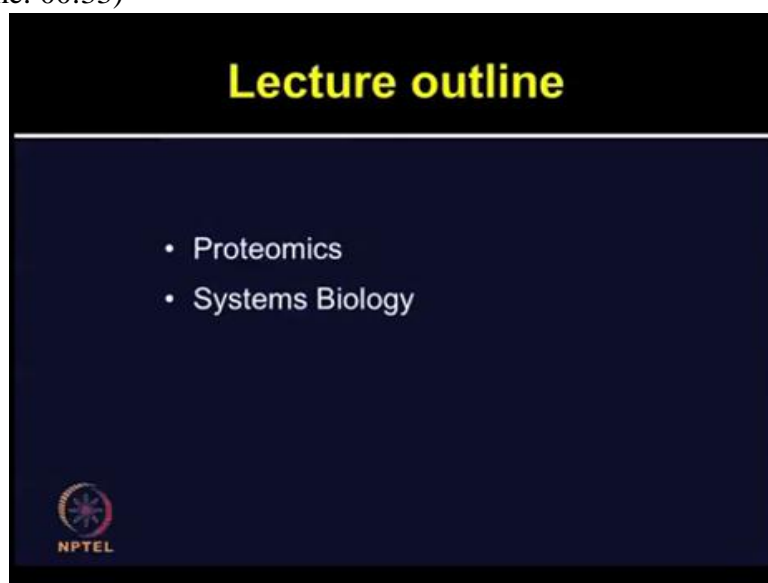
Proteins and Gel-Based Proteomics
Professor Sanjeeva Srivastava
Department of Biosciences and Bioengineering
Indian Institute of Technology, Bombay
Mod 03 Lecture Number 7

(Refer Slide Time: 00:15)



Welcome to the Proteomics course. Today's lecture is on Proteomics and Systems Biology.

(Refer Slide Time: 00:33)



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

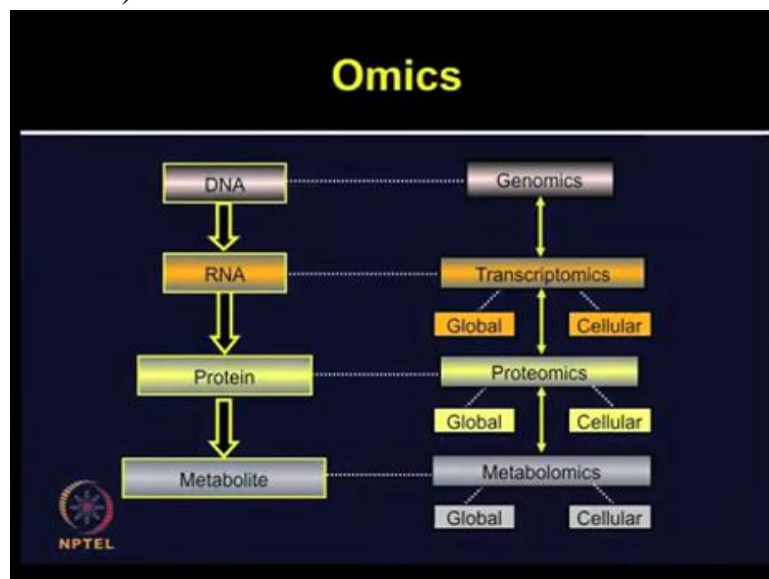
(Refer Slide Time: 00:53)

Section I

Understanding Proteomics

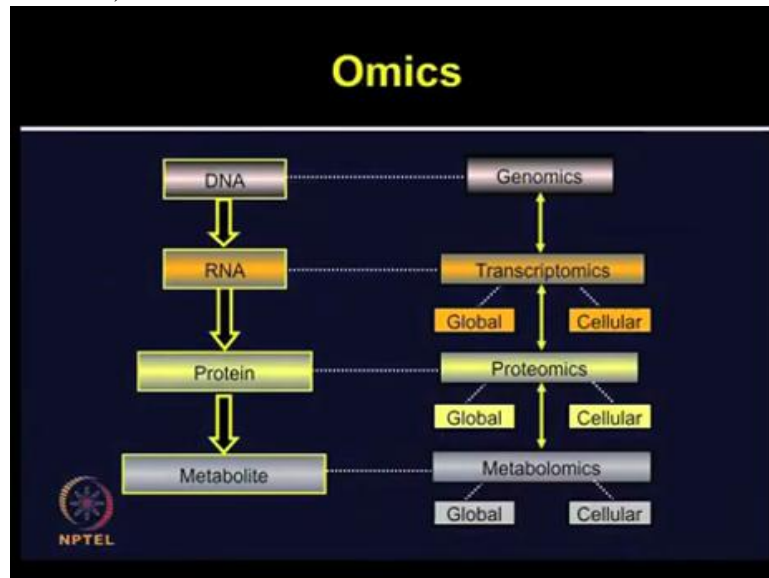
Today we will continue on first Proteomics and then how different types of Omic data can be applied for Systems Approach Analysis; that will be discussed.

(Refer Slide Time: 01:03)



So this slide, I have shown you how you can define different types of Omic platforms. Omic is a suffix derived from the Greek word “ome” means all or every. The Omics is used as a suffix which has enabled the explosion of terms; genomics, transcriptomics, proteomics and metabolomics and so on and so forth. Omics also implies an integration of biology with information science and conveys large scale biology using systems approach.

(Refer Slide Time: 01:41)



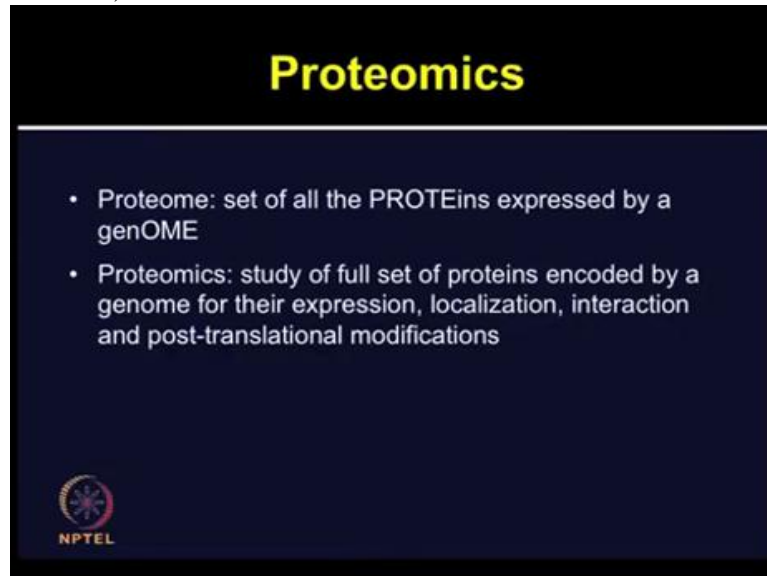
So as you can see the slide, if you are studying about DNA that will be in totality known as genomics, RNA study in transcriptomics, proteins in proteomics and metabolites in metabolomics. If you are looking at all of the cellular contents of the proteome that will be known as cellular proteomics or cellular genomics, and similarly all of the proteome of the organism will be known as global proteomics or similarly at the gene level, global genomics.

(Refer Slide Time: 02:23)



So let's first start with proteomics.

(Refer Slide Time: 02:31)



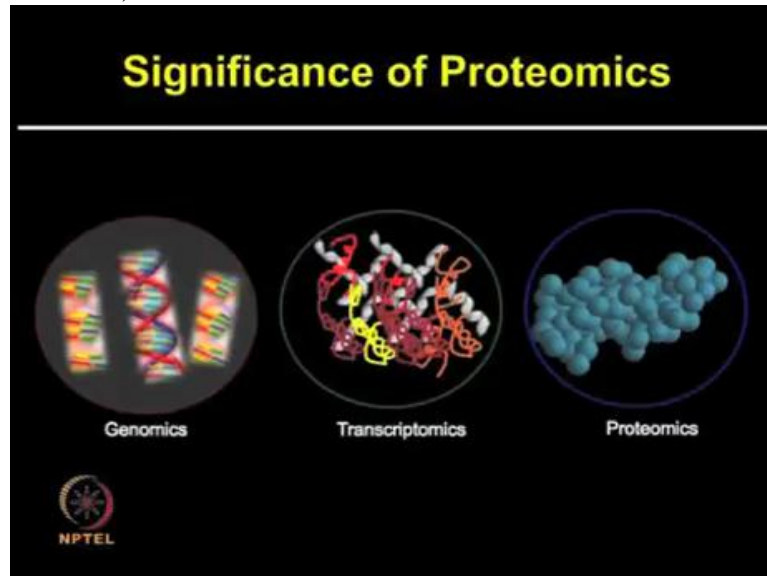
So if you remember from the previous lecture, proteome is set of all the proteins which are expressed by a genome. Proteomics is study of proteins and their properties to provide an integrated view of cellular processes.

What are these different properties? These properties include the extent of the protein expression, how different type of post-translation modification and co-modifications occurs in the cell, different types of enzyme regulation, whether it is activation or inactivation and then different types of intermolecular protein-protein interactions

The current goals of proteomics are very broad. It is including the diverse properties of proteins which we have discussed in the earlier lectures looking at the vast side chain chemistry of amino acids and different levels of protein structures.

So whether it is sequence, quantity, the state of modification, activity, interactions of proteins with other proteins and other biomolecules and sub-cellular distributions and structure analysis, all of these are broad goals of proteomics.

(Refer Slide Time: 03:59)

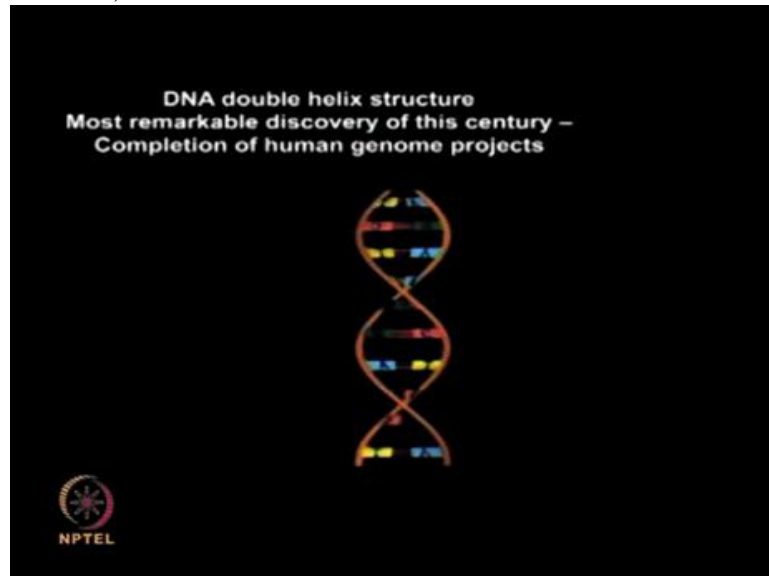


Central dogma concept was discussed earlier which is an orderly and uni-directional flow of information encoded in base sequences of cells passed on from DNA to RNA and then to the proteins. This is the simplest definition of central dogma.

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.

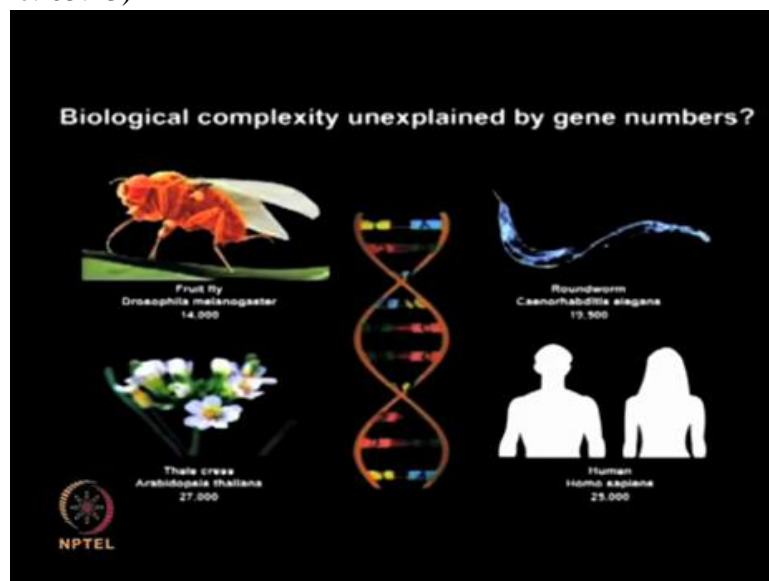
A gene can code for several types of proteins because of its alternative splicing and post-transcriptional and post-translational modifications. Therefore it suggests that studying proteins is more challenging than genome or transcriptome. Therefore proteomics has great significance to understand the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

(Refer Slide Time: 05:17)



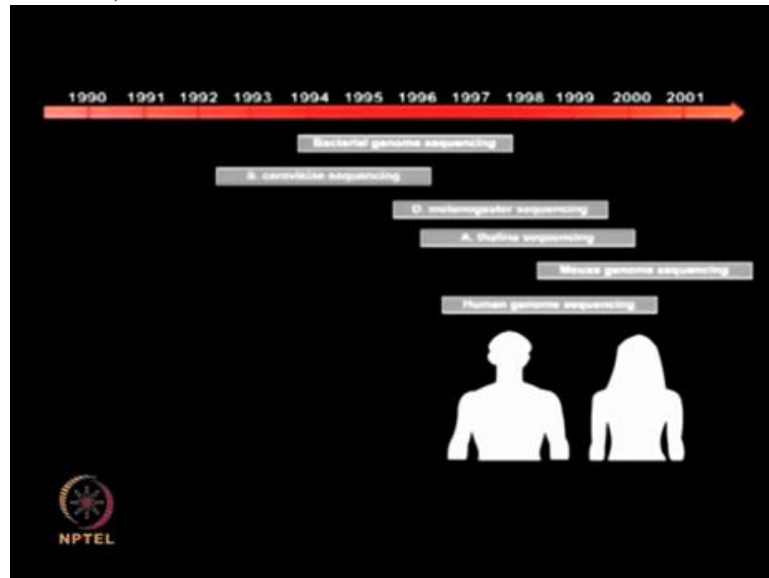
The completion of genome sequencing projects...

(Refer Slide Time: 05:23)



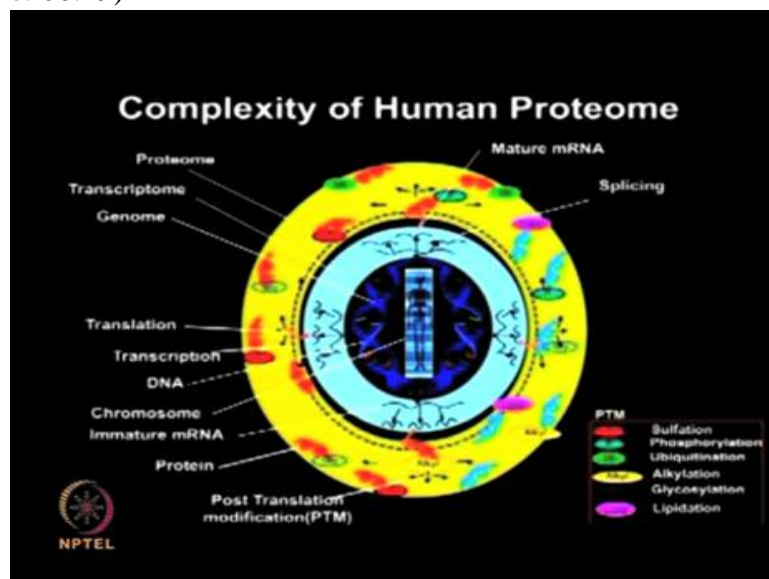
... of several organism including human has been one of the most remarkable achievements of this century. However, these have not been sufficient to unravel the mystery of complex biological processes. The similar gene numbers of many diverse group of organisms has failed to explain their varying biological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

(Refer Slide Time: 05:50)



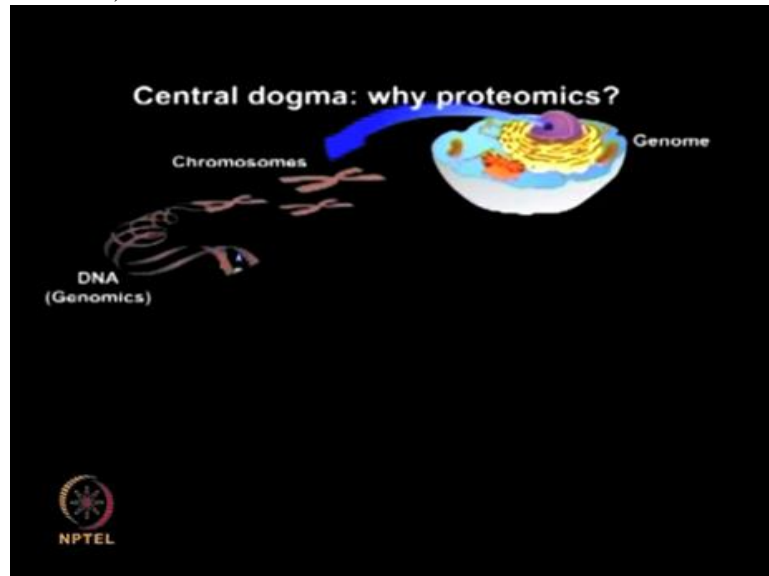
A more meaningful understanding of biological function can be obtained through the characterization of products of gene expression, the protein, which serve as ultimate effective molecule of biological systems

(Refer Slide Time: 06:29)



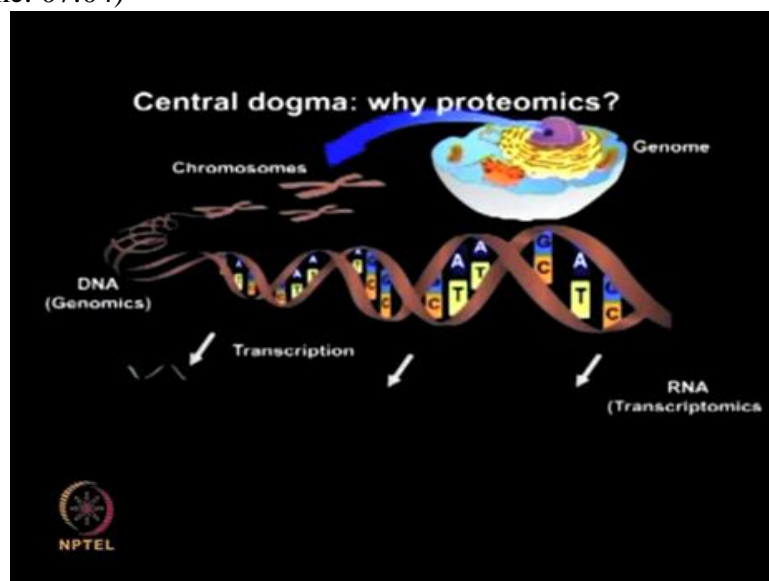
The proteomics refers to the study of entire protein complement expressed by an organism at any given time. While the genome of an organism is mostly static the proteome is dynamic and it changes with environment and time, thereby elevating its complexity level. The gene regulation is regulated by several post-transcriptional and post-translational modifications due to which the number of proteins expressed in a cell is much greater than its genomic counterpart.

(Refer Slide Time: 06:57)



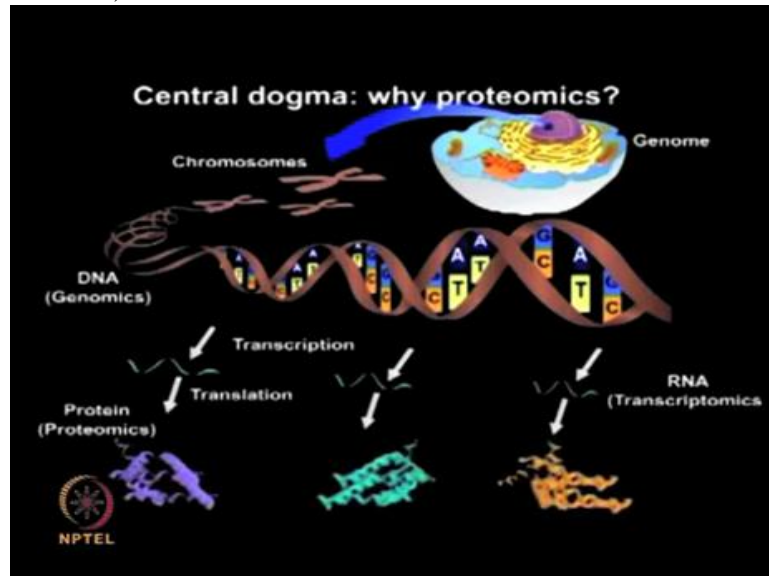
The proteomics aims to decipher the structure ...

(Refer Slide Time: 07:04)



... and function of...

(Refer Slide Time: 07:17)



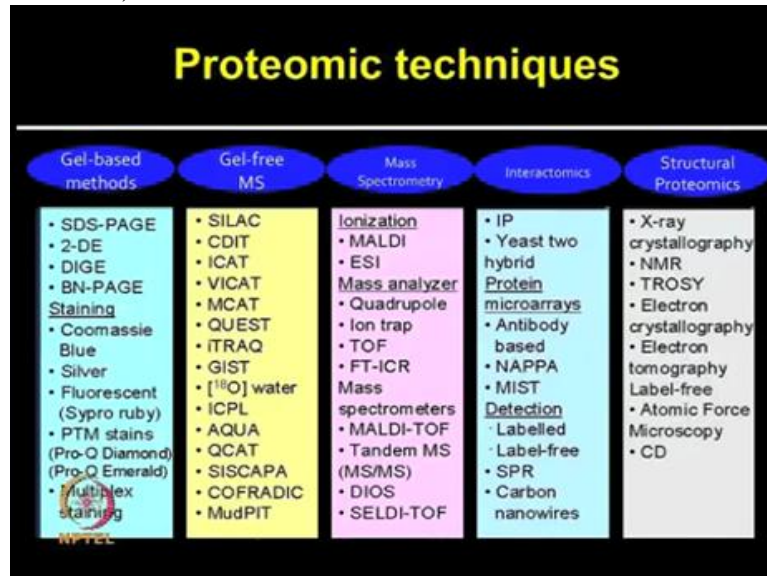
... all proteins in a given cell under specific conditions and to obtain a global view of cellular processes at the protein level Study at the DNA level is known as genomics, RNA level transcriptomics, and protein level proteomics.

(Refer Slide Time: 07:28)



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

(Refer Slide Time: 07:42)



There are various proteomic techniques which are currently employed in different applications. We will talk in detail about different proteomic technologies in the subsequent modules of this course but briefly, 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.

As I have shown some abbreviations in the slide, broadly different types of gel-based proteomics such as Sodium Dodecyl Sulfate PolyAcrylamide Gel Electrophoresis SDS PAGE, two-dimensional electrophoresis 2DE, Difference In Gel Electrophoresis DIGE, Blue Native PAGE as well as different types of staining methods such as Coomassie, silver, fluorescent dyes and phospho stains such as Pro-Q Diamond and multiplex staining methods. All these can be grouped under gel-based methods.

The gel-free methods especially in the mass spectrometry side includes SILAC which is Stable Isotope Labeling by Amino acids in Cell culture, CDIT Culture-Derived Isotope Tags, ICAT Isotope Coded Affinity Tagging, VICAT Visible Isotope Coded Affinity Tagging, MCAT is Mass Coded Affinity Tagging, and then QUEST which is Quantitation Using Enhanced Signal Tags, ITRAQ Isobaric Tagging for Relative and Absolute Quantitation, GIST Global Internal Standard Technology, ICPL Isotope Coded Protein Labeling , AQUA Absolute Quantitation, SISCAPA Stable Isotope Standards and Capture by Anti-Peptide Antibodies, COFRADIC Combined Fractional Diagonal Chromatography, and MudPIT

which is Multidimensional Protein Identification Technology. All of these are various new advancements in gel-free methodologies.

The basic mass spectrometry which is central to the proteomic application includes different types of ionization sources such as Matrix-Assisted Laser Desorption/Ionization MALDI, Electro Spray ionization ESI and different types of mass analyzers such as Quadrupole, Time Of Flight, Ion Trap and Fourier Transform Mass Spectrometry, then different types of Tandem MS based systems are also used. The Surface-Enhanced Laser Desorption/Ionization-Time Of Flight SELDI-TOF is also used for various clinical applications.

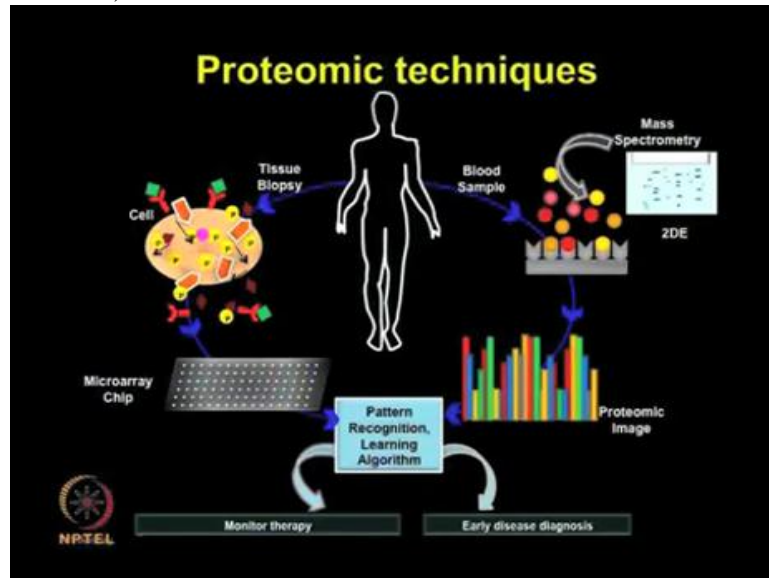
The protein interaction methodologies include immuno-precipitation, Yeast to hybrid methods and different types of protein Microarray platforms such as antibody arrays, Nucleic Acid Programmable Protein Arrays, Multiple Spotting Techniques and various other cell-based and cell-free expression based protein microarrays.

The detection can be either based on the label methods using fluorescence, chemiluminescence or radioactivity, or it could be different type of label-free methods such as Surface Plasmon resonance, interferometry based methods or different type of conductance-based methods employing nanotubes and nanowires.

The Structural Proteomics, it involves X-ray crystallography, Nuclear Magnetic Resonance NMR, Transverse Relaxation Optimization Spectroscopy TROSY, Circular Dichroism CD and different types of microscopy methods including Atomic Force Microscopy and Electron Microscopy.

So, as we have seen there are large numbers 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.

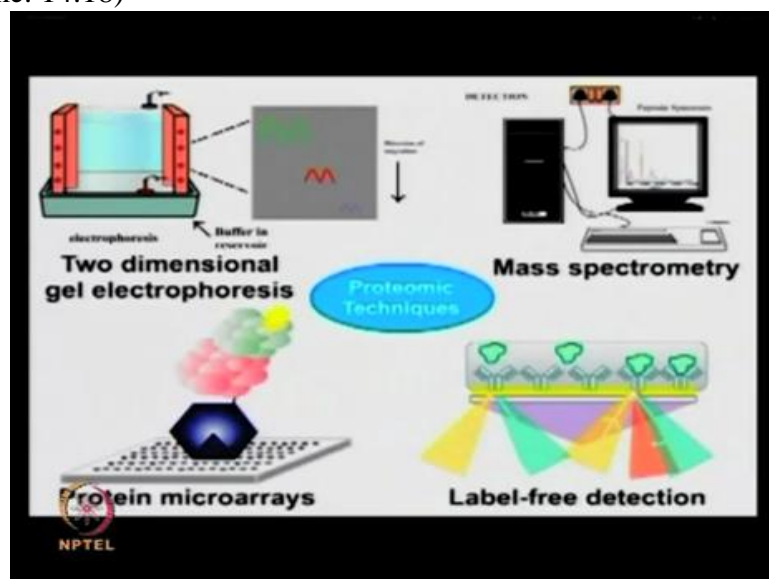
(Refer Slide Time: 12:49)



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.

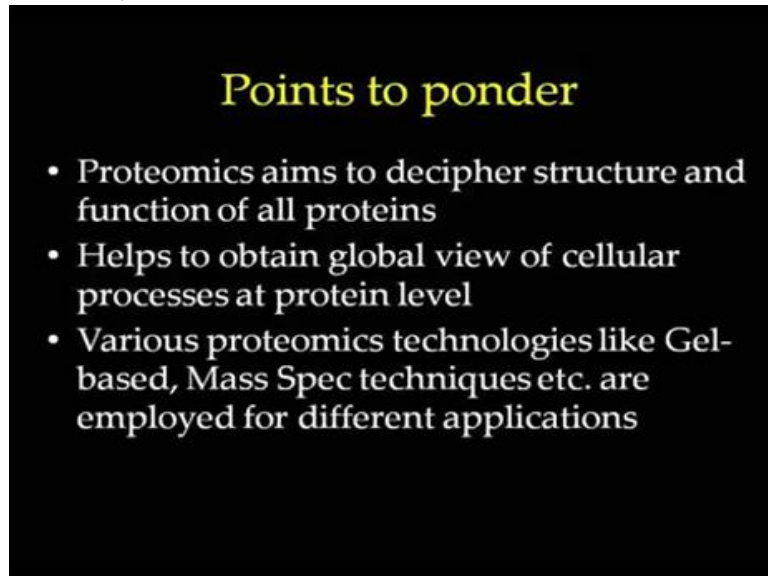
Eventually these types of results will enhance knowledge for the monitoring the therapy response as well as identification of early disease diagnosis. This is just one example. Similarly multiple type of proteomic technologies can be used for different applications.

(Refer Slide Time: 14:18)



There are several proteomic techniques which are employed for studying these proteins such as two dimensional gel electrophoresis, mass spectrometry, protein microarrays as well as some label-free detection techniques such as surface plasma resonance SPR.

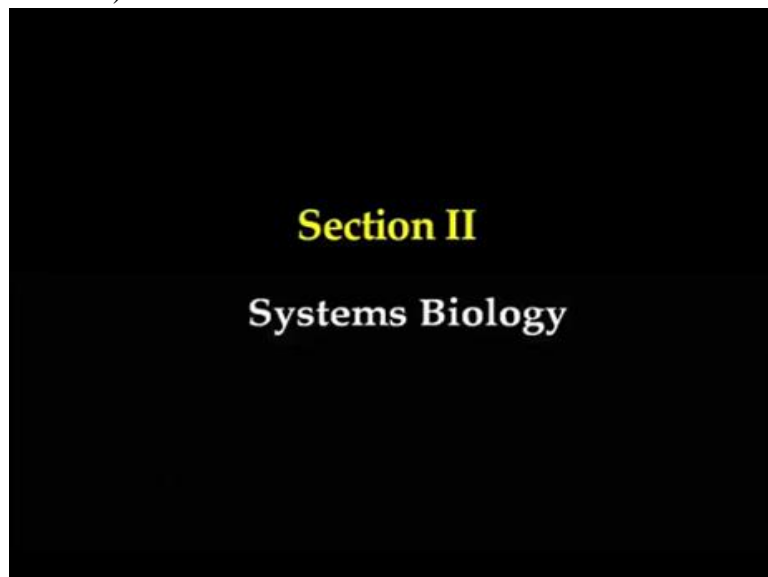
(Refer Slide Time: 14:29)



Points to ponder

- Proteomics aims to decipher structure and function of all proteins
- Helps to obtain global view of cellular processes at protein level
- Various proteomics technologies like Gel-based, Mass Spec techniques etc. are employed for different applications

(Refer Slide Time: 14:33)

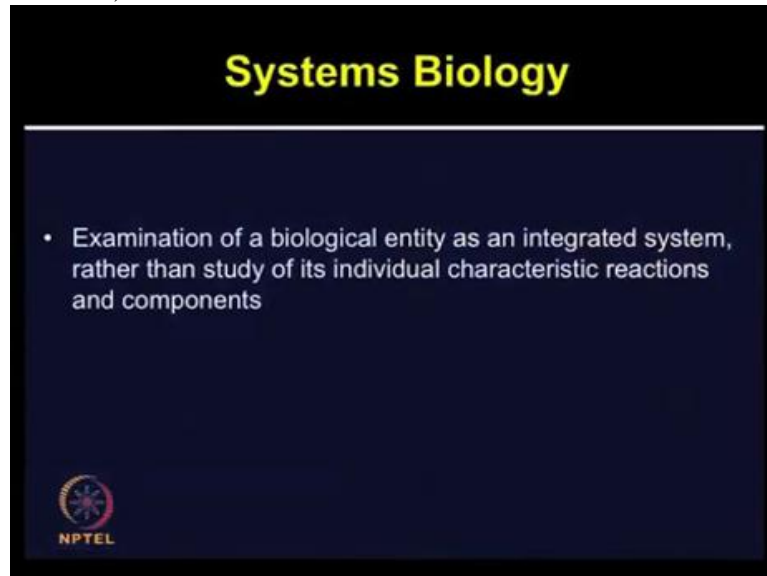


Section II

Systems Biology

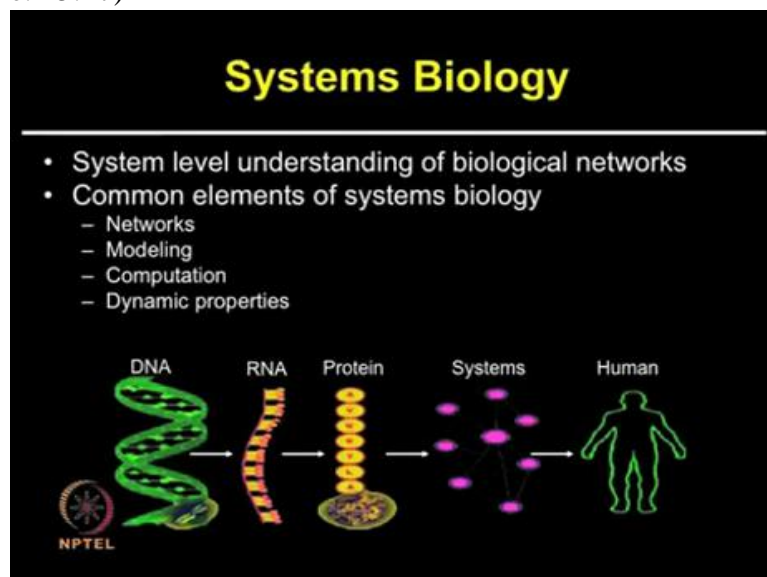
After discussing proteomics, now let's talk about Systems Biology. So what is Systems Biology? Is that an historic knowledge, a method to understand biological systems or a tool to solve the practical problems?

(Refer Slide Time: 14:56)



Systems Biology is the examination of a biological entity as an integrated system rather than 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: 15:27)

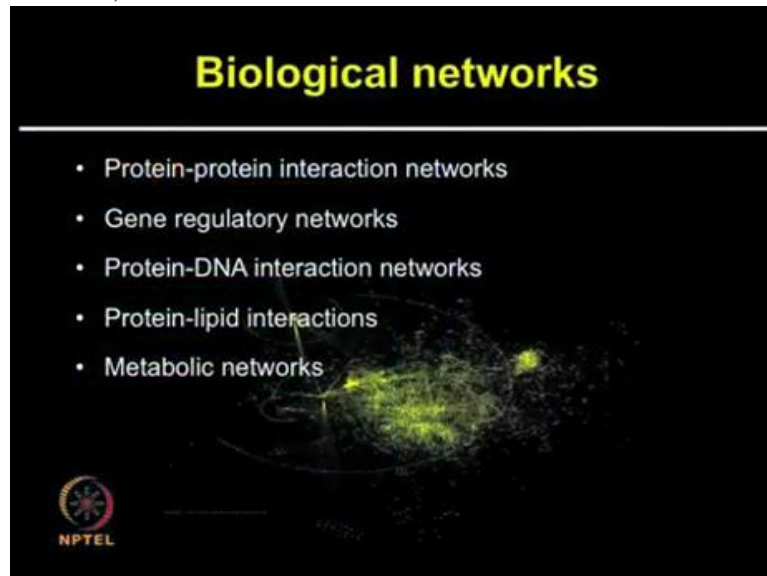


The systems level understanding of biological networks requires information from different levels. As you can see, from DNA, to RNA, to proteins forming systems and then that information can be applied to understand a complex system for different organism.

The biological information is represented by the networks of interacting elements and dynamic responses to the perturbations. These networks provide insights which cannot be

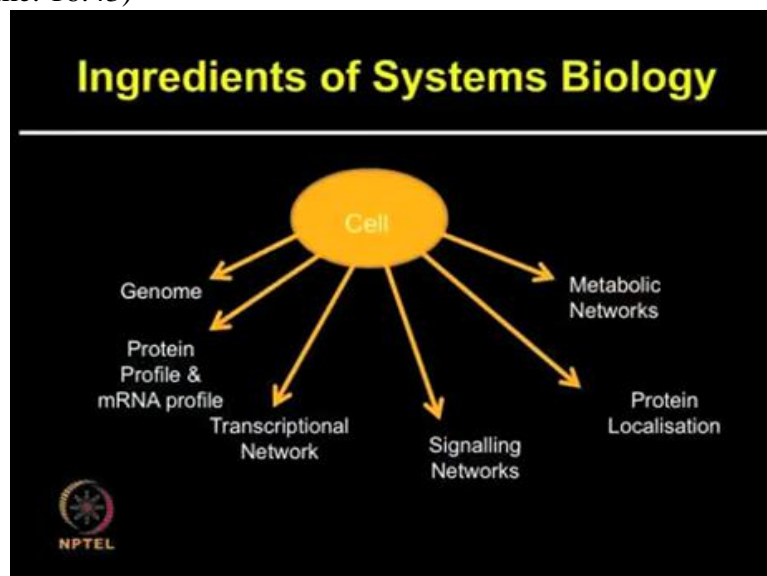
analyzed from the isolated components of the system. The common elements of the Systems Biology include networks, modeling, computation and dynamic properties.

(Refer Slide Time: 16:27)



There are different types of biological networks such as protein-protein interaction networks, gene regulatory networks, protein and DNA interaction networks; protein lipid, protein-other biomolecule networks and metabolic networks.

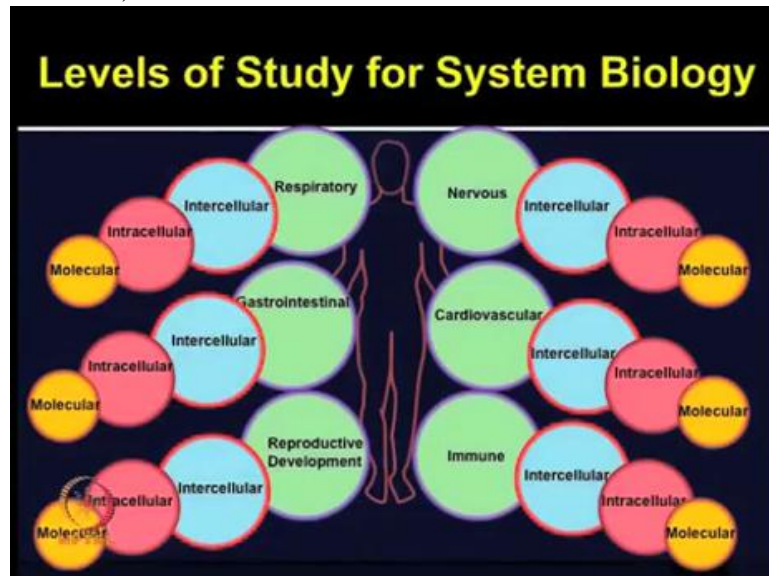
(Refer Slide Time: 16:43)



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

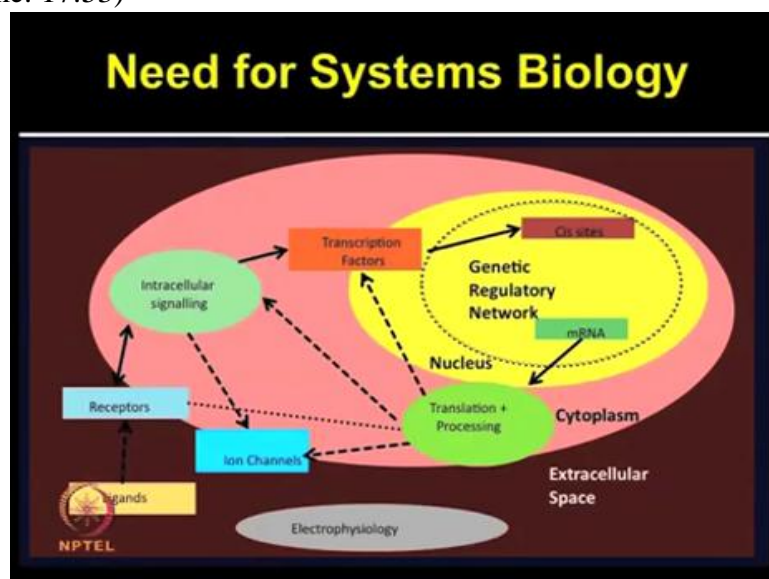
is localized in the intracellular dynamics and metabolic networks. All of these are ingredients of study about a system.

(Refer Slide Time: 17:21)



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: 17:53)

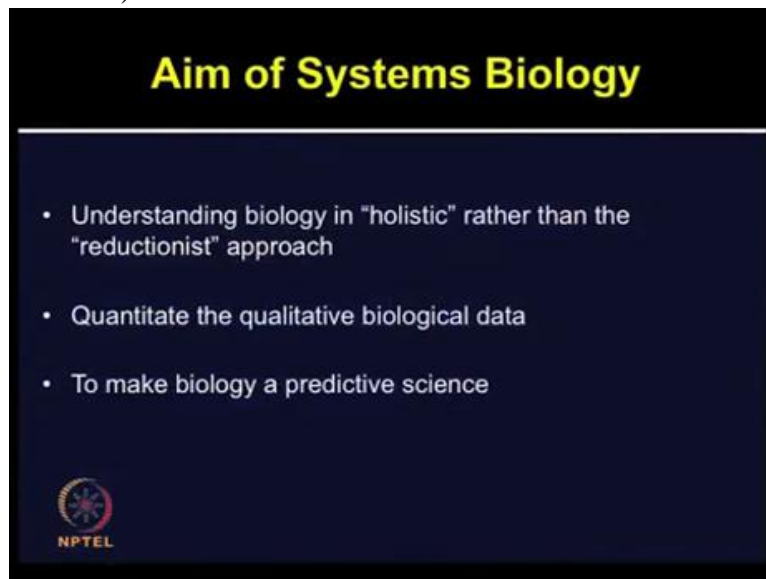


So, why there is need for Systems Biology? The study of biology at the system and subsystem level for understanding the biological processes and networks is very much required.

As you can see, to understand even a simple system of a cell how it is regulated with the extracellular space, the cytoplasm 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: 18:51)



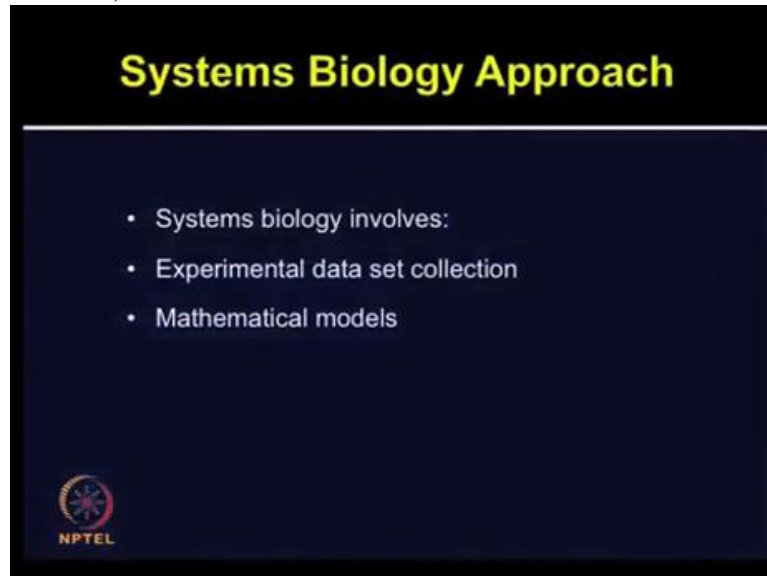
Aim of Systems Biology

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

 NPTEL

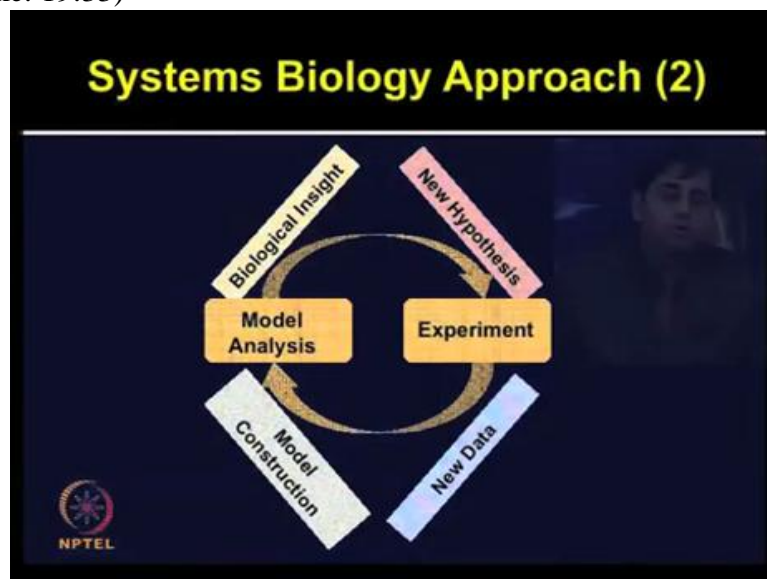
The Systems Biology aims to quantitate the qualitative biological data and provide some level of predictions by applying different types of computational modeling.

(Refer Slide Time: 19:09)



The systems biology approach involves, first of all, collection of large experimental datasets and then mathematical models to provide insight of some significant aspects of the datasets.

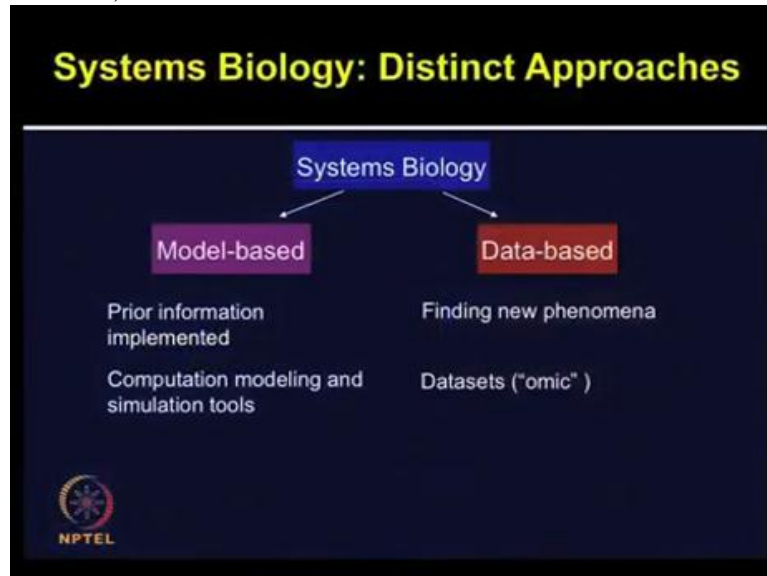
(Refer Slide Time: 19:33)



The simple systems biology approach would involve experiments by adding new datasets 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: 20:10)

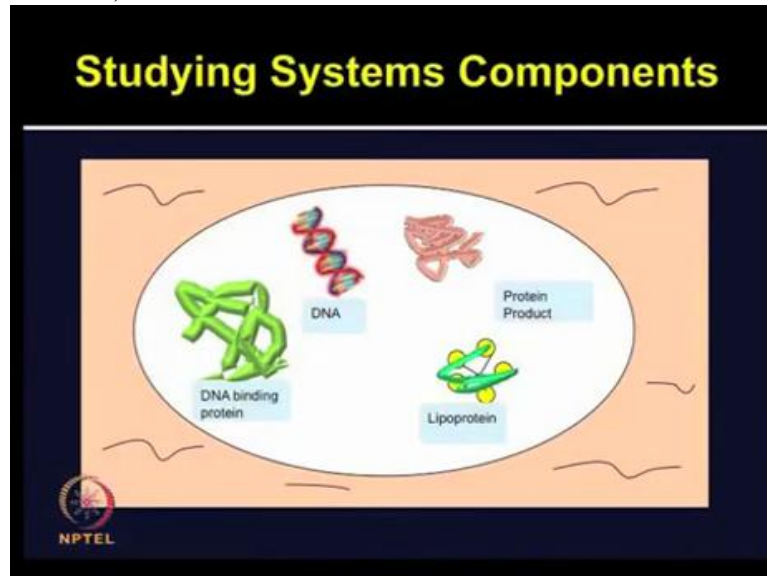


So what all different approaches have been taken to study the systems biology? The distinct approaches of systems biology include the model-based and data-based methods.

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

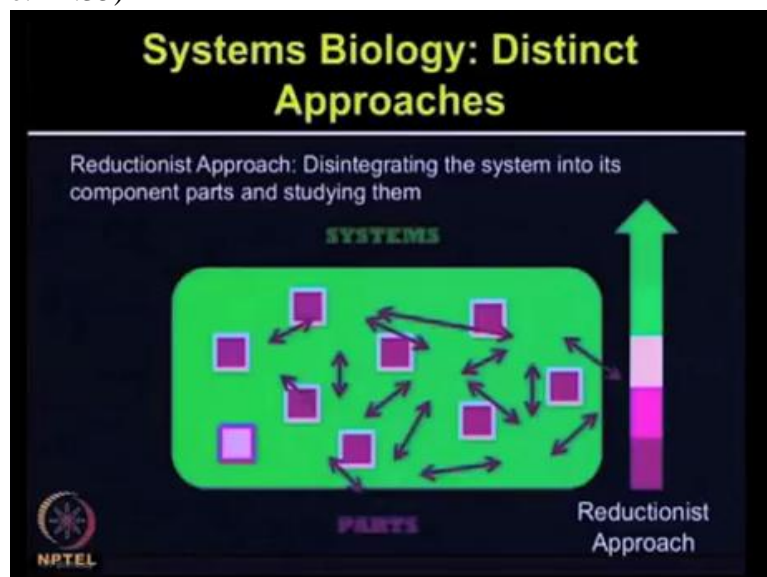
So in model-based systems biology approach, it is difficult to build the detailed kinetic models but in data-based system the complex relationship among various types of Omics information and metabolic pathways and networks can be created.

(Refer Slide Time: 21:07)



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: 21:35)



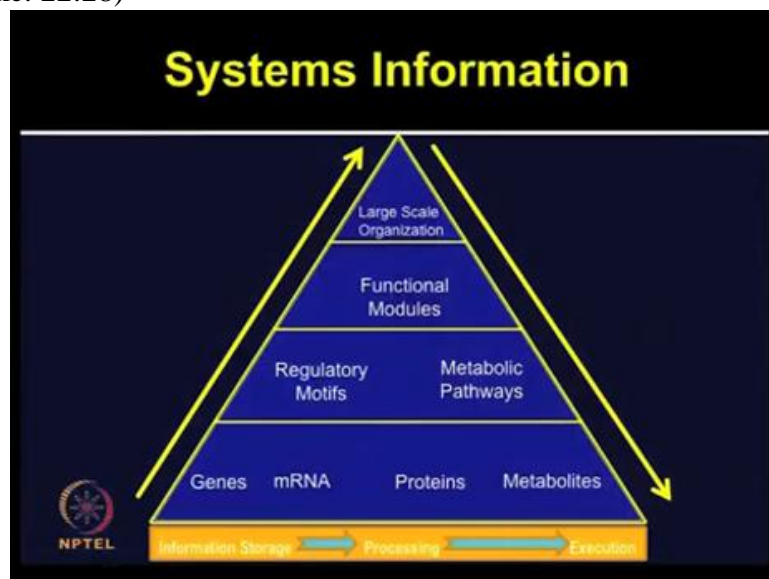
The reductionist approach involves disintegrating the system into its component parts and studying them where as the integrated approach involves integrating the study of individual components to form conclusions about the system.

(Refer Slide Time: 21:57)



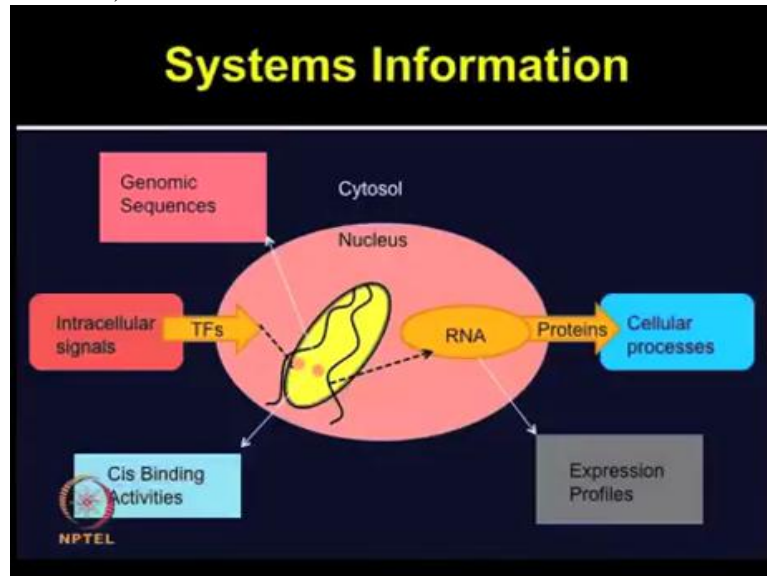
What is Systems Biology triangle? So first of all the systems information is generated at various levels.

(Refer Slide Time: 22:28)



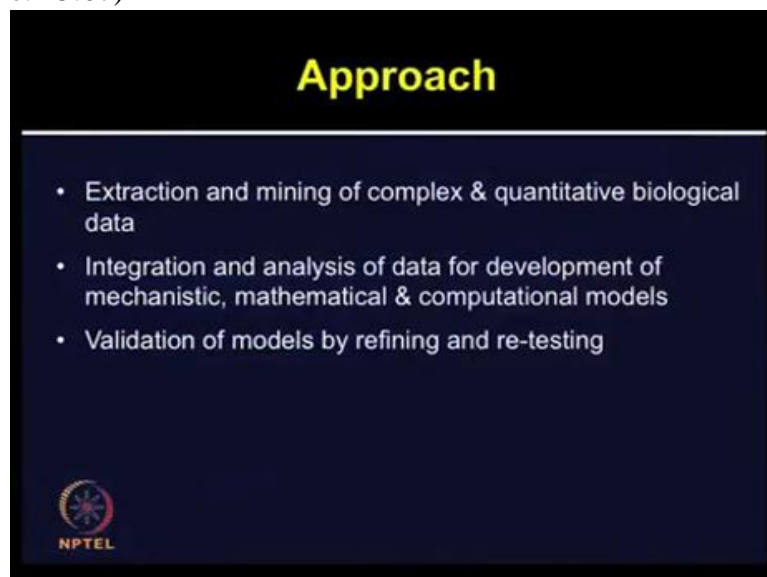
As we have discussed starting from genes to mRNA to proteins to metabolite or identifying regulatory motors, metabolic pathways, functional modules and different large scale organizations. This information has to be stored, processed and further executed to identify the system level information.

(Refer Slide Time: 22:44)



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

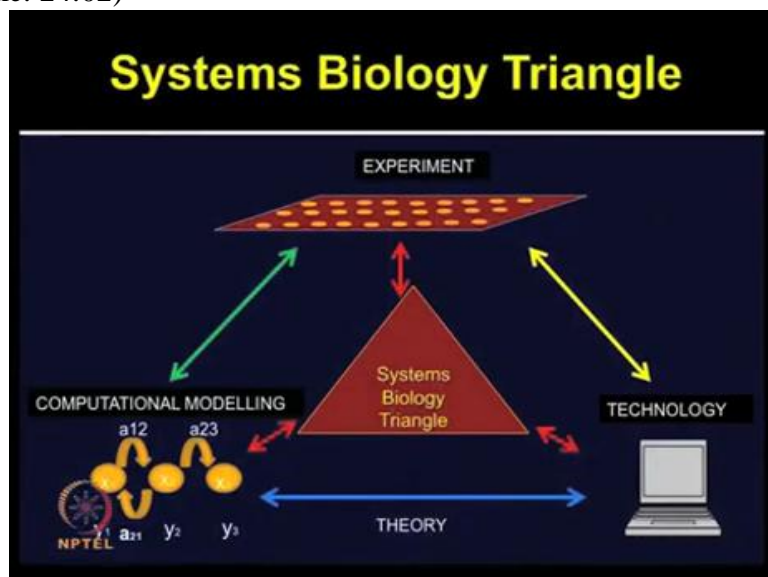
(Refer Slide Time: 23:07)



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

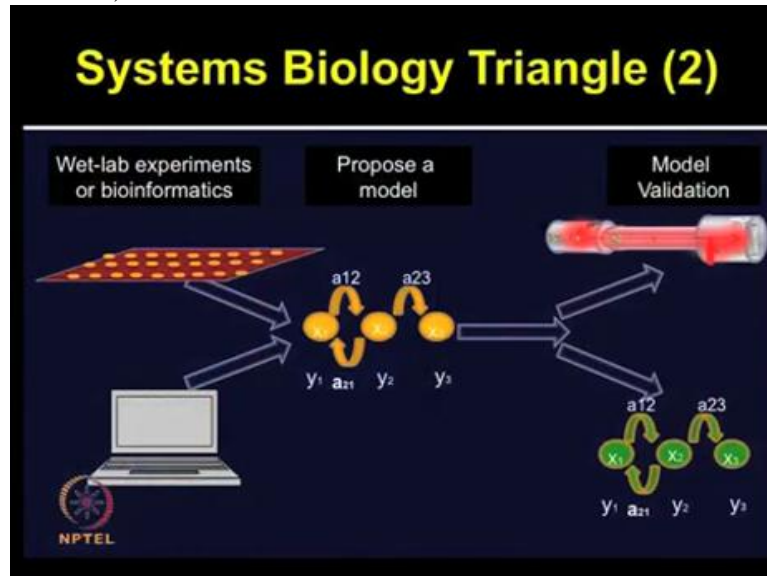
(Refer Slide Time: 24:02)



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.

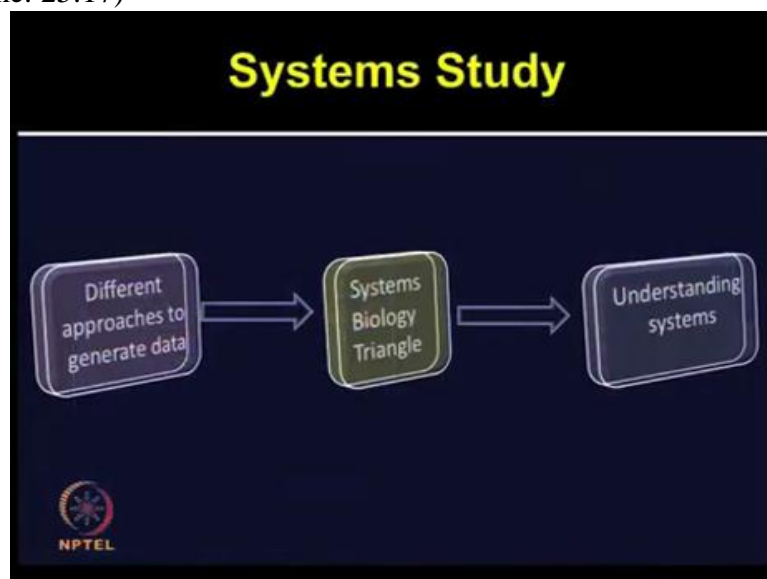
The synergistic application of the experiment, theory and technology with modeling to enhance the understanding of biological processes as a whole system rather than the isolated part is termed as Systems Biology Triangle.

(Refer Slide Time: 24:46)



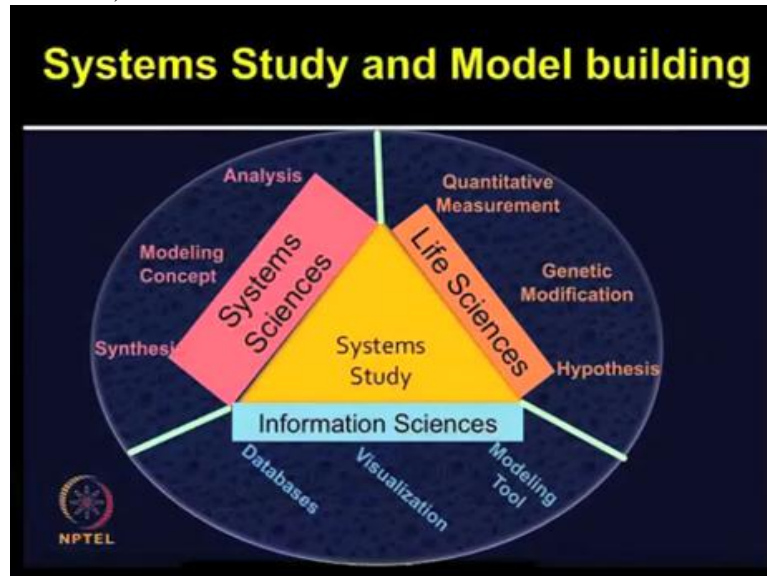
So the Systems Biology Triangle, the wet labs experiments or the bioinformatics based data analysis can be used to propose a model. The model building as an aid to understand the complex system and some hypothesis can be generated which could be used further to propose more quantitative models or predictive models and also it can be used for independent techniques for the model validation.

(Refer Slide Time: 25:17)



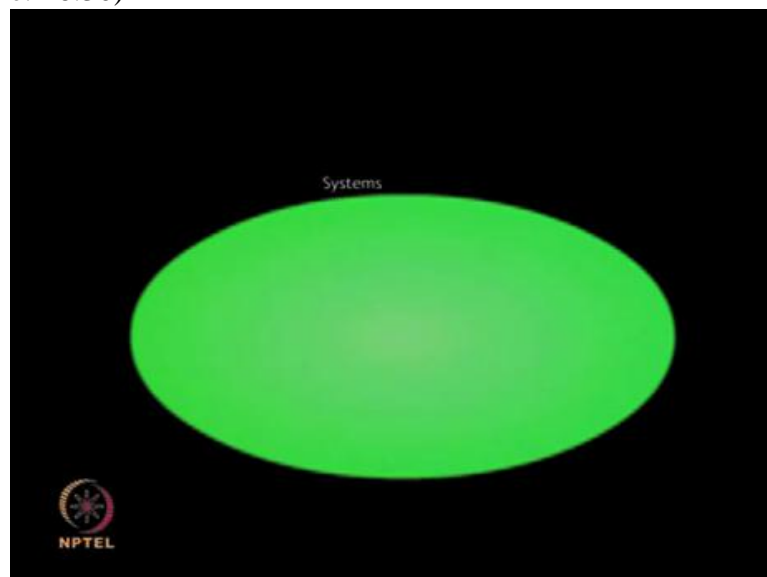
So what is Systems Study? First of all, the difference between systems study and components study, one needs to understand and this is what we have tried to emphasize in the previous slides. After generating the datasets and creating systems biology triangle, then this information can be used for understanding the systems in the more complex and mechanistic level.

(Refer Slide Time: 25:47)



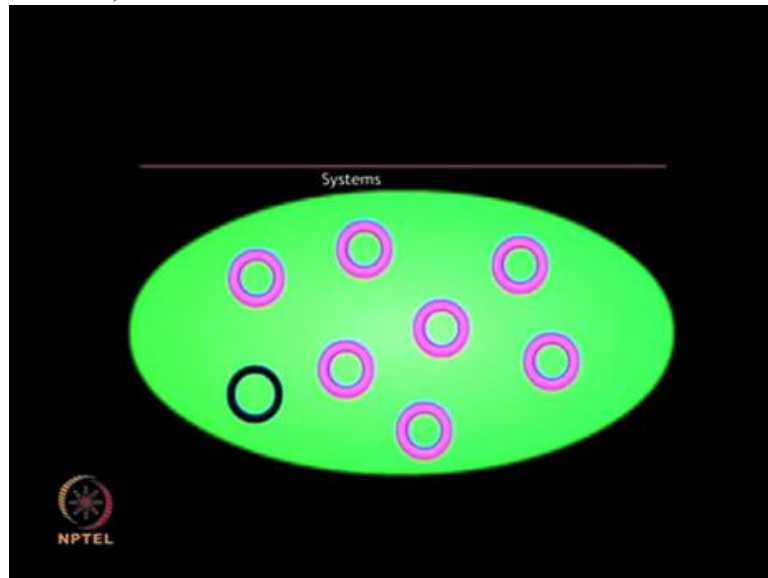
Systems study and Model Building: The systems science includes synthesis, modeling concepts, analysis, life sciences discipline, provides quantitative measurements, genetic modifications and deriving some hypothesis. The information sciences enable the visualization, the modeling tools and different databases. So this model-building as an aid to understand the complex system is very useful for systems level investigation.

(Refer Slide Time: 26:30)



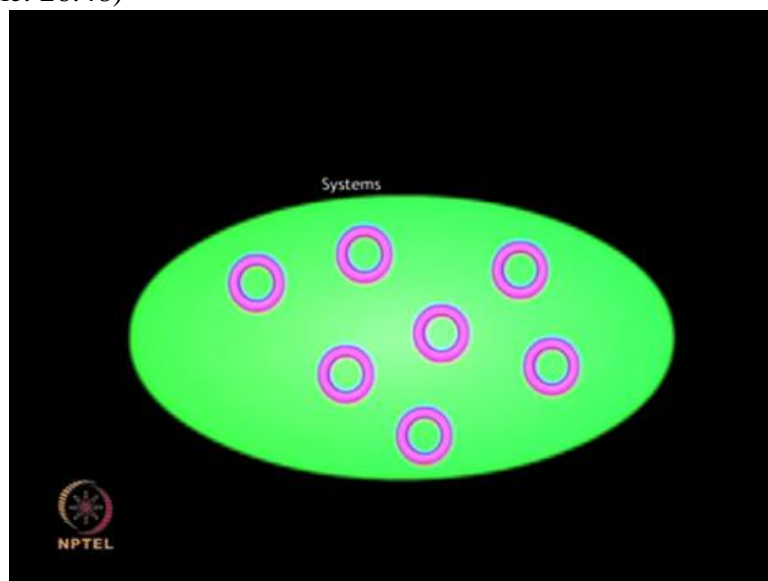
Systems is an entity which maintains its existence through mutual interaction of its constituent parts.

(Refer Slide Time: 26:42)



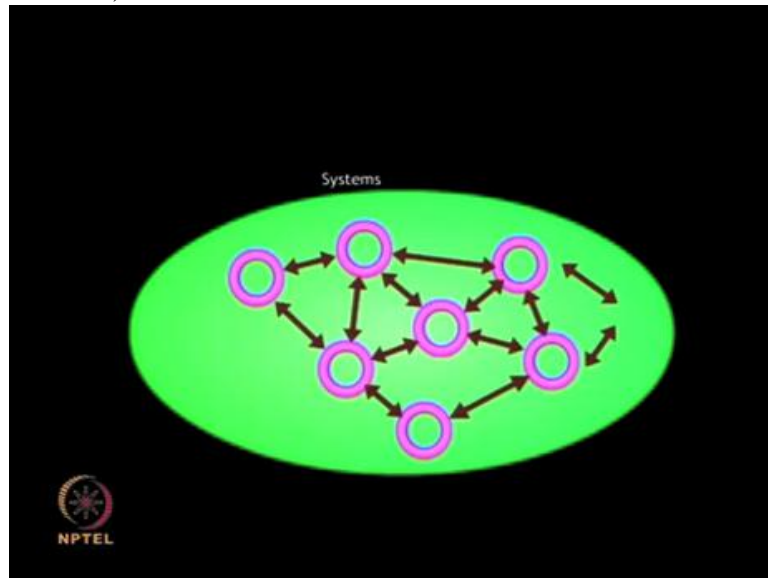
The systems biology research consists of identification of the parts, characterization of the components....

(Refer Slide Time: 26:48)



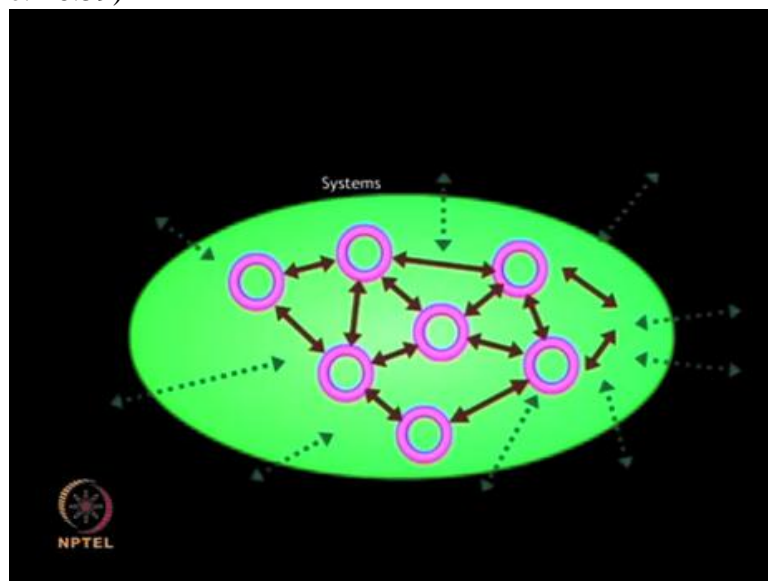
...exclude the ones which are not a part of the system...

(Refer Slide Time: 26:52)



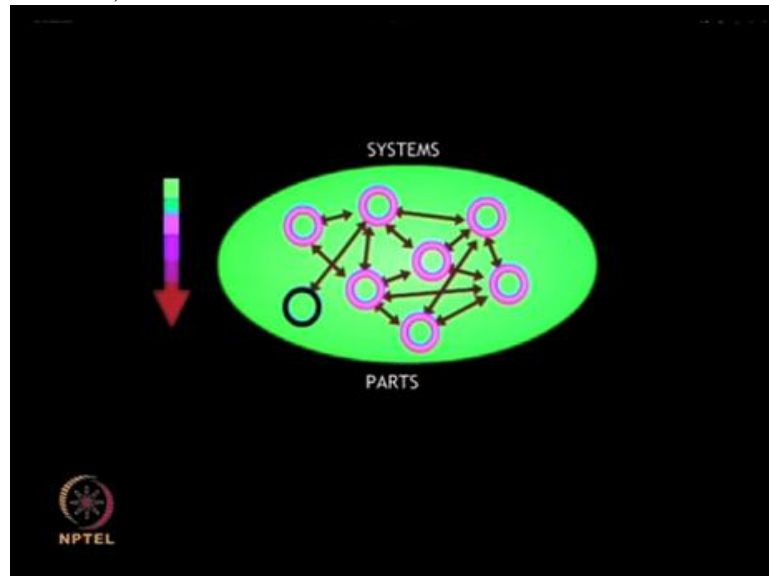
... identify the interaction of the components with each other...

(Refer Slide Time: 26:59)



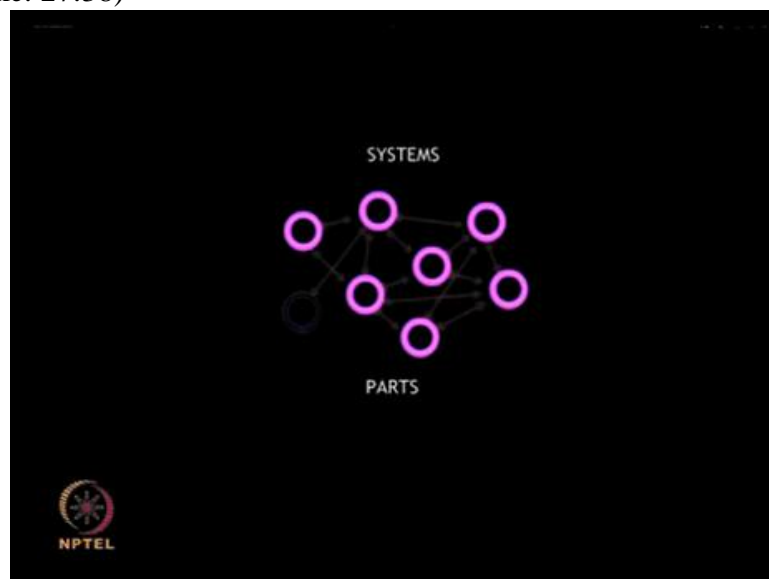
...and identify the interaction of the components with the environment which modulate the parts either directly or indirectly through modulation of internal interactions. The systems biology concept can be understood with help of two approaches such as reductionist approach and integrative approach.

(Refer Slide Time: 27:29)



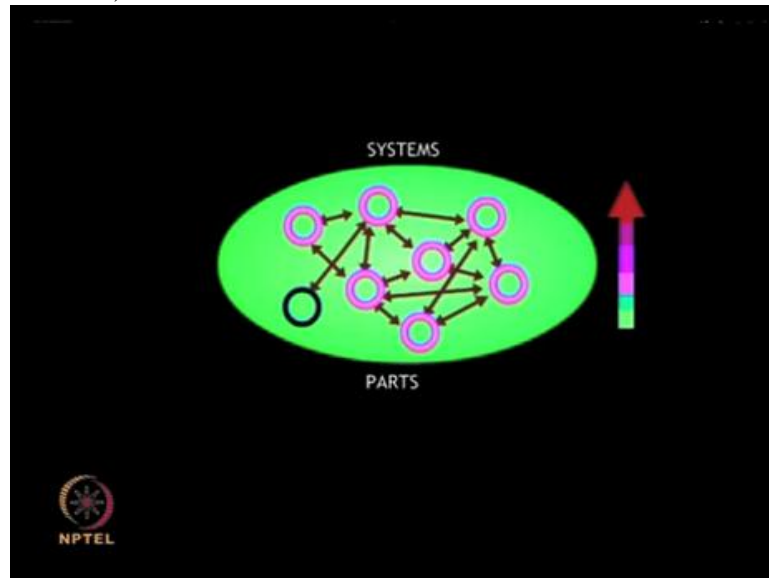
The reductionist approach focuses on disintegrating the system into its component parts and studying them...

(Refer Slide Time: 27:38)



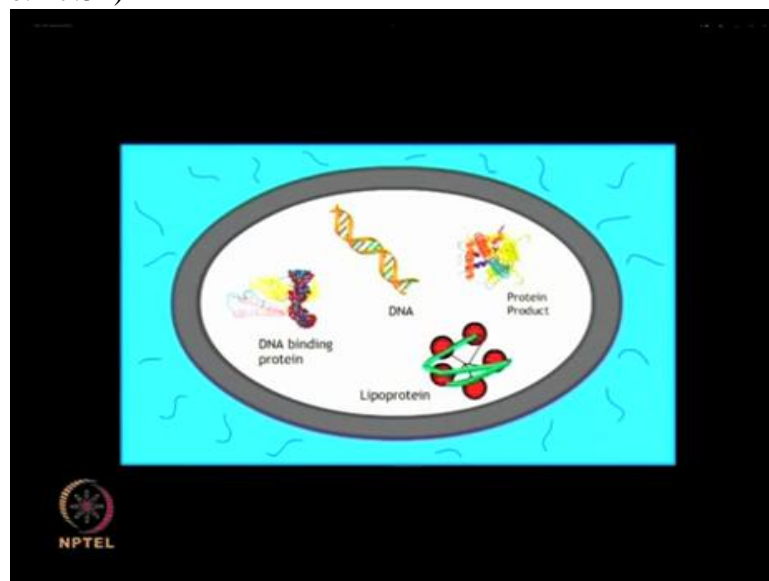
...where as the integrative approach focuses on integrating the study of individual component...

(Refer Slide Time: 27:48)



... to form conclusions about the system

(Refer Slide Time: 27:51)



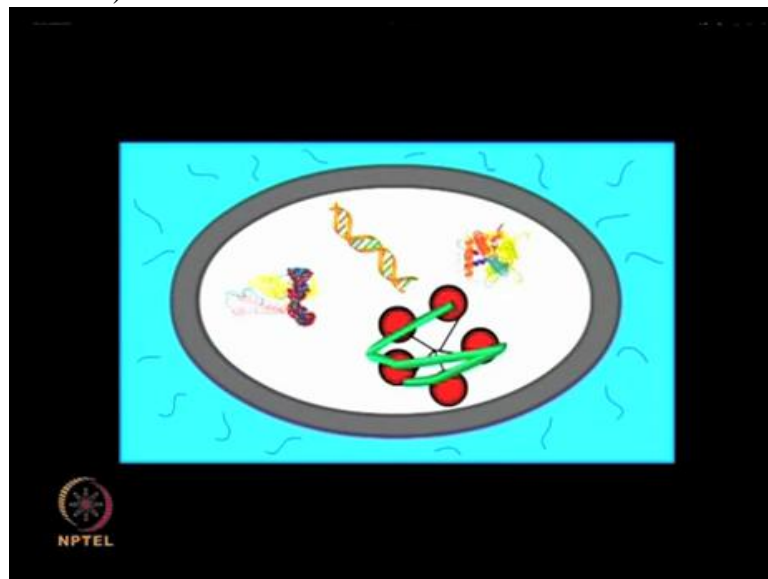
Consider a cell with its component molecules. Let's say we want to study the metabolic pathway as a biological system. When the environment of the cell is perturbed a little, the individual components undergo unique changes...

(Refer Slide Time: 28:11)



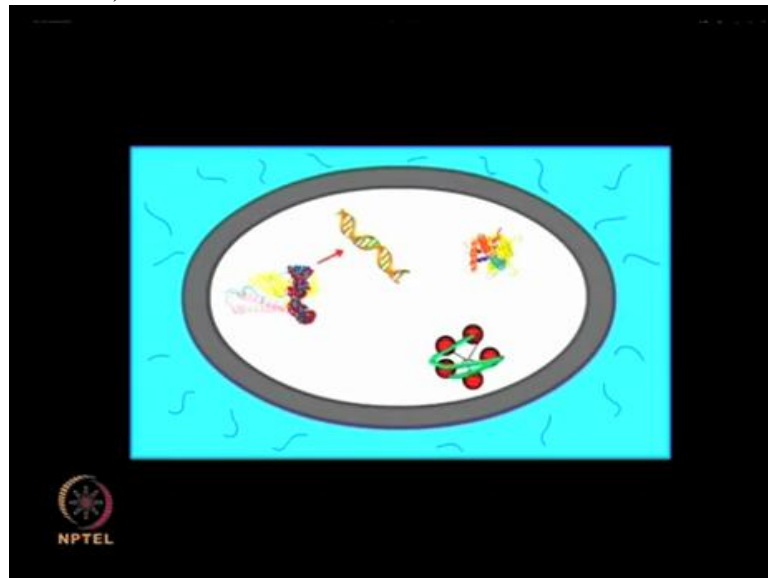
... such as increase in the production rate...

(Refer Slide Time: 28:16)



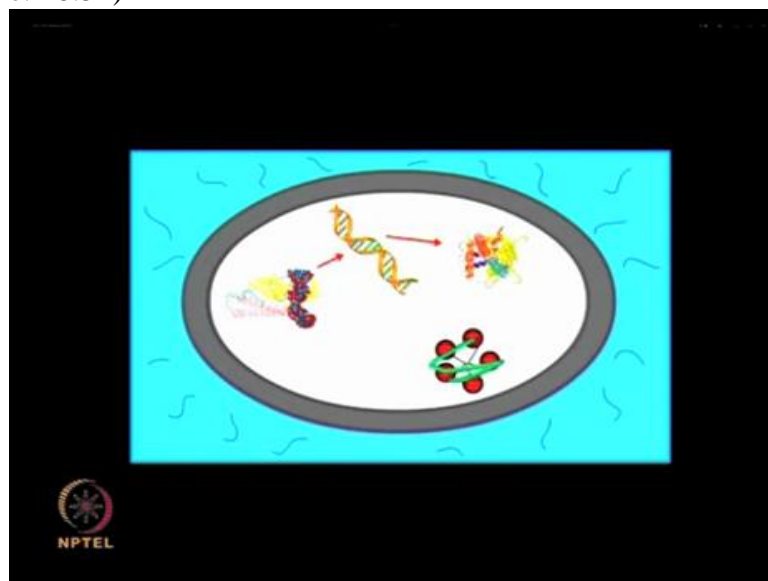
... or decrease in their amount. At this stage, due to lack of knowledge of the nature of interaction of the proteins, we cannot interpret how the system gets affected. But when study the interaction of one component with the other we can conclude that the increase in rate....if DNA binding protein leads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

(Refer Slide Time: 28:44)



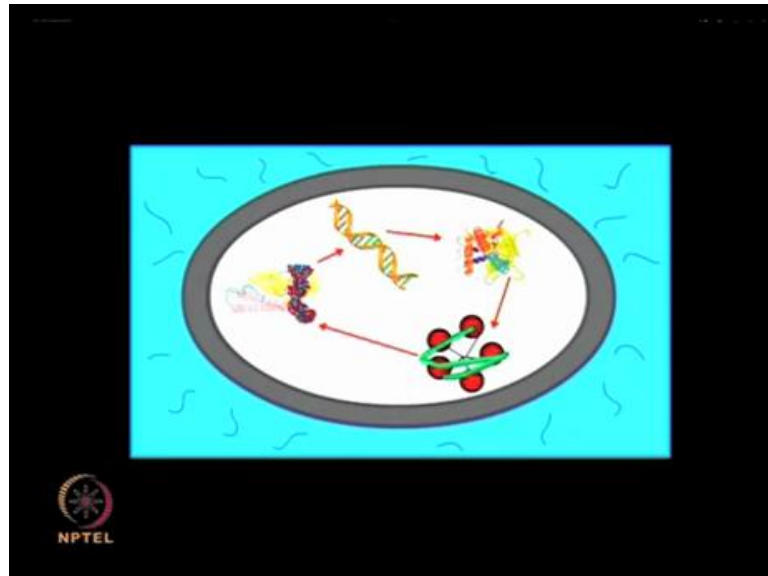
... increase in the synthesized amount of DNA which further changes...

(Refer Slide Time: 28:52)



...the final amount of lipoproteins produced.

(Refer Slide Time: 29:00)



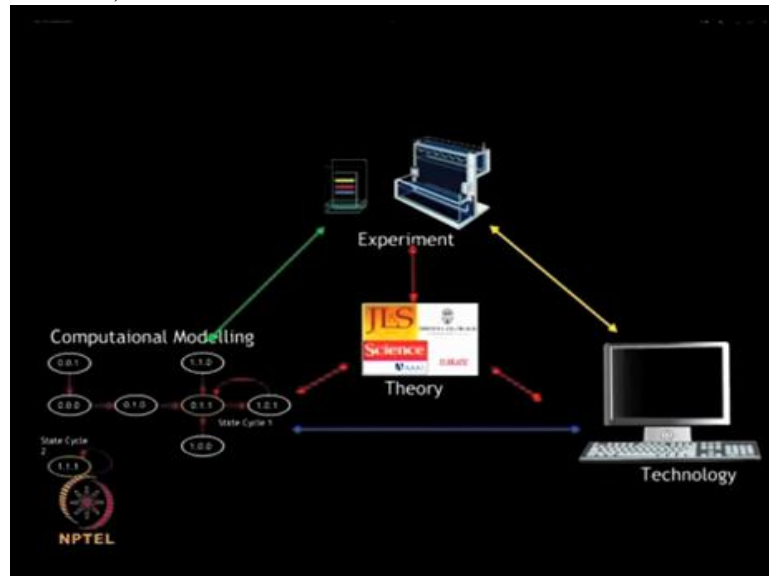
Thus we can see that, to study a system we need to analyze not just the components but their interactions. These biological systems can be protein-protein interaction networks, gene regulatory networks, protein DNA networks, protein lipid networks and metabolic networks. To study the system we need to know about the components and its interactions.

(Refer Slide Time: 29:33)



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

(Refer Slide Time: 30:02)



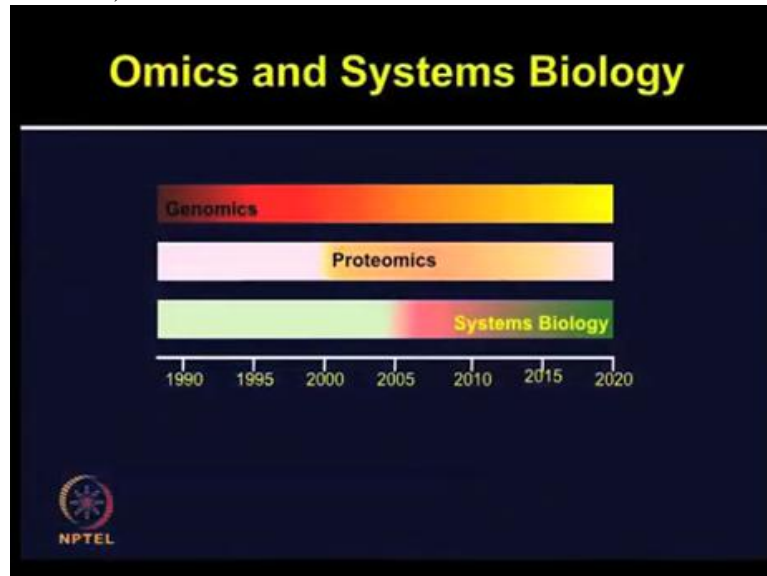
Here it is shown, that in a systems approach, experiment, technology and computational modeling this triangle is very important which has to be linked with the theory to form a Systems Triangle.

(Refer Slide Time: 30:05)



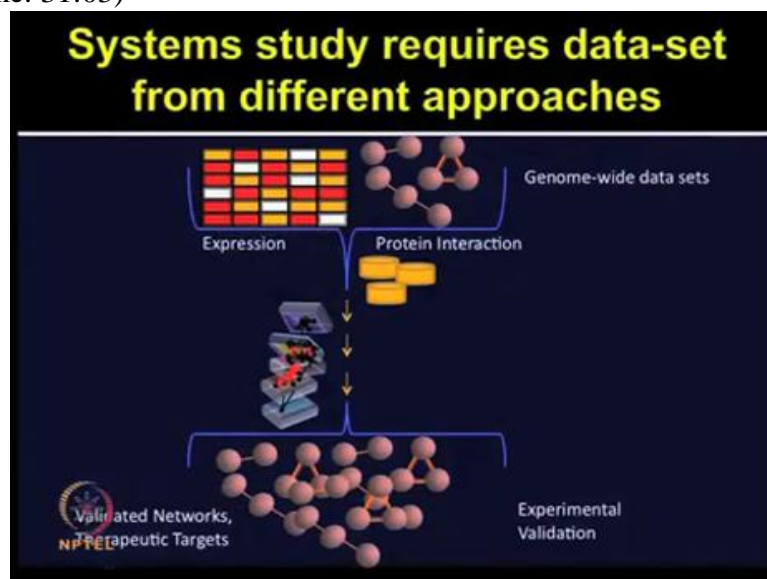
Now let's try to integrate Omics approaches with Systems Biology. So Genome sequencing projects in Genomics era from 1990s to 2000 accelerated the pace of Omics research.

(Refer Slide Time: 30:24)



Then from 2000 onwards, proteomics field also got accelerated and new methodologies, new tools came into the place for studying proteome and the data derived from genomics, transcriptomics, proteomics, metabolomics and other Omics approaches have now brought the integration of the datasets in the systems biology field. The systems study requires obtaining datasets from different approaches and analyzing them.

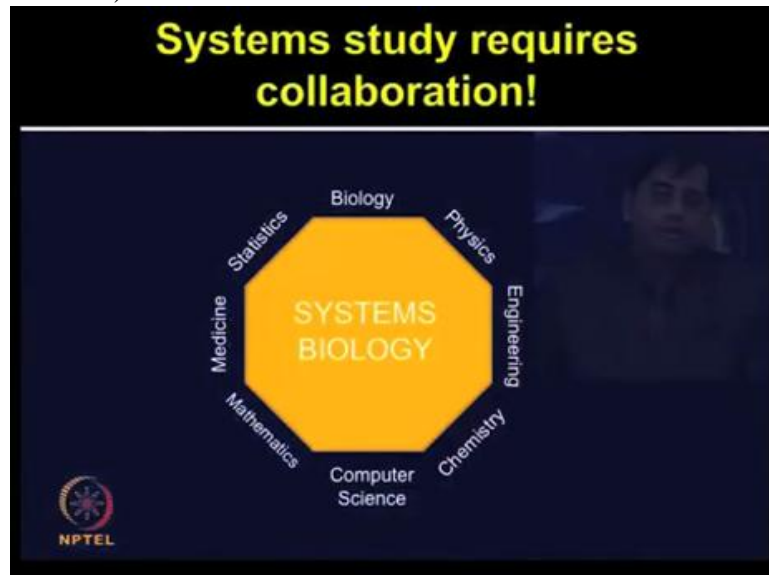
(Refer Slide Time: 31:03)



For example as shown in the slide, the genome-wide datasets can be derived at the genome level and looking at the expression of the different transcripts or at the proteome level looking at the different types of protein interactions.

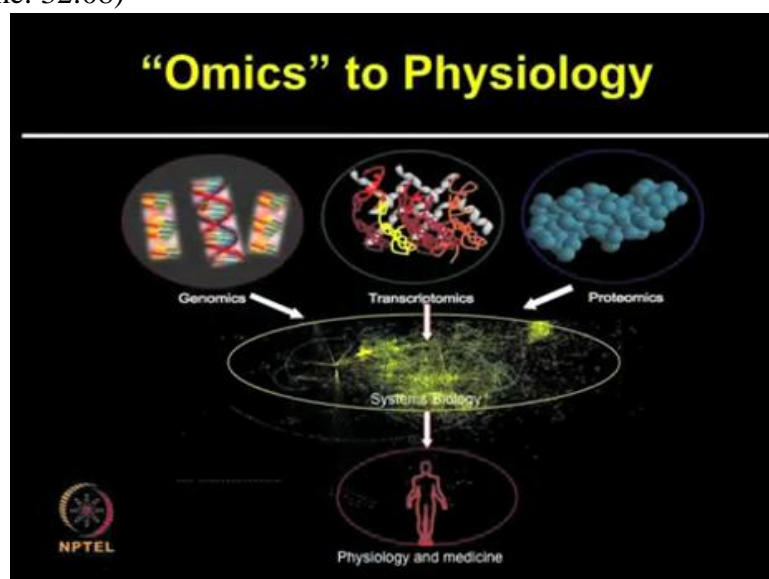
These datasets can be stored in the clinical databases and also can be mined from the literature, literature manual curation. Then integration of the orthogonal datasets further can be used for validating the networks and deriving identifying therapeutic targets. Further it can be used for experimental validation.

(Refer Slide Time: 31:44)



Studying systems cannot be done in isolation in individual labs. It requires different expertise and collaborations from scientists from different disciplines of biology, physics, engineering, chemistry, computer science, mathematics, medicine, statistics and many more.

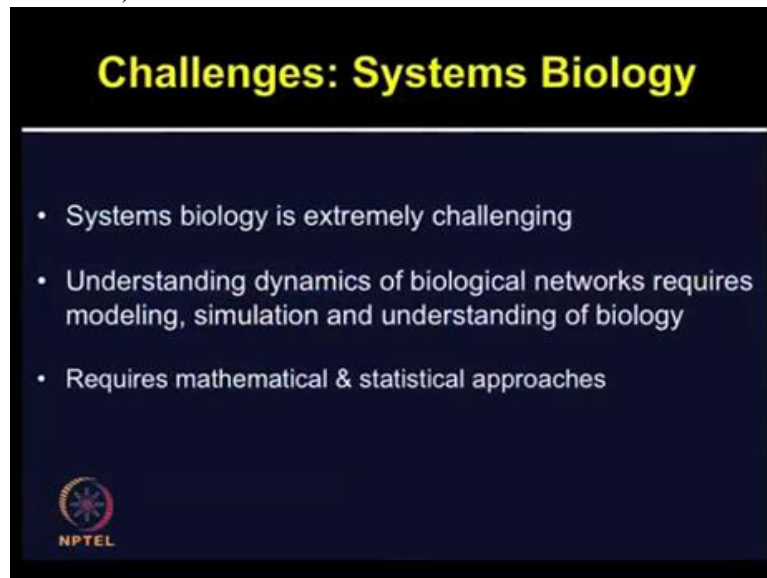
(Refer Slide Time: 32:08)



So eventual aim of this goal of this current Systems Biology field is to employ the Omics level information obtained from different levels, from genome, transcriptome and proteome,

derive that information at the systems level; integrate, quantitate some models and then propose and use it for the understanding the physiology and apply that in medicine. So this Omics physiology, this flow can be well maintained by employing systems biology tools.

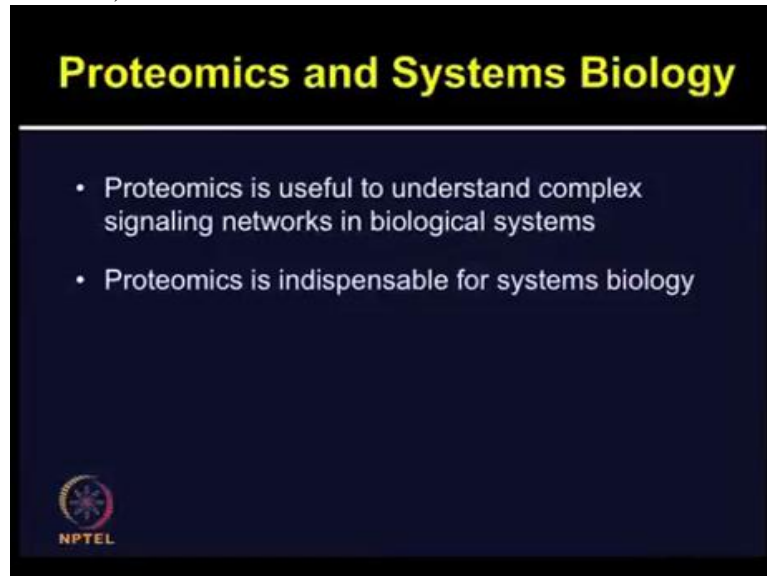
(Refer Slide Time: 32:45)



What are the challenges of Systems Biology? Systems Biology is extremely challenging. The emphasis is to understand a system. Understanding dynamics of even simplest biological networks not only requires only the understanding of biology but also its modeling and simulation.

The disintegrative study can be used for studying from cells to proteins to genes or integrative study could be used for putting these pieces back together again and then understanding and doing the prediction and control of functional biological processes. All of these are very challenging but currently being addressed by applying various systems level tools.

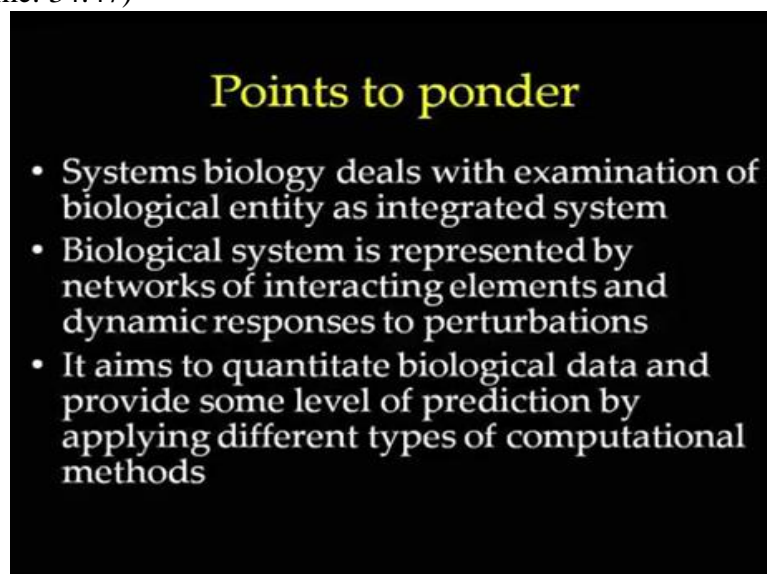
(Refer Slide Time: 33:43)



So how proteomics and Systems Biology are integrated? Proteomics as we have studied, it is useful to understand complex signaling networks in biological systems. It is very indispensable tool for Systems Biology. The global analysis of proteome is important. However there are many limitations. In each experiment only thousands of proteins can be studied. Therefore new approaches and systems level investigation and predictions are required.

The system investigation is required to study the complex dynamic structure interaction with the biological systems, whether it is at cellular level or at the organism level and ultimately it is responsible for their function and behavior.

(Refer Slide Time: 34:47)



(Refer Slide Time: 34:51)

Summary

- Proteomics is useful to understand complex signalling networks in biological systems.
- Proteomics is indispensable tool for systems biology
- The system investigation is required to study the complex dynamic structure interaction with the biological system at cellular level or at the organism level.

So in summary, today we discussed that how Omics era the technological advancement in genomics, proteomics and metabolomics have generated large scale datasets in all the aspects of biology. These large datasets have motivated the computational biologists and systems approaches with objective of understanding the biological system as a whole.

While proteomics continues to generate the quality data at the proteome level, the systems biology approach characterizes and predicts these dynamic properties of biological networks. Now in the next module we will focus in more detail different types of proteomic technologies. Thank you.