Introduction to Proteomics
Dr. Sanjeeva Srivastava
Department of Biosciences and Bioengineering
Indian Institute of Technology – Bombay

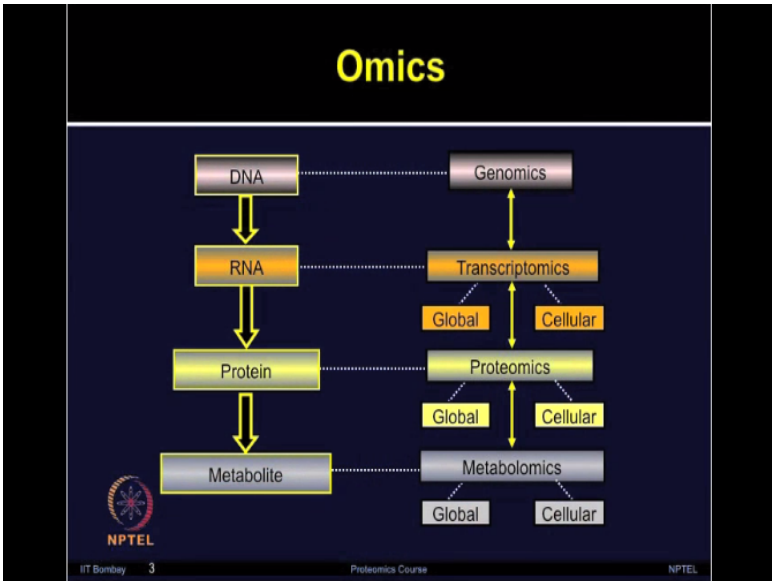
Lecture - 19
Systems biology and proteomics - I

Welcome to the Proteomics course. Today's lecture is on Proteomics and systems biology, the lecture outline.

(Refer Slide Time: 00:34)

First I will start on proteomics and then we will move on to the systems biology. Today we will continue on first proteomics and then how different type of Omic data can be applied for systems approach analysis that will be discussed.

(Refer Slide Time: 00:55)



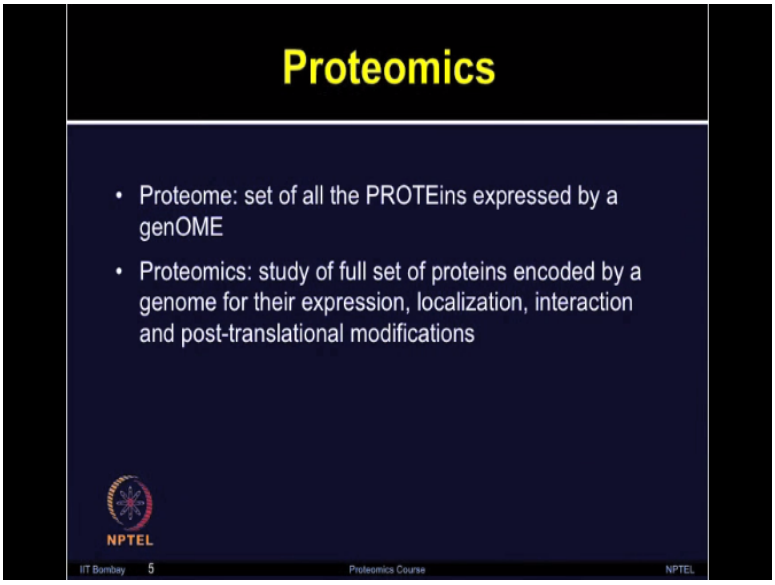
So this slide I have shown you how you can define different type of Omic platform. So, Omic is the suffix derived from the Greek world omega means on or every. The Omics is used as a suffix, which has enabled the explosion of terms Genomics, Transcriptomics, Proteomics and Metabolamites and so on and so far. The Omic also implies an integration of biology with information science and conveys large scale biology using systems approach.

Such you can see the slide if you are studying about DNA, that will be in totality known as Genomics, RNA study and Transcriptomics, Proteins and Proteomics and Metabolites in Metabolamites. If you are looking at all of the cellular contents of Proteome that will be known as cellular Proteomics or cellular Genomics and similarly all of the proteome of an organism will be known as global Proteomics or similarly at the gene level global Genomics.

(Refer Slide Time: 02:19)

So let us first start with Proteomics.

(Refer Slide Time: 02:30)

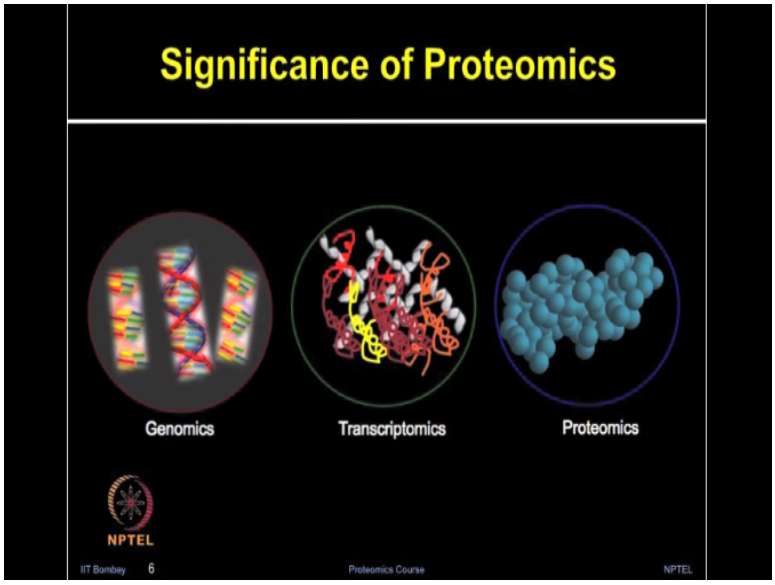


So if you remember from the previous lecture Proteomics set of all the proteins which are expressed by a genome. Proteomics is a study of proteins and their properties to provide an integrated view of cellular processes. What are these different properties? These properties include the extent of the protein expression. How different type of post translational modification and co modifications occurs in the cell.

Different type of enzyme regulation whether it is activation or inactivation and then different type of intermolecular protein-protein interactions. The current goals of proteomic are very broad. It is in including the diverse properties of proteins which we have discussed in the earlier lectures looking at the (()) (03:33) chemistry of amino acids and different levels of protein structures.

So with that sequence, quantity, the state of modification, activity, interactions of proteins with other proteins and other bio molecules and sub cellular distributions and structural analysis. All of these are broad goals of Proteomics.

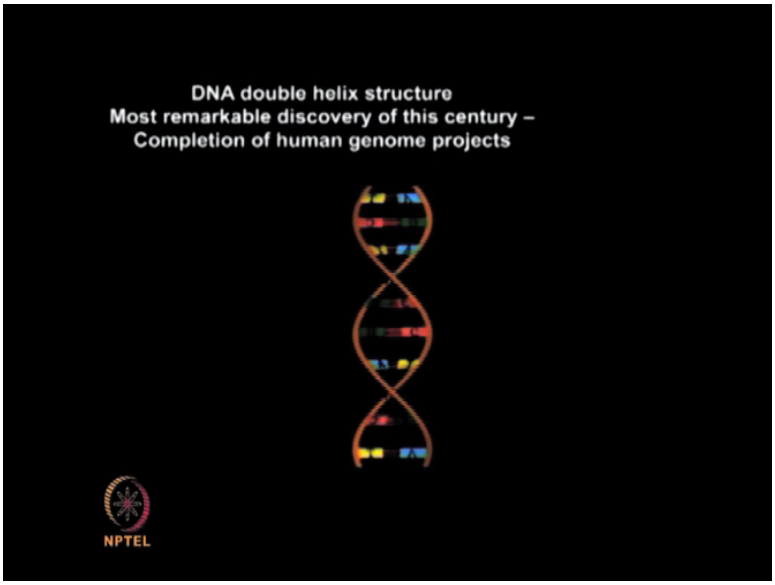
(Refer Slide Time: 04:02)



Concept was discussed earlier which is an orderly and unidirectional flow of information encoded in base sequential 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 an un presented information of genome sequences. However, their numerous proteins which can be encoded by the genome, therefore analysis of the static genome by doing sequencing alone is not sufficient.

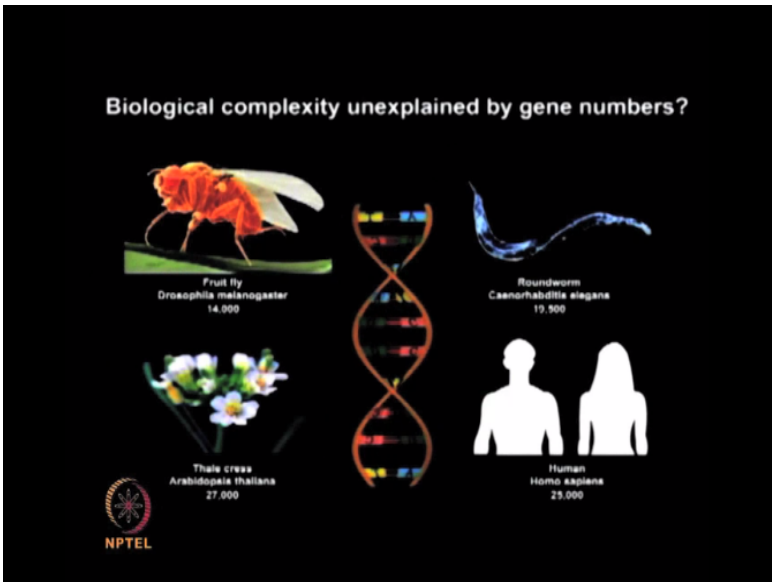
A gene can quote for several type of proteins because of an alternative splicing and post transcriptional and post translational modifications. Therefore, it suggests that a study in proteins is more challenging than geno or transcript. Therefore, Proteomics has great significance to understand the biological systems.

(Refer Slide Time: 05:20)



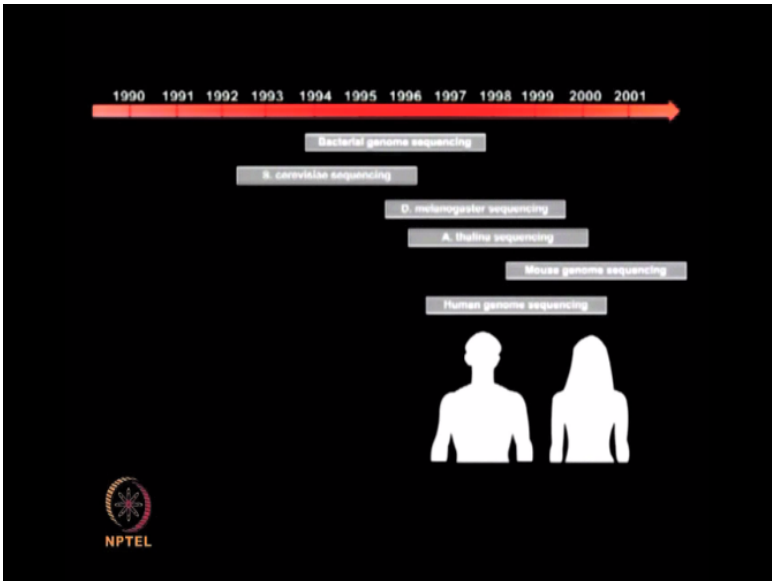
The completion of genome sequencing projects of several organism including human has been one of the most remarkable achievements of this century.

(Refer Slide Time: 05:30)



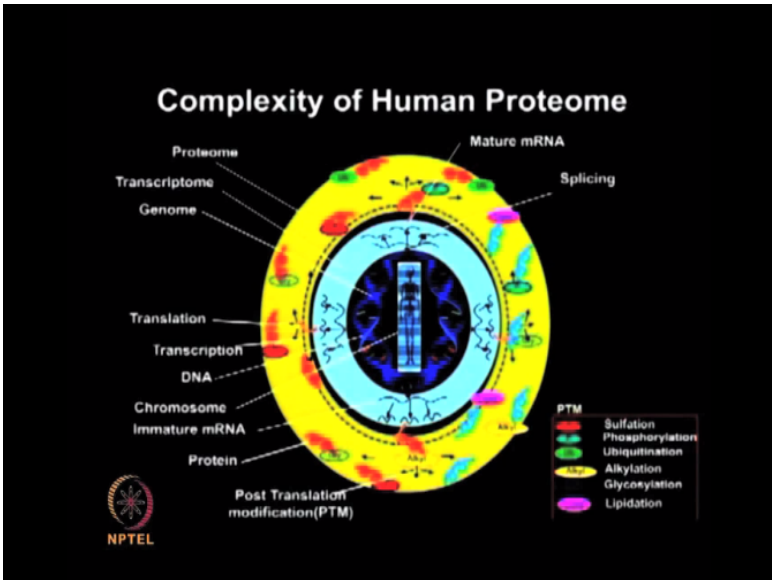
However, these have not been sufficient to unravel the mystery of complex biological processes.

(Refer Slide Time: 05:40)



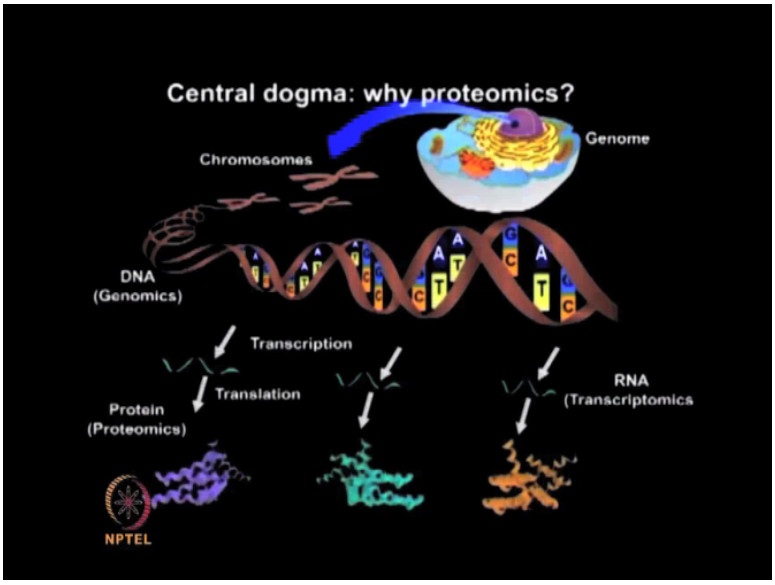
This similar gene numbers of many diverse group of organisms has failed to explain their varying biological complexity. In more meaningful understanding of biological function can be obtained through the characterization of products of gene expression. The protein which there was ultimate effective molecule of biological systems.

(Refer Slide Time: 06:14)



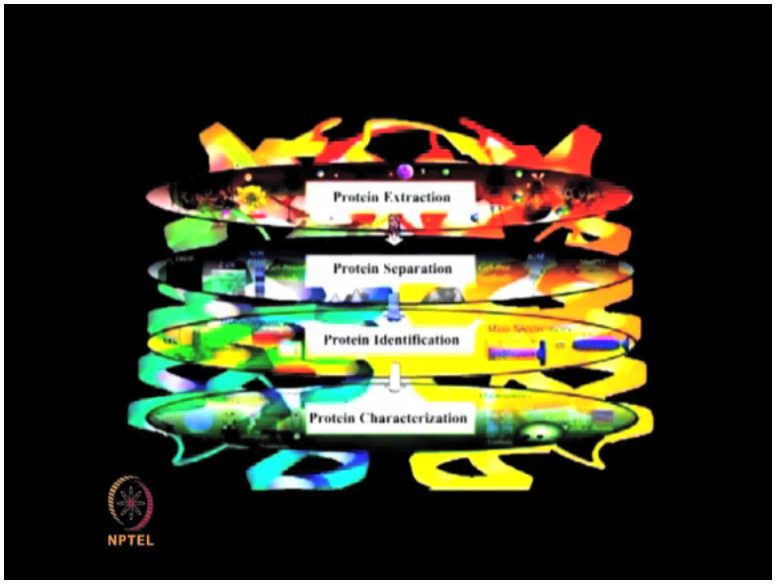
The Proteomics refer to the study of entire protein complement expressed by an organism at any given time, while the genome of an organism is mostly static, the proteome is dynamic and it changes with environment and time. Thereby elevating its complexity level. The gene regulation is regulated by several 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: 06:58)



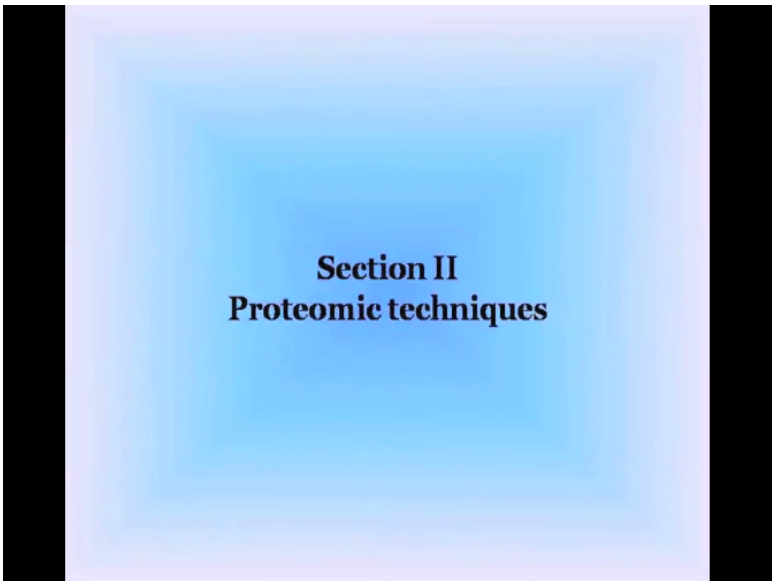
The Proteomics aims to decipher the structure and function of all proteins in a given cell under a specific conditions and to obtain a global view of cellular processes at the protein level. A study at the DNA level known as genomics.

(Refer Slide Time: 07:25)

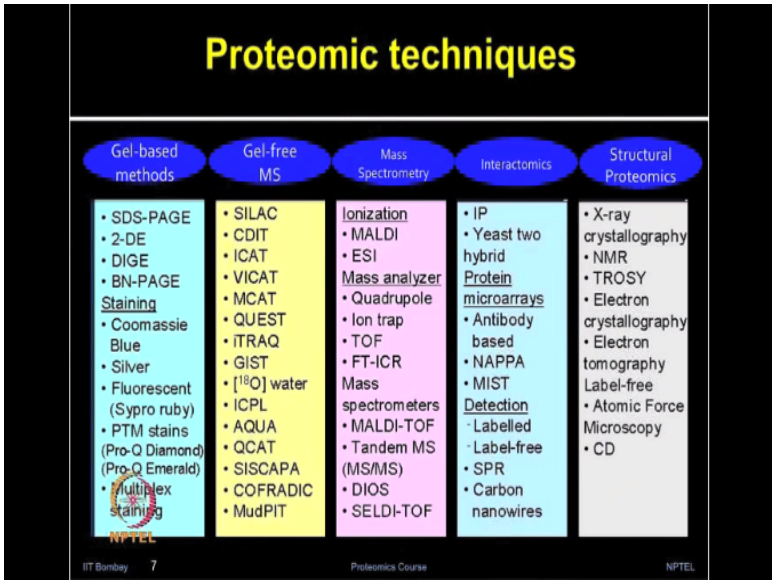


RNA level transcriptomics and protein level Proteomics. Analysis of the proteome involves protein extraction, separation, identification and finally characterization of various proteins.

(Refer Slide Time: 07:45)



(Refer Slide Time: 07:52)



The various Proteomics 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 which study the protein interactions and the structural proteomics.

As I have shown some abbreviation in the slide broadly different type of gel based Proteomics such as Sodium Dodecyl Sulphate, poly Acrylamide Gel Electrophoresis or SDS PAGE, 2 Dimensional Electrophoresis or 2DE, Difference in Gel Electrophoresis or DIGE, Blue Native PAGE as well as different type of staining methods such as coomassie silver fluorescent dyes and (()) (08:55) and multi plus training methods.

All these can be grouped under gel based methods. The gel free methods especially in the mass spectrometry side includes SILAC which is a 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 quantization using (()) (09:39) signal tags.

ITRAQ isobaric tagging for relative and absolute quantization, GIST global internal standard technology, ICPL isotope coded protein labeling, AQUA absolute quantitation, SISCAPA a stable isotope standards encaptured by anti peptide antibodies, COFRADIC combined fractional diagonal chromatography and MudPIT which is multi dimensional protein identification technology.

All of these are various new advancement in the gel free methodologies. The basic mass spectrometry which is central to the proteomic application includes different type of ionization sources such as matrix-assisted laser desorption ionization MALDI, electrospray ionization ESI and different type of mass analyzers such as quadrupole, time of flight, ion trap and Fourier transform mass spectrometry then different type of Tandem MS base systems are also used.

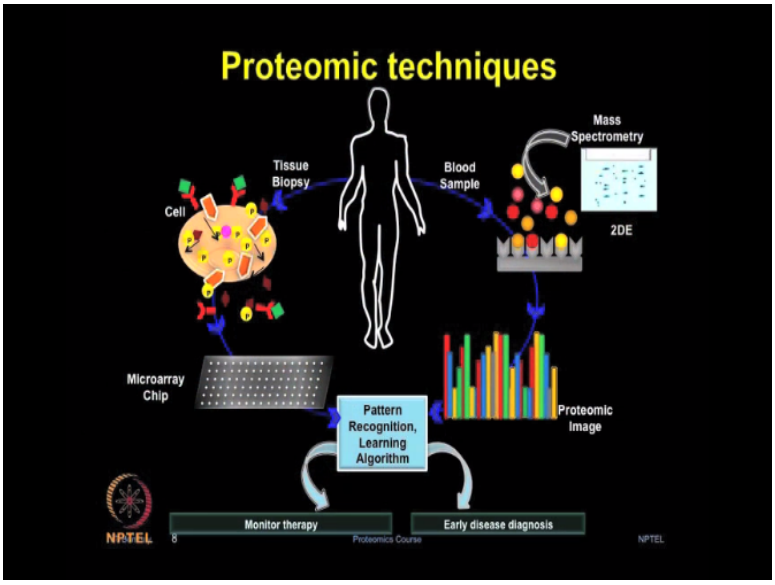
The surface enhanced laser desorption ionization time of flight SELDI-TOF is also used for various clinical applications. The protein interaction methodologies include immunoprecipitation, yeast-two-hybrid methods and different type of protein microarray platforms such as antibody arrays, nucleic acid 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 labeled methods using florescent, chemiluminescence or radio activity 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 implying nano tubes and nano wires.

The structure Proteomics, it involves X-ray crystallography, nuclear magnetic resonance NMR, Transverse relaxation optimized spectroscopy TROSY, Circular dichroism CD and different type of microscopy methods including atomic force microscopy and electron microscopy. So far we have seen the large number of Proteomics technologies which are currently available for various applications.

Many times to address one biological question, different type of methodologies come together and then provide solutions to that problem.

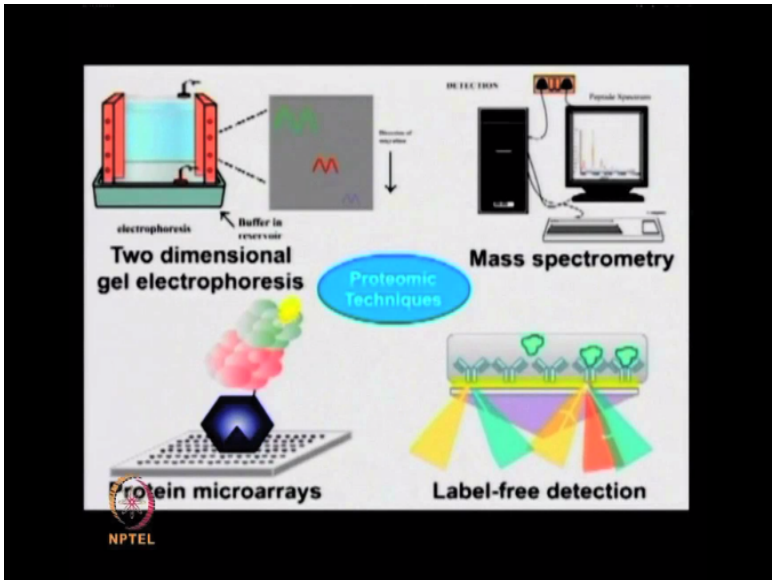
(Refer Slide Time: 12:56)



For example, looking at some clinical sample for identification of bio markets 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 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.

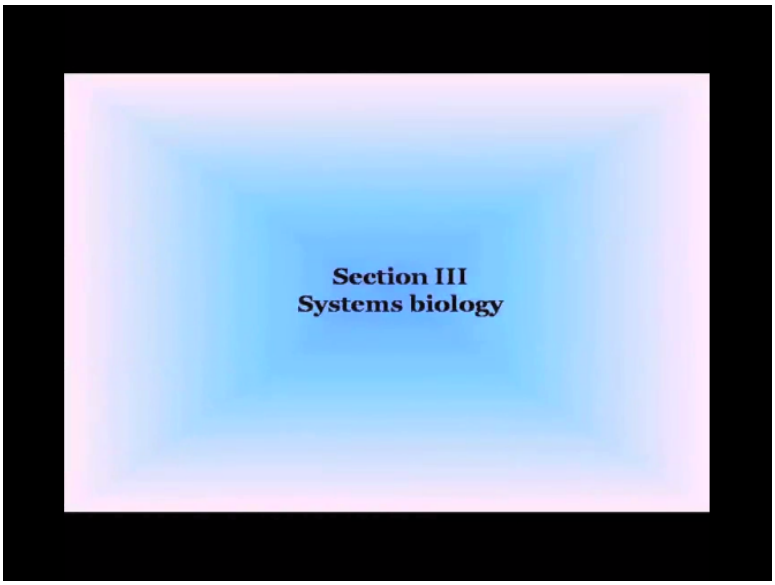
Eventually these type of results will, it has knowledge for the monitoring the therapy response as well as identification of early disease diagnosis. This is just one example. Similarly, multiple type of Proteomics technologies can be used for different applications.

(Refer Slide Time: 13:58)



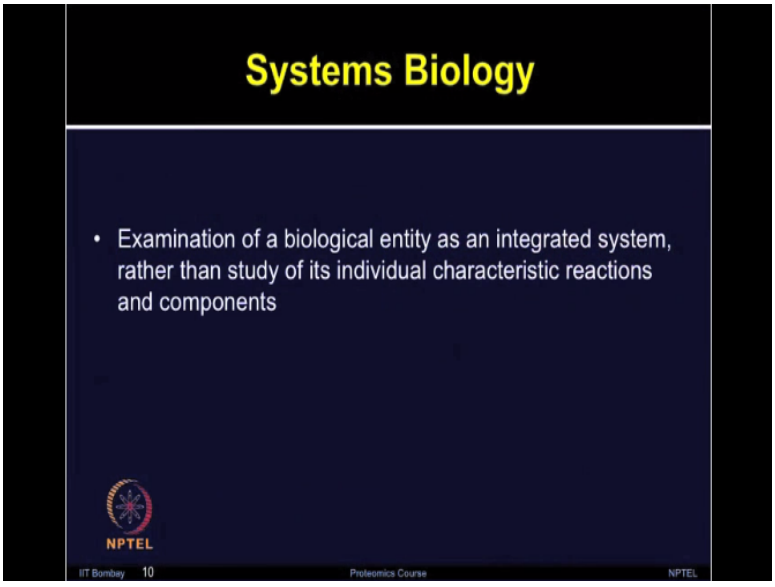
There are several Proteomic techniques which are employed for studying these proteins such as two dimensional gel electrophoresis, mass spectrometry, protein microarrays as well as some label free detection techniques such as Surface plasmon resonance SPR.

(Refer Slide Time: 14:29)



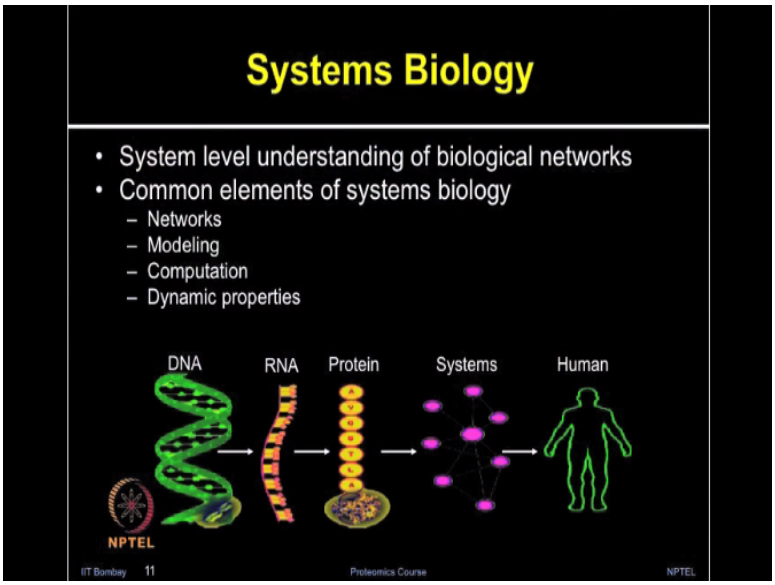
After discussing Proteomics, now let us talk about Systems biology. So what is systems biology, is that an historic knowledge, a method to understand biological systems or a tool to solve the practical problems.

(Refer Slide Time: 14:52)



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. A study of all the mechanisms underline the complex biological processes in the form of integrated system of many interacting components is studied under system biology.

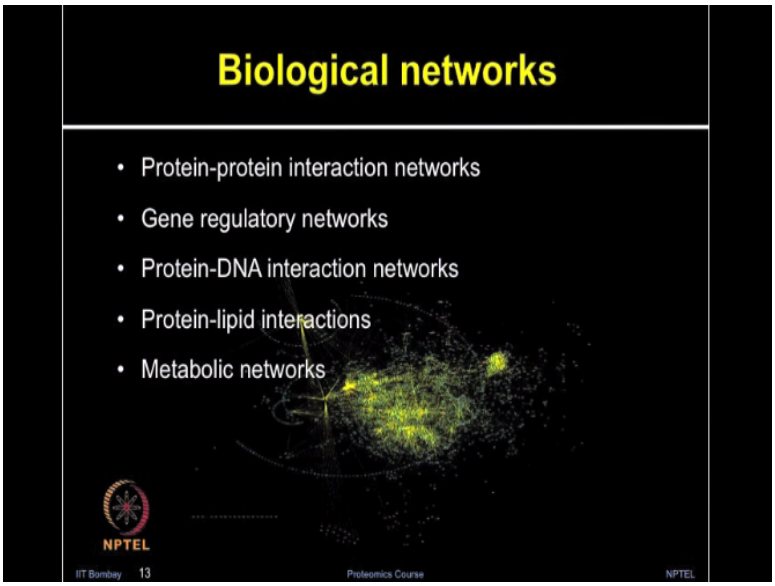
(Refer Slide Time: 15:25)



The system 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.

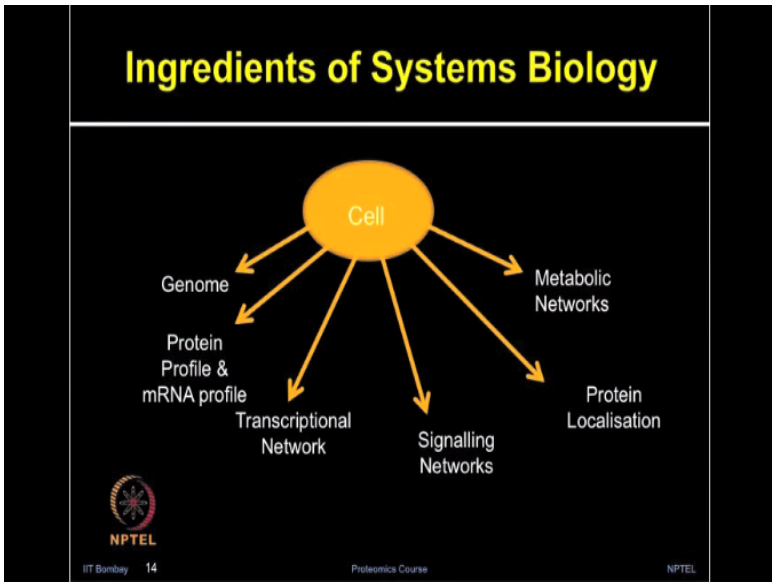
These networks provide insights which cannot be analyzed from the isolated components of the system. The common elements of the systems biology include networks, modeling, competition and dynamic properties.

(Refer Slide Time: 16:23)



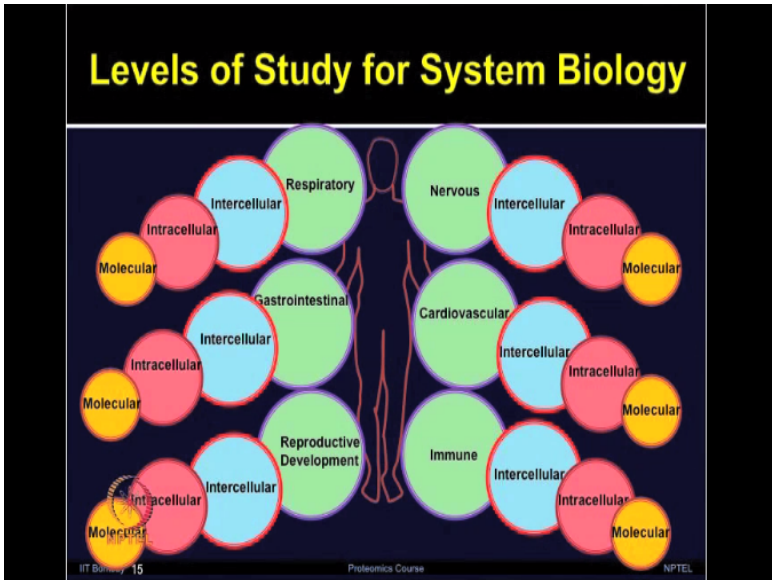
The different type of biological networks such as protein-protein interaction networks, gene regulatory networks, protein and DNA interaction networks, protein lipid, protein other bio molecules network and metabolic networks.

(Refer Slide Time: 16:40)



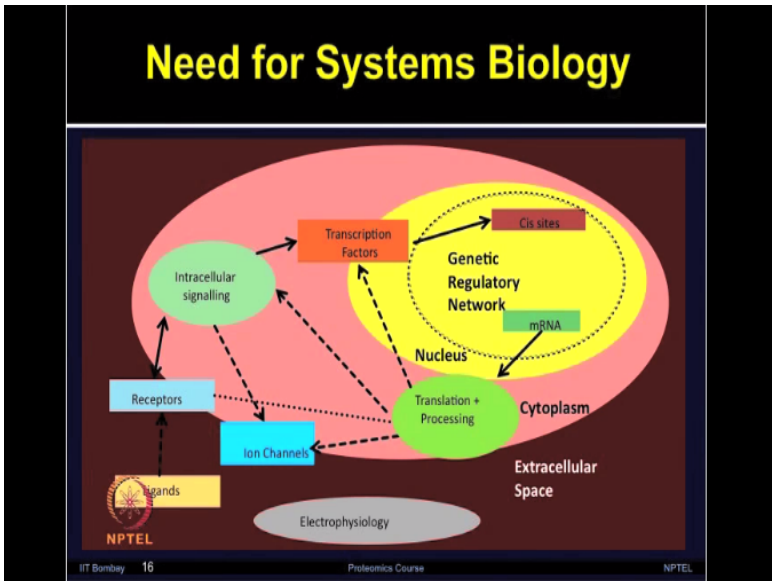
The various ingredients of systems biology for example if you are studying about a cell and the systems behavior, one need to look at the genome, its transcriptome profile, proteome profile, how protein DNA and different type of transcriptional networks are altered, protein-protein signaling networks, multi array complexes how they are formed, how protein is localized in the intra cellular dynamics and metabolic networks. So all of these are ingredients of a study about a system.

(Refer Slide Time: 17:18)



Now systems biology study can be done at different levels. For example, to study the complex physiology of human, one can look at individual systems such as respiratory systems, nervous systems or other physiological systems. The studies can be done at the inter cellular or intra cellular level and finally at the molecular level involving genomics, transcriptomics and Proteomics.

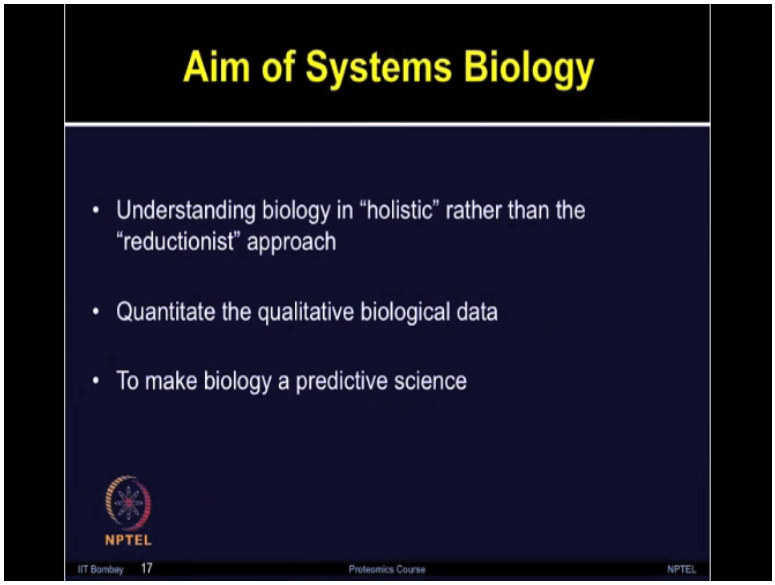
(Refer Slide Time: 17:50)



So, why there is lead for system biology. The study of biology at the system and sub system level for understanding the biological processes and network is very much required. As you can see, to understand even simpler system of a cell how it is regulated with its extra cellular space, the cytoplast and different other components.

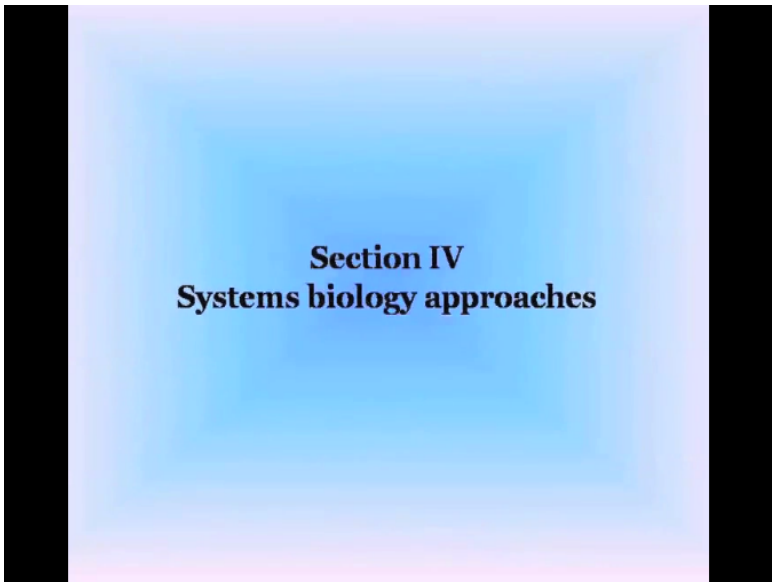
Examination of this structure and dynamics of cellular and organismal function is very much required for understanding the systems rather than the characteristics of isolated parts of the cell or organism. So what is the aim of systems biology? To understand the biology in holistic approach rather than the reductionist approach.

(Refer Slide Time: 18:47)

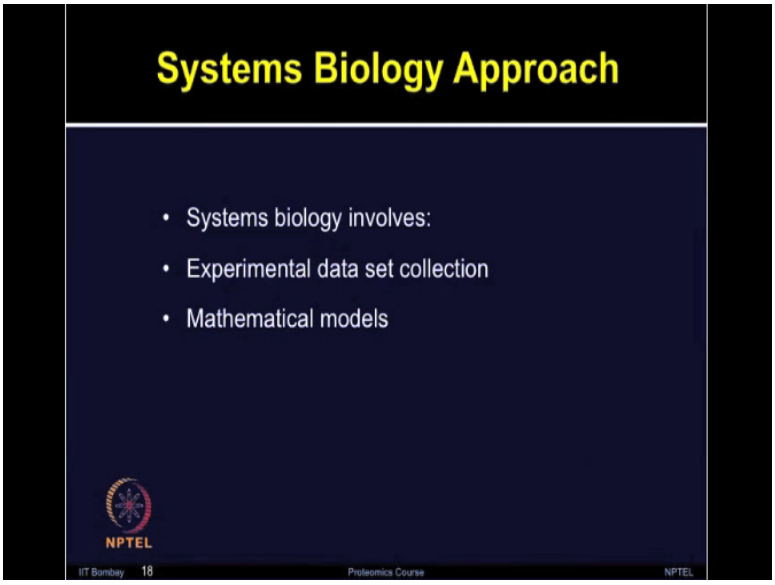


The systems biology aims to quantitate the qualitative biological data and provide some level of predictions by applying different type of computational modeling.

(Refer Slide Time: 19:02)

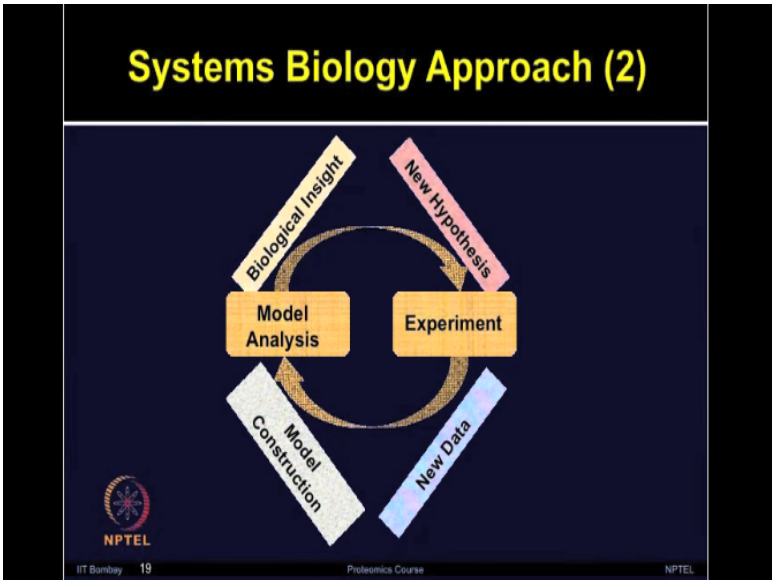


(Refer Slide Time: 19:12)



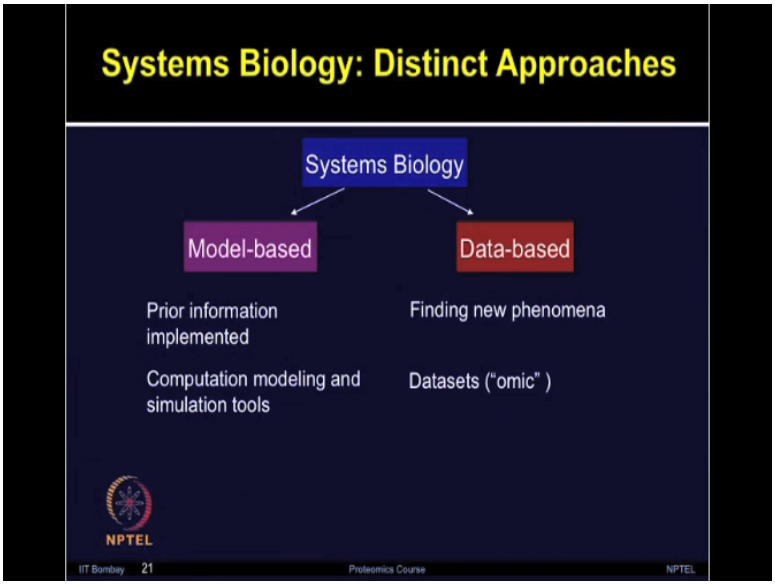
The systems biology approach involves first of all collection of large experimental data sets and then mathematical models to provide insight of some significant aspects of the data sets. The simple systems biology approach would involve experiments by adding new data sets.

(Refer Slide Time: 19:34)



Which will be used for model constructions and model analysis and the biological insight derived from these models can be used to propose new hypothesis. So the properties of a systems 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. So what are different approaches have been taken to study the systems biology.

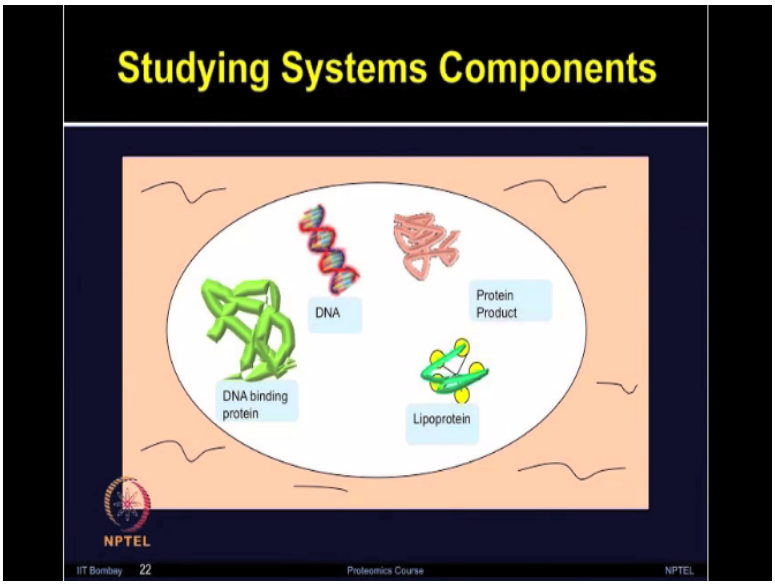
(Refer Slide Time: 20:13)



The distinct approaches of systems biology include the model based and data based methods. In model based approach involves some prior information which can be implemented in the models. Whereas in the data based, objective is to find a new phenomenon. The model based relies on computational modeling and simulation tools where as the data based method relies on the omics data sets.

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 the various type of omics information and metabolic pathways and networks can be created.

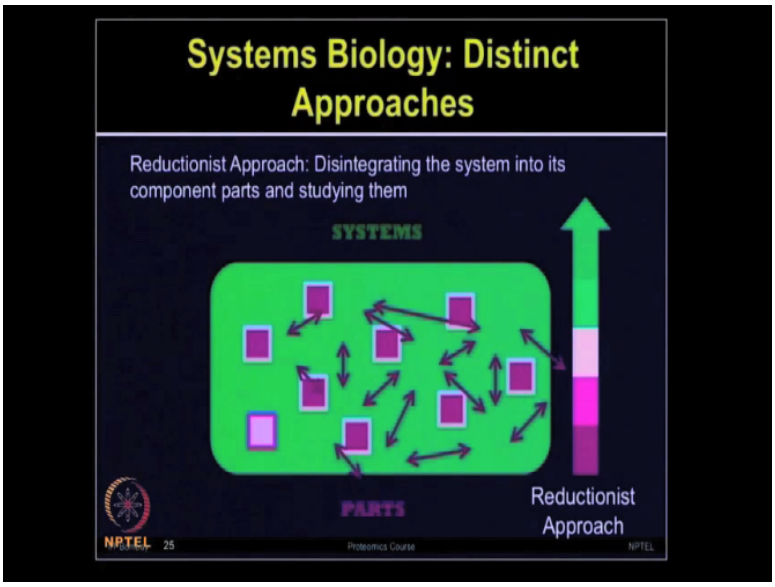
(Refer Slide Time: 21:09)



Studying systems component is very challenging. Systems biology and biological network modeling aims to understand the systems structure and function for better understanding of

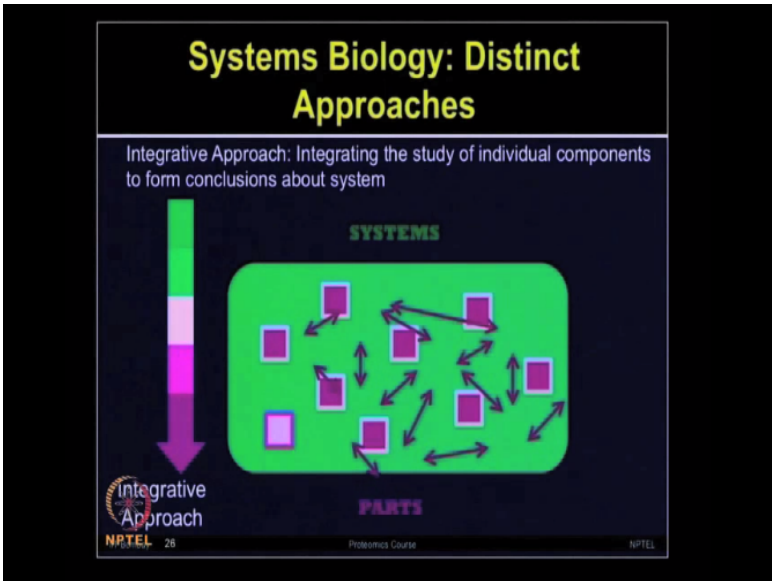
system properties like its robustness as well as used for the prediction of systems behavior in response to the perturbations.

(Refer Slide Time: 21:37)



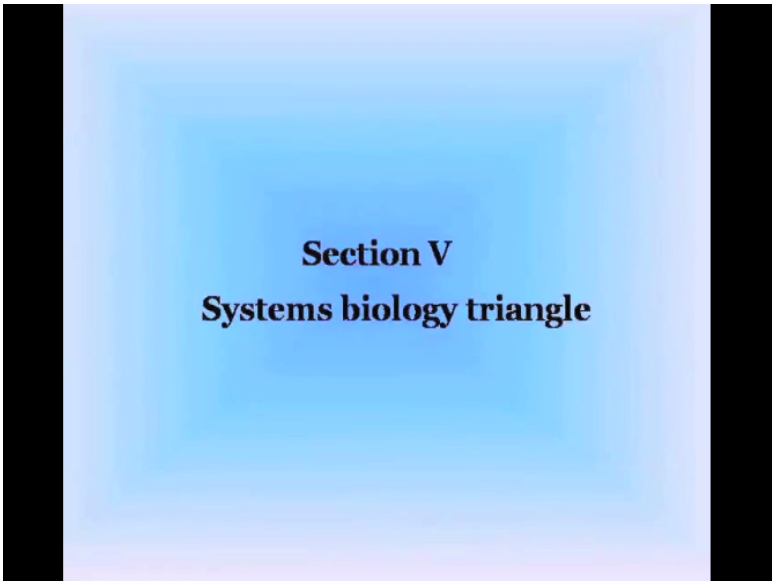
The reductionist approach involves disintegrating the system into its component parts and studying them.

(Refer Slide Time: 21:47)



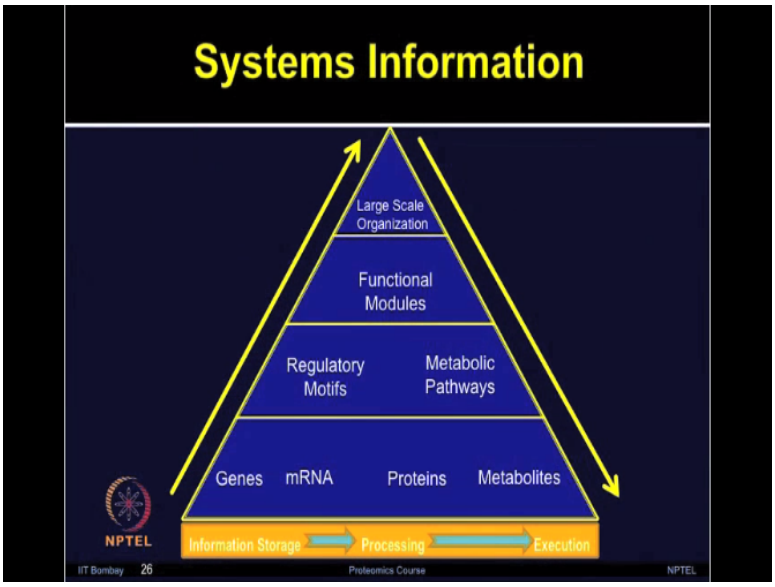
Whereas the integrative approach involves integrating the study of individual components to form conclusions about the system.

(Refer Slide Time: 21:58)



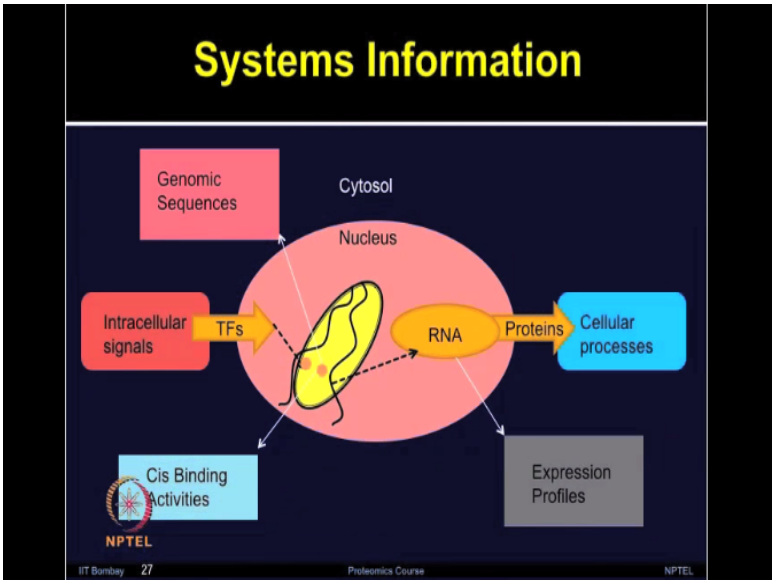
What is systems biology triangle? So, first of all the systems information is generated at various levels.

(Refer Slide Time: 22:24)



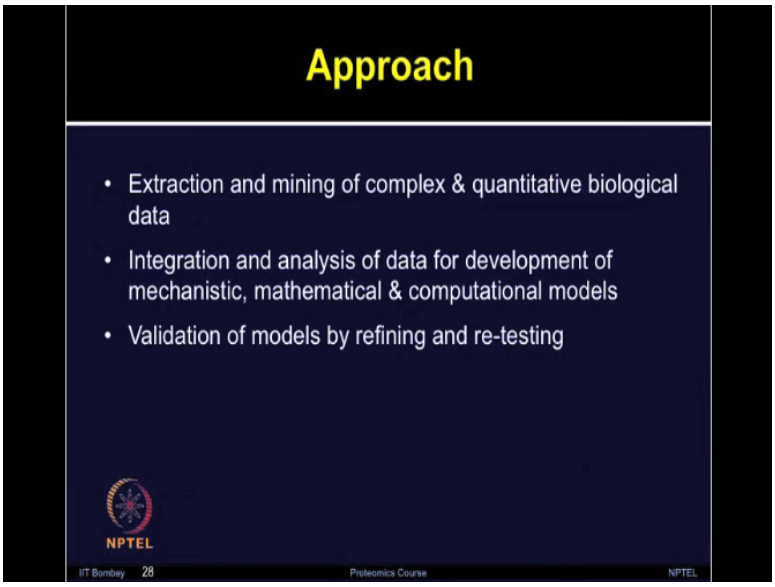
As we have discussed starting from genes to MRNA to proteins to metabolite or identifying regulatory motifs, metabolic pathways, functional modules and different large scale organizations. This information has to be stored, processed and further executed to identify the system level information. Even simpler systems such as cell can be linked with various properties.

(Refer Slide Time: 22:51)



Its genome sequences of different molecules intracellular signals, transcription factors different type of Cis binding activities, the expression profiling of RNA and proteins and different type of cellular processors. So what approach one can take to study about the system?

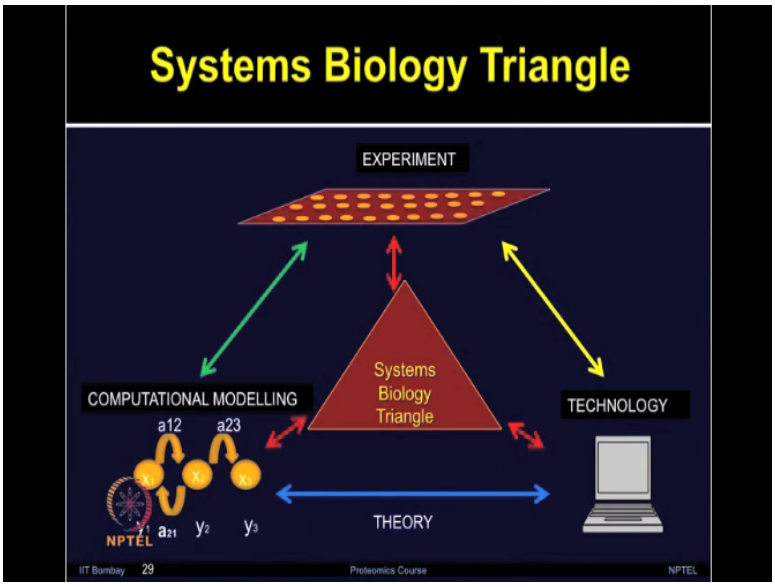
(Refer Slide Time: 23:14)



Extraction and mining of complex and quantitative biological data. Integration and analysis of these data sets for the development of mechanistic, mathematical and computational models. Validation of these models by retesting and refining after proposing some hypothesis. Different online data bases and repositories are nowadays developed for sharing larger assets 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.

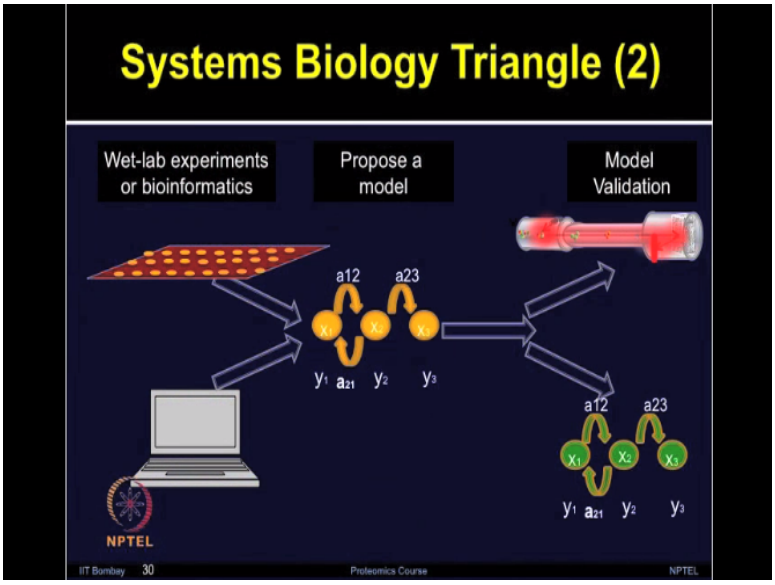
(Refer Slide Time: 24:10)



The systems biology triangle as you can see here involves the experimental data set could be derived from different type of omics platforms, technologies how the computational analysis can be performed, different type of bioinformatics software's and tools and then computational modeling's by obtaining some theoretical concepts.

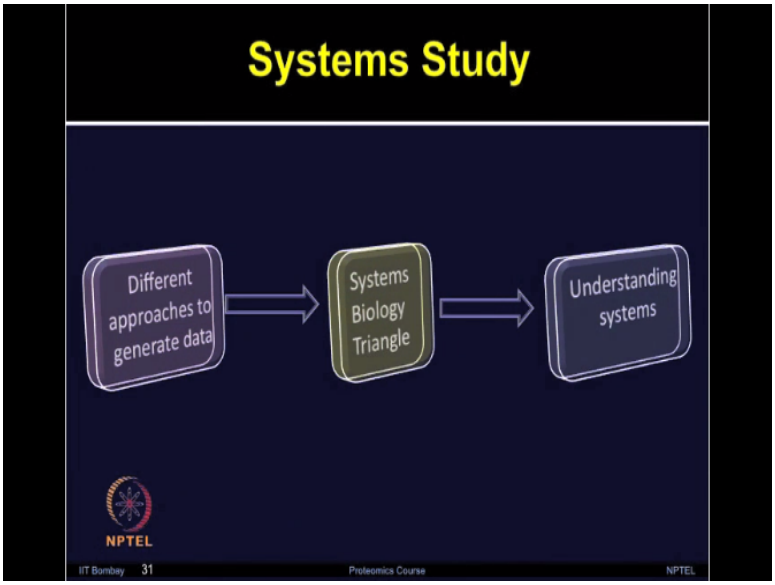
Thus energistic application of the experiment, theory and technology with modeling to enhance the understanding of biological processes as whole system rather than the isolated part is termed as systems biology triangle.

(Refer Slide Time: 24:52)



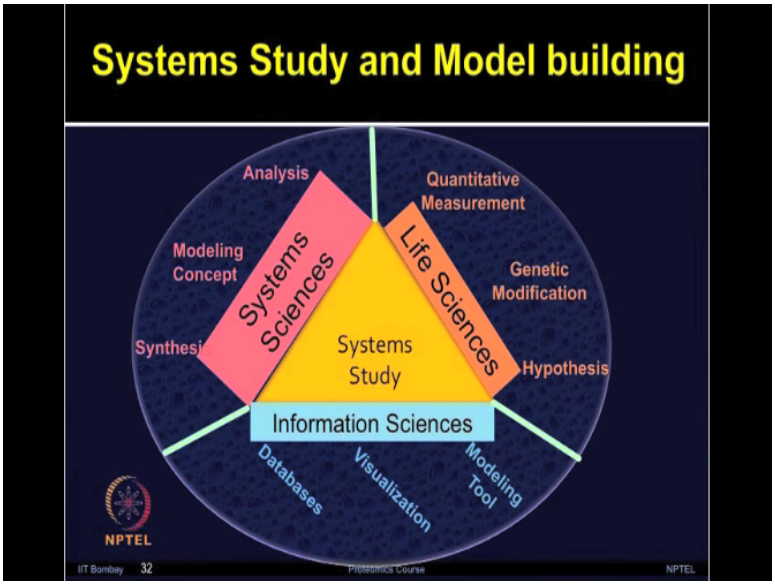
Systems biology triangle, the wet lab experiments are bioinformatics based data analysis can be used to propose a model. The model building as an aid to understand the complex system and some hypothesis can be generated which could be used further to propose more quantitative models or predictive models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

(Refer Slide Time: 25:25)



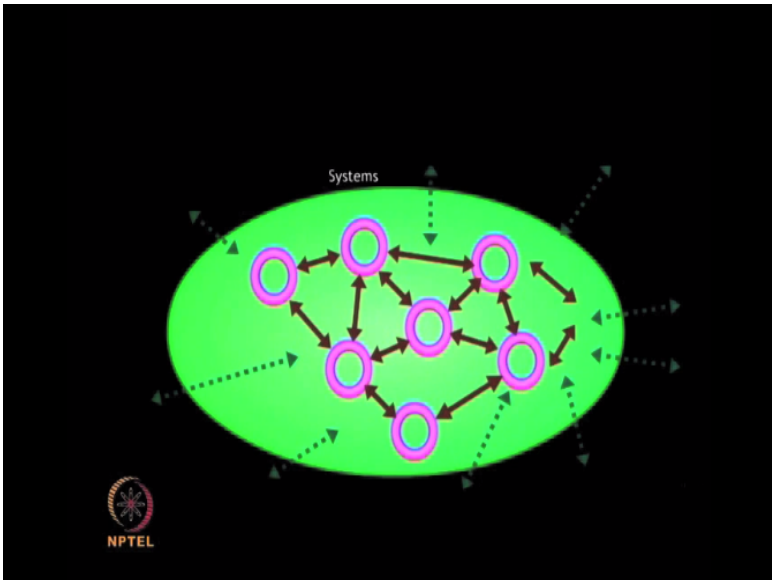
So what are systems study? First of all, the difference between systems study and the component study one need to understand and what we have try to emphasize in the previous slides after generating the data set and creating the system biology triangle then this information can be used for understanding the systems in the more complex and mechanistic level.

(Refer Slide Time: 25:54)



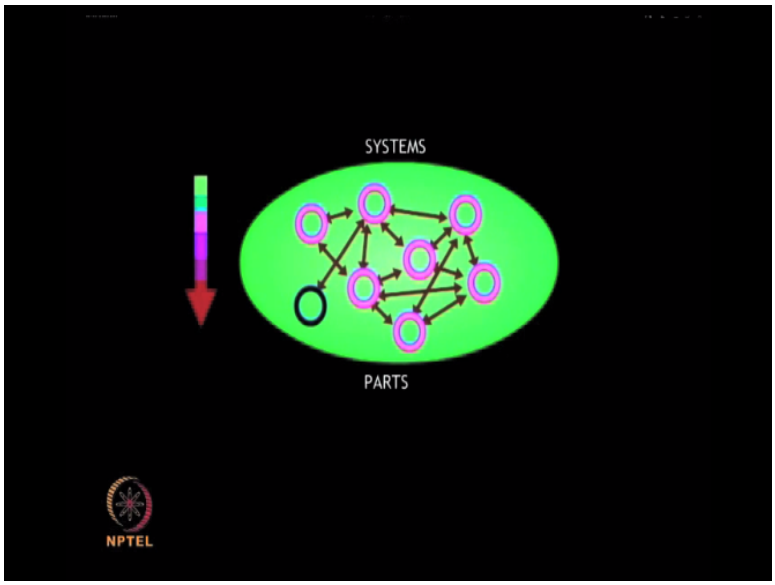
Systems study and model building. The system science includes synthesis. Modeling concepts, analysis, life sciences provide quantitative measurements, genetic modifications and deriving some hypothesis. The information sciences enable the visualization, the modeling tools and different databases. So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

(Refer Slide Time: 26:32)



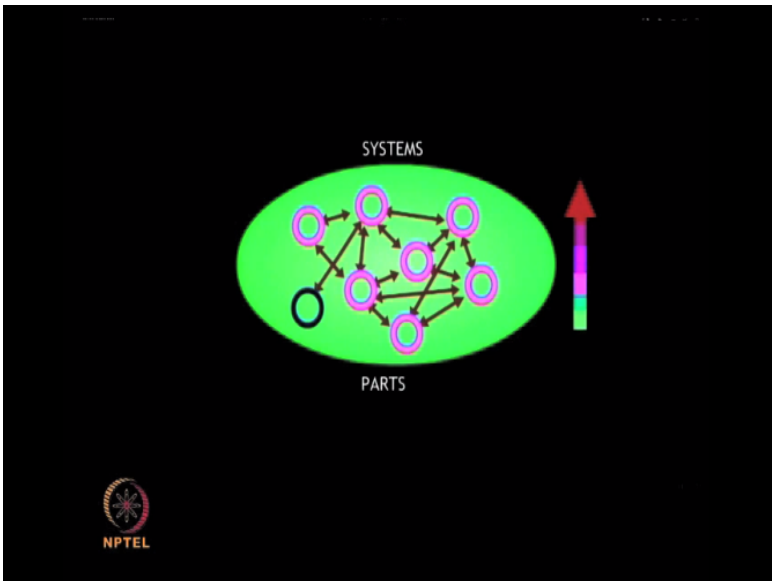
System is an entity which maintains its existence through a mutual interaction of its constituent parts. The systems biology research consists of identification of the parts, characterization of the components, exclude the ones which are not a part of the system identify the interaction of the components with each other and identify the interaction of the components with involvement which modulate the parts either directly or indirectly through modulation of internal interactions.

(Refer Slide Time: 26:17)



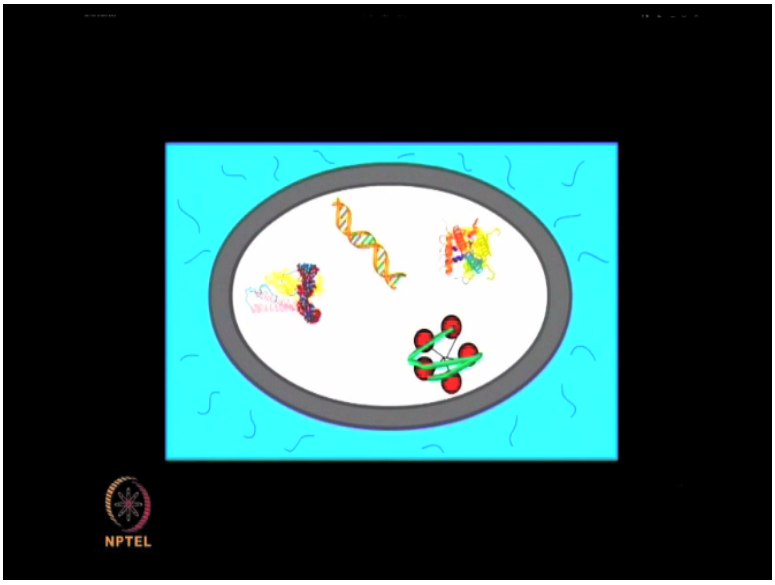
The system biology concept can be understood with help of two approaches such as reductionalist approach and integrative approach. The reductionalist approach focuses on disintegrating the system in to its component parts and studying them.

(Refer Slide Time: 27:45)



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

(Refer Slide Time: 27:57)

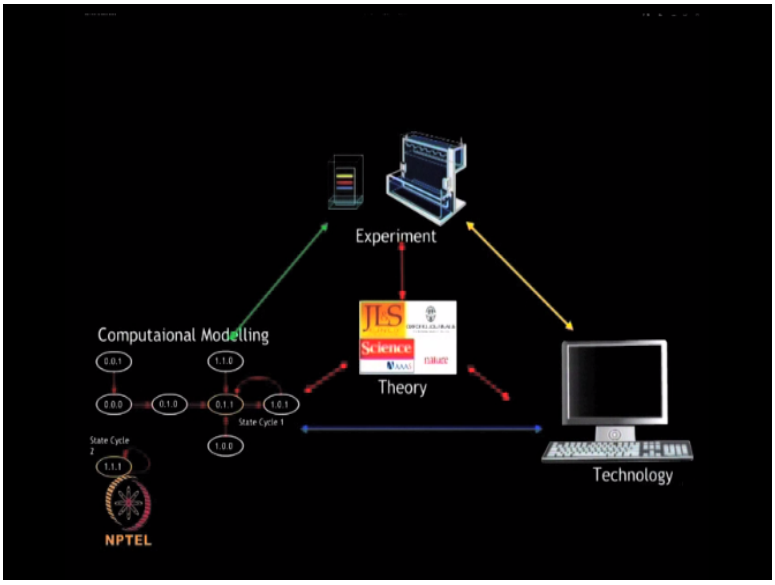


Consider a cell with its component molecules. Later we want to study the metabolic pathway as a biological system. When the environment of the cell is perturbed a little, the individual components undergo unique changes such as increase in the production rate or decrease in their amount. At this stage, due to lack of knowledge of the nature of interactions of proteins, we cannot interpret how the system gets affected.

But when we study the interaction of one component with the other, we can conclude that the increase in rate, if DNS binding protein leads to increase in the synthesized amount of DNA. Which further changes the fungal amount of lipo proteins produced. Thus we can see that we study a system, we need to analyze not just the components but their interactions.

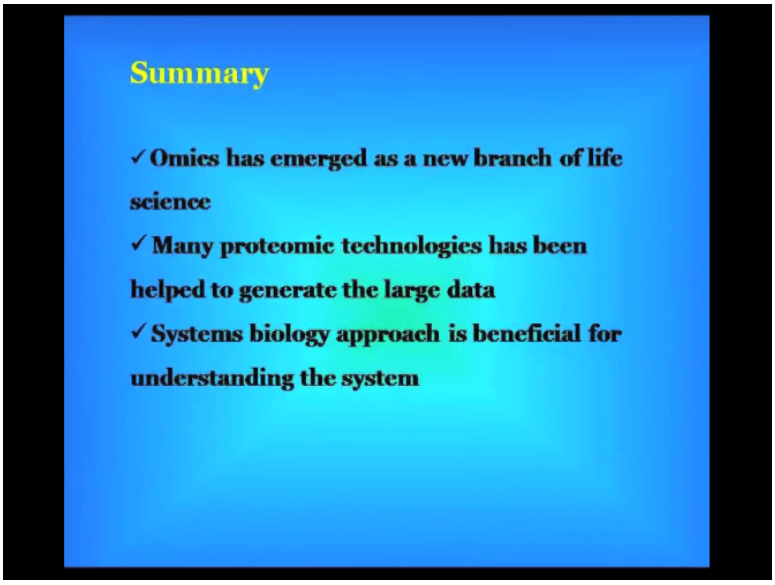
These (()) (29:13) systems can be protein-protein interaction networks, gene regulatory networks, protein DNA networks, protein lipid networks and metabolic networks. To study the system, we need to know about the components and its interactions. The data about the components comes from genomic and Proteomics studies. The information about the molecular interactions comes from the interactomic studies.

(Refer Slide Time: 29:51)



Here it is shown, that is 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: 30:14)



(Refer Slide Time: 30:22)

References

✓Systems Biology: A Brief Overview, Hiroaki Kitano,, 1662 (2002); 295 Science

✓Simpson R. J. Proteins and Proteomics: a laboratory manual ed., Cold Spring Harbor laboratory press, 2002. ISBN: 0879695544.

✓Trey Ideker, L. Raimond Winslow, A. Douglas Lauffenburger. Bioengineering and Systems Biology. Annals of Biomedical Engineering. February 2006, Volume 34, Issue 2, pp 257-264

✓Trey Ideker, et al., Integrated Genomic and Proteomic Analyses of a systematically Perturbed Metabolic Network, Science, 2001.

✓Sarath Chandra Janga and Julio Collado-Vides. Structure and evolution of gene regulatory networks in microbial.