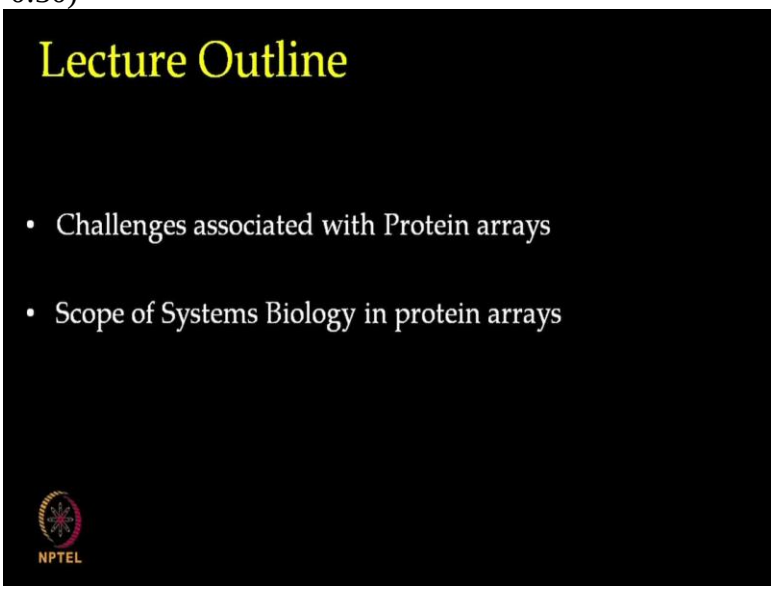
Interactomics: Protein Arrays & Label Free Biosensors
Professor Sanjeeva Srivastava
MOOC NPTEL Course
Indian Institute of Technology Bombay
Module 8
Lecture No 39
Application of cell-free expression protein microarrays in biomarker discovery

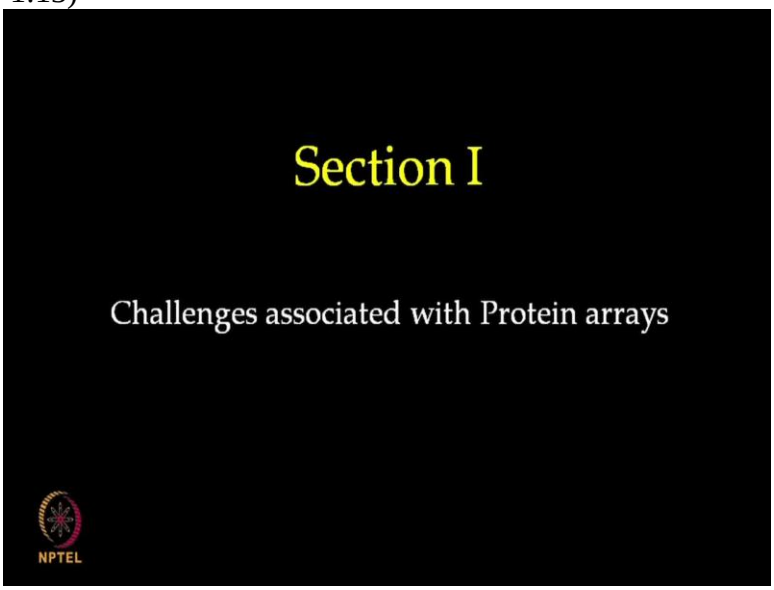
Welcome to mooc interactomics course. In this module you have gone through the details of protein microarray platform, its work flow and applications.

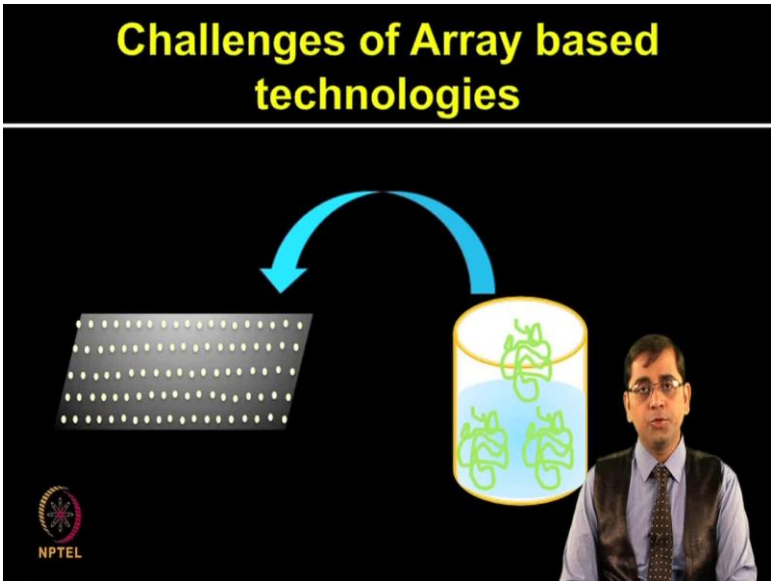
(Refer Slide Time: 0:30)



In today's lecture we will first go through some of the challenges associated with protein microarray platforms. It is essential to get a perspective of how this high throughput data set can be used to obtain a holistic understanding of biological systems. In today's lecture, we will therefore move from protein microarrays and see how any proteomics data be it from microarray or other platforms like mass spectrometry, gel based proteomics can be used in a discipline known as systems biology.

(Refer Slide Time: 1:13)





Let us begin with the challenges of array based proteomics protein and antibody microarrays. Protein microarrays is a platform on which thousands of proteins can be printed. As you have seen in the last previous lectures it provides an important platform for large scale functional analysis of proteins, although due to the high throughput capabilities array based proteomics attracted tremendous attention in clinical research.

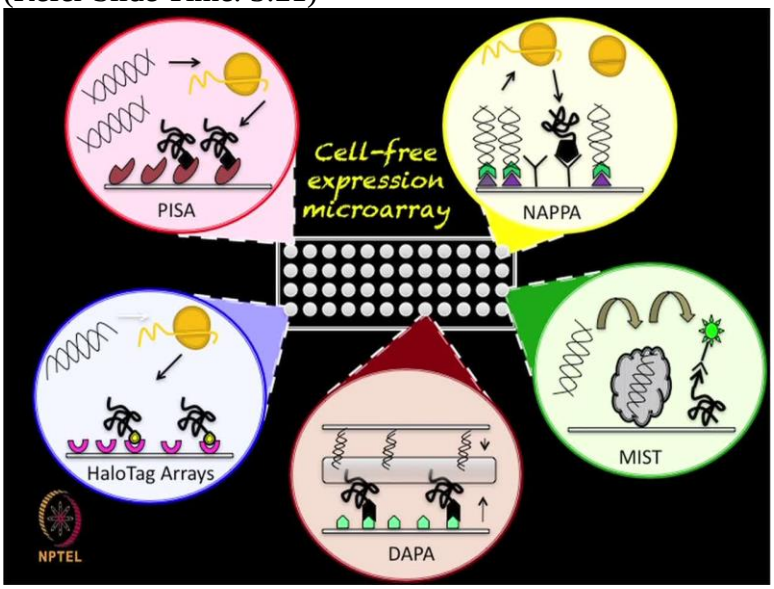
It has quite a few technological challenges as well, protein array designing difficulties include acquisition, arraying and stable attachments of proteins to array surfaces and detection of

interacting proteins. Further, miniaturization of assays, protein dehydration, nonspecific binding, unavailability of highly specific antibodies against all the proteins that comprise the complex proteome and lack of direct correlation between protein abundance and its activities.

The complex nature of proteins has posed many hurdles in the area of array fabrication, printing, scanning and data analysis. DNA microarrays had laid the foundation in developing experimental and analysis strategies. However, there is an enormous scope for innovation and making protein microarrays more robust and user friendly.

Diversity in protein size, spatial location of proteins in cell, post transition modifications and hydrophobic nature of the proteins post challenges for protein purification, maintenance of activity and orientation for functional assays using protein microarrays.

(Refer Slide Time: 3:21)

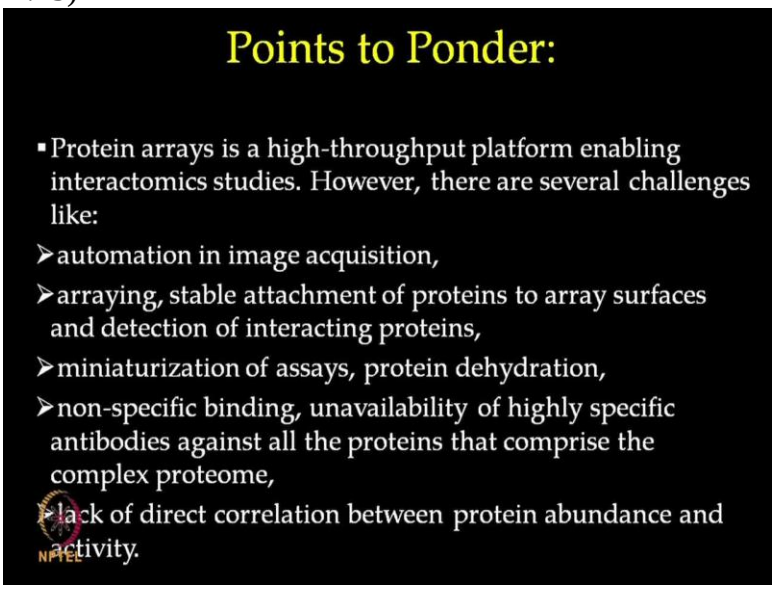


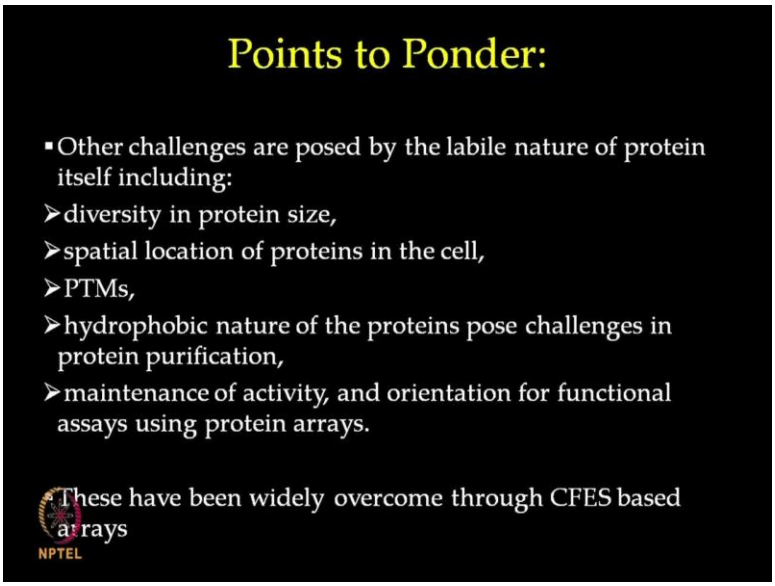
Technologies like cell-free expression system have helped to overcome the challenges of high throughput expression and purification for diverse class of proteins. Cell free expression aids in generation of near native configuration of proteins, nucleic acid probable protein arrays and halotag technologies allow the presentation of hidden or buried epitopes in a protein for development of better assays.

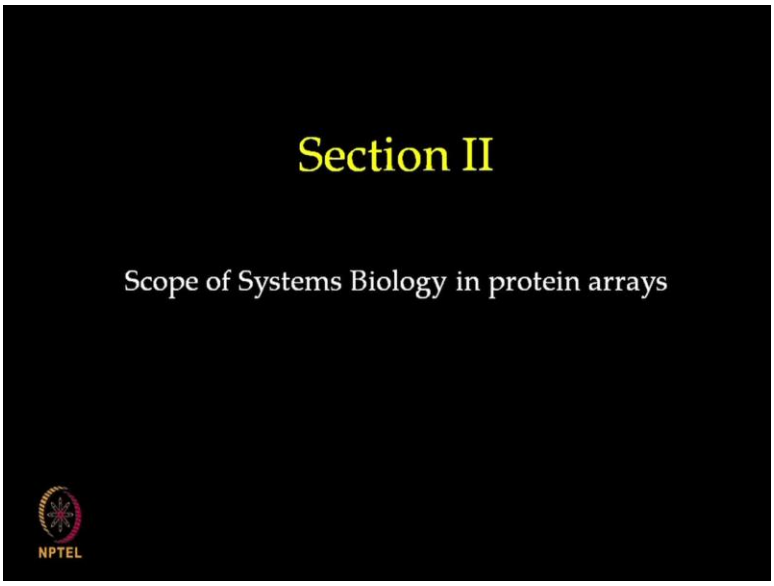
Many commercial biochips are now available for the study of post translationally modified proteins. Thus protein microarrays have overcome many of the inherent challenges through

innovation in last few years. However, data analysis continues to be an area which could be worked on further to establish novel methods for stringent data analysis.

(Refer Slide Time: 4:25)

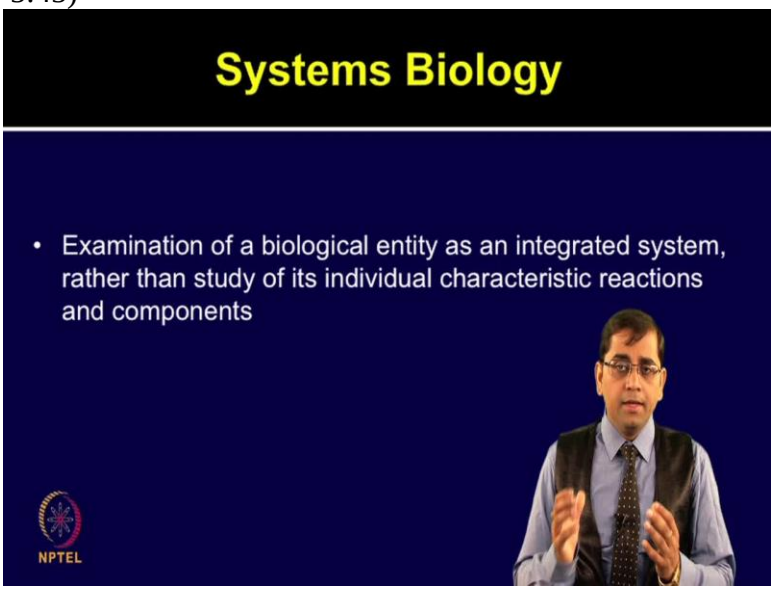






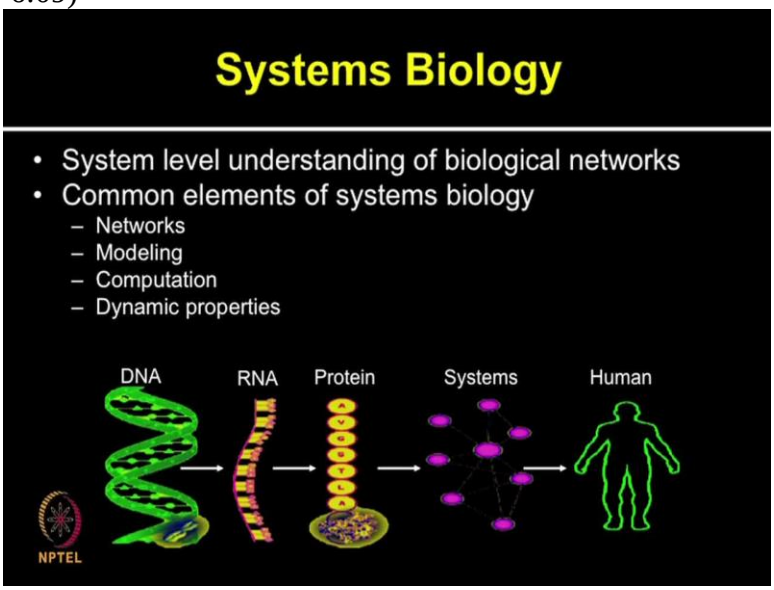
There are several challenges associated with these high throughput technologies, however the data from such platforms are indispensable resource for integrative understanding of living systems. This brings us to the last segment of this module which is known as systems biology.

(Refer Slide Time: 5:45)



System biology is an 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. It is a study of all the mechanism underlying complex biological processes in the form of its interacting components.

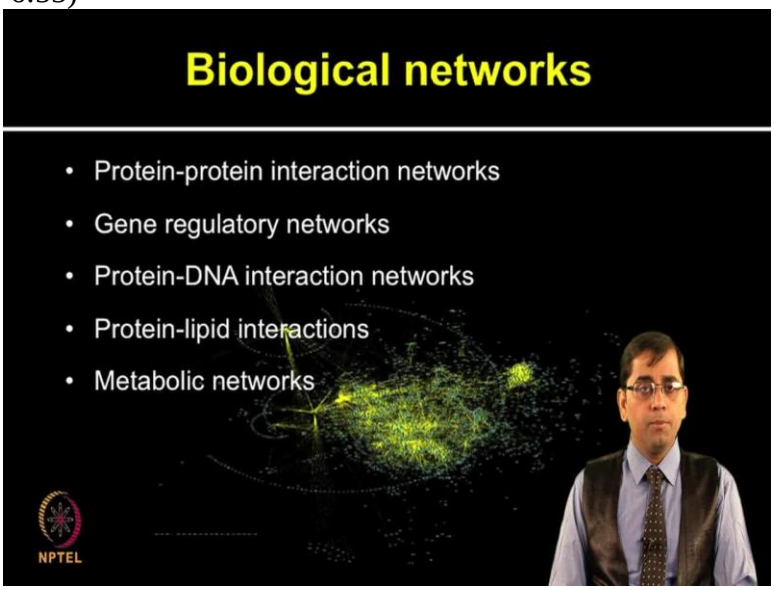
(Refer Slide Time: 6:09)



Systems level understanding require information from different levels at the gnome, transcriptome, or proteome level which can then be applied to understand the complex system and also be applied to different organisms. The biological information is represented by the network of interacting elements and dynamic response to the perturbations. These networks provide insights which cannot be analyzed from the isolated components of the system.

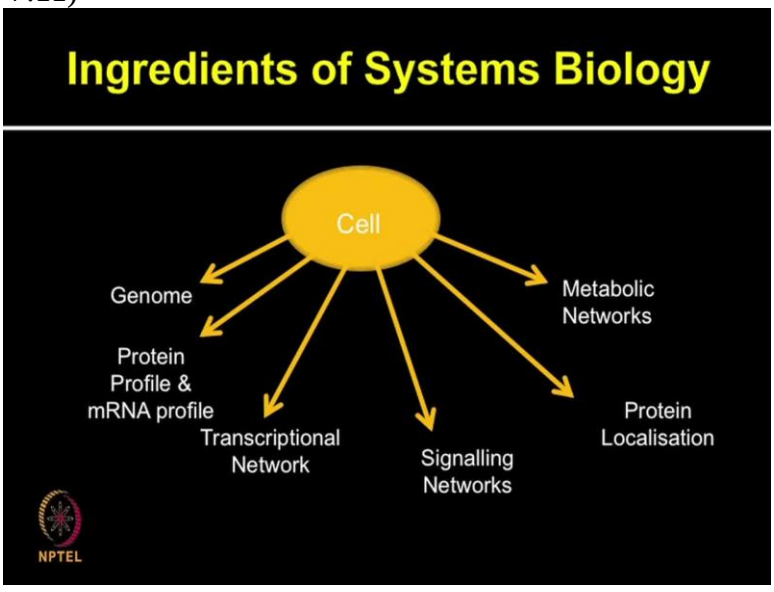
The common elements of systems biology include networks, modeling, computation and dynamics properties.

(Refer Slide Time: 6:55)



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

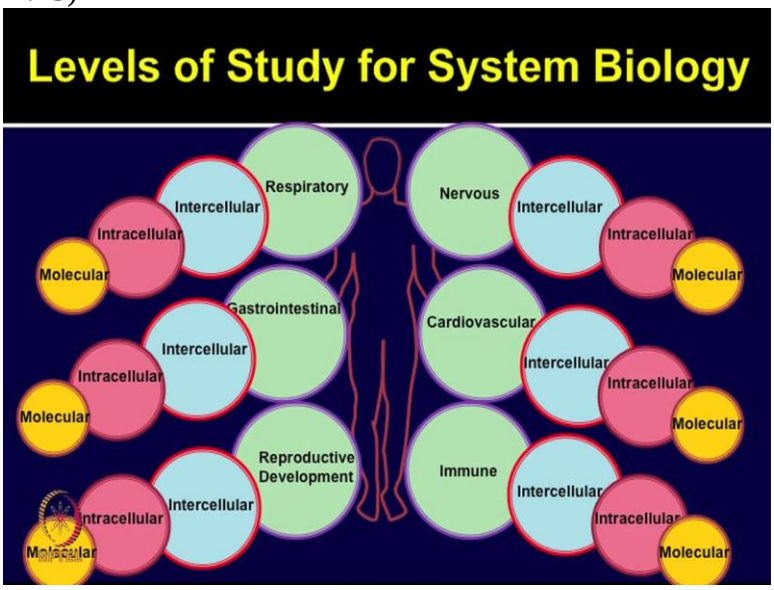
(Refer Slide Time: 7:11)



If you are studying a cell and its systemic behavior you need to look at the genome, its transcriptome profile, proteome profile, how protein-DNA and different transcription networks are altered, protein-protein signaling networks, multimeric complexes, etc. We need to know

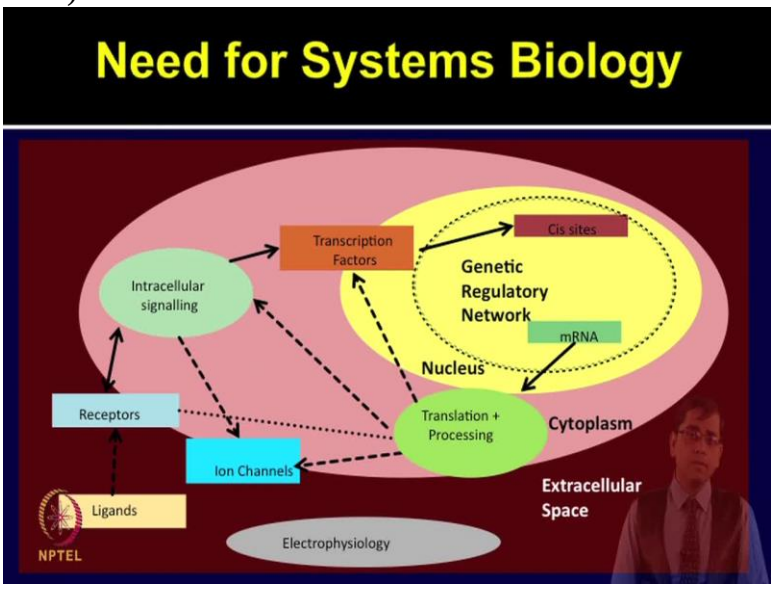
how they are formed, how protein is localized in the intracellular dynamics and metabolic networks, these are the vital entities of studying systems biology.

(Refer Slide Time: 7:43)



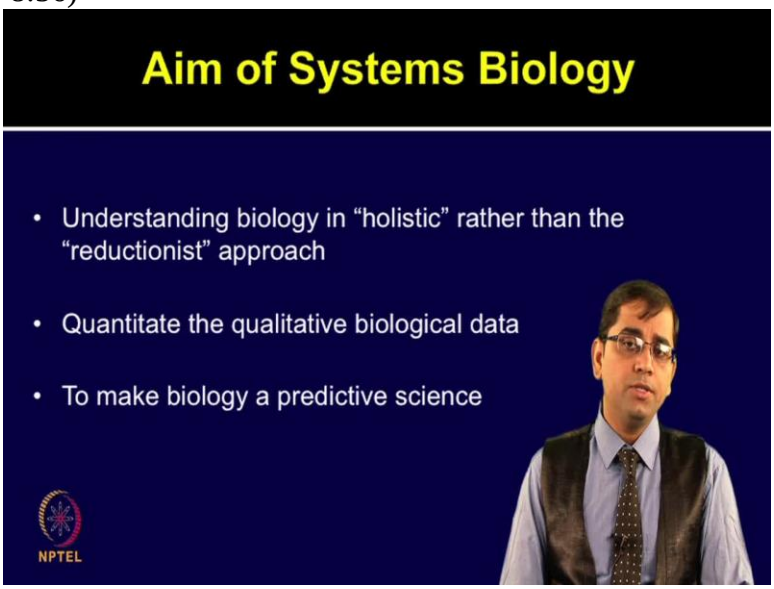
Now systems biology study can be done at different levels. For example, to study the complex physiology of human one could look at individual systems such as respiratory, nervous system or other physiological systems. Studies can be done at the intracellular or intercellular level and finally at the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

(Refer Slide Time: 8:12)



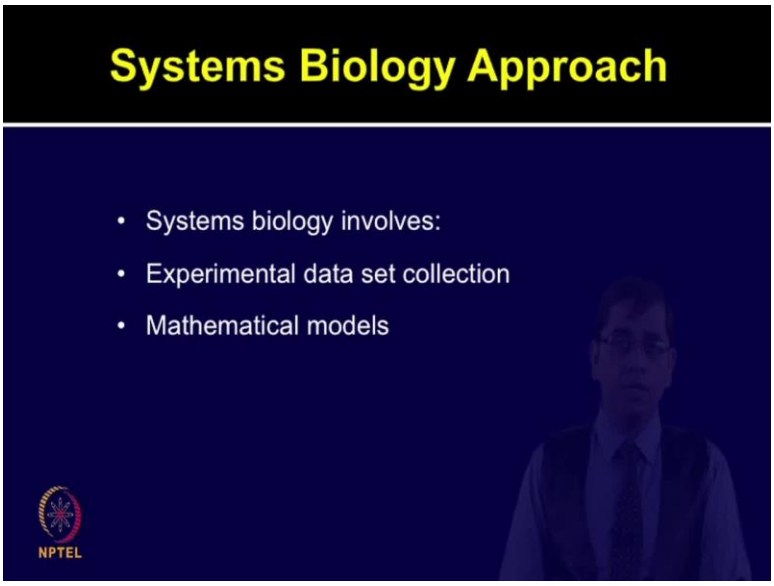
So why there is need for systems biology? The study of biology at the system or subsystem level for understanding the biological processes and network is very much required. As you can see in this slide to understand even simpler system of a cell how it is regulated with the extra cellular space and the cytoplasm and different other components. Examination of structure and dynamics of cellular and organism function is very much required for understanding of systems rather than characteristics of isolated parts of the cell or the organism.

(Refer Slide Time: 8:56)



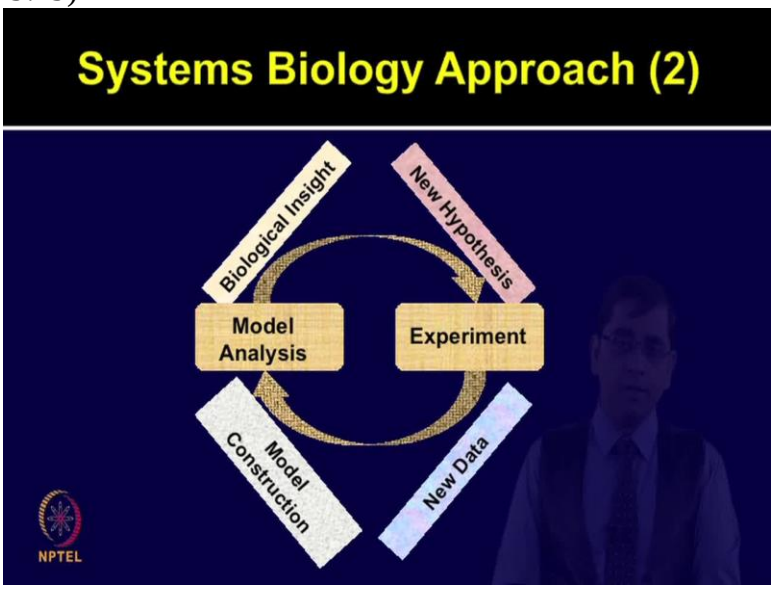
So what is the aim of systems biology? To understand the biology in holistic approach rather than the reductionist approach. The systems biology aims to quantitate the qualitative biological data and provide some level of prediction by applying different types of computational methods.

(Refer Slide Time: 9:17)



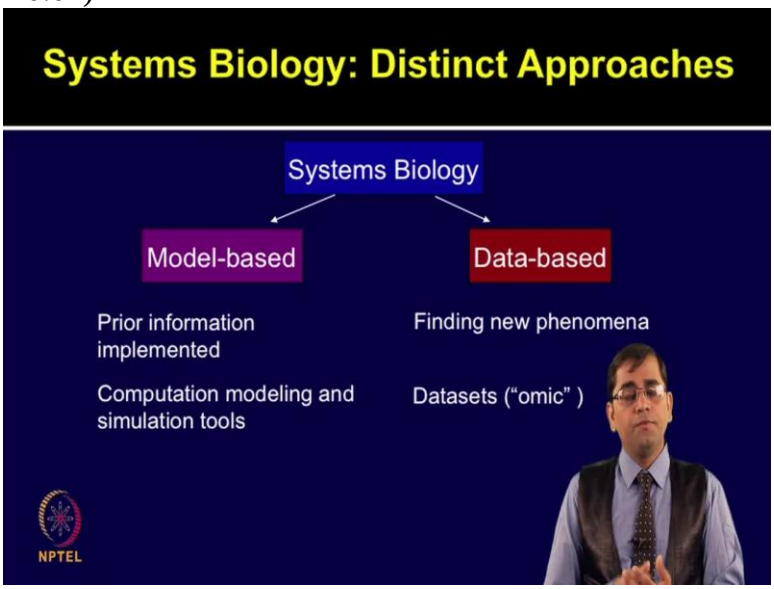
The systems biology approach involves first of all collection of large experimental data sets and then mathematical models to provide insight of some significant aspects of data set. The simple system biology approach would involve experiments by adding new data sets which will be used for model constructions and model analysis and the biological insight derived from these models could be used to propose new hypothesis

(Refer Slide Time: 9:49)



The properties of systems are probably more than just the sum of all of its individual components, therefore it is possible that system may have its own property by applying all the components.

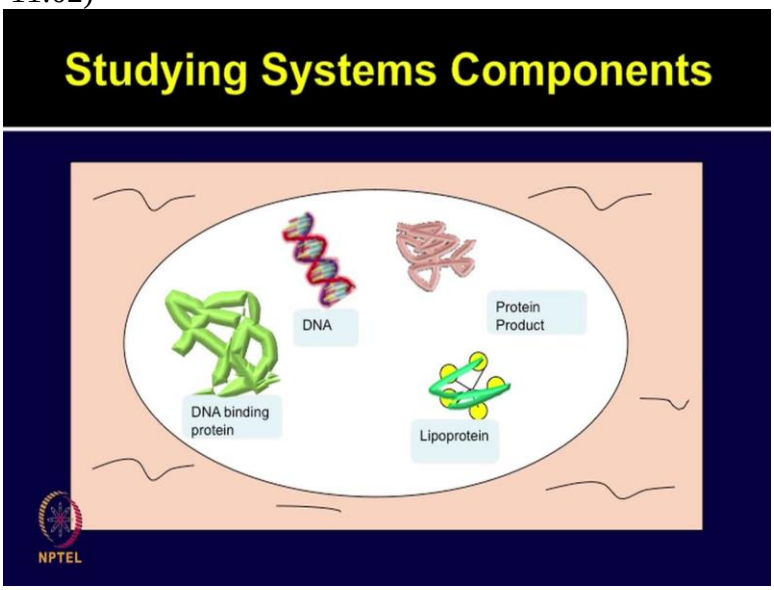
(Refer Slide Time: 10:04)



So what are the different approaches which have been taken to study systems biology? The different approaches of systems biology includes the model based and data based methods. The model based approach involves some prior information which can be implemented in the models, whereas the data based the objective is to find a new phenomenon.

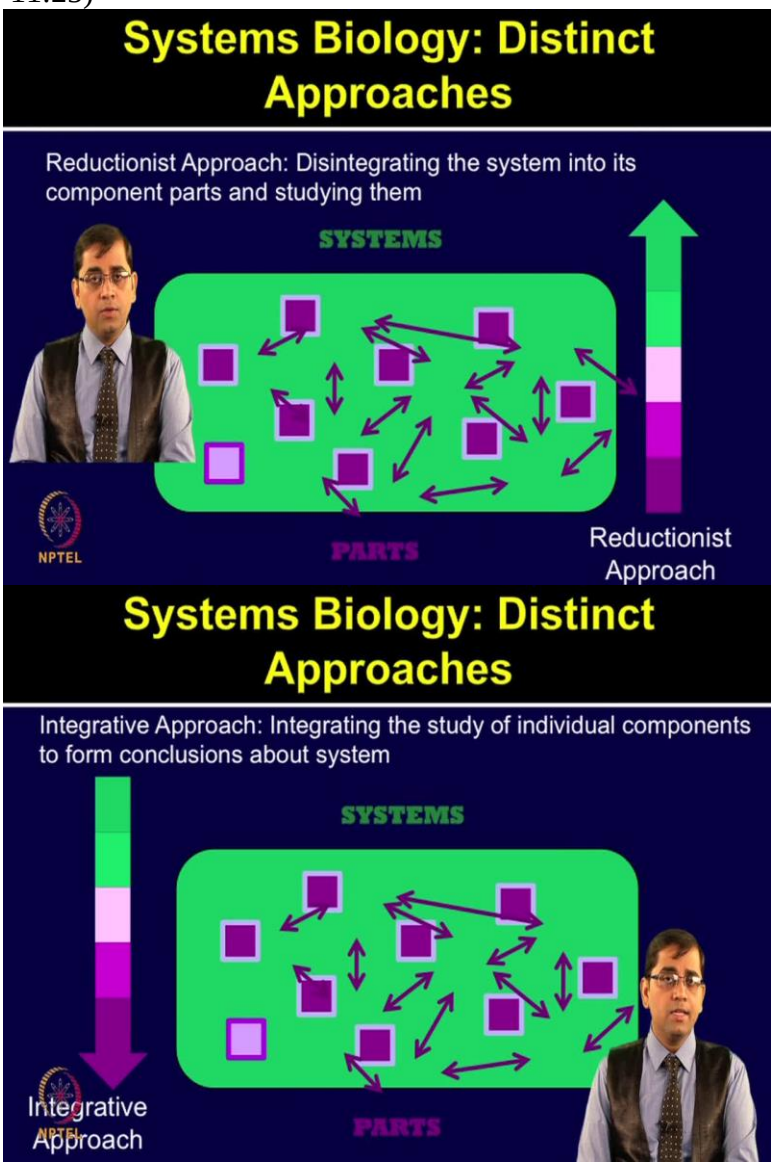
The model based relies on computational modeling and simulation tools, whereas the data based methods rely on the omics data sets. In model based systems biology approach it is difficult to build detailed kinetic models but in data based system the complex relationship among various types of omics information metabolic path ways and networks can be created.

(Refer Slide Time: 11:02)



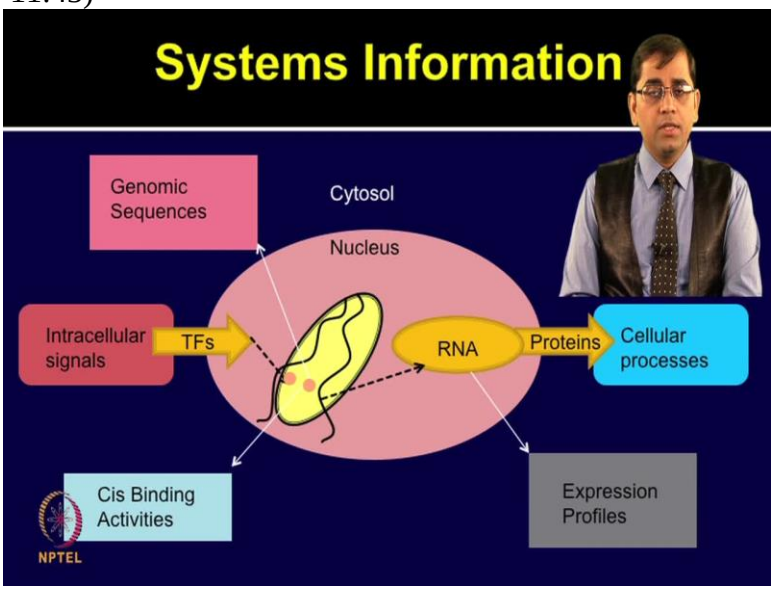
Studying systems biology is very challenging, systems biology and biological networks modeling aim to understand the system structure and function for better understanding of system properties like its robustness as well as its use for prediction of system behavior in response to the perturbations.

(Refer Slide Time: 11:25)



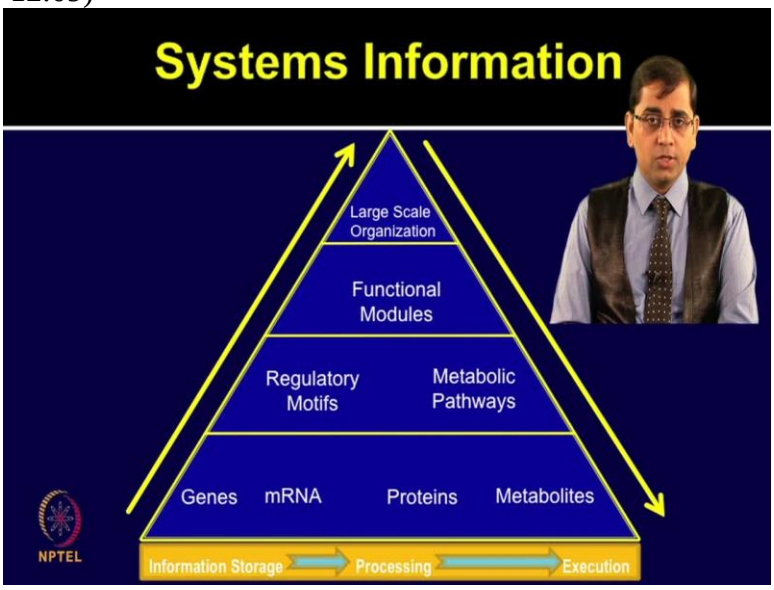
The reductionist approach involves disintegrating the system into its component and studying them, whereas the integrative approach involves integrating the study of individual components to form conclusions about the system.

(Refer Slide Time: 11:43)



Even simpler systems such as cell can be linked with various properties its genome sequences of different molecules intracellular signals transcription factors different type of sis binding activities the expression profiling of RNA and proteins and different type of cellular processes.

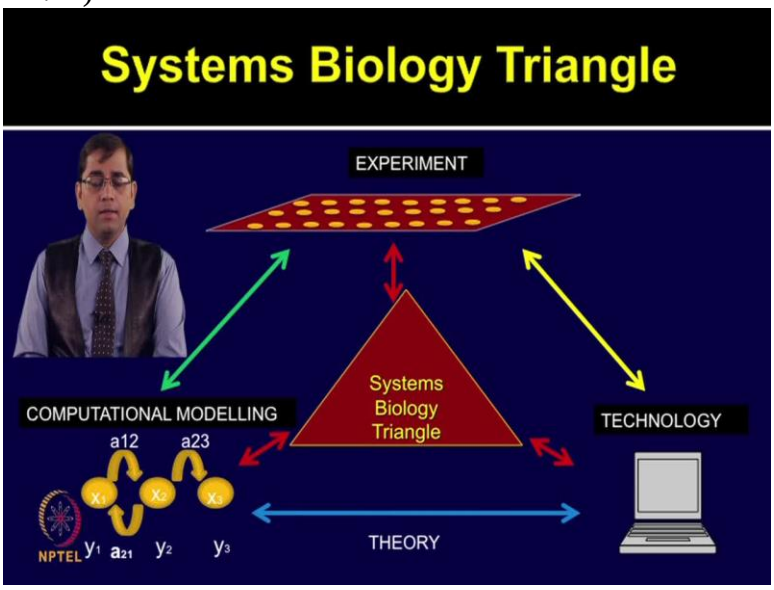
(Refer Slide Time: 12:09)



What is system biology triangle? First of all the systems information is generated at various level as we have discussed starting from genes to mRNA to proteins to metabolites, or even identifying regulatory motives, metabolic path ways, functional modules and different large scale

organizations. This information has to be stored and processed and further executed to identify the system level information.

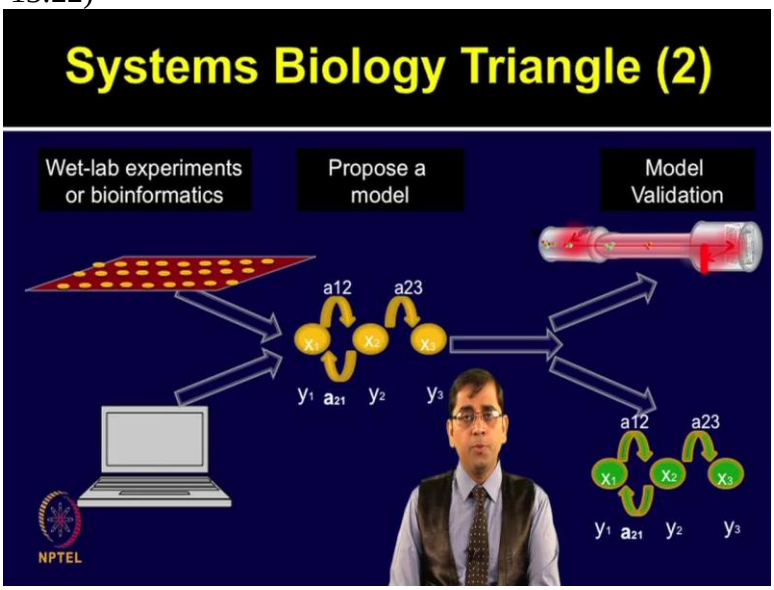
(Refer Slide Time: 12:42)



The systems biology triangle as can been seen here involves the experimental data sets which could be derived from different types of omic technologies. How the computational analysis can be performed different type of bioinformatics software and tools and then computational modeling by obtaining some theoretical knowledge.

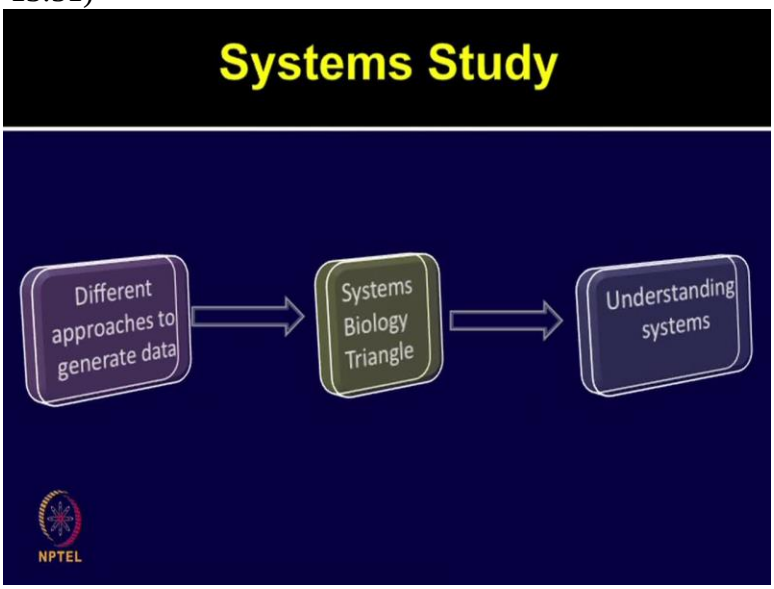
The synergistic application of the experiment, theory and technology with modeling to enhance the understanding of biological processes as whole system rather than the isolated components is termed as systems biology triangle.

(Refer Slide Time: 13:22)



A systems biology triangle the wet-lab experiments or bioinformatics based data analysis could be used to propose a model. The model building as an aid to understand complex system and some hypothesis could 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 model validation.

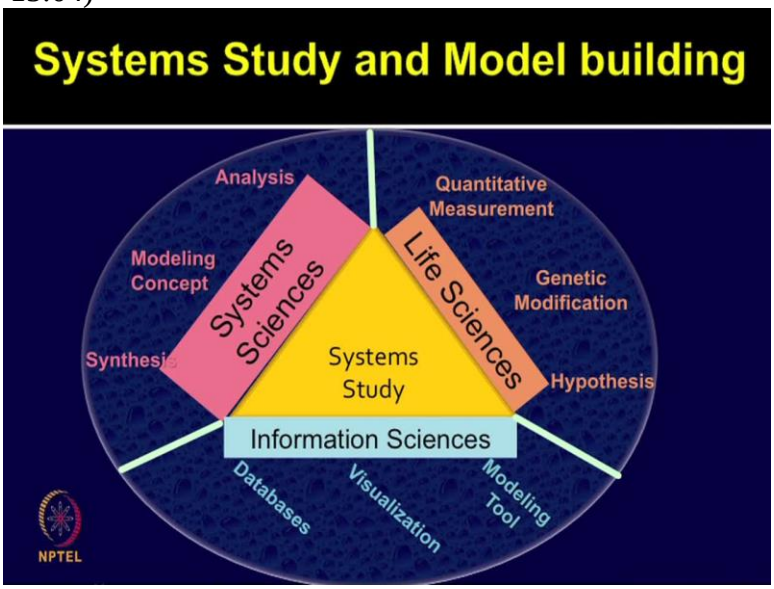
(Refer Slide Time: 13:51)



What is systems study? First of all, the difference between system study and component should be understood as discussed earlier. After generating the data set and creating the systems biology triangle this information could be used for understanding the systems in more complex and mechanistic level. So what approach one could take to study about the system extracting and mining the complex and quantitative biological data.

Integration and analysis of these data sets for development of mechanistic, mathematical and computational models. And validation of these models by retesting and refining after proposing some hypoxias. Different online databases and repositories are nowadays developed and available for sharing large datasets and various systems model. The systematic approach how molecules act together within the network of interaction that make up life is defiantly going to be useful to understand the systems biology.

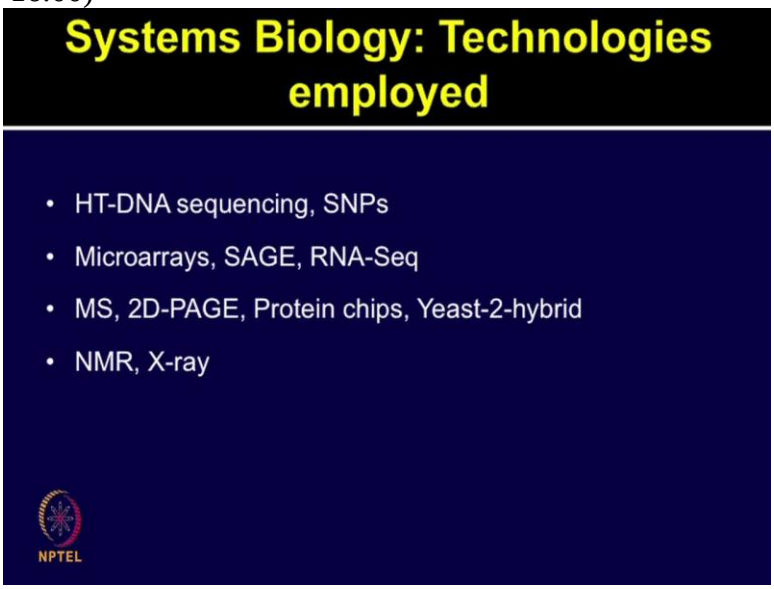
(Refer Slide Time: 15:04)



Systems study and model building, the systems science includes synthesis, modeling, concepts, analysis. Life science disciplines provide quantative measurements, genetic modifications, and deriving some hypoxias. The information science enables the visualization, the modeling tools and different databases. This model building as an aid to understand the complex system is very useful for system level investigation.

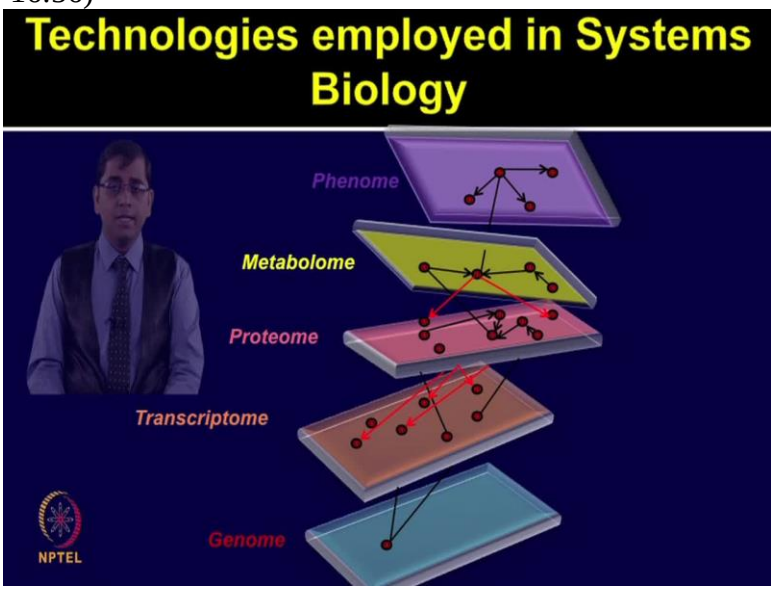
There are different technologies which have been employed to study the systems biology obviously, we need high throughput data set which could be derived from microarray platform or RNA deep sequencing different configurations of mass spectrometry different type of proteomic tools and protein interaction datasets.

(Refer Slide Time: 16:00)



Some of the technologies which are commonly employed in systems biology could be classified broadly under the following techniques. For genomics the high throughput DNA sequencing methodologies, mutation detection using SNP method. For transcriptomics the transcript measurements cam include serial analysis of gene expression, gene chips, microarrays and RNA sequencing. For proteomics mass spectrometry, two dimension electrophoresis, protein chips and yeast-2-hybrids and different structure proteomic tools such as x-ray and nuclear magnetic resonance. X-ray and NMR are mainly employed for metabolomic analysis in the field of metabolomics.

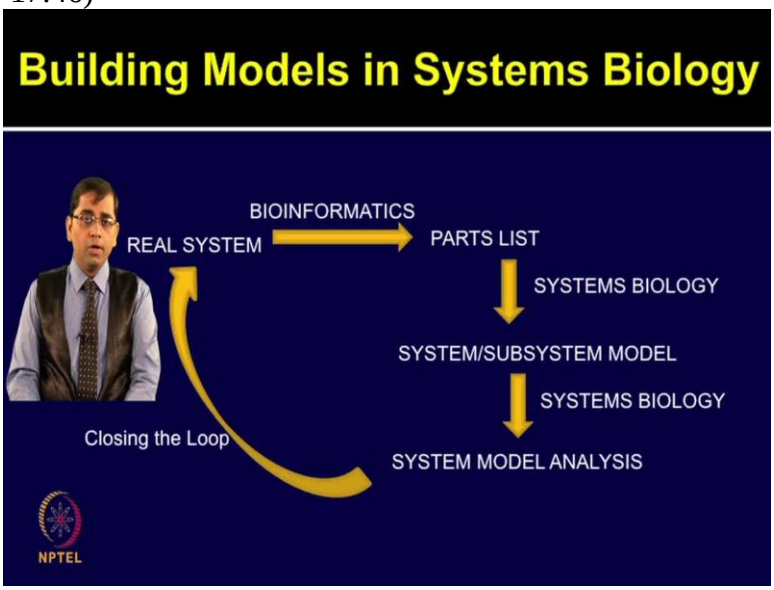
(Refer Slide Time: 16:56)



So as you can see here to generate the systems level information the systems study requires different technologies which could be employed at different levels in biological systems. At genome level by studying different type of technologies such as high throughput sequencing, high throughput arrays, transcriptomics, transcript analysis using RNA sequencing and microarrays proteome we have discussed many methodologies.

Metabolome could be performed using either by NMR or (())(17:30). And in phenome studying the images by using NMR method each level of this omic technologies could be useful for studying systems biology.

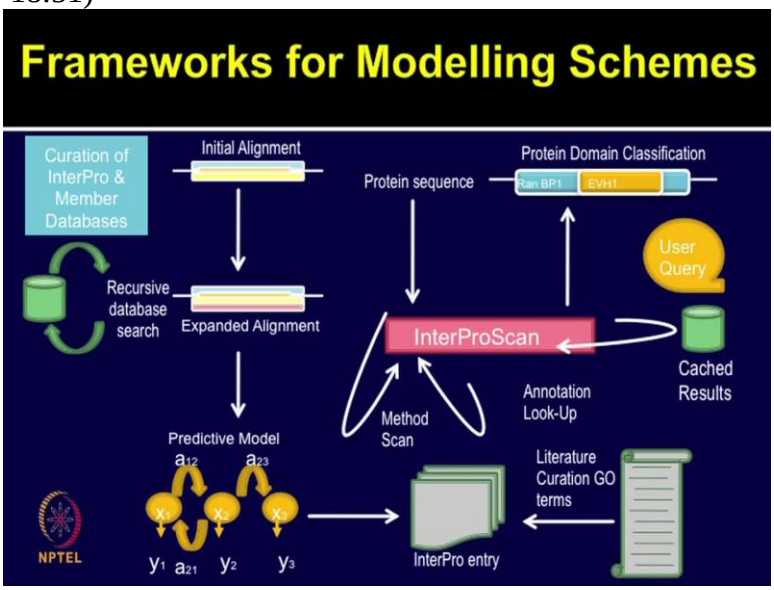
(Refer Slide Time: 17:46)



Let us now talk about how to model biological networks. To build a model in systems biology, first of all, all parts can be generated by the datasets derived from systems biology approach. This system or sub system model can be generated which could be used for systems model analysis. Now this could be applied for real systems and by applying knowledge using bioinformatics tool it could be applied again back to the original components which can be used to derive some hypothesis and validation of these datasets could be performed.

It will work like a closed loop. To build the models in systems biology information is generated at different level, 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.

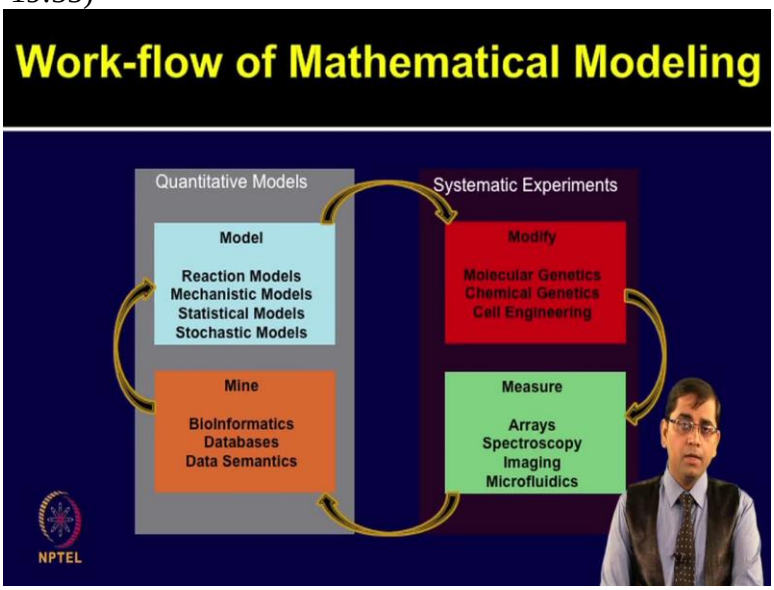
(Refer Slide Time: 18:51)



What are the frameworks required for modeling schemes? Different types of deterministic or stochastic models have been repost. The compartmental variables or individuals or function variables have been studied especially homogenous or especially explicit models are generated which could be applied in the uniform time scale or separate time scale.

This framework could involve single scale entities or cross scale entities. As you can see in this slide, this framework requires different level of information in very complex manner whether it is curation of the databases, how to align this information using bioinformatics tools to generate the predictive models which could be also developed by using the literature curated datasets or experimental datasets. And finally it could be used to study the system level properties.

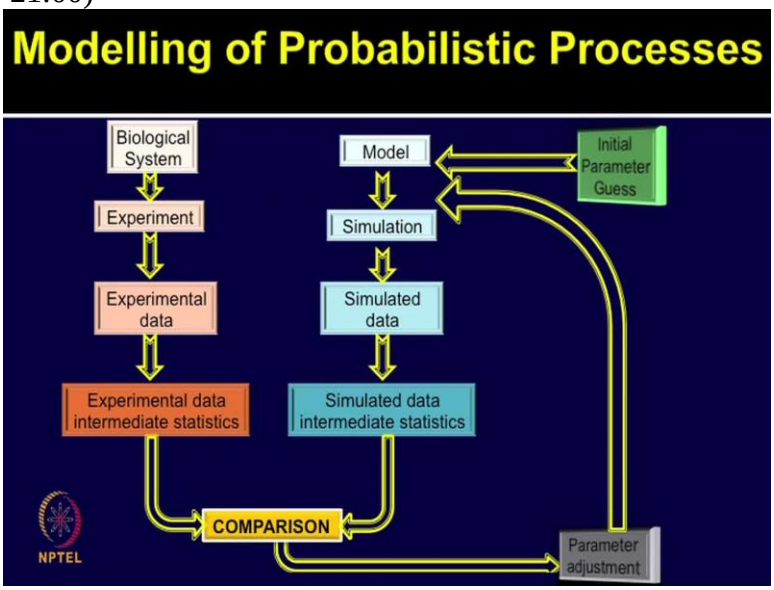
(Refer Slide Time: 19:55)



Let us discuss the work flow of mathematical modeling. A paradigm can be proposed based on modify, model, measure, mine. Systematic experiments different type of molecular genetics, chemical genetics and cell engineering approaches could be used to modify and different level of measurements by applying microarrays, spectroscopy imaging, microfluidics based approaches from proteomics and genomics which could be used further for mining which involves bioinformatics databases and data semantics.

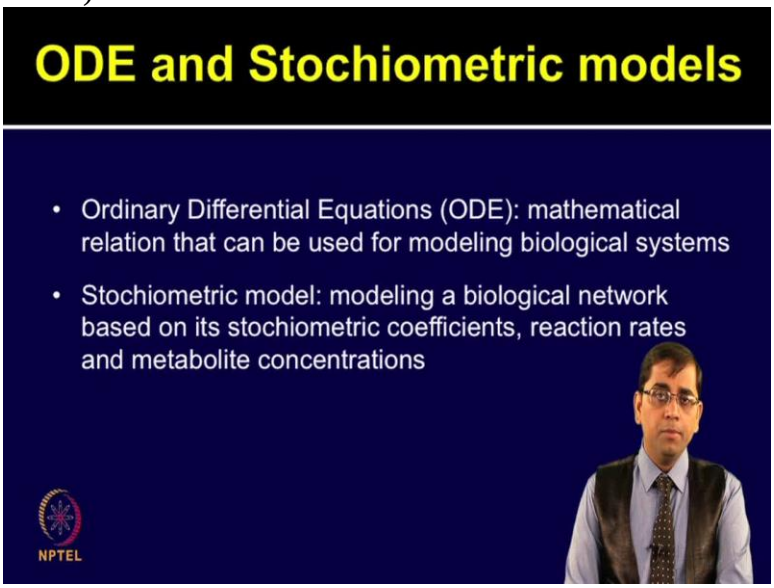
Now these datasets could be used to derive the model which could be reaction mechanistic, statistical or stochastic models. Staring from systematic experiments to reaching and deriving the quantitve models this work flow can be applied.

(Refer Slide Time: 21:00)



The modeling of probabilistic processes involves let us say you want to study a biological system, so some experiments has to performed. The experimental dataset will be generated from which some statistics could be applied which can be used for comparison. Now different type of models can be generated using simulations and simulation datasets which can be used for intermediate statistics. By comparing these two types of information and adjusting the parameters one can study the systems and derive probabilistic processes.

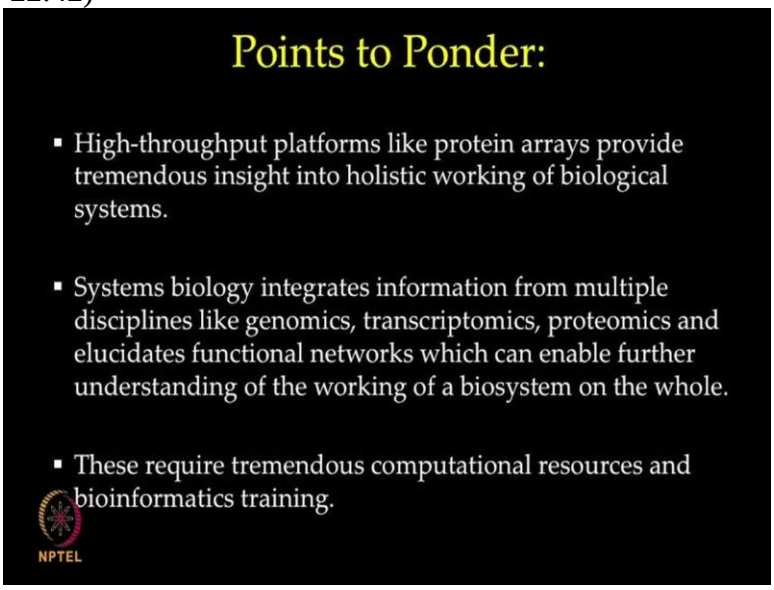
(Refer Slide Time: 21:43)

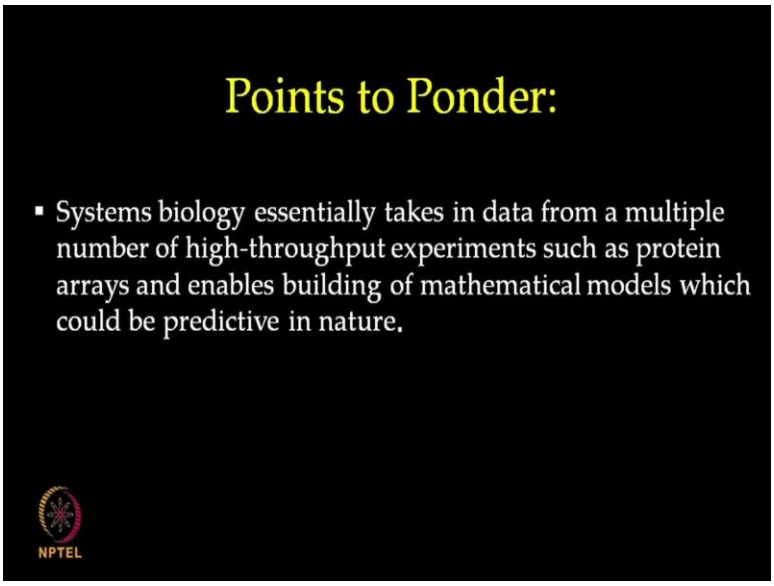


What are ordinary differential equations and stochiometric models? The quantitive analysis measures and names to make models for precise kinetic parameters of a systems network component. It also uses properties of network connectivity. ODE is a mathematical relation that can be used for modeling biological systems. The quantitive models mostly use ODE to link reactants and products concentrations through the reaction rate constant.

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

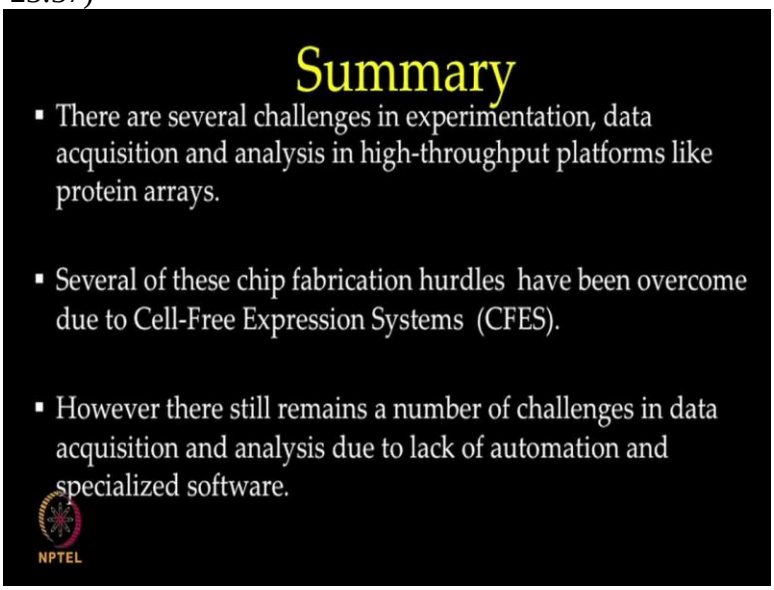
(Refer Slide Time: 22:42)

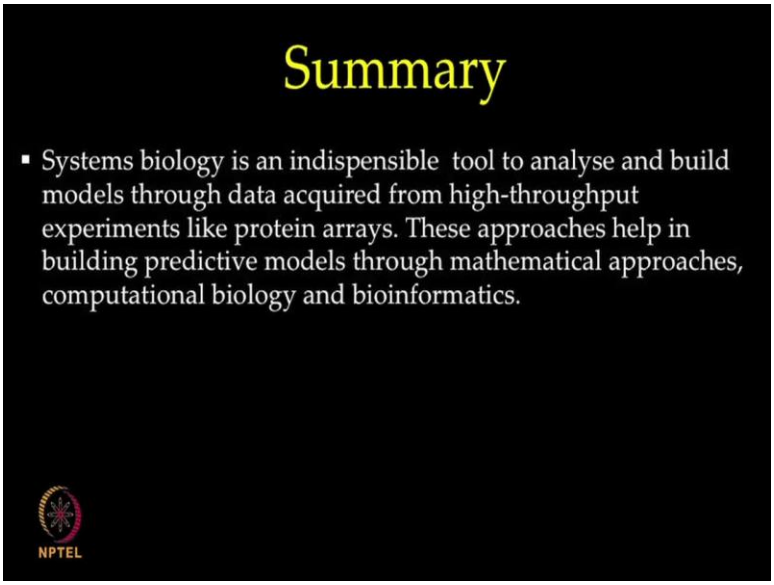




So in this lecture we understood the various challenges associated with the protein array work flow, you were introduced to the basic concept of systems biology which are helping new age researchers to integrate data from the multi omics resources and pure sciences to understand biological organism systematically. With this foundation concepts we will look at some advanced skills of systems biology in the next lecture, thank you.

(Refer Slide Time: 23:37)



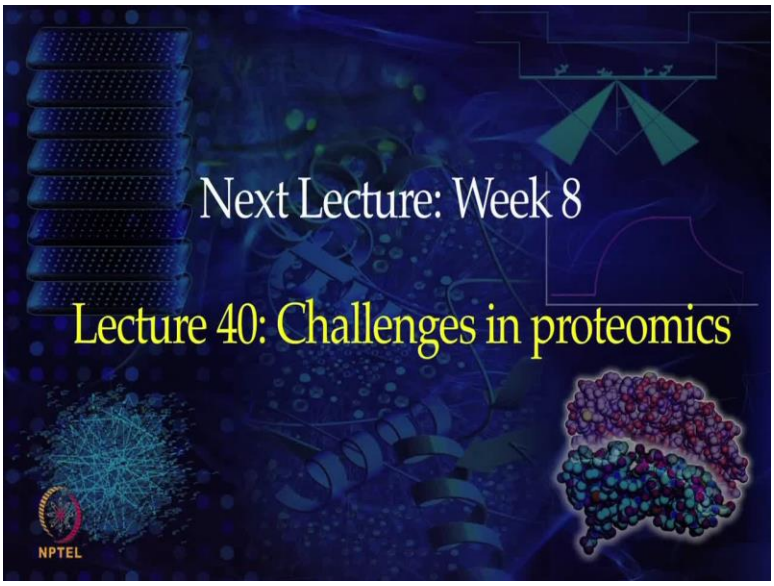


References

- Ideker, T. (2004). Systems biology 101--what you need to know. Nat. Biotechnol. 22, 473–475.
- Ideker, T., and Krogan, N.J. (2012). Differential network biology. Mol. Syst. Biol. 8, 565.
- Ideker, T., and Sharan, R. (2008). Protein networks in disease. Genome Res. 18, 644–652.
- Ideker, T., Thorsson, V., Ranish, J.A., Christmas, R., Buhler, J., Eng, J.K., Bumgarner, R., Goodlett, D.R., Aebersold, R., and Hood, L. (2001). Integrated genomic and proteomic analyses of a systematically perturbed metabolic network. Science 292, 929–934.

References

- Ideker, T., Winslow, L.R., and Lauffenburger, A.D. (2006). Bioengineering and systems biology. Ann. Biomed. Eng. 34, 257–264.
- Kitano, H. (2002). Systems biology: a brief overview. Science 295, 1662–1664.


Next Lecture: Week 8
Lecture 40: Challenges in proteomics
NPTEL