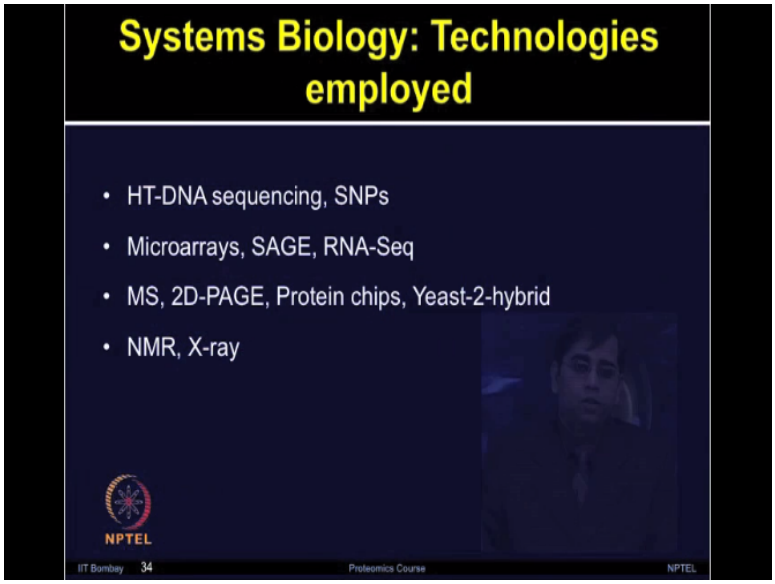
Introduction to Proteomics
Dr. Sanjeeva Srivastava
Department of Biosciences and Bioengineering
Indian Institute of Technology – Bombay

Lecture - 20
Systems biology and proteomics - II

So the different technologies which have been employed to study the systems biology. Obviously we read high throughput data sets which could be derived from microarray platforms or in a deep sequencing different configurations of mass spectrometry, different type of a structural Proteomic tools and protein interaction data sets.

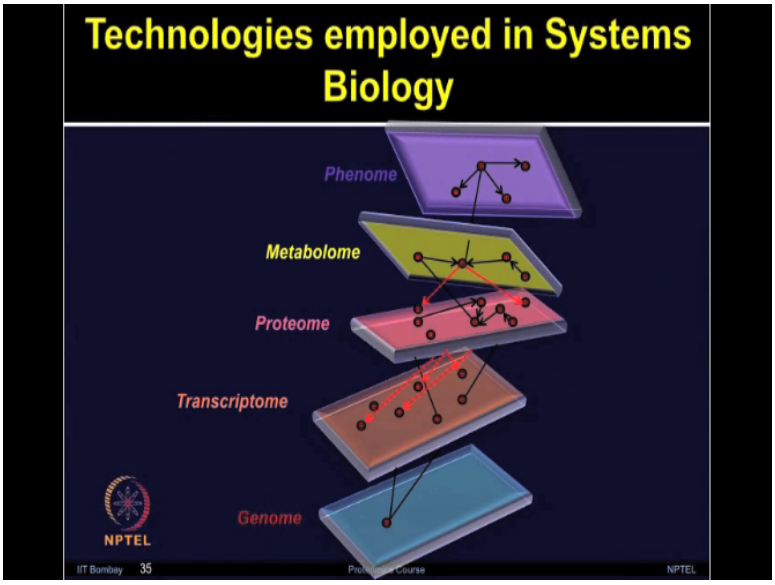
(Refer Slide Time: 00:51)



So some of the technologies which are commonly employed in systems biology can be classified broadly under the following techniques for genomics, the throughput DNA sequencing methodologies, mutation detection using SNP methods. For transcriptomics, the transcript measurement can include serial analysis of gene expression, SAGE, gene chips, microarrays and RNA sequencing.

For Proteomics, mass spectrometry, two dimensional electrophoresis, protein chips, yeast-2-hybrids. X-ray and MR is mainly employed for the metabolic analysis, the metabolomics.

(Refer Slide Time: 01:41)



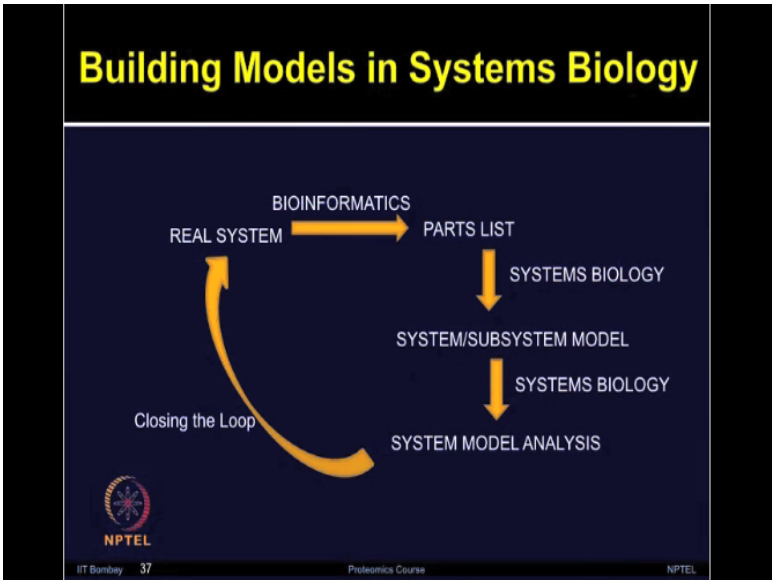
You can see here to generate the systems level information, the systems study requires different type of technologies which could be employed in the biological systems at genome level by studying different type of technologies using high throughput sequencing, high density arrays, transcriptomics, different type of transcriptome analysis using RNA sequencing and microarrays, proteome.

We discuss many methodologies metabolome could be either using NMR or mass spectrometry and then phenome which is studying about the images by using patch or NMR methods. So each level of these omic technologies can be useful for studying the systems biology.

(Refer Slide Time: 02:36)

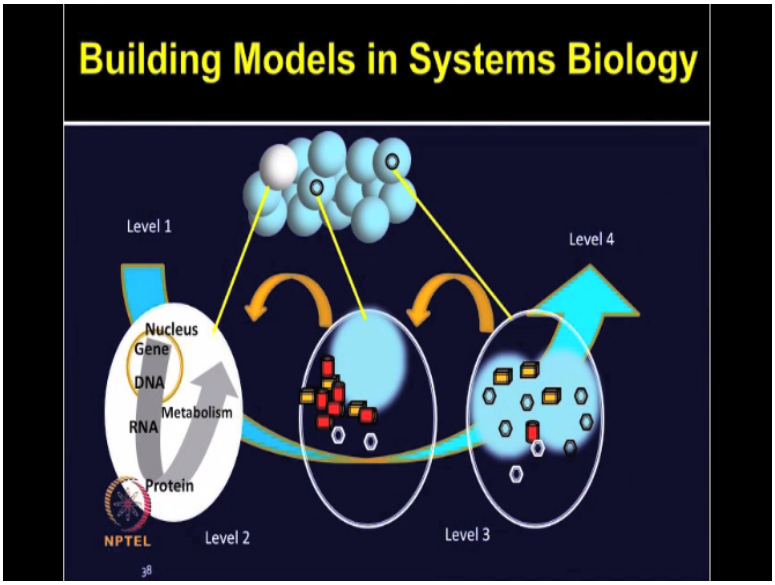
Lets now talk about how to model the biological networks. To build a model in systems biology, first of all the parse list can be generated by using datasets derived from the systems biology approaches. The systems or sub systems model can be generated which can be used for the systems model analysis.

(Refer Slide Time: 03:05)



Now this could be applied for the real systems and by applying the knowledge using bioinformatics tools it could be again applied back to the original components which could be used to derive some hypothesis and validation of the data sets. So it will work like a closed loop. To build the models in systems biology, information is generated at different levels.

(Refer Slide Time: 03:39)

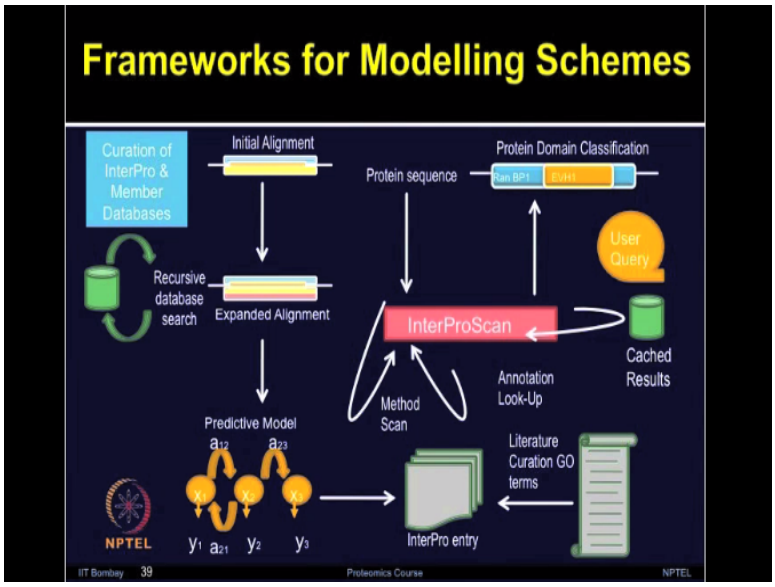


Level 1 such as DNA and gene expression. Level 2, the intra cellular networks. Level 3, cell-cell and trans membrane signals and level 4 integrated organ level information. What are the

frameworks required for the modeling schemes? Different type of deterministic or stochastic models have been proposed, the compartmental variables or individual or functional variables have been studied.

The specially homogeneous or specially explicit models are generated which could be applied in the uniform timescale or separated time scales. This framework could involve single scale entities or cross scale entities.

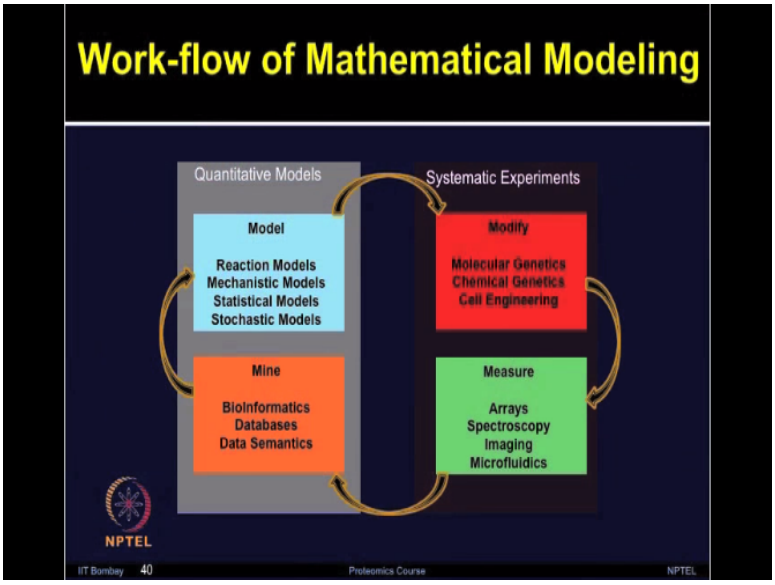
(Refer Slide Time: 04:32)



As you can see here, this framework requires different level of information in very complex manner. Whether it is curation of the databases. How to align these information using bio informatics tools to generate the predictive models which could be also developed by using the literature curated datasets or experimental data sets and finally it could be used to study the systems level properties. Let us discuss the work flow of mathematical modeling.

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

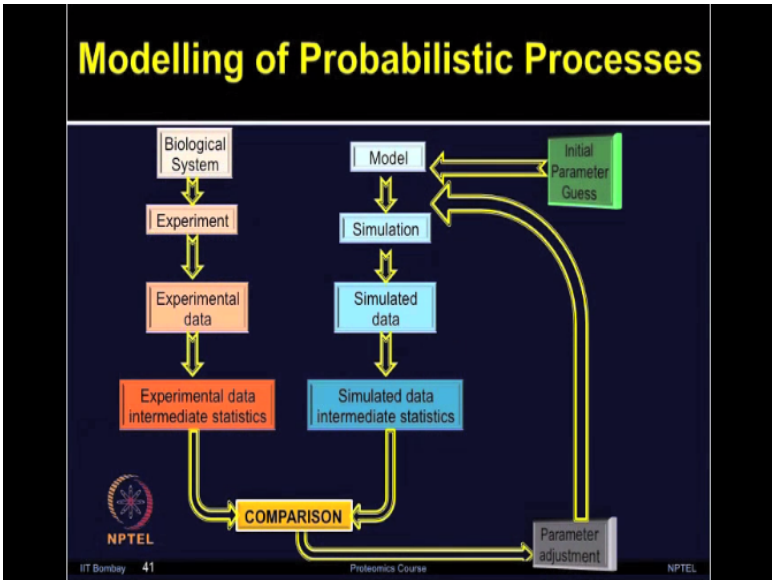
(Refer Slide Time: 05:21)



So systematic experiments different type of molecular genetics, chemical genetics and cell engineering approaches can be used for modifying and different level of measurements by a applying microarrays, spectroscopy, imaging and microfluidics based approaches from Proteomics and genomics can be used further for mining which involves bio informatics, data bases and data semantics.

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

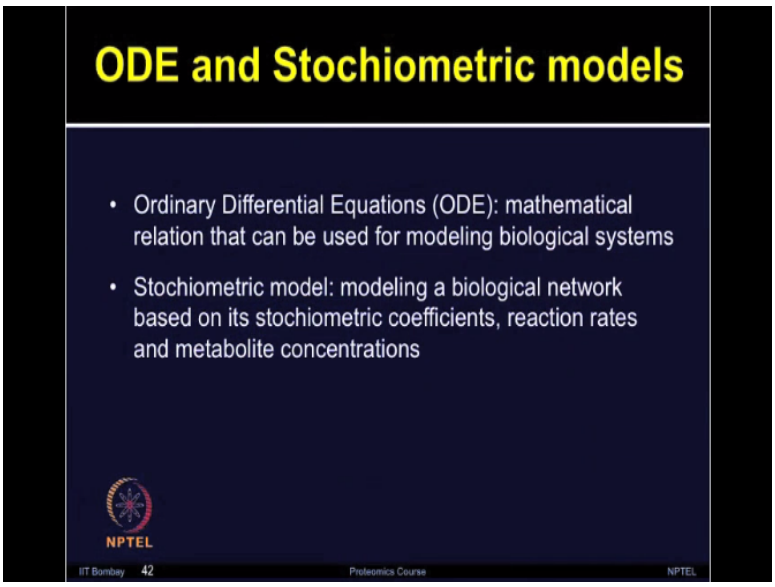
(Refer Slide Time: 06:17)



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

Now, different type of models can be generated by using simulations and simulation datasets which can be used for intermediate statistics and by comparing these two type of information and adjusting the parameters one can study the systems and derive the probabilistic processes.

(Refer Slide Time: 07:05)



So what is Ordinary differential equations and stochiometric models. The quantitative analysis measures a names to makes a models for precise kinetic parameters of the systems network component. It also uses the properties of network connectivity. The ODE is a mathematical relation that can be used for modeling the biological systems. The quantitative models mostly use ordinary differential equations or ODE to link the reactants and products concentration through the reactions, reaction rate constants.

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

(Refer Slide Time: 08:16)

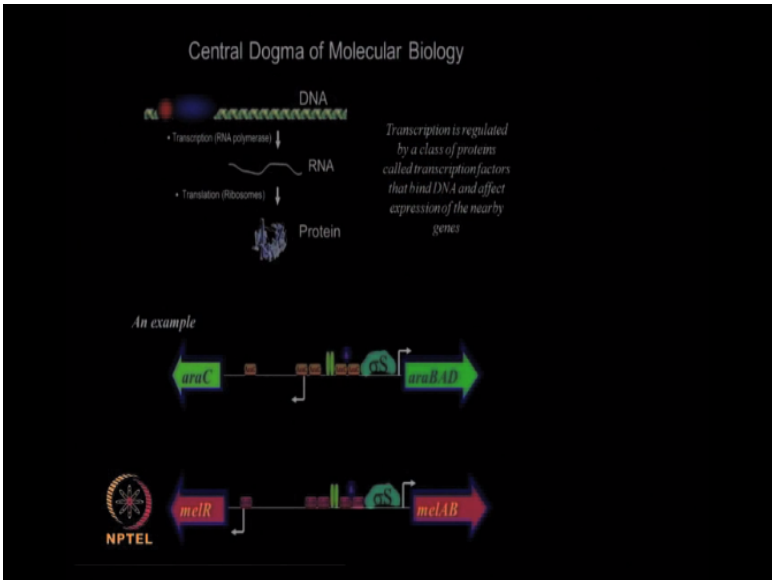
This is my pleasure to introduce Dr. Sarath Chandra Janga from Indiana University and Purdue University Indianapolis. He is in school of informatics and school of medicine. So we will be discussing about need for studying Proteomics and system biology. The lot of information available at the transcription and translation level and often there is not good correlation between RNA level and the protein level.

So today, it will be interesting to talk about systems approaches for studying biological networks from post transcriptional control towards the drug discovery. So I have invited Professor with, for having a discussion and short talk on this topic. Thank you Dr. Srivastava. It is my pleasure to be here to talk about some of the work that we have been doing and more generally the principles of regulation and how you can use systems approaches for understanding biological networks more generally.

As some of you might be familiar with the use of, the concept of networks is increasingly becoming prominent in not just Proteomics but also in genomics data and all kinds of high throughput data. So today what we will be talking about is some basic introduction to the application of networks in a biological systems and how it can be applied to understanding transcription regulation, post transcription regulation and as well as to the Proteomics data and at large, how this can be used to understand the drug discovery.

How it can be applied to the drug discovery concept. According to the central dogma of molecule biology, DNA gives rise to RNA through the process of transcription.

(Refer Slide Time: 10:22)



And this processes facilitated by the binding of the RNA polymerase as well as a number of other transcription factors which bind to the upstream regions of the DNA as you can see and control the expression. And RNA can give rise to protein through the process of translation. And this happens through the process of translation with the help of the ribosomes.

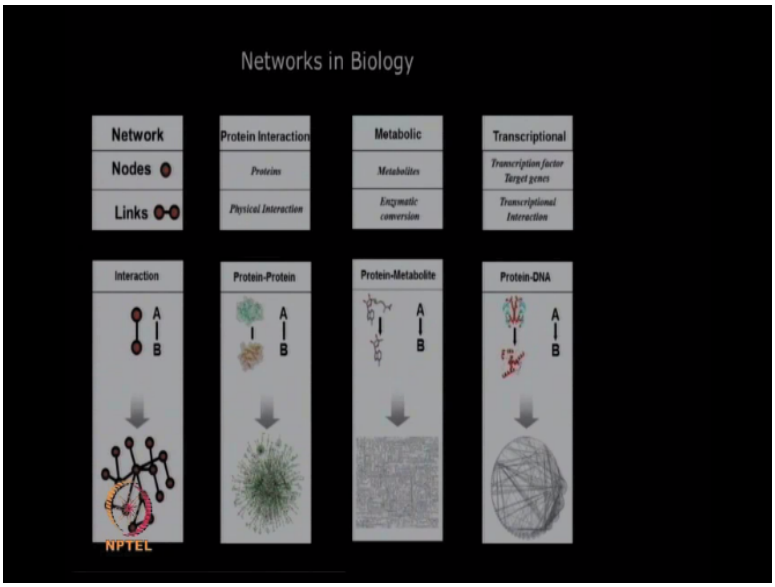
Now in this process the proteins which are produced, some of them can be classified as transcription factors which bind to the DNA and some others are classified as RNA binding proteins which can bind to the RNA and control the expression at the post transcription level as oppose to at the transcription level where transcription factors bind to the DNA. Now as an example, let us see, that case of araC transcription factor in a bacterial genome such as Ecoli.

This particular transcription factor binds to the upstream regions of araBAD operon which encodes for the enzyme and the transporter responsible for the uptake of arabinose from the environment. Now, the transcription factor araC, not only binds to the upstream of araBAD, but it can also bind to itself and control the expression. As you can see from the small orange boxes which are shown as the representation for the binding sides.

Now what decides is, transcription factor can auto regulate bind and regulate its own expression or it can also bind to other genes controlling their expression. There are also many cases actually where multiple transcription factors bind to the upstream regions. As you can see in this case, its represented with the orange box as well as the blue box, a blue circle where other transcription factors bind.

Now, in addition to this binding of this transcription factors, as I mentioned earlier, RNA polymerase also binds shown with the green box, a green circle, green whole box of them, so that they can control the expression. Now, there are other examples also shown in this figure with a melR regulator also doing something similar.

(Refer Slide Time: 12:29)



Now, this is what we just discussed is an idea of how regulation happens from a biological view point. Now, an increasing thing, increasing amount of literature now supports the idea of networks in biology. So what exactly are networks? A network simply is represented as nodes and links are edges. These nodes can be biological entities and the links are edges are actually the associations between these entities.

Now there are number of ways you can talk about the nodes or the entities. So one common form of representation are protein interaction networks where the proteins form the nodes and the physical interaction between these proteins forms the edge as you can see in the figure below. You can have a representation of these networks in a fashion that is shown in this figure below. Now, an alternate kind of network which is also studied in the literature over the last 10 years or so are metabolic networks.

In metabolic networks, the metabolites form the nodes and the conversion of one metabolite to the other forms the edge in this case. Now as you can imagine, the conversion f one metabolite to the other is actually facilitated by the enzyme. So the particular protein enzyme converts a metabolite A to B and when you look at on a global scale and when you are

looking at the conversion of number of metabolites one to the another and sometimes one metabolite can give rise to more than one set of metabolites.

Such complex set of associations can be called as a metabolic network. Now the third kind of networks which I will be elaborating in more detail in the next slides are transcriptional networks. In transcriptional networks, transcription factors form one set of nodes and the target genes form other set of nodes.

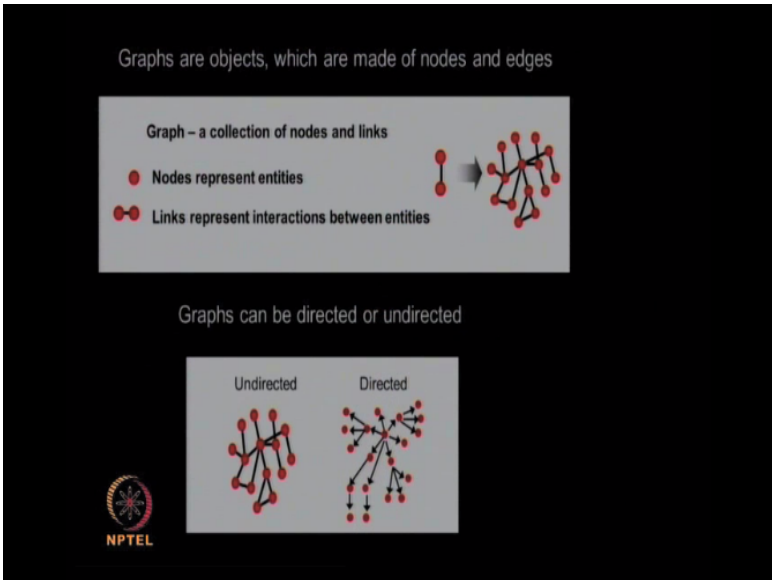
So as you can imagine, what you are actually looking at in this case from a biological view point is the interaction of the transcription factor with the DNA and controlling of the expression of the downstream genes but in the context of networks, what we are showing is the transcription factor and the target gene or operon whose expression is controlled. Again in this case, you can see that the protein A which is a transcription factor controls B but it may or may not be that B is the transcription factor and it also controls A.

So that might be a case to case specific and it may or may not be having a reciprocal interaction. As we just discussed, these networks are actually, the concept of networks has been borrowed from physics and computer science where often these kind of networks are referred to as graphs and graphs are objects which are collection of nodes and entities. The nodes are representing the entities.

It could be these entities could be genes, proteins, small molecules, cells, organs or at any level you can represent these entities. The interactions are associations between them or the links. Now as I am just mentioned there are different kinds of networks that protein-protein interaction networks, metabolic networks, transcriptional networks.

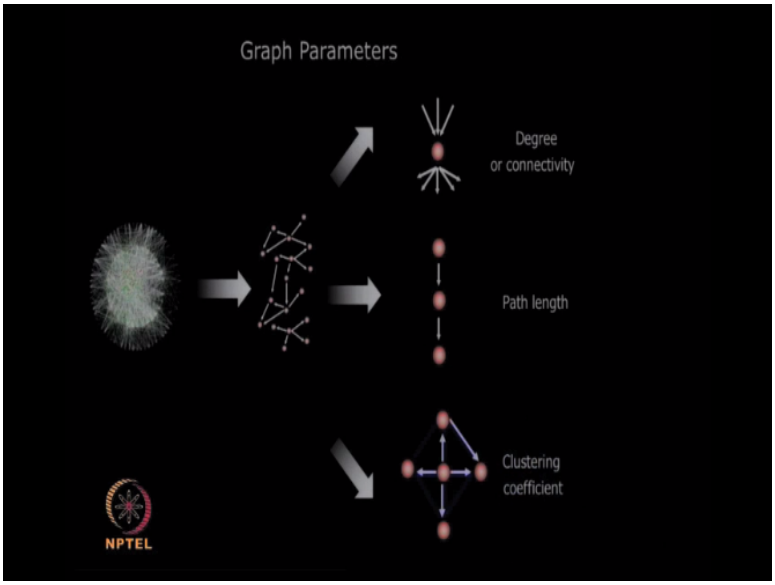
In the case of protein-protein interaction networks, what we are looking at often is no directionality in such interactions and these are called as undirected networks.

(Refer Slide Time: 15:44)



However, there are also directed networks such as transcriptional networks or metabolic networks. In these cases, there is a flow of information IE where A controls B, which should mean A is controlling, A is regulating B. So, therefore there is a directionality and these are often commonly studied as regulatory networks and we will be talking in more detail in the next slides.

(Refer Slide Time: 16:08)

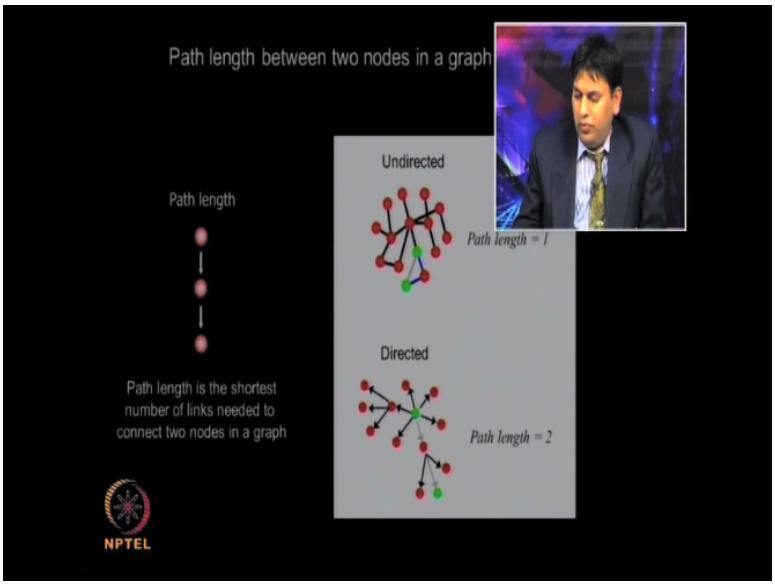


However, before we get into the more specific observations about the properties of this networks. One set of common properties which are studied when you are looking at biological networks are degree, path length and clustering coefficient. Now often when you look into a network as such, you do not have a clear understanding of the properties of the different nodes.

But when you look into the specific aspect such as shown in this case as degree, what it tells you is how many connections or particular gene protein or node has in your network. So what we can say from the first example on the top is the degree of the node is 8, that means it is connected to 8 other proteins. On the second property is the path length, what its showing in this case, if you cal ease that the number of edges that you need to travel from one node to the other.

So, if I ask you what is the path length between that first and the bottom node in this figure, you would say the path length is = 2. The third kind of property which often studied is the clustering coefficient, clustering coefficient tells how often the neighbors of a given node are connected to what you would see in a completely connected graph. Let us look at a more detailed examples. For instance, if you are studying the degree of a node.

(Refer Slide Time: 17:28)



In the case of an undirected network such as in the example shown on the top, the florescent node that is shown in florescent color has a degree = 2. On the other hand, a directed node examples shown at the bottom has a degree = 4 because It is called connected to 4 other nodes. However, what you can also say is, here is an in degree and out degree. And in degree is the number of incoming connections of a particular node.

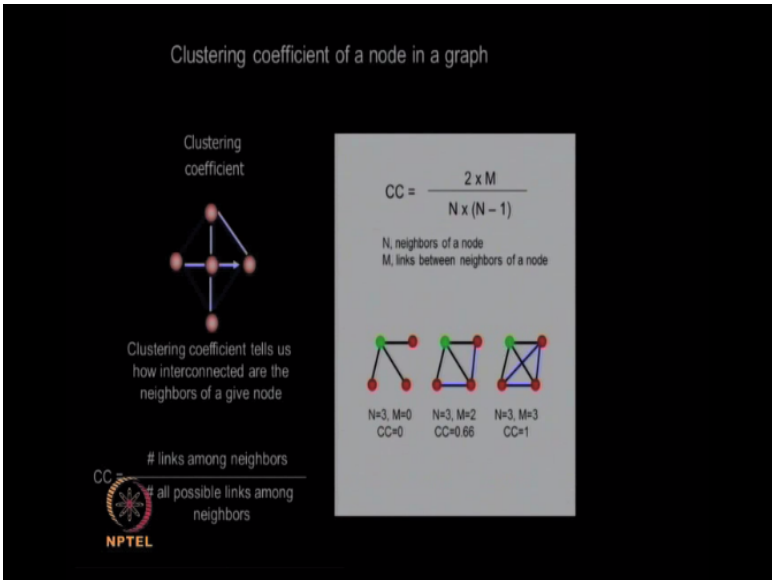
So the green or florescent node here, has an in degree of 1. It also has an out degree = 3, because it is directing 3 other nodes shown in red color out there. So It is out degree = 3. Now, you can also extend this idea of un directed and directed graphs and ask what is the

path length of a node. Now as I mentioned, the path length is referred to as the number of edges that one need to travel between two different nodes that you are interested.

In the top network that you are seeing, the path line between the two green or florescent nodes is = 2 as well as = 1. Because, the path that you can take can be different than the shortest path that you are looking at. However, almost often unless otherwise specified when you are talking about the path length between two nodes, It is a shortest path length. So the 2 florescent nodes have a path length = 1.

However, if you are asked what are all the path lengths, you would say, it has 2 different paths, one with a path length of 1 the other with the path length of 2. In the un directed networks, your definition of path length essentially does not change. So in the example that you see at the bottom, the path length between the two florescent nodes is = 2.

(Refer Slide Time: 19:14)



The other property that I was referring to previously is the clustering coefficient of a node. And clustering coefficient refers to the number of connections between the neighbors of a given node of interest to what you would see in a completely connected graph. Now, let us look at an example in this figure that you see, there is that in the first example the florescent node, the green node has 3 connections, 3 red dots are connected to it.

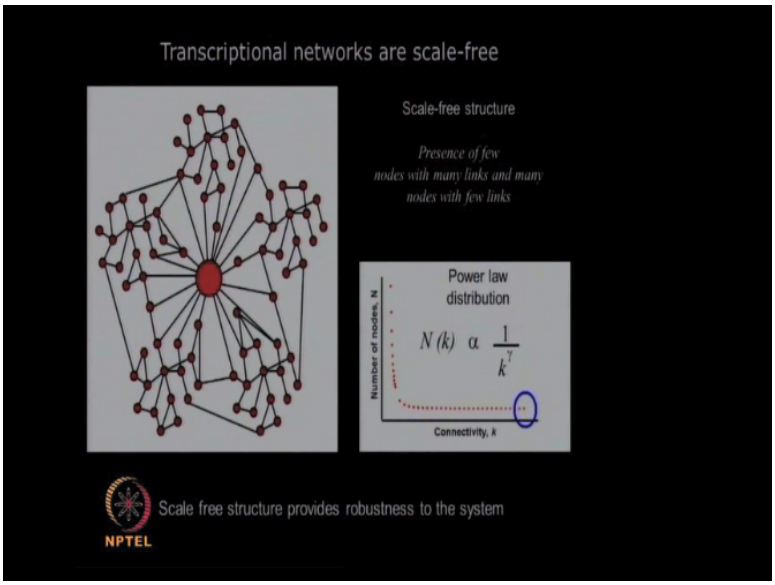
However, if you ask the number of connections between the red dots, it is 0. There are 0 connections between the red dots. But if they were fully connected, you would see that they will have 3 edges between them. So the clustering coefficient of the florescent node right now

is 0 upon 3. Let us look at the second toy network. In the second toy network, the florescent node has a clustering coefficient of 2 upon 3.

In the third case, the clustering coefficient of the florescent node is 3 upon 3, which is completely connected. So the clustering coefficient is = 1. Now, more generally formula can be brought up and it can be written as if there are M number of interactions between the neighbors of a node of interest and there are N number of neighbors of a given node of interest, then it can be written as M upon N* N - 1 / 2.

So that would be defined as the clustering coefficient of that particular node. And when average the clustering coefficient of a node on a whole network scale, it gives you an essence of modularity of the network. The higher the average clustering coefficient, the more likely is the network cluster, it can be decomposed into specific modules. Another property that is of great interest in understanding biological networks is a scale free structures.

(Refer Slide Time: 21:04)



And lot of biological networks are documented and shown to be scale free, transcriptional networks are also documented to be scale free structures. So what exactly are scale free networks? Scale free networks are corresponded to the structure of a network where there are free nodes which are highly connected for instance in the figure, in the network figure that you see to the left, there is a big red dot, big red node which is highly connected.

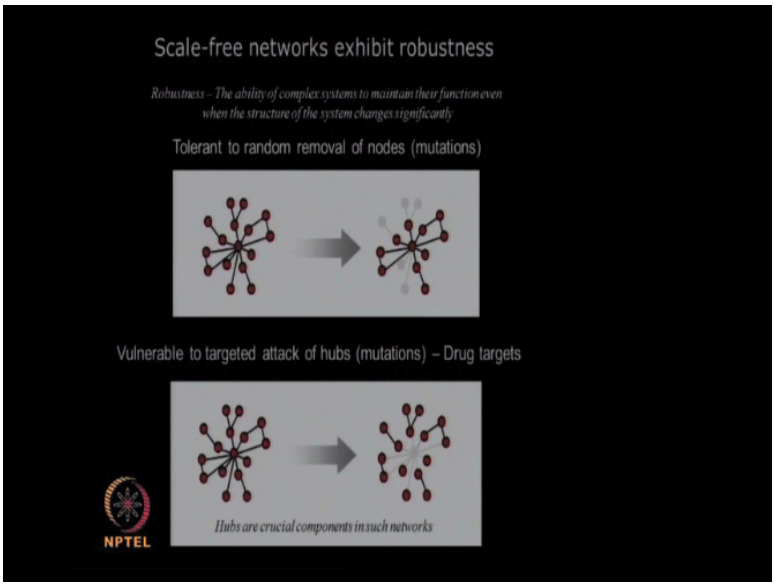
So there are not many such highly connected nodes. And there are many nodes which are very poorly connected. So in other words, a scale free structure refers to network structure

where there are few nodes which are highly connected and most nodes are poorly connected. Or more mathematically, if you plot the connectivity of a node versus the number of nodes with a given connectivity you should see a power law distribution or otherwise, if you plot the log law plot of the connectivity versus the number of nodes with a given connectivity.

You should see a negative slope of gamma as shown in this figure where gamma lies between 2 to 3. That is when you can call the structure to be scale free and the distribution to be a power law distribution. Now, what is so special about this scale free structure? Scale free structures have been postulated to provide robustness to the biological system. Now, what exactly is robustness?

So robustness is the ability of a complex system, a complex system such as a biological system to maintain its function even when the structure of the system changes significantly. Now, let us look at an example.

(Refer Slide Time: 22:42)

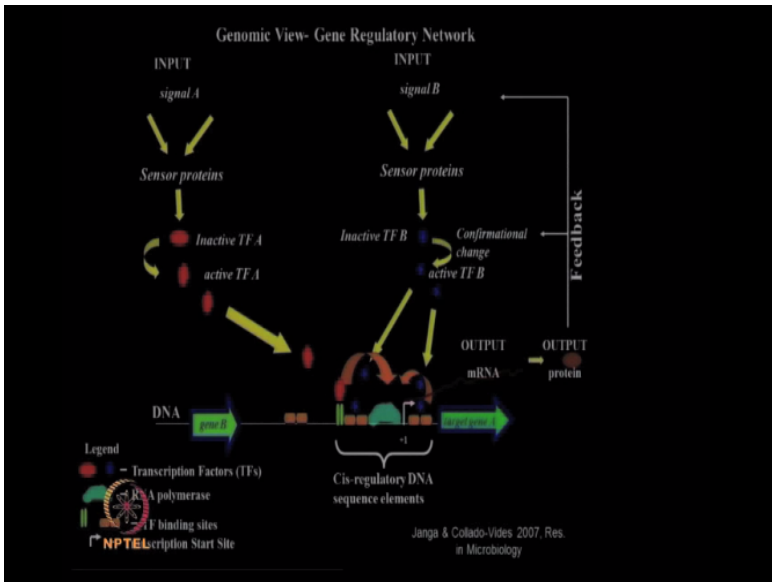


So, in the network figure that you see, if you randomly perturb any of these nodes, you are likely to affect a small fraction of the network. However, if you target the highly connected node, that is the central node which is highly connected, you are going to disrupt a major fraction of this network suggesting that, these highly connected nodes can be vulnerable to be the targets.

So if you are trying to incubate the growth of a pathogen, you are likely to target this highly connected nodes because you are more likely to be able to crumble the biological system of

the pathogen. So and this has been increasingly gaining attention as a method of targeting a drugs to these classes of proteins.

(Refer Slide Time: 23:26)



So as mentioned earlier, we have been talking about regulation of a single transcription factor. But, in the context of networks regulation is much more complex and what we are referring to is a combinatorial regulation by many different transcription factors. Let us look at a specific scenario. So the slide shown here shows a typical regulatory system in a bacterial organism. What you usually have is a set of signals, which are sensed by the cell and these signals are sensed by sensor proteins.

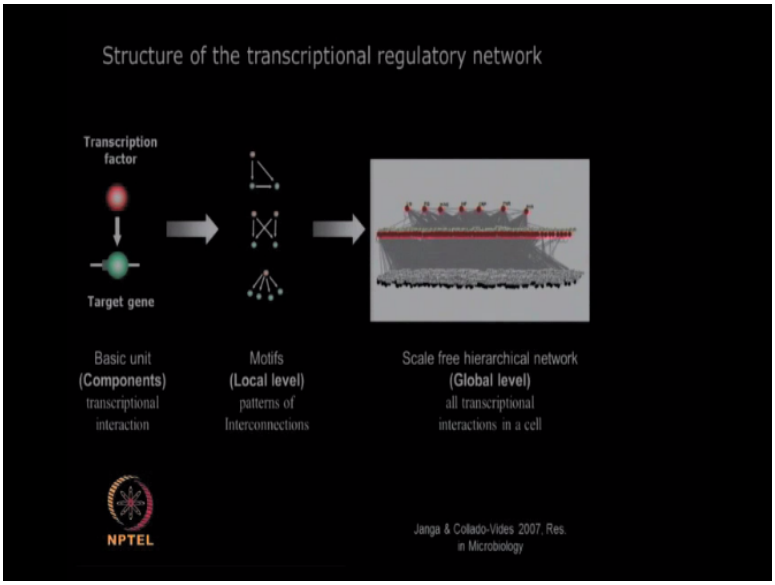
These sensor proteins could be transporters or this could also be listed in (()) (24:06). And once these sensor proteins sense the signals from the exterior or even sometimes interior of the cell. They can cascade the information to transcription factors. The transcription factors upon receiving the signals can change from active to inactive or inactive to active state. And when this happens, because of multiple sensor proteins these transcription factors can change the conformation and bind to the upstream regions.

And shown at the bottom, is a stretch of DNA where these transcription factors can bind in a combinatorial fashion often and control the expression of the target gene or operon. As the rule of thumb, if transcription factors bind to the upstream regions, in the upstream of transcription start side shown as + 1, That is where the transcription actually starts. You often are simulating the polymerase and enhancing the expression.

However, when you bind to the downstream of transcription start side, you typically repress the expression of the target gene there by blocking the transcription by the polymerase shown in the oval shaped polymerase symbol in green. So based on these principles and together with the interplay with the transcription factors and the polymerase, your transcript is produced. And once transcript is produced, you can have MRNA and protein levels regulation which is not what we will be talking immediately now.

But all these levels together contribute to provide feedback and this is typically a system or simple regulatory system that you encounter in bacterial organisms. But more complex systems, more complex eukaryotic gene regulation is much more complex and beyond the scope of our current discussion.

(Refer Slide Time: 25:50)



As discussed in the previous slides, the basic unit of regulation is a transcription factor and a target gene whose expression is being controlled. Now on a different scale, if you increase, or if you put together all the set of regulatory events between transcription factors and the target genes are operons, you construct a global transcription regulatory network. And as I mentioned earlier, this network is a scale free structure, scale free network.

But in addition to these, It is also hierarchical structure where in what we are actually referring to in a hierarchical structure is, there are set of transcription factors, which are able to regulate a large number of genes, and there are set of genes, other transcription factors which are also controlled by this global transcription factors shown at the top of this network

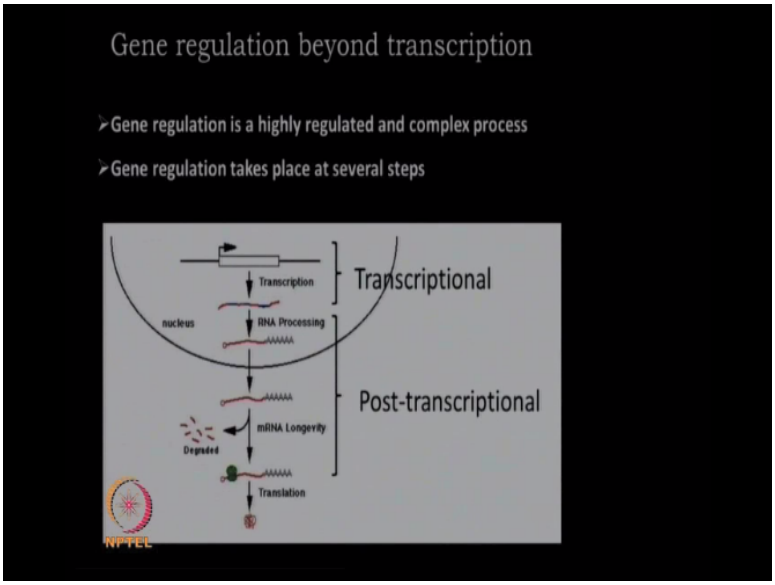
structure. And both the top layer and the second layer, all of them together regulate the set of genes which are not essentially encoding for the transcription factors.

So in a way there are transcription factors which are at the top of this system, there are transcription factors which are controlled by this top layer and there are subsequent layers and the number of layers in such a hierarchical structure depends on the complexity of the system. Now, in between the top and the bottom layer, in between the left most figure of the basic unit and the right most figure.

There are set of sub structures or sub graphs within the regulatory network which we call as motifs. Motifs are the set of sub graphs which occur more often than expected by chance and there were 3 kinds of regulatory motifs that are identified in regulatory networks. One is the feet forward loop where there are two transcription factors, the first transcription factor regulates the other 2 genes.

The second transcription factor regulates the target gene. The second kind of motif is multiple input module where there are 2 different transcription factors, both of them regulate 2 different target genes. The third is a single input module where a single transcription factor regulates a set of target genes. Now each of these set of regulatory motifs have been shown to a specific functions and which would be beyond the scope of our current discussion.

(Refer Slide Time: 28:07)

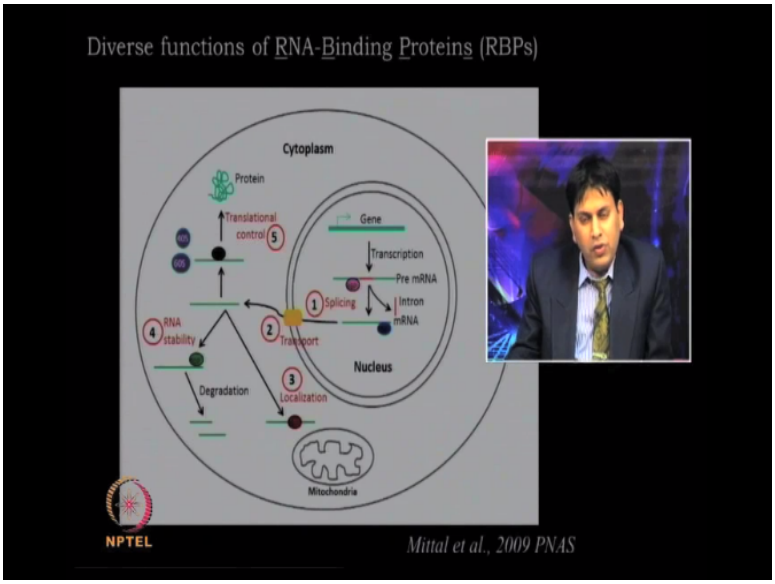


Now although there are idea of regulation of a gene expression, the level of transcription has been documented for several years and we have extensive understanding, very little is known

about the regulation of gene expression beyond transcription. And it is only been recently being appreciated about the role of regulation at the post transcription level. Now most of this evidence for the reason why post transcription regulation is becoming important, It is coming from the lack of correlation between MRNA and protein pools in model systems.

Now, there is now enough evidence to suggest that this post transcriptional processors are actually controlled by a class of proteins called RNA binding proteins. Among non protein coding components such as micro RNAs, non coding RNAs. So RNA binding proteins are now known to be involved in controlling the RNA processing, RNA longevity as well as in the translation.

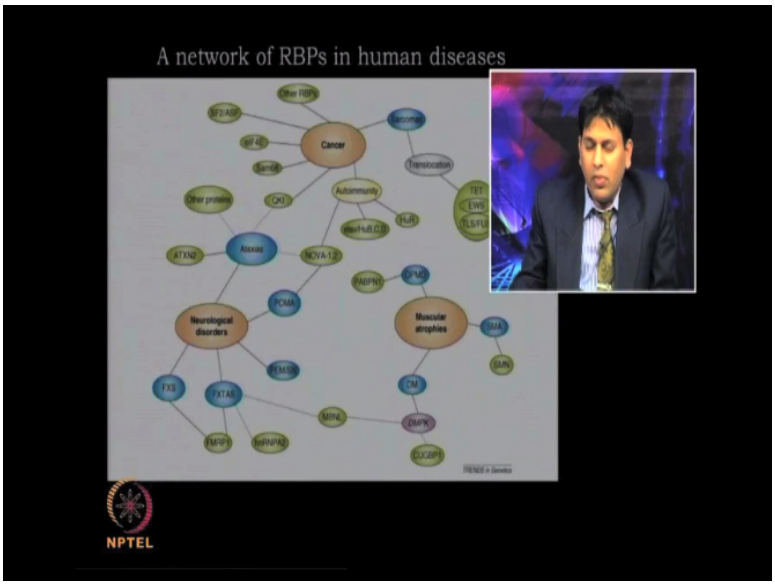
(Refer Slide Time: 29:06)



Now, in particular as soon as a gene is transcribed and pre MRNA is produced, splicing associated RNA binding proteins bind to the pre MRNA and convert into mature MRNA by splicing of the intrans. Now the produced RNA not necessarily only MRNA needs to be exported from the nucleus into the cytoplasm and this is carried out by a class of RNA binding proteins, which can be termed as transport RNA binding proteins shown with number 2 in the figure.

RNA binding proteins have also been implicated in the specific sub cellular localization of these transcripts. RNA binding proteins are documented also in controlling the stability of the transcripts there by promoting or degrading the expression of these transcripts. As expected, RNA binding proteins, a number of them are associated with the ribosomal proteins to control the regulation of expression at the translational level.

(Refer Slide Time: 30:05)



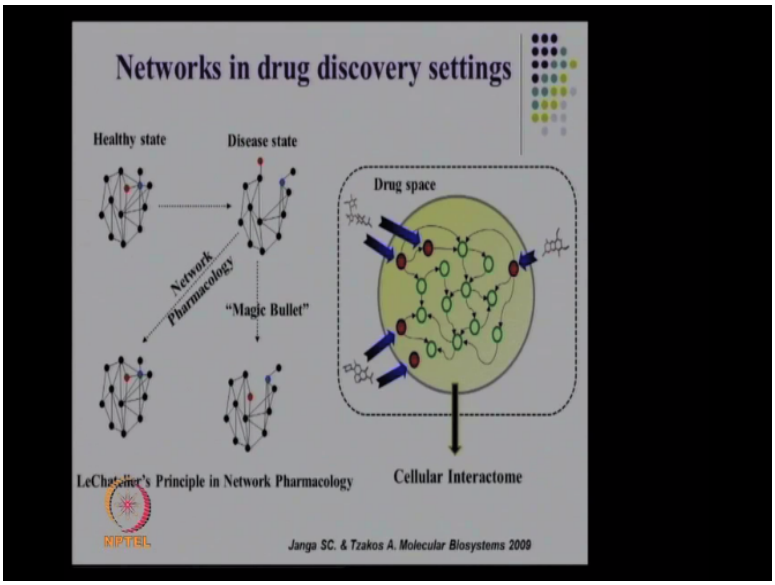
Now, another aspect of understanding regulation in the post transcription level is that number of RNA binding proteins are involved in major class of human diseases such as cancer, muscular atrophies and neurological disorders. In this network diagram shown here, the major class of diseases are shown in orange, while the sub types of diseases which can be sub classified are shown in blue and the specific RNA binding proteins which are been documented are implicated in these disorders are shown in green.

Now, let us take it a specific example of a muscular atrophy called myotonic dystrophy. In this particular kind of disorder, a CUG RIP it binding protein called CUGBP1 binds to the 3 prime on translator region of a DM protein kinase and because of the sequence of this CUG RIP it binding protein on to the trinucleotide repeat expansion in the three prime untranslated regions. This particular disease phenotype is observed.

Another example of mis-regulation of an RNA binding protein happens in OPMB, which is another kind of muscular atrophy. In this particular kind of disease, there is a GCD repeat expansion in the xon1 of an RNA binding protein which is a poly binding protein called PABPN1. Another example we can observe which is heavily documented in the literature is a brain specific splicing factor called nova whose mis-expression is known to cause a disease called POMA, which is a subtype of neurological diseases.

So what I am trying to add over it here is that if there is a change in expression of either RNA binding protein or any of its targets, it can be associated to a disease phenotype.

(Refer Slide Time: 31:56)



And all these studies basically suggest that it is not just the effect of a single gene or protein, its rather a combination of different set of genes and proteins which contributes to a disease phenotype. Now, while this observation is not very new while we knew that this is common for a number of complex diseases, what we have still being not able to achieve is, the able to cure diseases for this complex diseases.

Now let me introduce to you the tradition notion of how drug discovery is usually happening in most places. Let us represent the healthy state of an individual with a network of interactions shown in this figure on to the left. Now a disease state could be studied as a perturbation such a network where some of these nodes are actually not properly connected compared to the healthy state.

Now according to the idea of Paul Ehrlich others, the magic bullet approach suggests that, the conversion of disease state to the healthy state should involve one or most likely one particular drug which is non promiscuous and specific to a particular drug target. So that you have minimal half target effects. Now often such magic bullet approach can only yield only a semi recovery from the disease state.

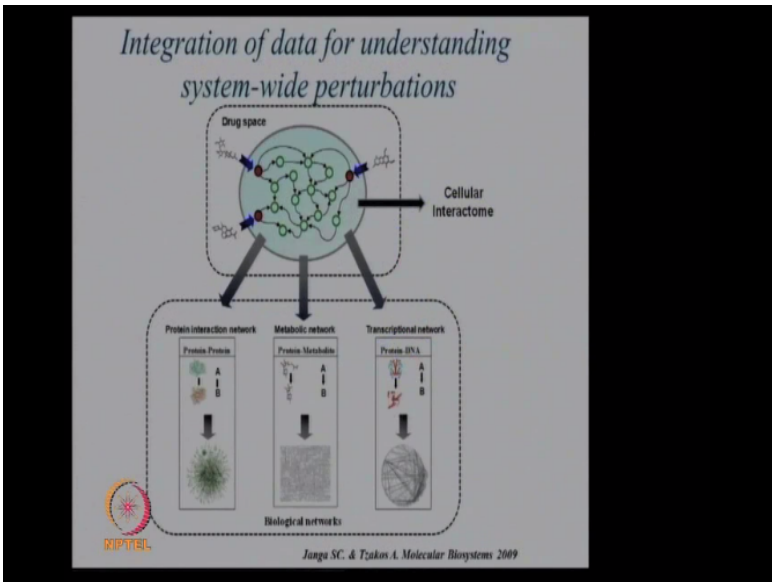
Now what network pharmacology are network medicine approaches are trying to arrive at is use a combination of perhaps promiscuous drugs but which do not cause negative side effects, which do not cause side effects with a lethal and can still convert the disease state

into healthy state as close as it is to the original one. Now how would you achieve such an approach?

To understand this particular idea, let us look at a network representation of how the different entities in the cell are interacting. In the figure to the right, you can see that, a number of drugs each of them can be perturbing different nodes. Now all of these nodes are actually interconnected to each other because we are looking into the cellular context and there are protein-protein interactions, there are metabolic interactions, there are also regulatory interactions.

Perturbation of 1 cannot be seen in isolation. It has to be seen in the context of other perturbations. Now a combination of these perturbations is going to yield a phenotype which we hope can be treating the complex disease. That is the concept behind this idea of network pharmacology.

(Refer Slide Time: 34:24)

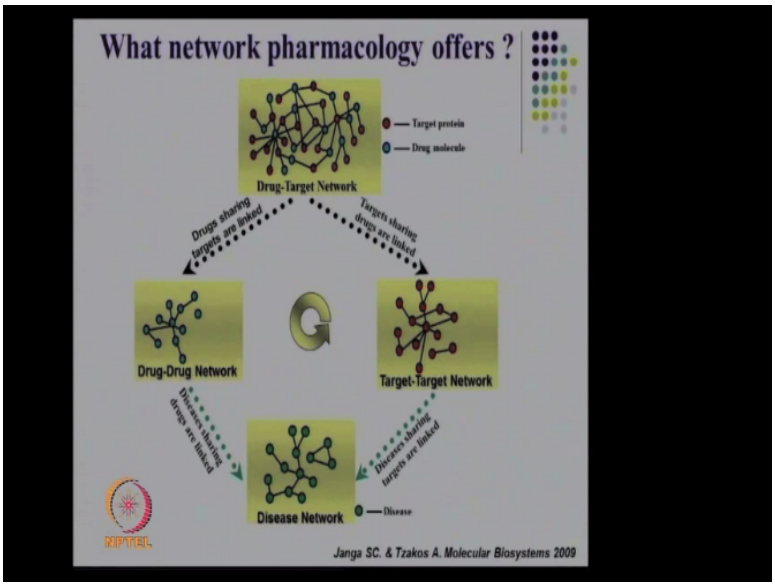


Now, how do we achieve such bigger goal? So usually when you have such kind of a complex phenotype, you have to put together data such as knowledge on the current metabolic network in the human genome, knowledge on the transcription network, knowledge in the protein-protein interaction network and knowledge on the post transcription network.

And together with the current knowledge of the drugs and the targets and the target pathways, one can start looking at how these perturbations can be studied in the context of specific

diseases and what particular drugs can be used to identify potential new therapies for existing diseases.

(Refer Slide Time: 35:04)

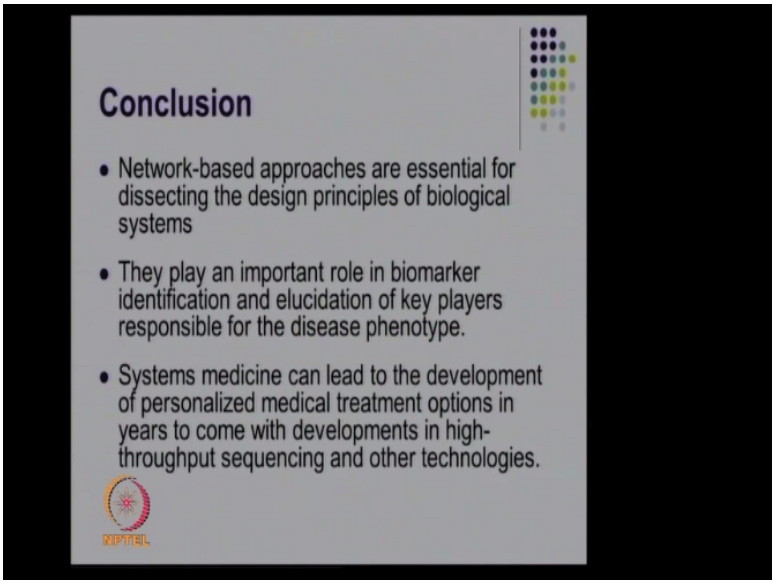


An alternative set of approaches which are being used in the context of drug discovery is that, if you have a target, a drug target network for all the approved drugs in the literature, one can start understanding what are the drugs which are sharing the targets. Can we use the drugs which share the targets as alternatives to existing our drugs, if there is an resistance acquired for a particular drug.

Can you comprehend the current drug with another drug which is having the same set of targets or one can start studying the set of drug relations if their drugs are sharing the targets, can we start studying what other profiles of the two drugs which are linked. Are they similar in the structure, are they similar in the final phenotypes or what are the common principles of these drugs which are connected to each other.

Likewise, one can also study disease-disease associations by linking any pair of drugs which are used for the same disease. Likewise, one can study target-target network to construct a disease-disease association network. So these are some of the ideas which where the field is moving to understand or to even repurpose existing drugs for novel therapies.

(Refer Slide Time: 36:18)

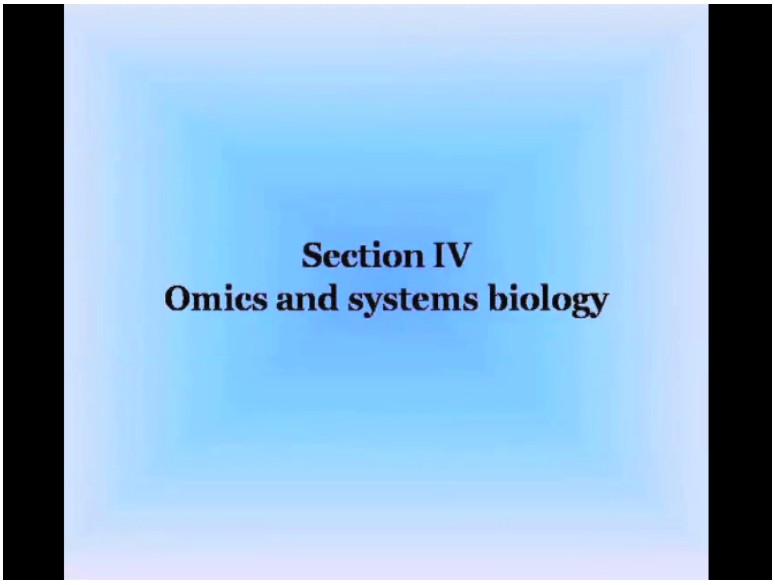


So to conclude what we have tried to cover in the past set of slides is that the network based approaches are essential and a powerful paradigm for dissecting the design principles of biological systems. They play an important role in biomarker identification and even in the elucidation of key players responsible for the disease phenotype.

Systems medicine can lead to the development of personalized medical treatment options in years to come with developments in high throughput sequencing and other technologies which can yield a lot of data in a very short time so that clinical relevance can be achieved based on these kind of a techniques, application of these network based approaches in the context of clinical settings.

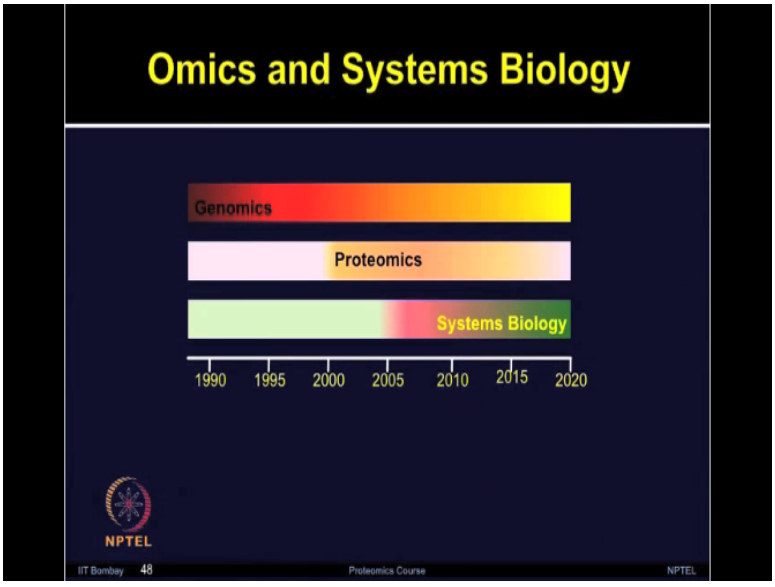
Thank you very much Sarath for giving very nice talk and giving some of the basic concepts as well as illustrating how a systems level network studies can be employed for little various type of problems including in the drug discovery as well as in pharmacology and it could be extended for even biomarker discovery and many other applications. So thank you very much. Thank you.

(Refer Slide Time: 37:28)



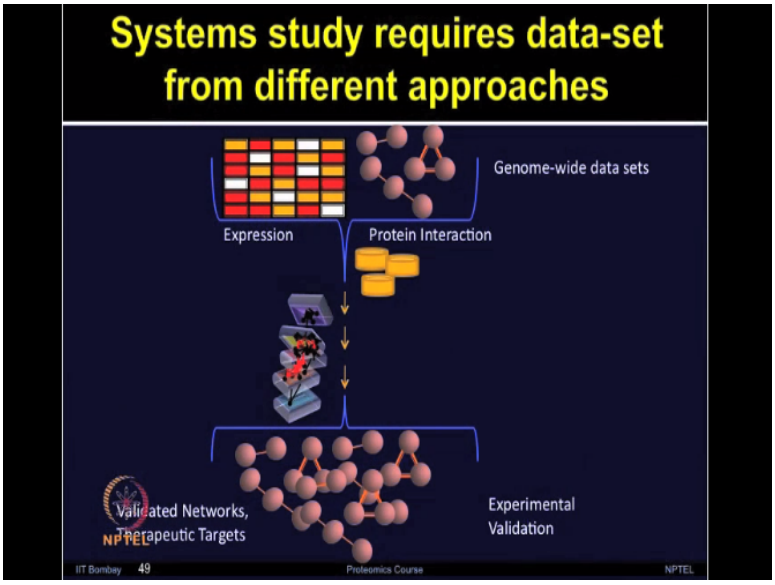
Now let us try to integrate the omics approaches with systems biology. So genome sequencing projects in genomics era from 1990s to 2000 accelerated the phase of omics research.

(Refer Slide Time: 37:49)



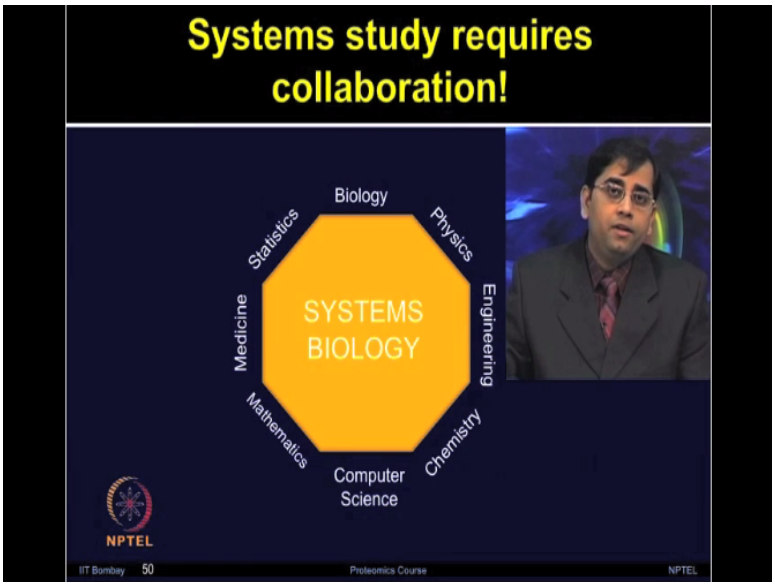
Then from 2000 onwards, Proteomics field also got accelerated and new methodologies, new tools came into the place for studying the proteome. And the data derived from genomics, transcriptomics, Proteomics, metabolomics and other omics approaches has now brought the integration of the datasets in the systems biology field. The systems study requires obtaining datasets from different approaches and analyzing them. For example, as shown in the slide.

(Refer Slide Time: 38:27)



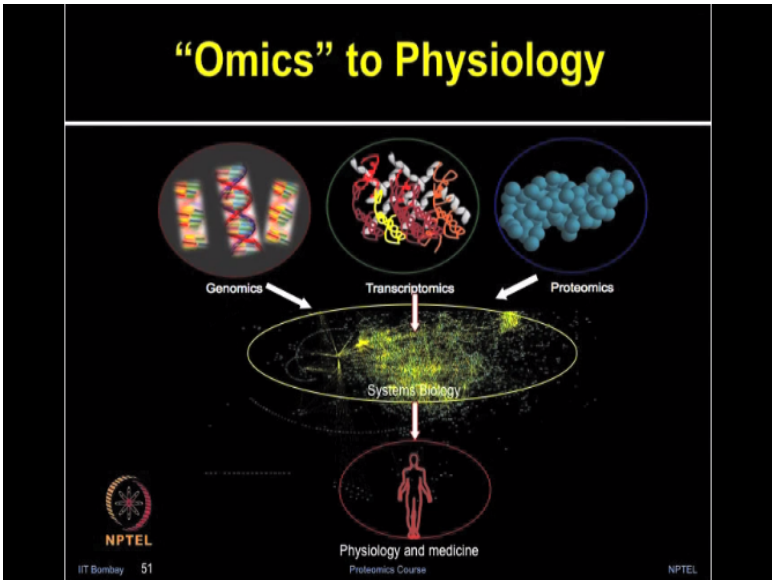
The genome wide datasets can be derived at the genome level and looking at the expression of the different transcripts or at the proteome level looking at different type of protein interactions. These datasets can be stored in the clinical databases and also it can be mined from the literature, literature 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: 39:09)



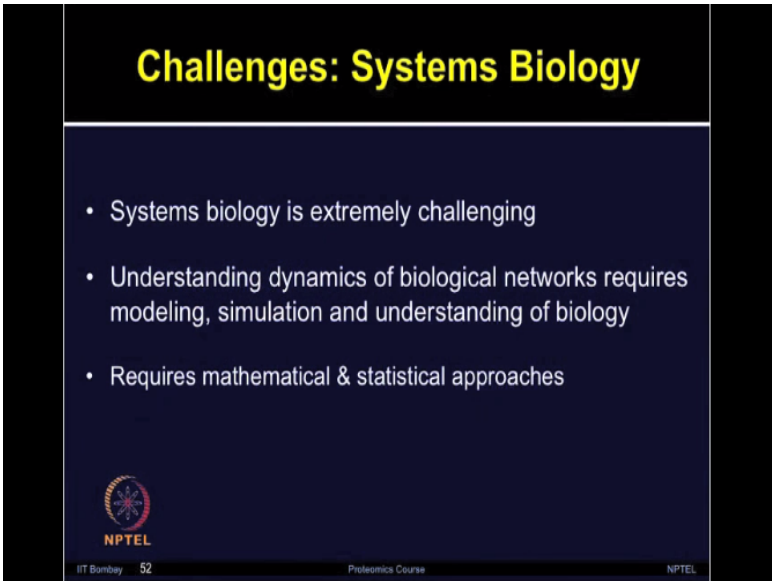
A study in systems cannot be done in isolation in individual labs. It requires different expertise and collaborations from scientists from different dispense of biology, physics, engineering, chemistry, computer science, mathematics, medicine, statistics and many more.

(Refer Slide Time: 39:32)



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 quantitative the models and then propose and use it for the understanding the physiology and apply that in medicine. So this omics to physiology, this flow can be well maintained by employing systems biology tools.

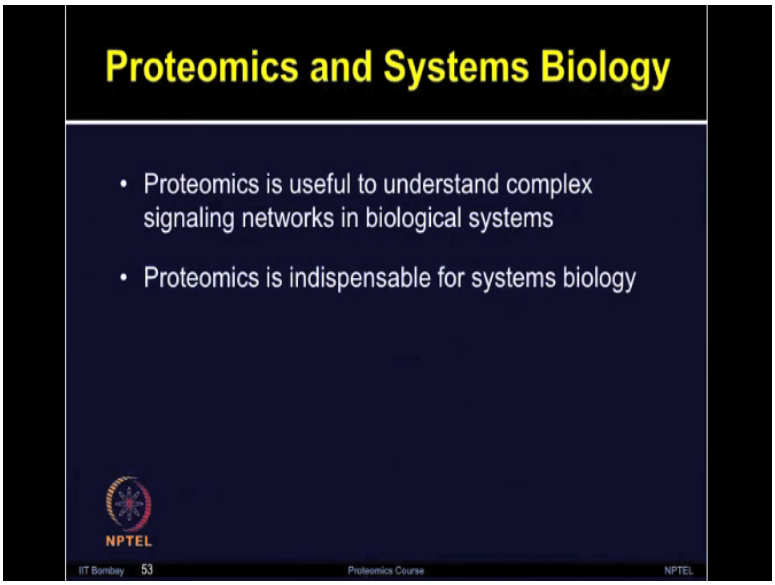
(Refer Slide Time: 40:11)



What are the challenges of systems biology? Systems biology is extremely challenging, the emphasis 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 gene 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.

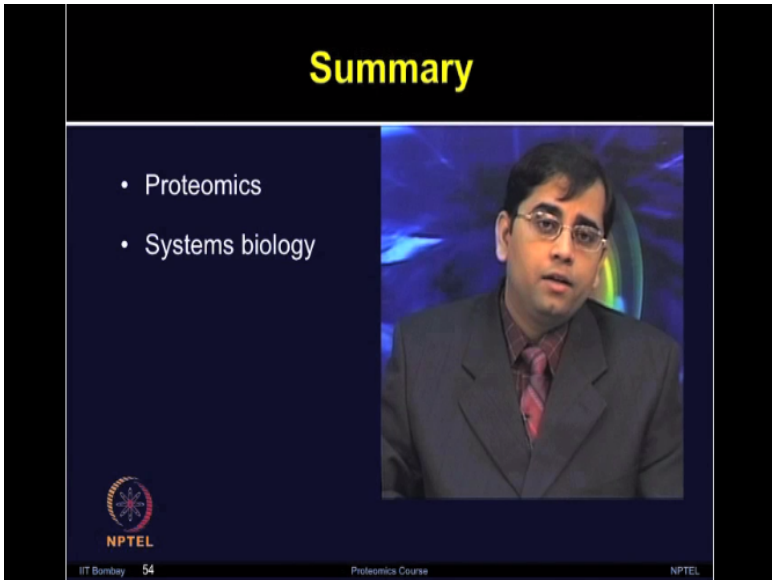
(Refer Slide Time: 41:07)



So how Proteomics and systems biology are integrated? Proteomics as we have studied, it is useful to understand the complex signaling networks in biological systems, it is very indispensable tool for systems biology. The global analysis of Proteomic 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 structural 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: 42:05)



So in summary, today we discussed that how omic 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.

So systems biology approach characterizes and predicts these dynamic properties of biological networks. Now the next module, we will focus in more detail different type of Proteomic technologies. Thank you.

(Refer Slide Time: 43:02)