

# NPTEL VIDEO COURSE – PROTEOMICS

## PROF. SANJEEVA SRIVASTAVA

---

### HANDOUT

### LECTURE-04

### PROTEOMICS AND SYSTEMS BIOLOGY

#### Slide 1

Today we will talk about proteomics and systems biology.

#### Slide 2

The lecture outline

- First I will start on proteomics
- then we will move on to systems biology.

#### Slide 3

##### **Omics platforms.**

- So 'omic' is a suffix derived from the Greek work 'ome' meaning all or every.
- Omics is used as a suffix which has enabled the explosion of terms genomics, transcriptomics, proteomics and metabolomics and so on.
- The omics also implies an integration of biology with information science and conveys large scale integrated information about biological systems.
- So as you can see in this slide, if you are studying about DNA that will be in totality known as genomics, RNA study is transcriptomics, proteins in proteomics and metabolites in metabolomics. If you are looking at all of the similar 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.

#### Slide 4

So let's first start with proteomics.

#### Slide 5

##### **Proteomics**

- So if you recall from the previous lecture, proteome is the set of all the proteins expressed by a genome, proteomics is thus the study of full set of proteins and their properties to provide an integrated view of cellular processes.

# NPTEL VIDEO COURSE – PROTEOMICS

## PROF. SANJEEVA SRIVASTAVA

---

- What are these different properties? These properties include the extent of protein expression, how different type of post-transcriptional modifications occurs in the cell, different type of enzyme regulation whether it is activation or inactivation and then different type of protein-protein interaction.
- The current goals of proteomics are very broad, including the diverse properties of proteins, such as the side chain chemistry of amino acids and different levels of protein structures including sequence, quantity, the state of modification, activity, interactions of proteins with other proteins and other molecules and sub-cellular distributions and structure analysis.

### Slide 6

- The Central dogma concept was discussed earlier which is an orderly and unidirectional flow of information encoded in base sequences of cells passed on from DNA to RNA and then to proteins- this is simplest definition of central dogma.
- The genome sequence projects have provided researchers with unprecedented amount of information of genome sequences. However there are numerous proteins which can be encoded by the genome, therefore the analysis of static genome by doing sequencing alone is not sufficient. A gene can code for several type of proteins because of its alternative splicing and post-transcriptional and post-translational modifications therefore it suggests that studying proteins is more challenging than genome or transcriptome therefore proteomics has great significance to understand biological systems.
- The completion of genome sequencing projects of several organisms including human has been one of the most remarkable achievements of this century however these have not been sufficient to unravel the mystery of complex biological processes. The similar gene numbers of many diverse groups of organisms has failed to explain their varying biological complexity. A more meaningful understanding of biological function can be obtained through the characterization of products of gene expression, the protein which serves as the ultimate effector molecule of biological systems.
- The proteome refers to the 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. Proteomics aims to decipher structure and function of all the proteins in a given cell under specific conditions and to obtain a global view of cellular processes at the protein level. Analysis of the proteome involves protein extraction, separation, identification and finally characterization.

# NPTTEL VIDEO COURSE – PROTEOMICS

## PROF. SANJEEVA SRIVASTAVA

---

### Slide 7

#### Various Proteomics technologies available

- There are various proteomic techniques which are currently being applied in different applications. We will talk in detail about different proteomic technologies in the subsequent modules of the course but briefly very broadly you can group these technologies as gel-based methods, gel-free MS based methods, Mass spectrometry techniques, techniques which study protein interactions and structural proteomics.
- As I have shown some abbreviations in this slide, broadly different type of gel based proteomics such as SDS-PAGE, 2-DE, DIGE, BN-PAGE as well as different type of staining methods such as coomassie, silver, fluorescent dyes, phosphor-stains such as Pro-Q diamond and multiplex staining methods all these can be grouped under gel-based methods.

Gel-free MS based methods includes

- SILAC which is stable isotope labeling by amino acids and cell culture,
- CDIT- culture derived isotope tags,
- ICAT- isotope coded affinity tags,
- VICAT- visible isotope coded affinity tagging,
- MCAT- mass coded affinity tagging,
- QUEST is quantitation using enhanced signal tags,
- ITRAQ- isobaric tagging for relative and absolute tagging,
- GIST- global internal standard technology,
- ICPL- isotope coded protein labeling,
- AQUA- absolute quantitation,
- SISCAPA- stable isotope standards captured by antipeptide-antibodies,
- COFRADIC- combined fractional diagonal chromatography
- MudPIT which is multi-dimensional protein identification technology.

All of these are various new advancements in the gel-free methodologies.

- Mass spectrometry which is central to the proteomic application include 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

# NPTTEL VIDEO COURSE – PROTEOMICS

## PROF. SANJEEVA SRIVASTAVA

---

different type of tandem MS-based systems are also used. The surface enhanced laser desorption ionization-time of flight- SELDI-TOF is also used for various clinical applications.

- The protein interaction methodologies include immune-precipitation, yeast two hybrid method 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 based on labeled using fluorescence, chemiluminescence or radioactivity or it could be different type of label-free methods such as SPR, interferometry based methods or different types of conductance based methods employing nanotubes and nanowires.
- The structural proteomics it involves X-ray crystallography, Nuclear magnetic resonance- NMR, transverse relaxation optimization spectroscopy-TROSY, circular dichroism-CD, and different type of microscopy methods including atomic force microscopy and electron microscopy.

### Slide 8

- So as we have seen there are large numbers of proteomic technologies which are currently available for various applications.
- Many times, to ask just one biological question, different type of methodologies come together and then provide solutions to that problem. For example, for looking at some clinical samples for identification of biomarkers of a disease, one can employ the samples such as tissue or blood or different type of body fluid and then either directly extract the protein and subject to the mass spectrometry or first resolve on 2-DE followed by identification on mass spectrometry or take these samples and directly apply on the microarray based platforms and then detect using label-based or label-free methodologies. Eventually these types of results will enhance knowledge for monitoring the therapy response as well as identification of early disease diagnosis. This is just one example similarly multiple types of proteomic technologies can be used for different applications.
- There are several proteomics techniques which are employed for studying these proteins such as 2-DE, mass spectrometry, protein microarrays as well as some label-free detection techniques such as surface plasmon resonance.

# NPTEL VIDEO COURSE – PROTEOMICS

## PROF. SANJEEVA SRIVASTAVA

---

### Slide 9

#### Systems biology.

- So what is systems biology? It is a method to understand biological systems or a tool to solve practical problems.

### Slide 10

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

### Slide 11

- Systems level understanding require integration of information from different levels- as you can see DNA to RNA to proteins form a system 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 component of the system. The common elements of systems biology include networks, modeling, computing and dynamic properties.

### Slide 12

- Systems biology does not investigate individual biomolecules such as DNA, RNA, proteins or look at only cells, networks or individual properties but rather it investigates the behavior and relationship of all of the elements in a particular biological system while it is functioning. These data sets can be integrated and finally graphically visualized and ultimately it can be modeled using computational tools.

### Slide 13

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

# NPTEL VIDEO COURSE – PROTEOMICS

## PROF. SANJEEVA SRIVASTAVA

---

### Slide 14

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

### Slide 15

- 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.
- Studies can be done intercellular or intracellular level and finally at the molecular level involving genomics, transcriptomics and proteomics.

### Slide 16

#### 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, to understand even the simpler system of a cell, how it is regulated with the extracellular space and the cytoplasm and different other components, an examination of structure and dynamics of cellular and organismal function is very much required.

### Slide 17

#### Aims of systems biology

- To understand the biology in a holistic approach rather than the reductionist approach. The systems biology aims to quantitate the qualitative biological data and provide some level of predication by applying different type of computational methods.

### Slide 18

- 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 sets.

# NPTEL VIDEO COURSE – PROTEOMICS

## PROF. SANJEEVA SRIVASTAVA

---

### Slide 19

- The simple systems biology approach would involve experiments by adding new datasets which will be used for model construction and model analysis and the biological insight derived from these models can be used to propose new hypothesis. So the properties of 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.

### Slide 20

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

### Slide 21

#### Approaches to study systems biology

- The distant approaches of systems biology include the model-based and data-based methods. The model-based approach involves some prior information which can be implemented in the models whereas in the data-based methodology, the objective is to find the new phenomenon.
- The model based relies on computational modeling and simulation tools whereas the data-based method relies on the omics data sets.
- So in model-based systems biology approach it is difficult to build detailed kinetic models but in data-based systems a complex relationship among various type of omics information, metabolic pathways and networks can be created.

### Slide 22

- Studying systems components 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 its use for prediction of systems behavior in response to the perturbations.

### Slide 23

- The reductionist approach involves disintegrating systems into its components and studying them.

### Slide 24

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

# NPTEL VIDEO COURSE – PROTEOMICS

## PROF. SANJEEVA SRIVASTAVA

---

### Slide 25

What is systems biology triangle?

### Slide 26

#### Systems Biology Triangle

- So first of all the systems information is generated at various levels as we have discussed starting genes to mRNA to proteins to metabolites or identifying regulatory motifs, metabolic pathways, functional modules and different large scale organizations. This information has to be stored and processed and further executed to identify the system level information.

### Slide 27

- Even simpler systems such as cell can be linked with various , such as 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 processes.

### Slide 28

- So what approach one can take to study about the systems?
  - Extraction and mining of complex and quantitative biological data
  - Integration and analysis of these data sets for development of mechanistic and computational models
  - Validation of these models by retesting and refining after proposing some hypothesis.
- Different online databases and repositories are nowadays developed and available for sharing large datasets and various systems models. The systematic approach to understand how molecules act together within the network of interactions that make up life is required the systems biology.

### Slide 29

- The systems biology triangle as you can see here involves the experimental data sets could be derived from different type of omics platforms, technologies how the computational analysis is being 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



# NPTEL VIDEO COURSE – PROTEOMICS

## PROF. SANJEEVA SRIVASTAVA

---

with modeling to enhance the understanding of biological processes as whole system rather than the isolated part is termed as systems biology triangle.

### Slide 30

- So in systems biology triangle, the wet lab experiments or bioinformatics-based data analysis can be used to propose a model. The model building serves as an aid to understand complex systems.
- Some hypotheses can be generated which could be used further to propose more quantitative or predictive models and also it can be used for independent techniques for the model validation.

### Slide 31

- After generating the data set and creating the systems biology triangle then this information can be used for understanding the systems at a more complex and mechanistic level.

### Slide 32

System study and model building-

- The system science include synthesis, modeling concepts and analysis.
- Life sciences discipline provides quantitative measurements, genetic modifications and is useful in deriving some hypotheses.
- The information sciences enables the visualization, the modeling tools and different databases. So this model building as an aid to understand the complex systems is very useful for systems level information.

The System is an entity which maintains its existence through a mutual interaction of its constituent parts. Systems biology research consists of the:

1. Identification of the Parts
2. Characterization of the components.
3. Exclude the ones which are not a part of the system
4. Identify the interaction of the components with each other
5. Identify the interaction of the components with the environment which modulate the parts either directly or indirectly through modulation of internal interactions

# NPTEL VIDEO COURSE – PROTEOMICS

## PROF. SANJEEVA SRIVASTAVA

---

The systems biology concepts can be understood with the help of two approaches, namely, the “Reductionist” approach and the “Integrative” approach.

### “Reductionist Approach”

- This focuses on disintegrating the system into its component parts and studying them, whereas, Integrative Approach focuses on integrating the study of individual components to form conclusions about system.
- Consider a cell with its component molecules. Here we study a “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 production rate, or decrease in their amount.
- At this stage, due to lack of knowledge of the nature of interaction 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 of DNA-binding protein leads to increase in the synthesized amount of DNA which further changes the final amount of lipo-protein produced. Thus we can see that to study a system, we need to analyze not just the components, but also their interactions. These biological systems can be:

1. Protein-protein interaction networks

2. Gene regulatory networks

3. Protein-DNA networks

4. Protein lipid interactions

5. Metabolic Networks

- To study the systems, we need to know about the components and their interactions. The data about the components comes from Genomic and Proteomic studies. The data about the molecular interactions comes from interactomics studies.
- Here it is shown, that is a systems approach- experiment, technology and computational modeling, this triangle is very important. This has to be linked with the theory to form a systems triangle. A typical work flow for approaching a biological problem with a systems biology perspective is selection of biological significant problem, creation of a model for the problem which represents computable set of assumptions and hypothesis that will be subjected to experimental validation.
- Compatibility of the model with established experimental facts will reveal the adequacy of assumptions made. If incompatible the model will be modified or

# NPTEL VIDEO COURSE – PROTEOMICS

## PROF. SANJEEVA SRIVASTAVA

---

rejected. The systems analysis is constructed on the consistent models to make predictions for the system. From all the predictions a small is selected to validate them using wet lab experiments which are designed for this purpose. the data from the experiments is analyzed and the inadequate models are discarded.

### Slide 33

So there are different technologies which have been employed to study the systems biology. Obviously you need high throughput data sets which could be derived from microarray platforms, RNA deep sequencing, different configuration of MS, different type of proteomic tools and protein interaction data sets.

### Slide 34

So some of the technologies which are commonly employed in systems biology can be classified broadly under the following techniques-

- For genomics, the high throughput-DNA sequencing, mutation detection using SNP method;
- for transcriptomics the transcript measurement can include serial analysis of gene expression-SAGE, gene chips, microarrays and RNA sequencing;
- for proteomics MS, 2-DE, protein chips and yeast two hybrids and different structural proteomics tools such as X-ray and NMR.
- X-ray and NMR are mainly employed for metabolic analysis, the metabolomics.

### Slide 35

So as you can see here the systems study requires information from different technologies that study biological systems. This information can be generated at the various levels:

- At genome level by studying different type of technologies using HT-sequencing and HT-arrays,
- At the Transcriptome level, with: different type of transcriptome analysis using RNA sequencing and microarrays,
- At the proteome level, as previously discussed
- At the level of the metabolome, which can be studied by using either NMR or MS.
- AT the level of thephenome which involves the NMR method.

So each level of these omic technologies can be useful for studying systems biology.

# NPTEL VIDEO COURSE – PROTEOMICS

## PROF. SANJEEVA SRIVASTAVA

---

### Slide 36

Let's now talk about how to model biological networks.

### Slide 37

- To build a model in systems biology, first of all, a list of the parts can be generated by data sets derived from the systems biology approaches.
- The systems or subsystems model can be generated which can be used for systems model analysis. Now this could be applied to a real systems and by doing sousing bioinformatics tools it can again applied back to original component which could be used to derive a hypothesis and validation of these data sets. So it will work like a closed loop.

### Slide 38

- To build the models in systems biology, information is generated at different levels. 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.

### Slide 39

Frameworks required for modeling schemes-

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

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 bioinformatical tools to generate the predictive models which could be also developed by using the literature curated data sets or experimental data sets and finally it could be used to study the systems level properties.

### Slide 40

Let's discuss workflow of mathematical modeling.

- A paradigm can be proposed based on modify- model- measure-mine.
- So systematic experiments, different type of molecular genetics, chemical genetics and cell engineering approaches can be used for modifying and different level of

# NPTEL VIDEO COURSE – PROTEOMICS

## PROF. SANJEEVA SRIVASTAVA

---

measurements by applying microarrays, spectroscopy imaging and microfluidics based approaches from proteomics and genomics can be used further for mining which involves bioinformatics databases and data semantics.

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

### Slide 41

- The modeling of probabilistic processes can be illustrated as follows. Let's say you want to study a biological system, so an experiment has to be performed.
- The experimental data sets will be generated, from which some statistics can be derived which this can be used for comparison.
- Now different types of models can be generated using simulations and simulations data sets which can be used for intermediate statistics and by comparing these two types of information and adjusting the parameters, one can study the systems and derive probabilistic processes.

### Slide 42

So what are ordinary differential equations and stoichiometric models?

- The quantitative analysis uses measures and names to make models that define precise kinetic parameters of a systems network component. It also uses properties of network connectivity.
- The ODE is a mathematical relation that can be used for modeling biological systems. The quantitative 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 underline gene regulatory networks, these ODE models can be used.
- The stoichiometric model- It is modeling a biological network based on its stoichiometric coefficients, reaction rates and metabolite concentrations. The basics of stoichiometric analysis and ordinary differential equations will be explained in following animation.

To model more complex networks stoichiometric modeling can be used which has 3 basic components:

1. The concentration of all the reactants and products in the form of a vector of Concentration values (S)

# NPTEL VIDEO COURSE – PROTEOMICS

## PROF. SANJEEVA SRIVASTAVA

---

2. The Stoichiometric Matrix that describes the production and consumption of each metabolite at each of individual reaction (N)
3. The vector for Rate of all the reactions involved (v)

Information on any two will help us in determining the third.

Let us consider first 3 reactions of glycolysis and model their stoichiometric matrix.

- The rows of this matrix represent each reactant and products of all the reaction in this system. The column represents individual reaction, ex- in this case, there are 3 reaction steps. So column 1 represents the first step, column 2 represents the second step, and column 3 the third step and so on. If a particular substrate is consumed in the reaction step, we fill -1 in the matrix. If it is produced we enter +1. And if it is not involved in the reaction, we write 0. For huge systems, we continue to fill the matrix in a similar form, by breaking the entire network into individual steps.
- Here, S is matrix of concentrations of each of the reactants/substrates. N is the matrix of the stoichiometric coefficient describing the production and consumption of reactants and products. V is the matrix of rates of the 3 reactions involved. Since v matrix needs to be multiplied with N, we will transpose it such that the number of rows in v equals the number of columns in N.
- The simple linear equation for modeling networks, takes the following shape in its vector form. Once such a model is constructed, if 2 variables are known, we can determine the unknown variables. The same concept of modeling systems can be applied to larger biological systems like:
  1. Metabolic Network
  2. Protein-Protein Interaction
  3. Gene regulatory Network
- Here we describe basic Ordinary Differential Equations (ODE) with the help of simple equations. A general ODE " $da/dt = f$ ", which signifies that 'the change in "a" during a short time interval 'dt' is "f" times of "dt"'. In the first equation " $da/dt = \text{zero}$ ", the concentration of "a" does not change with time. In the second equation, " $da/dt = \text{one}$ ", concentrations of "a" changes by dt in time "dt". In the third equation " $da/dt = -a$ ", Concentrations of "a" decreases with time. The larger the initial concentration of "a", the larger is the decrease. The "minus" sign shows the decrease in concentration.
- The Law of Mass-Action says that "The rate of a reaction is proportional to the product of the concentration of the reactants". The highlighted variables depict the

# NPTTEL VIDEO COURSE – PROTEOMICS

## PROF. SANJEEVA SRIVASTAVA

---

position of “reactant” in the “reactions” and the position of “concentration of reactants” in the “ODEs”.

- The only reactant involved in first reaction is “A”. As the equation shows, “A” reduces to zero with time. From the ODE, we infer that as the time increases, “A” reduces in concentration by a value which is proportional to its initial concentration.
- In the second reaction, “A” converts to “B” with time. From the ODE for “A”, we infer that as the time increases, “a” reduces in concentration by a value which is proportional to its initial concentration. From the ODE for “B”, we infer, that as the time increases, “b” increases in concentration by a value which is proportional to initial concentration of “a”.
- In the third reaction the concentration of “A” and “B” reduces and “C” increase with time by a value which is proportional to initial concentration of “A” and “B”.
- Here we display a summary of common reactions and their corresponding Ordinary Differential Equations. These equations are for individual reactions. In a metabolic network, we extrapolate these reactions to cover an entire network by summing the contribution from all individual reactions within the pathway.
- In this example we construct a model for the initial steps of Glycolysis with the help of ODE models. Here you can see the reactions in glycolysis. The several reactions are accompanied by the consumption of ATP into ADP. The rate of each individual reaction is represented by a variable  $V$ . These rates are determined by reaction kinetics.
- Now, to determine the concentration of each of the substrates at a given time, we need to account for all the reactions which involve this substrate. Thus, to determine the concentration of G6P at a given time, we need to determine its rate of production ( $V_1$ ) and its rate of consumption ( $V_2$ ,  $V_3$ ). The same process when extrapolated to all the reactants and products of a metabolic pathway, we can get the ODE Model for glycolysis.

### Slide 43

After discussing some of these computational modeling and ODE based methods now let's look at some of the systems biology applications.

### Slide 44

#### Systems Biology Applications

- The systems biology research focuses on data mining, data integration, its processing and representation. The data analysis involves the network structure and dynamics of molecular interactions.



# NPTEL VIDEO COURSE – PROTEOMICS

## PROF. SANJEEVA SRIVASTAVA

---

### Slide 45

- Systems profiling, as you can see in the slide it is quite complex process which involves different type of challenging test personalized integrative intervention, optimized interventions and the systems response profiling by applying different type of methodologies can be used for monitoring the patients response.
- So due to advancements in systems biology, it is becoming clear that it will accelerate the fundamental changes in the future of health care. It is also expected that a paradigm shift in medicine will take place due to the proteomics and systems biology tool in near future.

### Slide 46

- Different systems biology applications involve understanding metabolism, cellular networks and cell to cell communication and pathogenicity and toxicity while predicting phenotype from genotype and creating predictive model of cells, organs, biochemical processes and complete organisms.

### Slide 47

Now let's try to integrate omics approaches with systems biology.

### Slide 48

- So genome sequencing projects in genomics era from 1990s to 2000 accelerated the pace of omics research. Then from 2000 onwards the proteomics field also got accelerated with the advent of new methodologies and new tools. And the data derived from genomics, transcriptomics, proteomics, metabolomics and other omics approaches have now brought the integration of these data sets in the systems biology field.

### Slide 49

- The systems study requires data sets from different approaches and their analysis. For example as shown in the slide, the genome wide data sets can be derived at the genome level and looking at the different transcripts or at the proteome level looking at different type of protein interactions.
- These data sets can be stored in the clinical databases and also it can be mined from the literature, literature manual curation then integration of orthogonal data sets further can be used for validating the networks and deriving identifying therapeutic targets. Further it can be used for experimental validation.



# NPTEL VIDEO COURSE – PROTEOMICS

## PROF. SANJEEVA SRIVASTAVA

---

### Slide 50

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

### Slide 51

- So eventual aim of systems biology field is to employ the omics level information obtained from different levels such as the genome, transcriptome and proteome, derive that information at the systems level, integrate and quantitate some models and then propose and use it for the understanding the physiology and apply that in medicine. So this flow from omics to physiology can be well maintained by employing systems biology tools.

### Slide 52

What are the challenges of systems biology?

- Systems biology is extremely challenging. The emphasis is to understand a system. Understanding dynamics of even simplest biological networks not only requires only the understanding of biology but also its modeling and simulation. The disintegrating 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 understanding and doing the prediction and control of functional biological process. All of these are very challenging but currently being addressed by applying various systems levels tools.

### Slide 53

So how are proteomics and systems biology integrated?

- Proteomics as we have studied, is useful to understand complex signaling networks in biological systems. It is a very indispensable tool for systems biology. The global analysis of the proteome is important however there are many limitations in each experiment, only thousands of proteins can be studied.
- Therefore new approaches and systems level investigation and predictions are required.
- The system investigation is required to study the complex dynamic structure interaction with the biological systems whether it is at cellular level or at the organism level and ultimately it is responsible for their function and behavior.

# NPTEL VIDEO COURSE – PROTEOMICS

## PROF. SANJEEVA SRIVASTAVA

---

### Slide 54

In summary today we discussed that how omics era the technological advancements in genomics, proteomics and metabolomics have generated large scale data sets in all the aspects of biology. These large data sets have motivated the computational biologist to develop systems approaches with an objective of understanding the biological system as a whole, while proteomics continues to generate the quality data at the proteome level so the integrated systems biology approach characterizes and predicts dynamic properties of biological networks. In the next module we will focus on different type of proteomic technologies. .