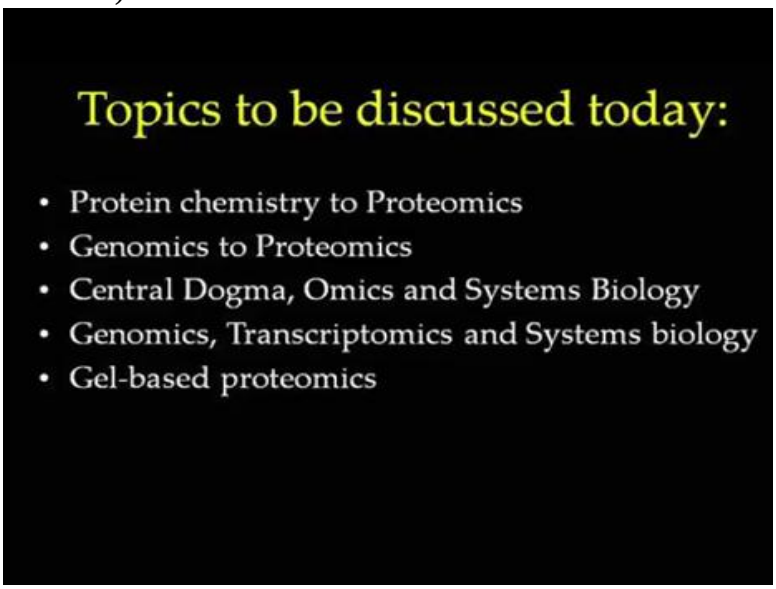
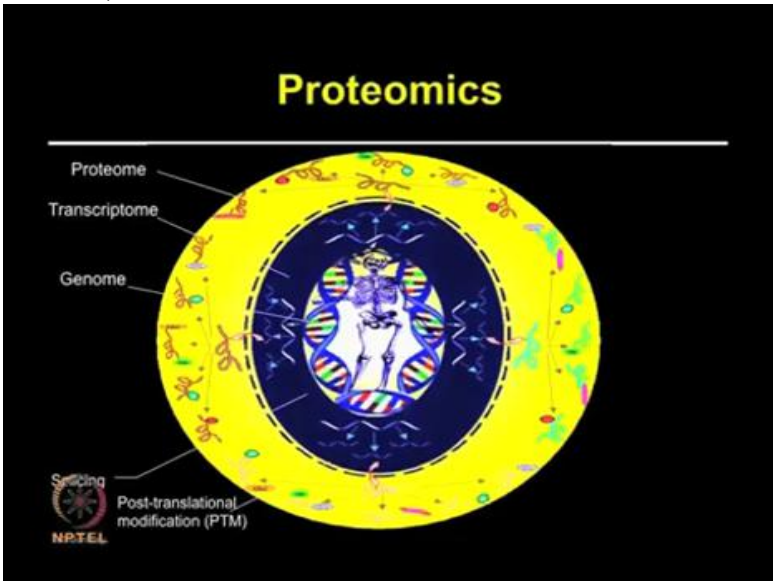
Proteins and Gel-Based Proteomics
Professor Sanjeeva Srivastava
Department of Biosciences and Bioengineering
Indian Institute of Technology, Bombay
Mod 02 Lecture Number 6

(Refer Slide Time: 00:16)



In this lecture, I will discuss about Proteomics.

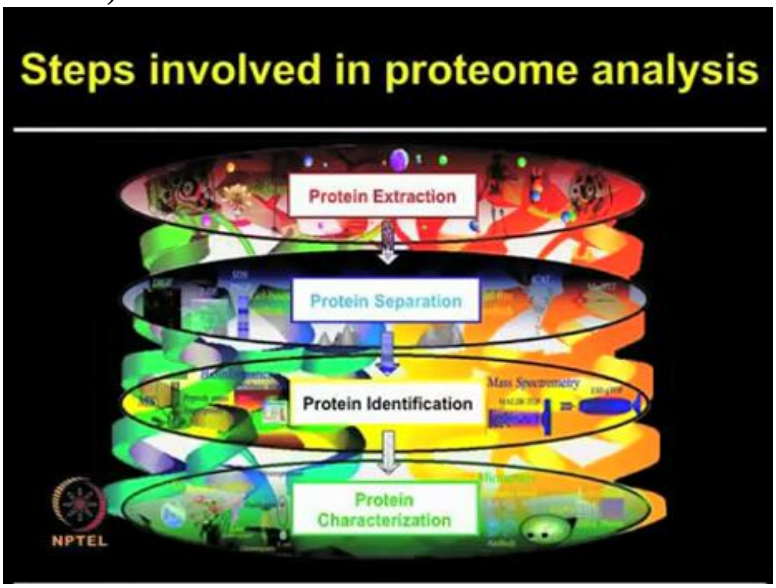
(Refer Slide Time: 00:22)



So first of all, what is Proteomics? The proteome describes the protein complement expressed by the genome or more precisely we can say, the protein complement of a given cell at a given time including the set of all the protein isoforms and its modifications. The study of entire compendium of proteins which are encoded by the genome is known as proteomics.

In this slide, I have illustrated the complexity of human proteome as compared to the genome or transcriptome. The extent of diversity and complexity due to alternative splicing and post translational modifications is tremendous. Therefore studying proteins and proteomes are very important.

(Refer Slide Time: 01:35)



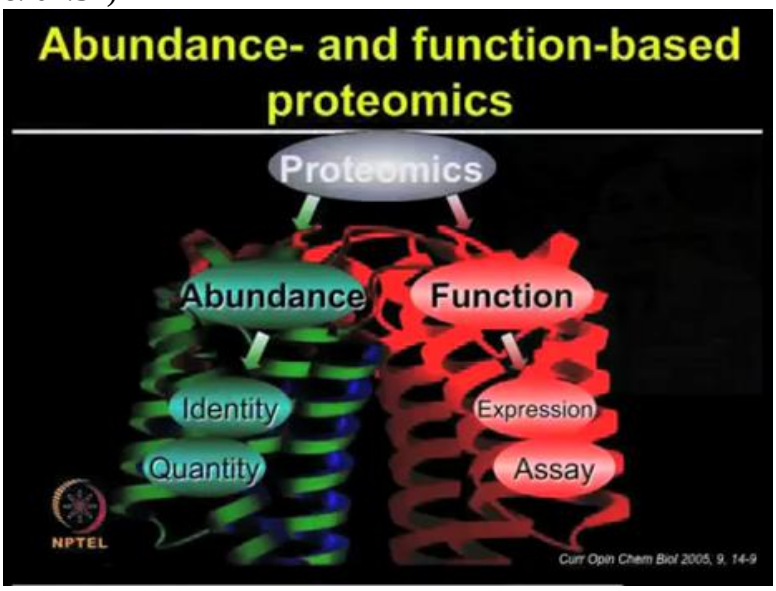
What are different steps involved in the proteome analysis?

As shown here, the protein extraction, protein separation, protein identification and protein characterization, these are the major steps which are involved in proteome analysis. The protein extraction from whole cells, tissues or organism is first requirement for proteome analysis.

Protein separation and quantification is achieved by various proteomic techniques including gel-based techniques such as two-dimensional electrophoresis and gel-free techniques such as iTRAQ Mass Spectrometry based techniques. The functional characterization of proteins using novel proteomics platforms open new horizon for exploration in biology.

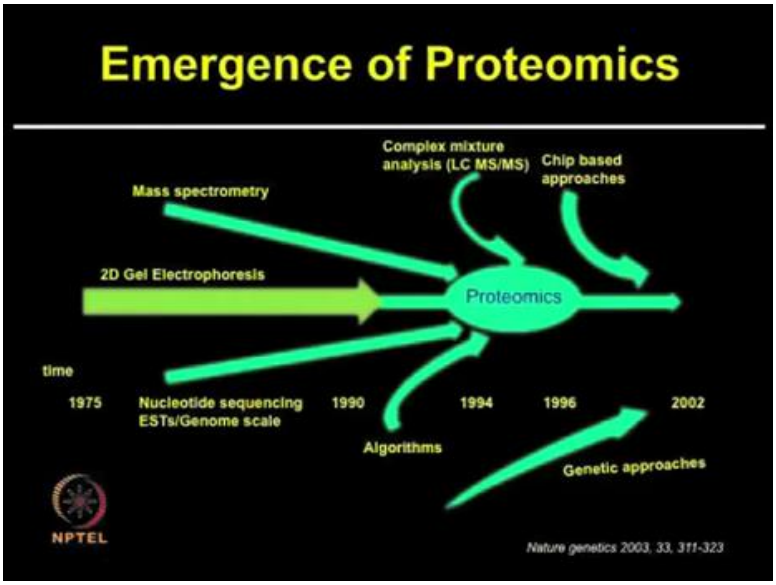
The proteomics discipline can be grouped under two major disciplines, abundance and function based proteomics.

(Refer Slide Time: 02:54)



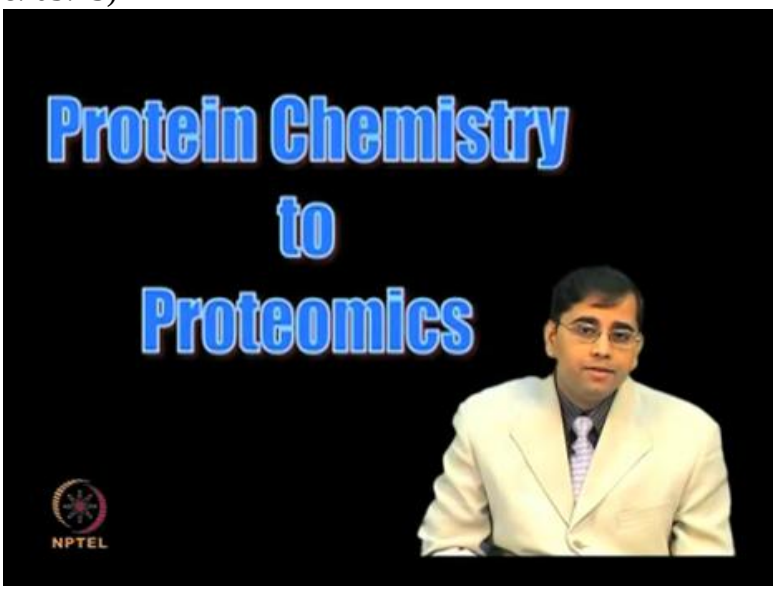
The abundance based proteomics aims to measure the abundance of protein expression where as the functional proteomics aims to determine the role of proteins by addressing protein interactions and their biochemical activities. So how did proteomics field emerge?

(Refer Slide Time: 03:24)



As you can see in the timescale here shown in the slide, advancement of various techniques such as two-dimensional electrophoresis and mass spectrometry, genome sequencing information and computational algorithm together led to the emergence of proteomics field

(Refer Slide Time: 03:45)

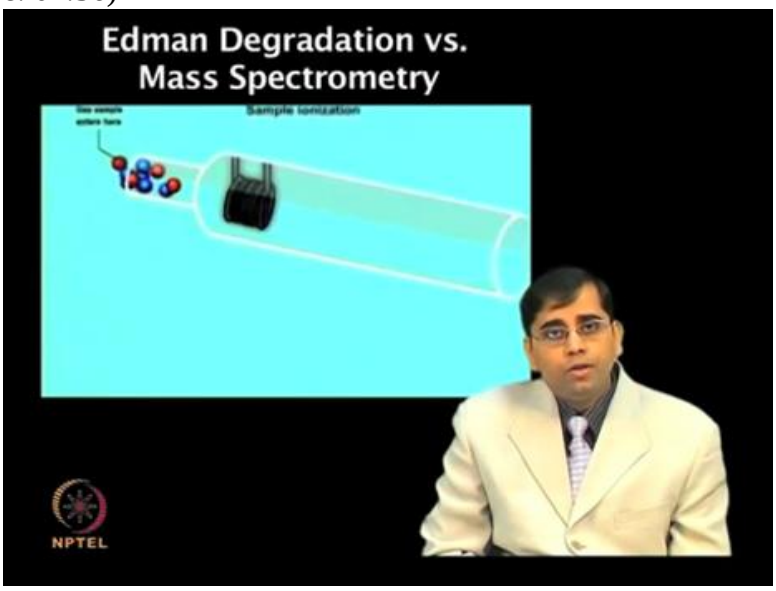


Proteomics research originates from classical protein chemistry and it has embraced new high throughput techniques to analyze complex samples. Many of the techniques used under the modern proteomic umbrella, for example two-dimensional electrophoresis, mass spectrometry have actually originated several years ago.

So what is new? The technological advancements in protein analysis with increased sensitivity, resolution and capability to carry out high throughput studies has led to the transition from protein chemistry to new field of proteomics.

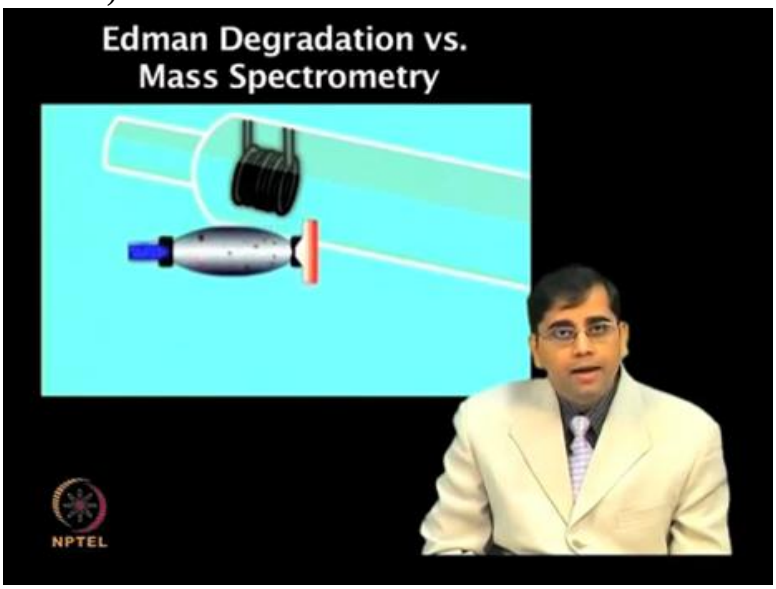
Protein analysis by mass spectrometry was challenging due to complete degradation of samples with available hard ionization techniques. This limitation was overcome…

(Refer Slide Time: 04:56)



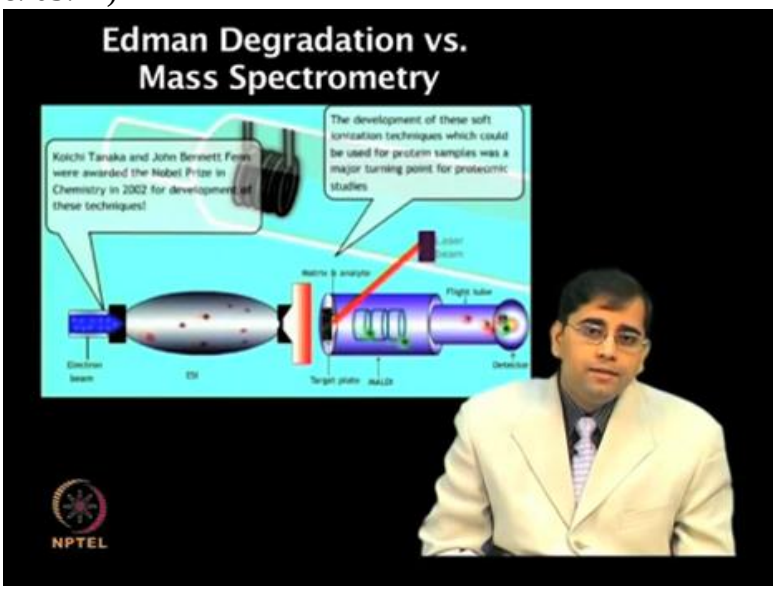
by soft ionization techniques…

(Refer Slide Time: 05:05)



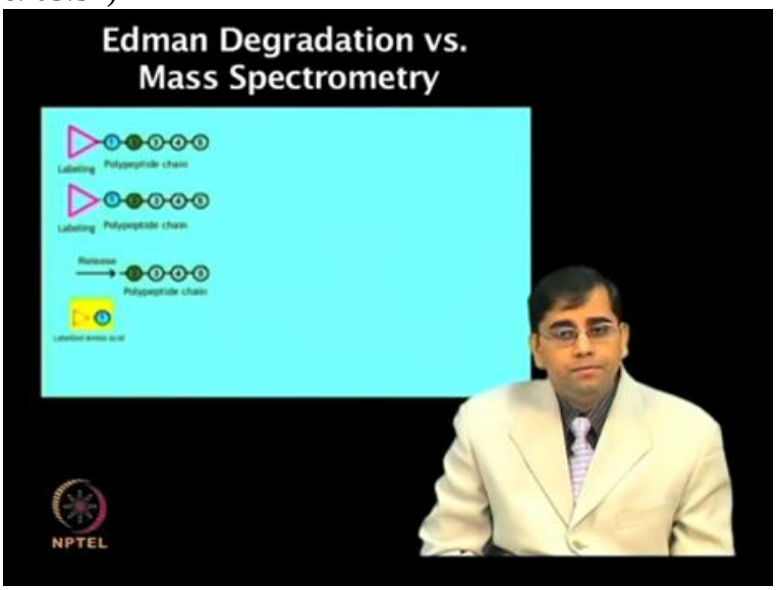
… such as MALDI and electospray ionization.

(Refer Slide Time: 05:14)



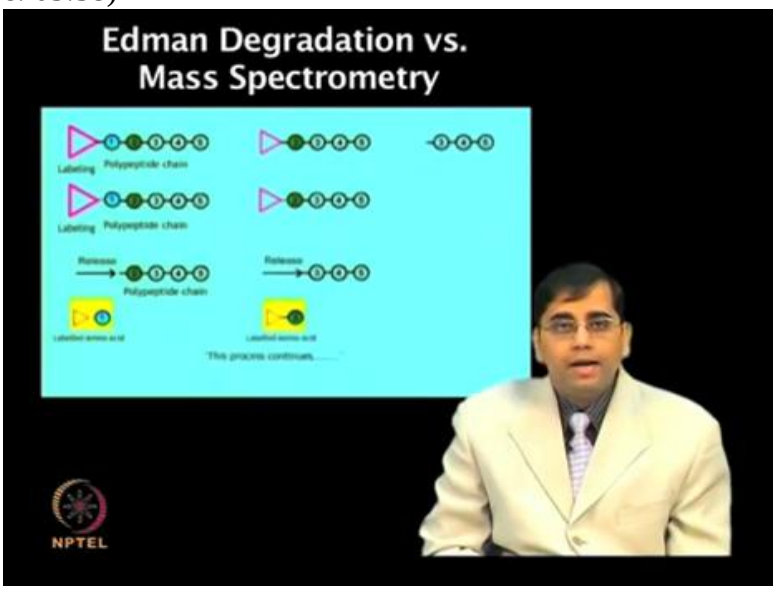
These techniques have greatly improved the proteomic studies as they facilitated mass spectrometry analysis of protein samples.

(Refer Slide Time: 05:31)



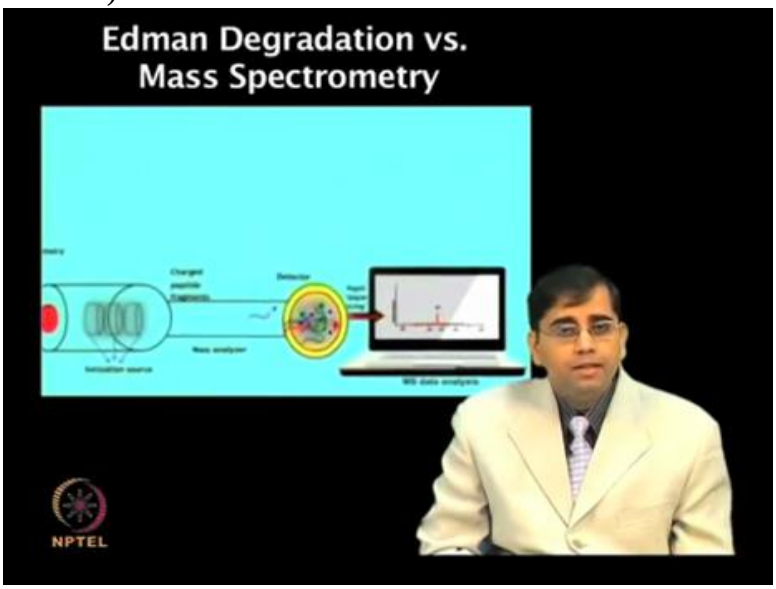
Protein sequencing by Edman degradation is time consuming and cumbersome.

(Refer Slide Time: 05:38)



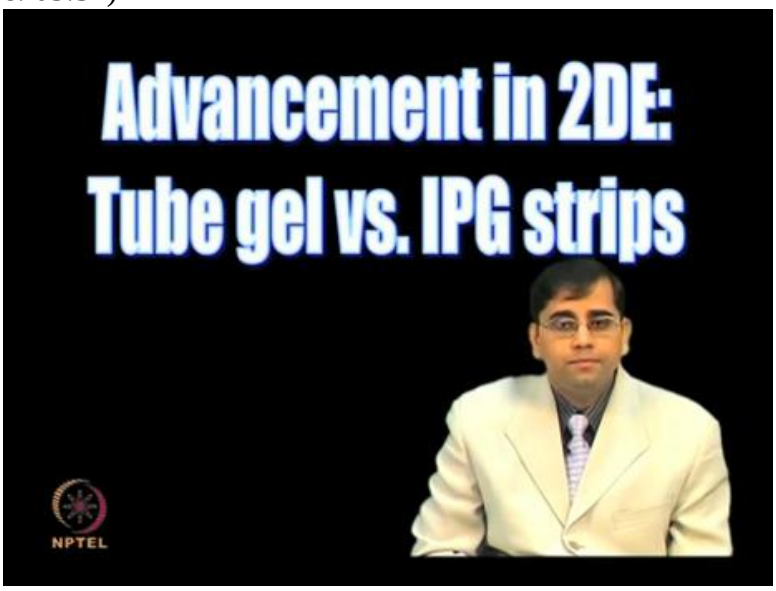
Several rounds of sequencing are required for analysis of polypeptide chains.

(Refer Slide Time: 05:42)



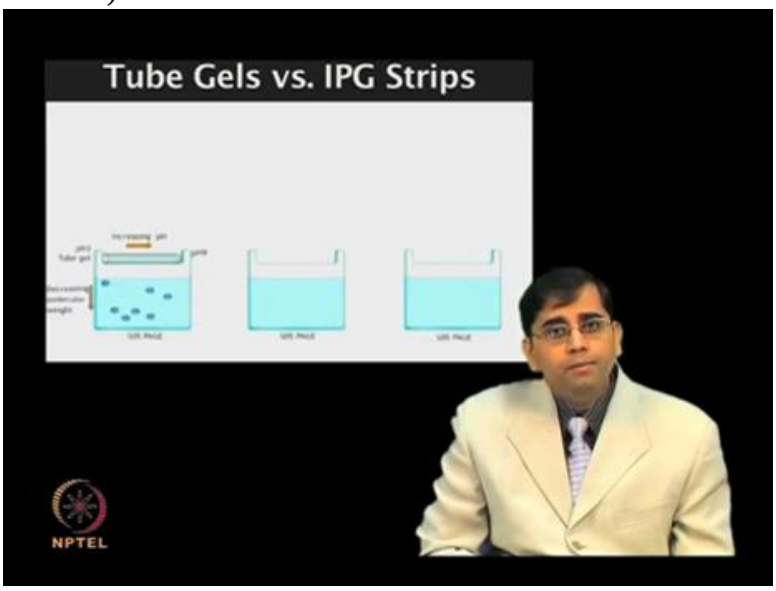
However peptide sequencing by mass spectrometry is much faster and allows large number of samples to be analyzed in a short time.

(Refer Slide Time: 05:57)



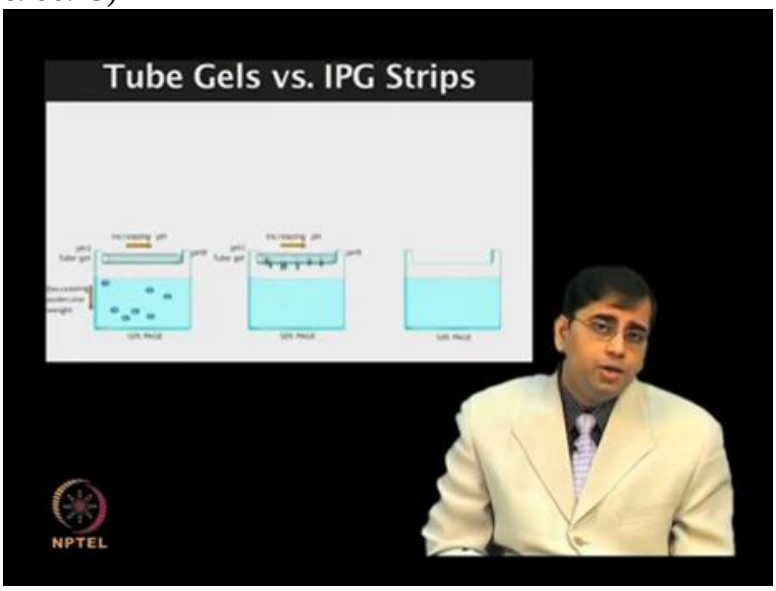
Another aspect, development of immobilized pH gradient strips facilitated proteomic analysis sing two-dimensional electrophoresis. The pH gradient in tube gels are established by ampholyte gradients which are not always very stable …

(Refer Slide Time: 06:23)



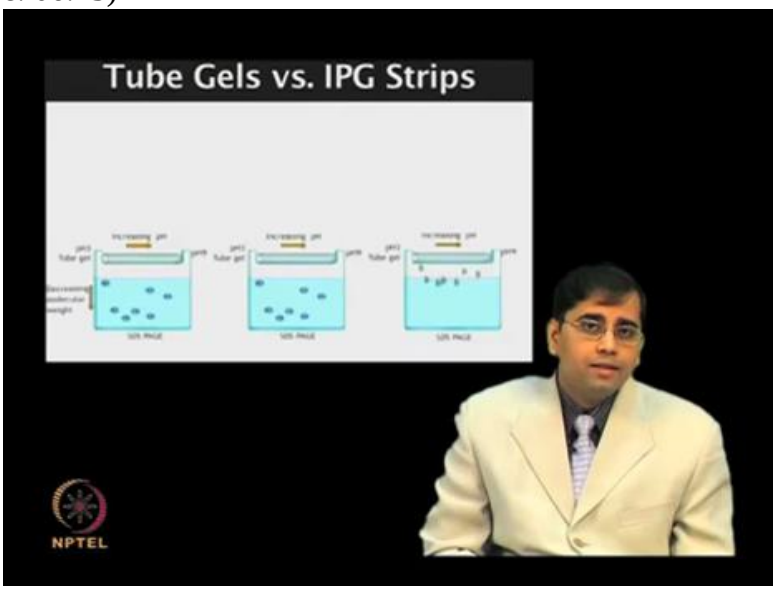
…and tend to break down upon addition of the concentrated samples.

(Refer Slide Time: 06:29)



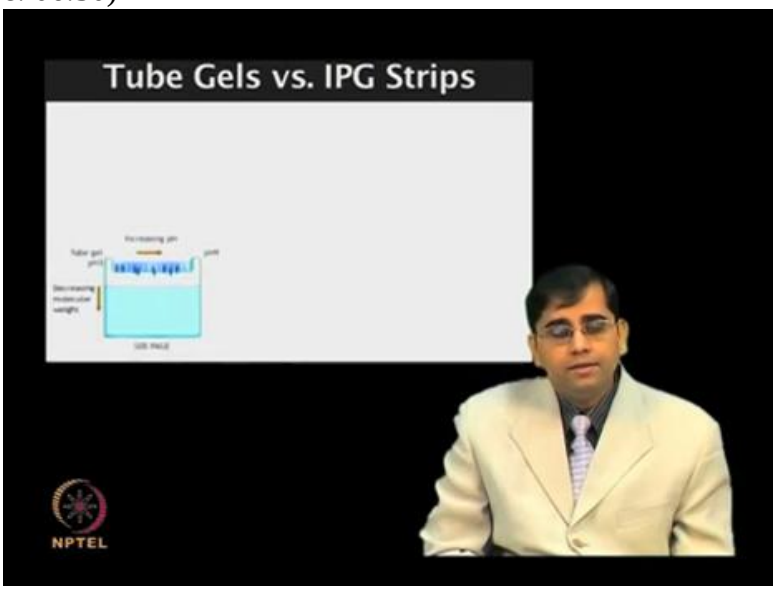
Analysis of protein mixture by two dimensional electrophoresis using tube gels often resulted into…

(Refer Slide Time: 06:43)



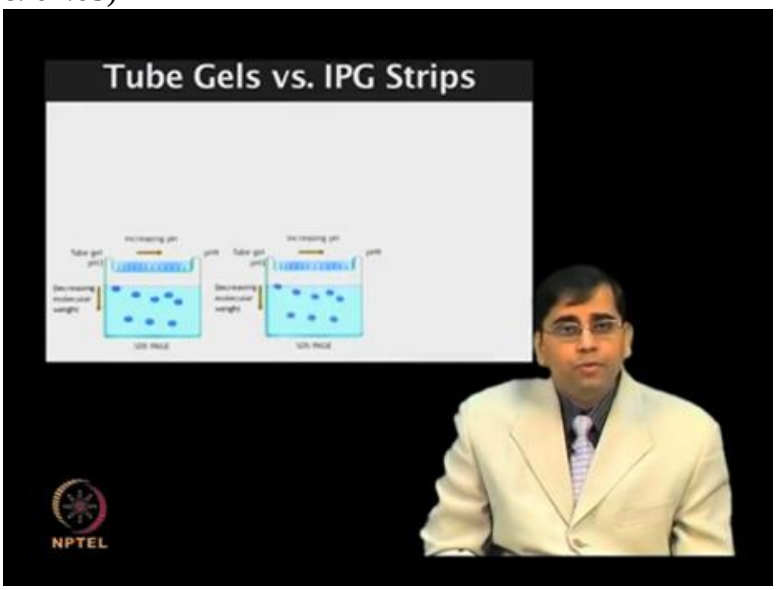
… variations in the gels.

(Refer Slide Time: 06:50)



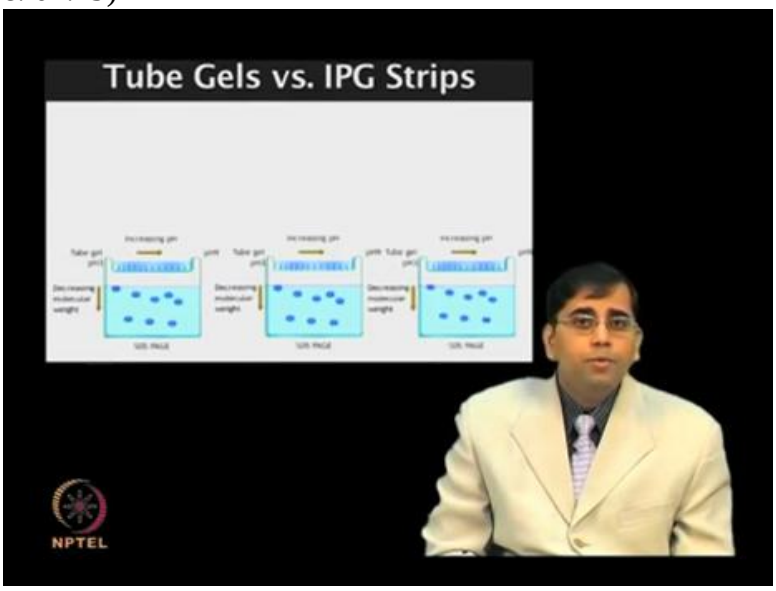
The problem of reproducibility was overcome to a large extent by the development of Immobilized pH Gradient strips or IPG strips.

(Refer Slide Time: 07:05)



Minimal gel to gel variation was observed when samples were run by…

(Refer Slide Time: 07:19)

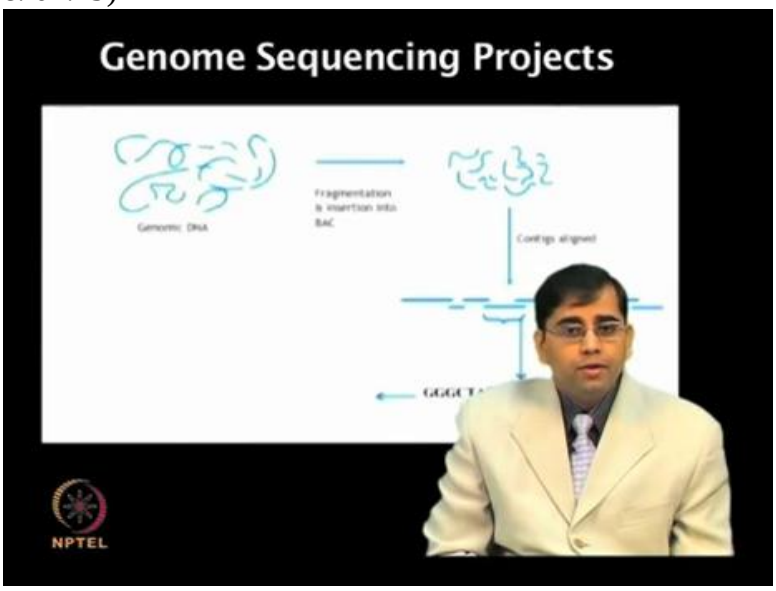


… two dimensional electrophoresis employing IPG strips which made this technique suitable for the large scale proteomic applications.

(Refer Slide Time: 07:22)

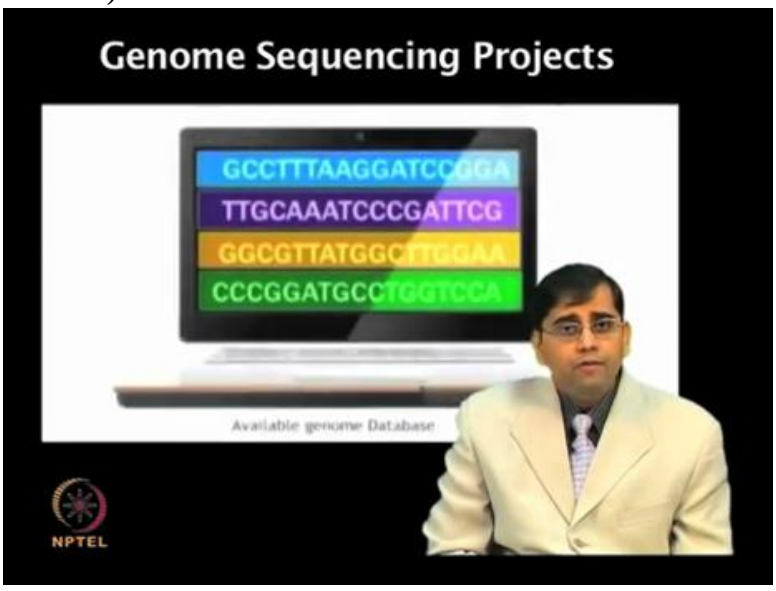
Completion of several genome sequencing projects

(Refer Slide Time: 07:29)



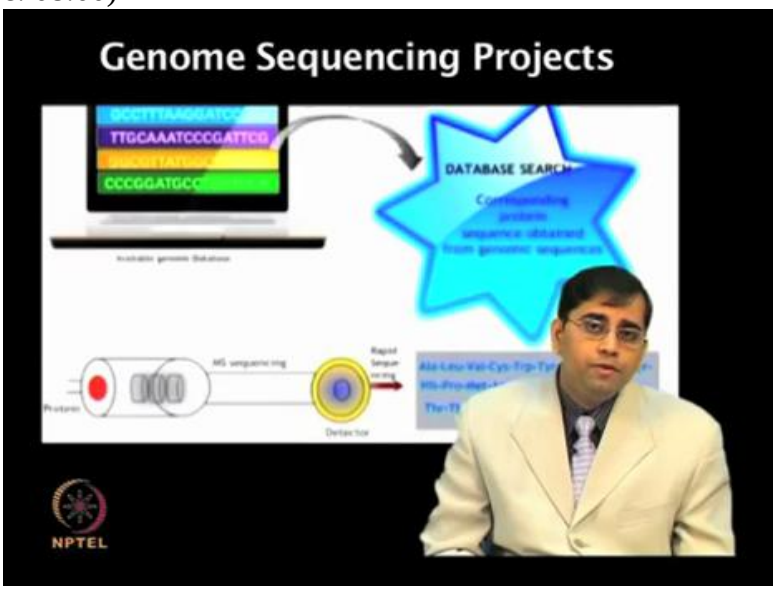
Genome sequencing of several organisms including humans have been successfully completed

(Refer Slide Time: 07:37)



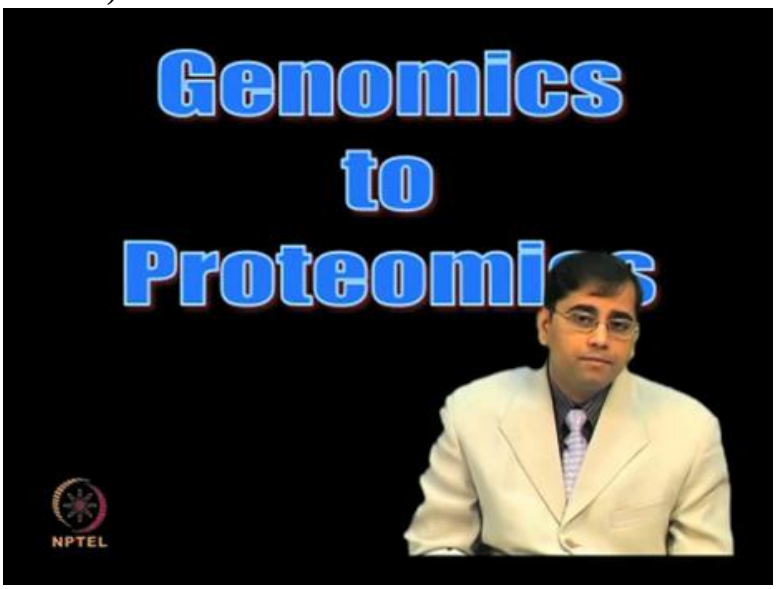
… and these genome databases are extremely useful in correlation of gene and protein sequences.

(Refer Slide Time: 08:00)



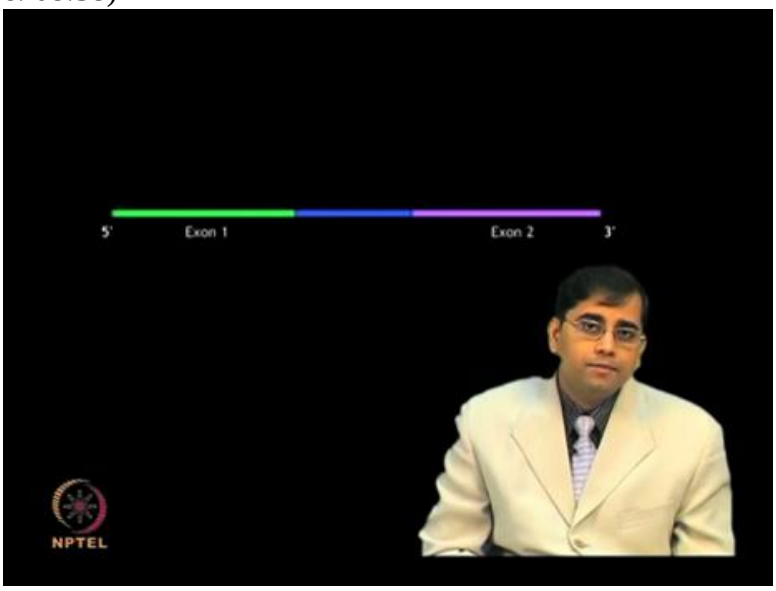
Several databases are now really available which can really help in identification of gene sequence of a protein which has been sequenced by mass spectrometry.

(Refer Slide Time: 08:08)



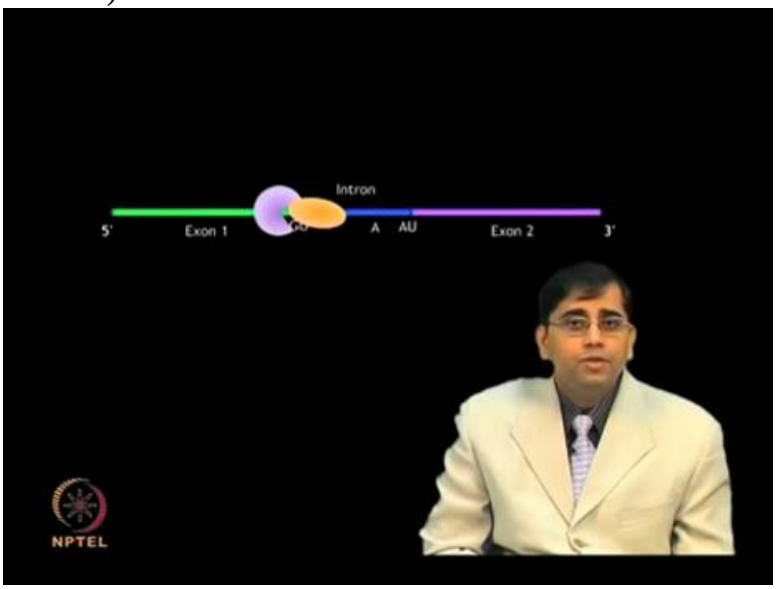
Genome represents an important starting point towards understanding complexity of biological functions. However proteins provide a much more meaningful insight into the mysteries of essential biological processes. To obtain better understanding of cellular processes and regulation there has been an increasing interest in studying proteome. There are several reasons why one needs to study proteomics.

(Refer Slide Time: 08:58)



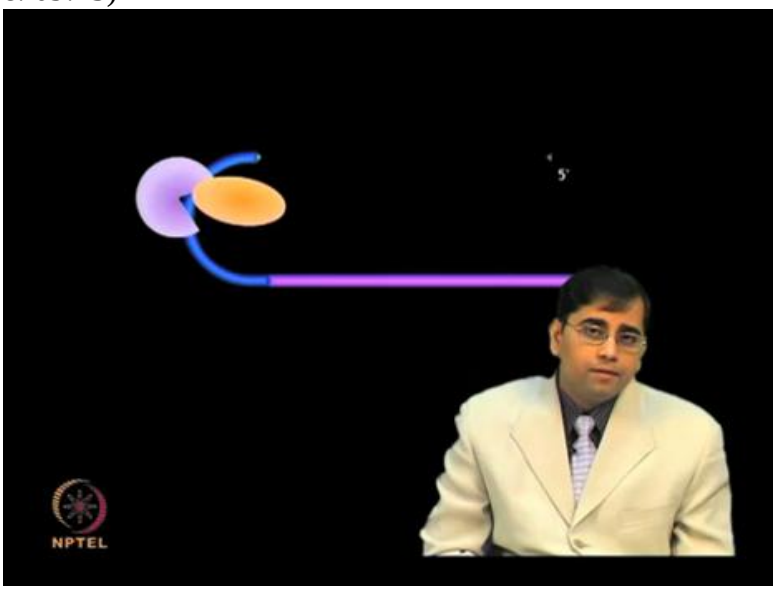
First, the genomic DNA contains large stretches of non-coding regions.

(Refer Slide Time: 09:05)



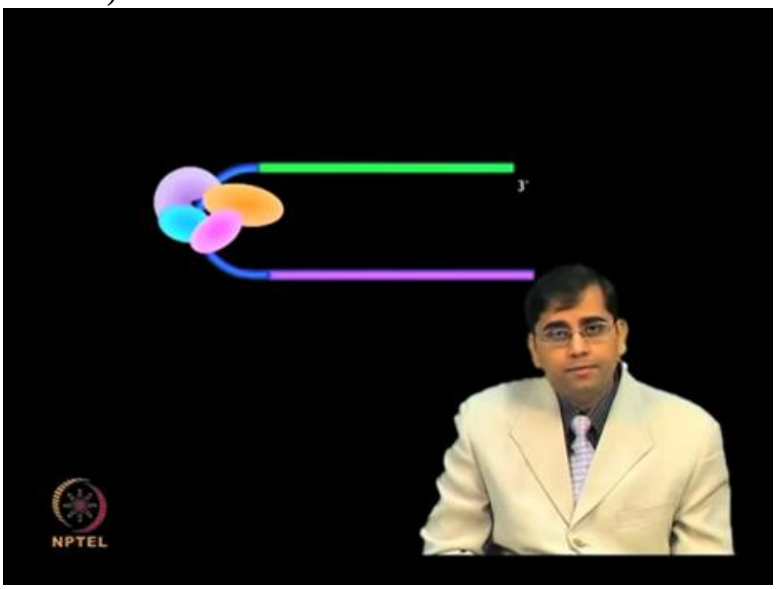
The pre-mRNA is synthesized from the genomic DNA by the process of transcription.

(Refer Slide Time: 09:15)



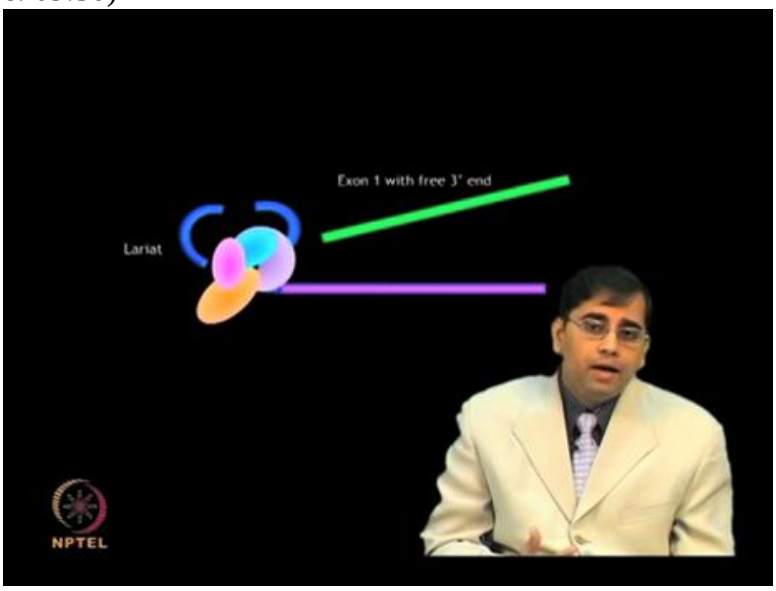
…mRNA contains both exons, the coding sequences

(Refer Slide Time: 09:23)



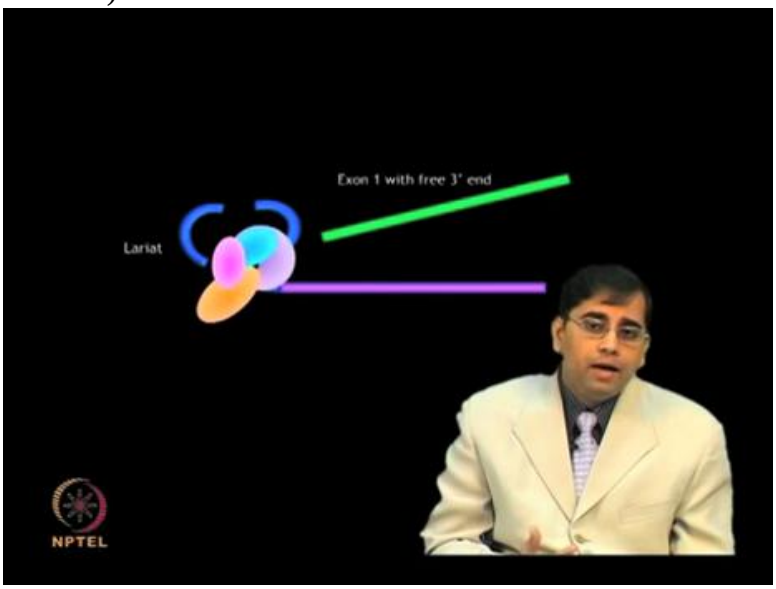
… as well as introns which are intervening non-coding sequences.

(Refer Slide Time: 09:30)



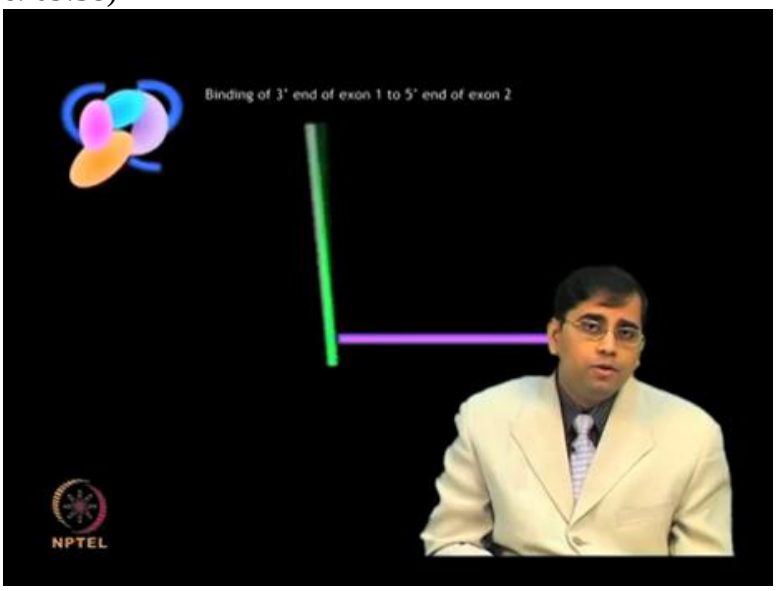
By involving series of steps

(Refer Slide Time: 09:34)



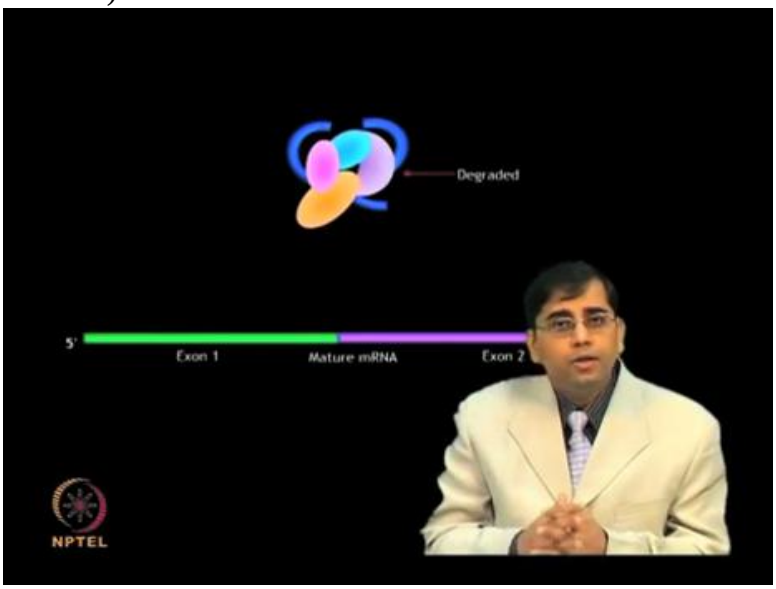
… finally the three prime hydroxyl group of the first exon attacks

(Refer Slide Time: 09:38)



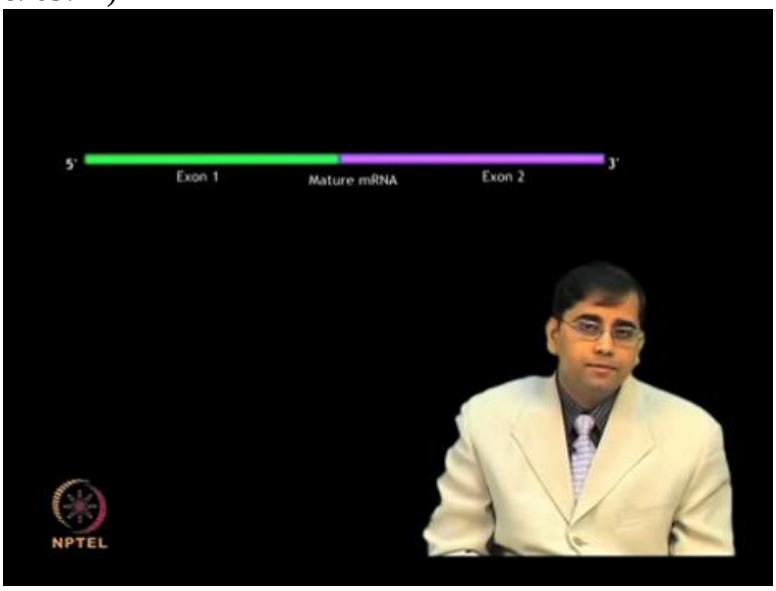
… the 5 prime end of the second exon such that

(Refer Slide Time: 09:41)



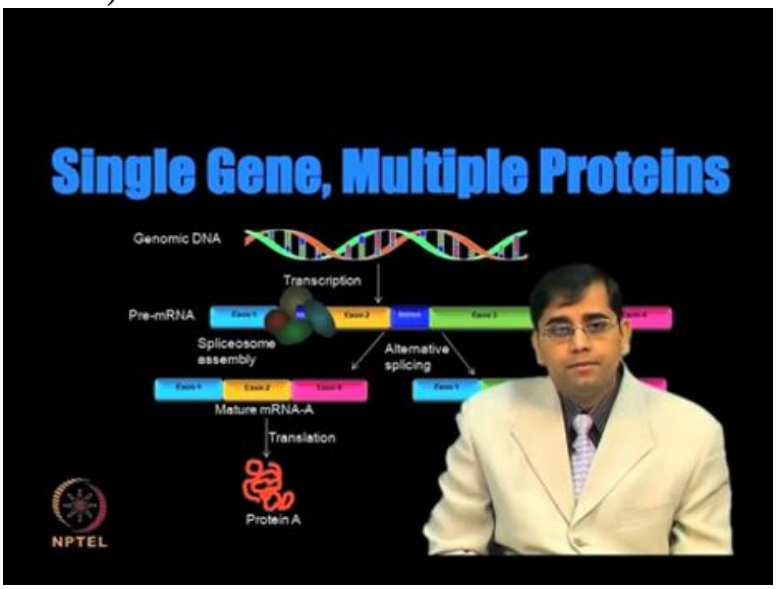
…they are joined together…

(Refer Slide Time: 09:44)



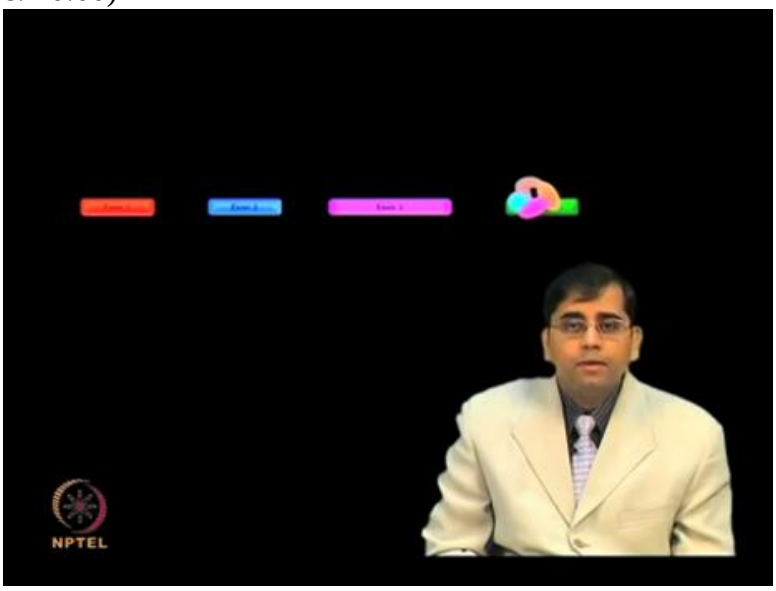
… to give the mature mRNA.

(Refer Slide Time: 09:47)



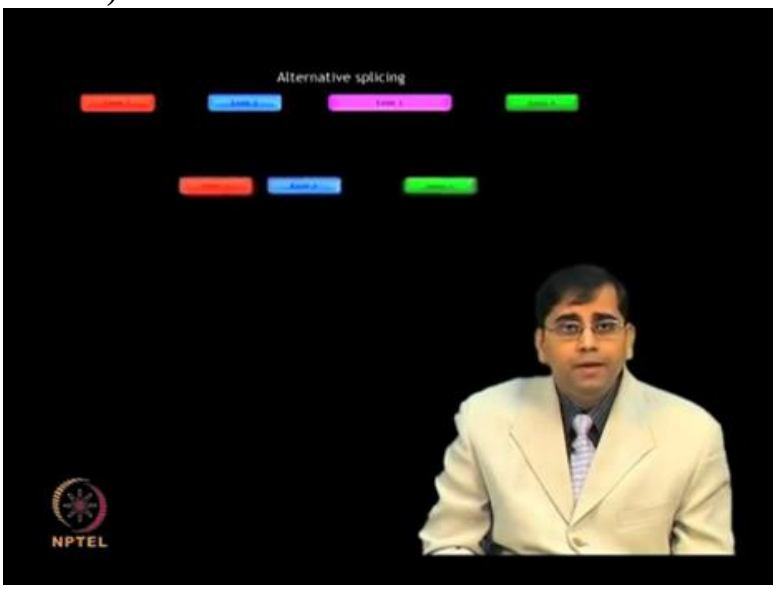
Second important factor is… single gene can give rise to multiple proteins.

(Refer Slide Time: 10:00)



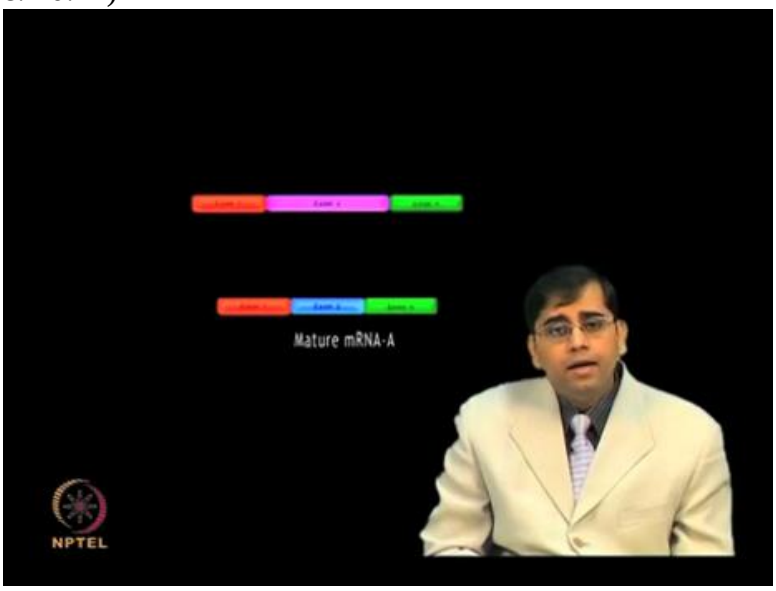
The alternative splicing is a process by which, exons or coding sequences of pre mRNA

(Refer Slide Time: 10:07)



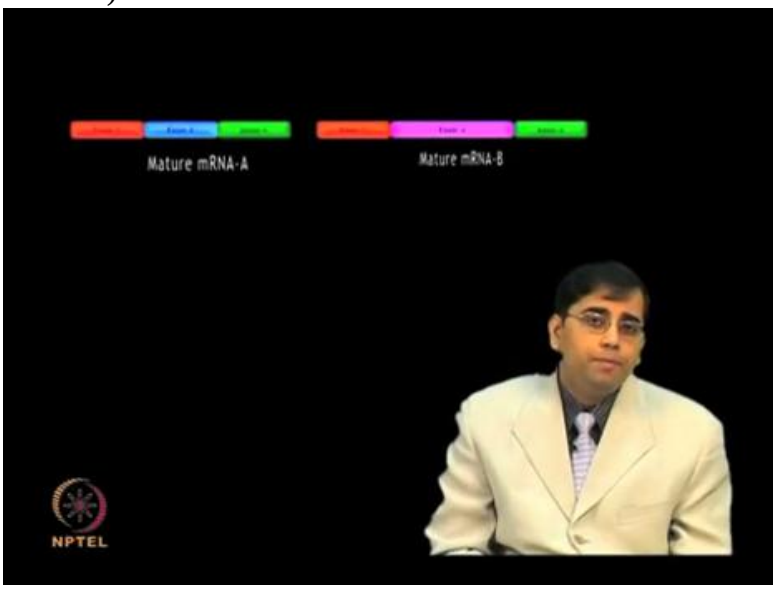
produced by transcription of a gene

(Refer Slide Time: 10:12)



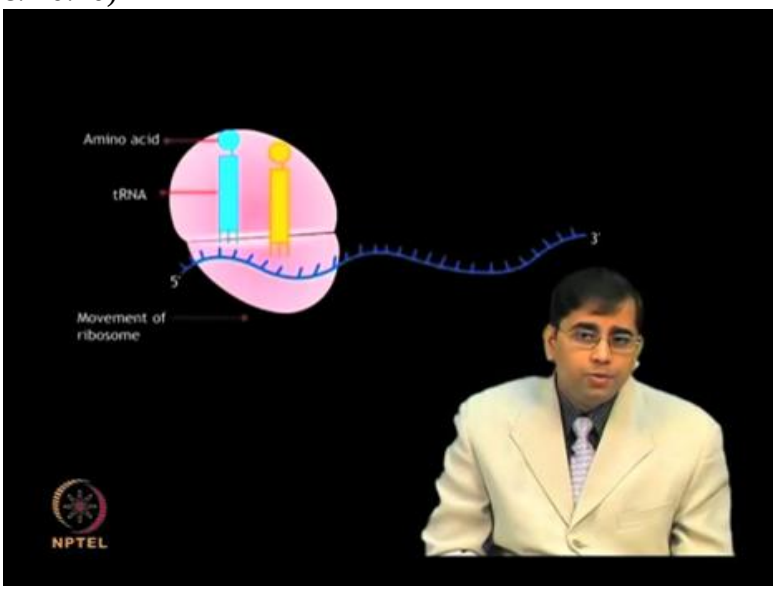
…are combined in different ways…

(Refer Slide Time: 10:14)



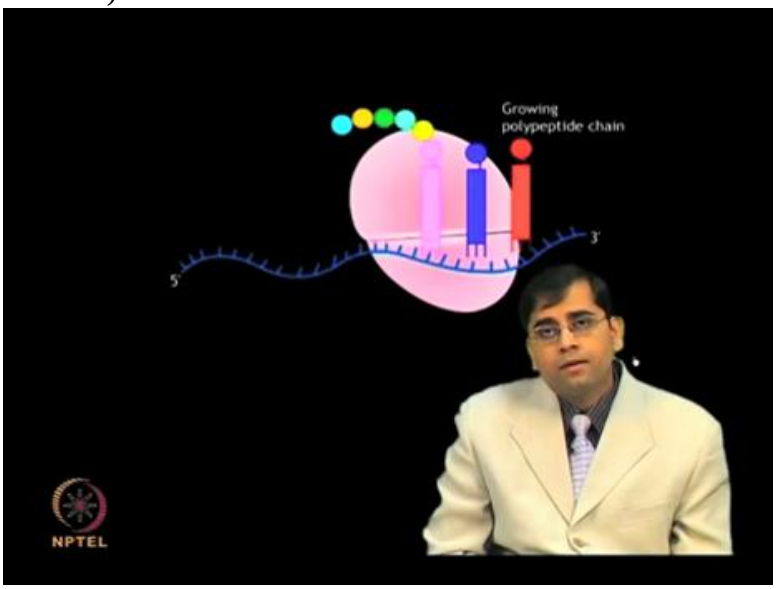
…during RNA splicing

(Refer Slide Time: 10:20)



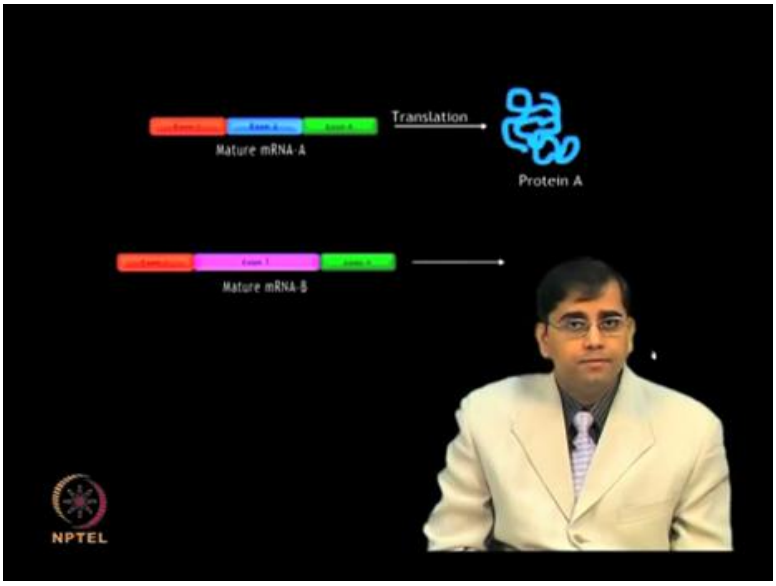
Resulting mature mRNA gives rise to

(Refer Slide Time: 10:32)



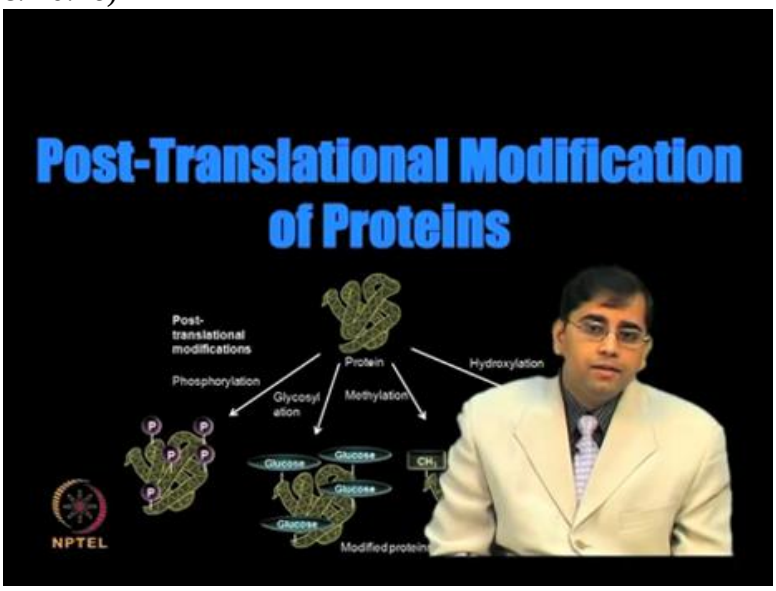
different products by translation, most of which are isoforms of one another.

(Refer Slide Time: 10:43)



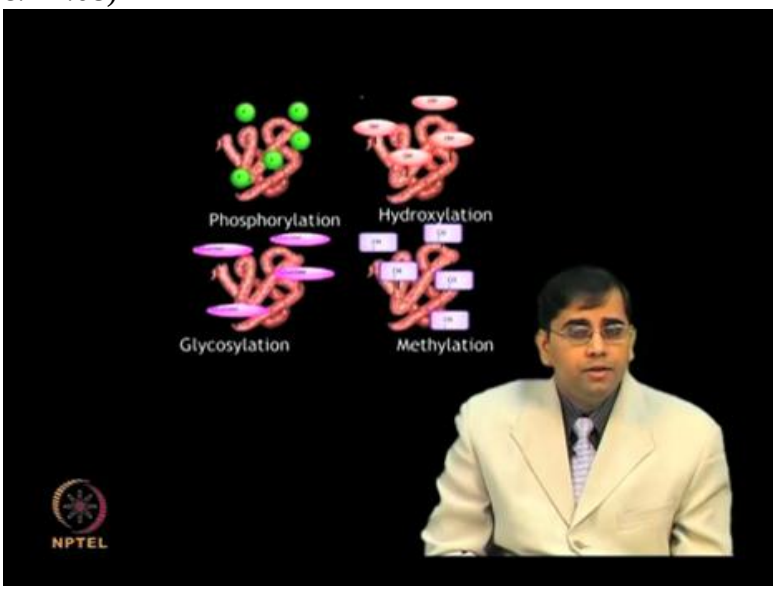
The diversity of proteins encoded by a genome is greatly increased due to alternative splicing.

(Refer Slide Time: 10:46)



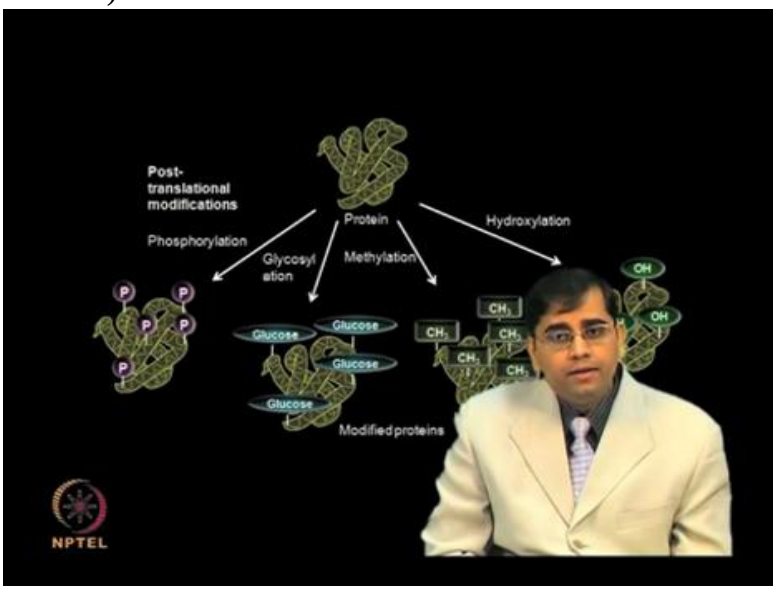
Third important factor is post translational modification of proteins.

(Refer Slide Time: 11:08)



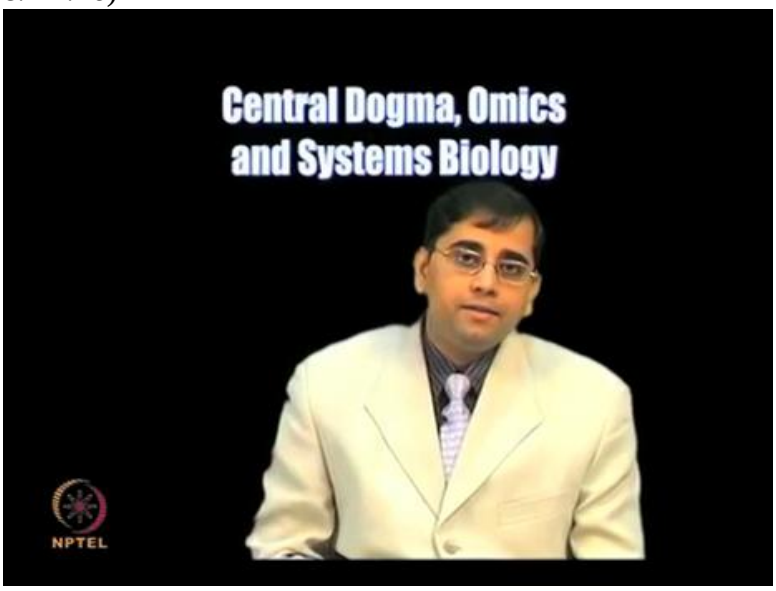
The proteins obtained by translation undergoes folding and various post-translational modifications such as phosphorylation, glycosylation, alkylation, hydroxylation etc. to give the final functional proteins.

(Refer Slide Time: 11:19)



PTMs generate diversity, complexity and heterogeneity of gene products and its functional consequences can be modulation in protein dynamics and alteration of its functional activity.

(Refer Slide Time: 11:46)

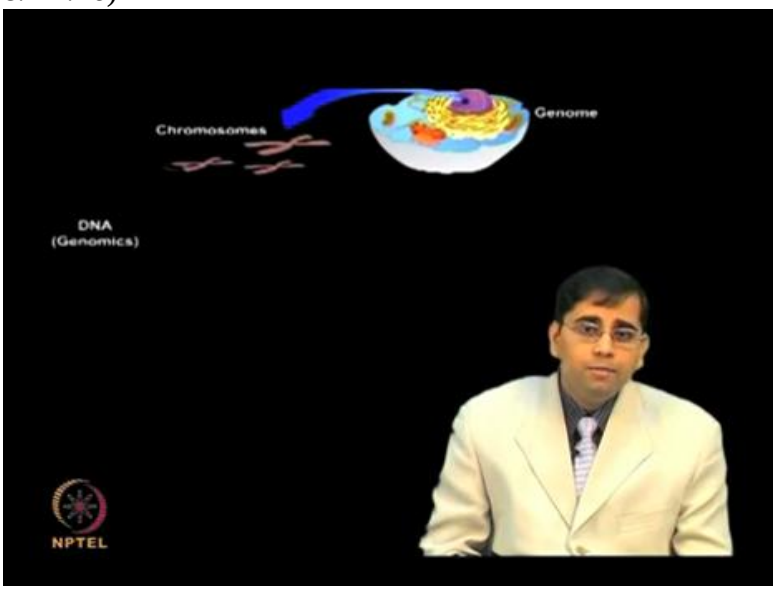


During the last decade, we have witnessed the revolution in biology as this discipline has fully embraced Omics tools. The emergence of genome-wide analysis to understand cellular DNA, RNA and protein content by employing genomics, transcriptomics and proteomics at Systems level has revolutionized our understanding of control networks that mediate the cellular processes. These concepts will be discussed in first module.

(Refer Slide Time: 12:33)

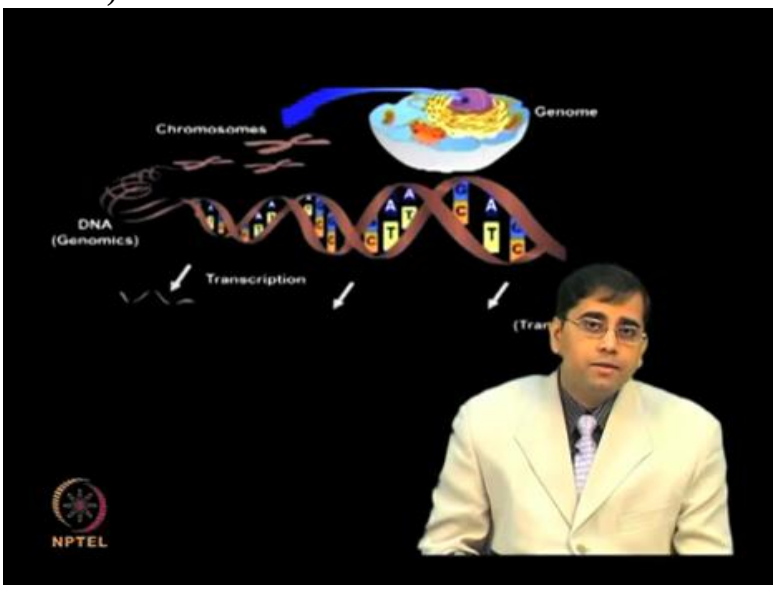
Genes are the blueprint for life and proteins are the effector molecules.

(Refer Slide Time: 12:46)



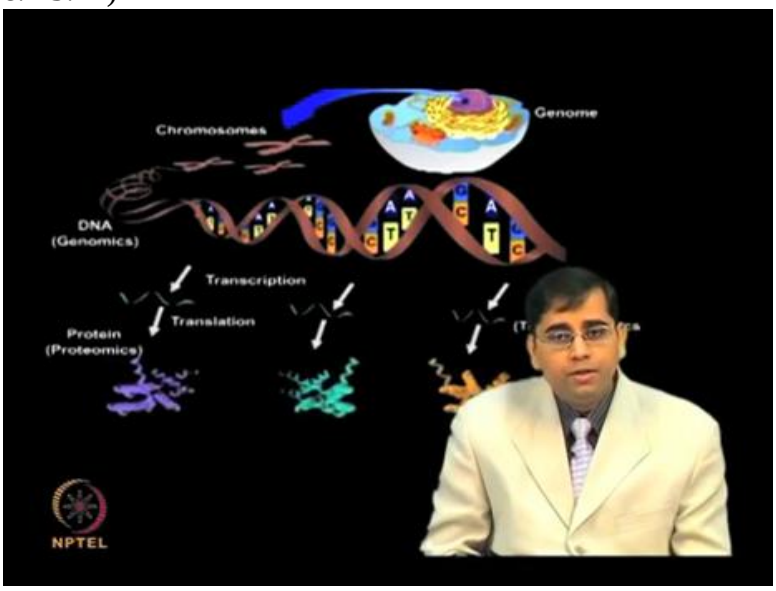
Due to this fact, the central dogma has guided research at the Systems level.

(Refer Slide Time: 12:59)



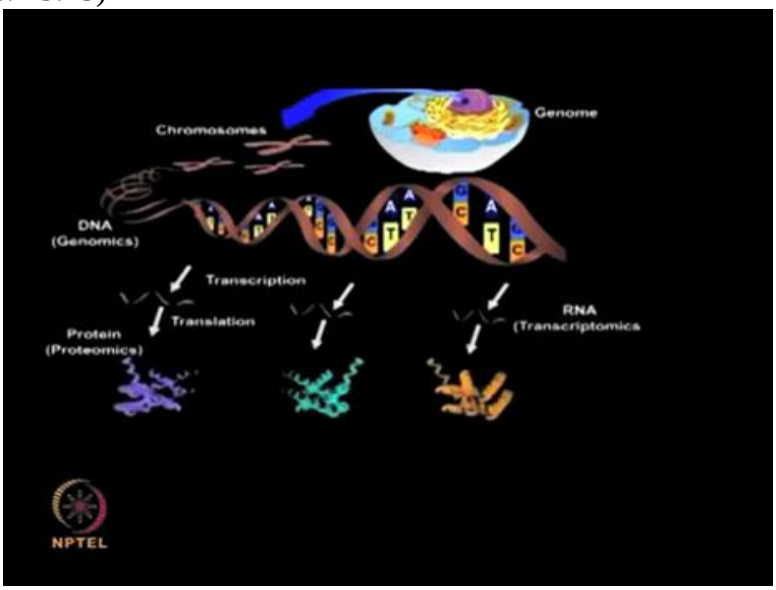
After completion of human genome sequence, numbers of genes around 25000s are surpassed by an estimated number of proteins in millions.

(Refer Slide Time: 13:17)



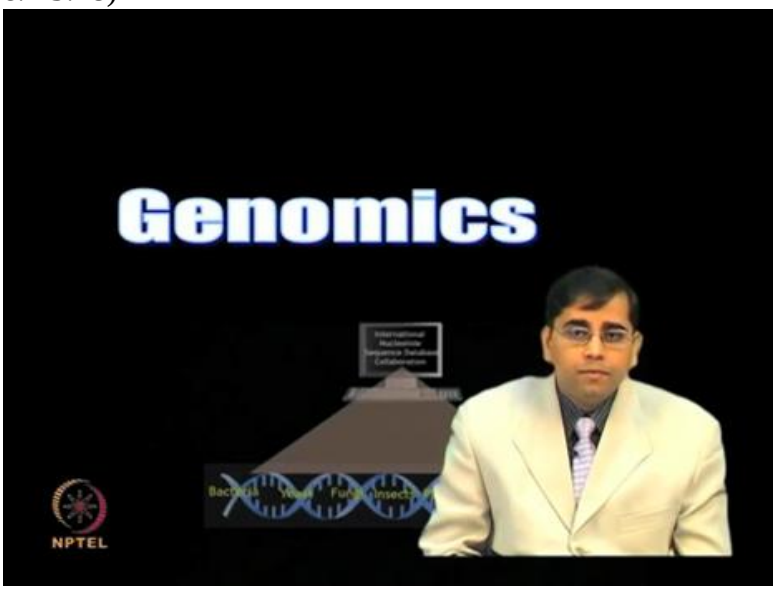
Studying large scale study of protein structure and function requires a thorough understanding of protein composition and their various structural levels

Refer Slide Time: 13:25)



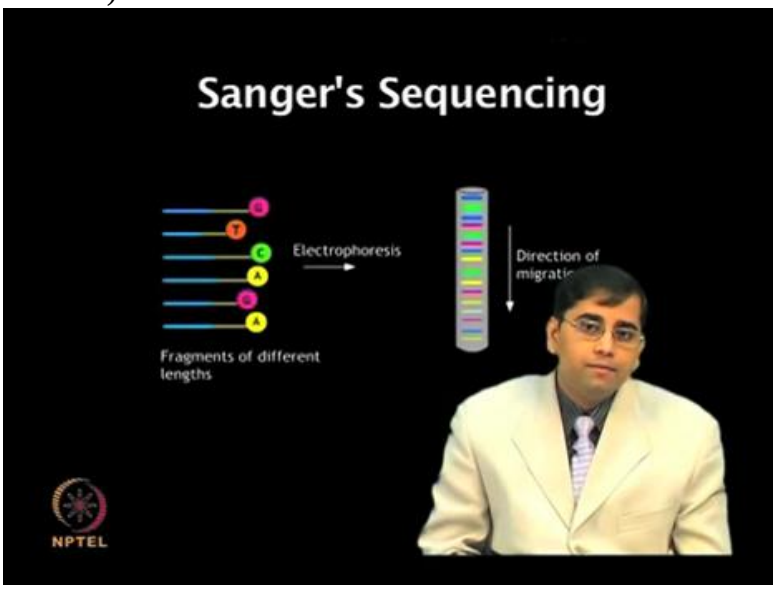
by employing high throughput tools.

(Refer Slide Time: 13:28)



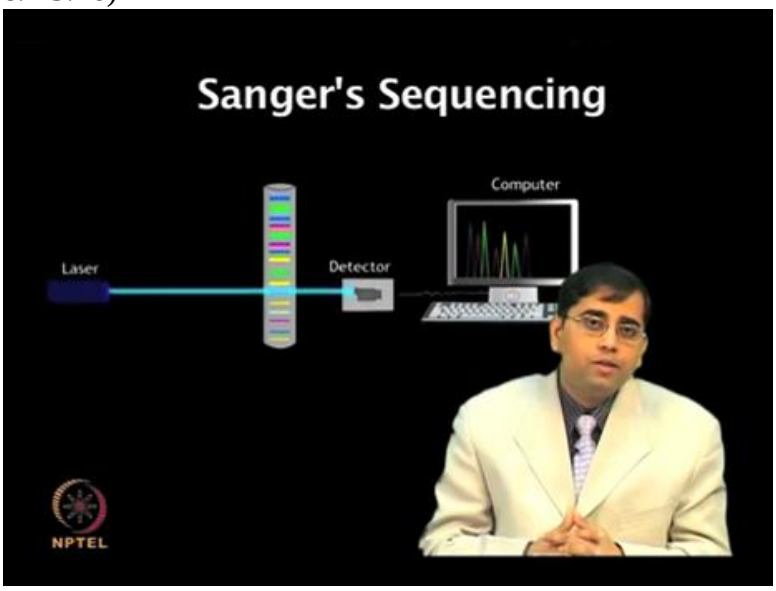
Studying genome of an organism by employing sequencing and genome mapping is known as genomics.

(Refer Slide Time: 13:41)



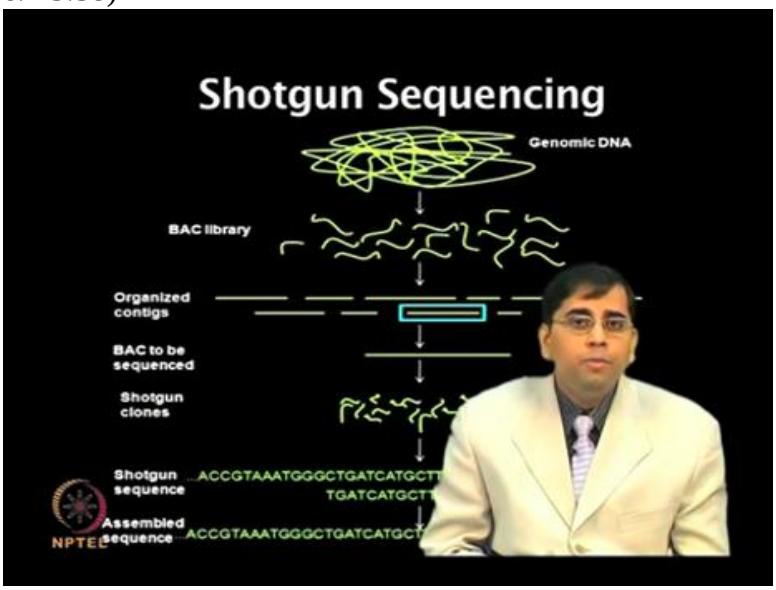
Several genome sequencing projects that aim to elucidate…

(Refer Slide Time: 13:46)



…the complete genome sequence of organisms have been undertaken by several research groups all over the world.

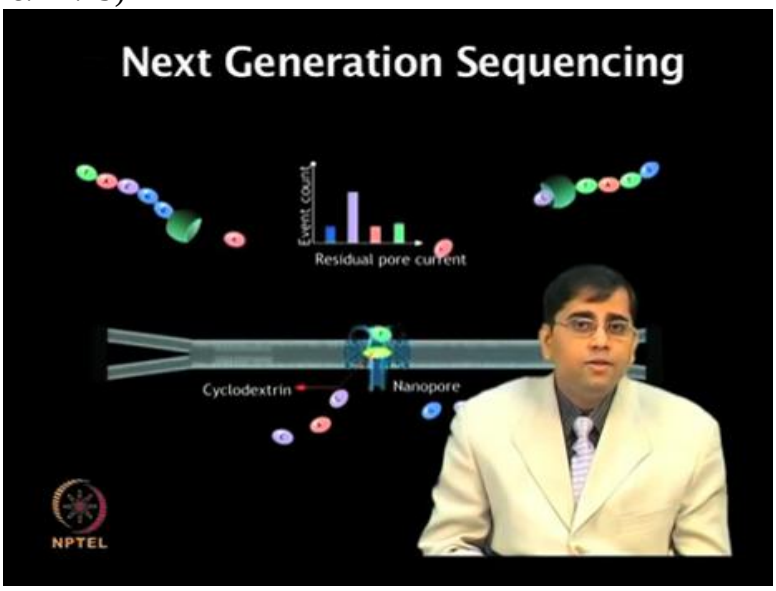
(Refer Slide Time: 13:58)



From a genomic library, clones were isolated and ordered into a detailed physical map.

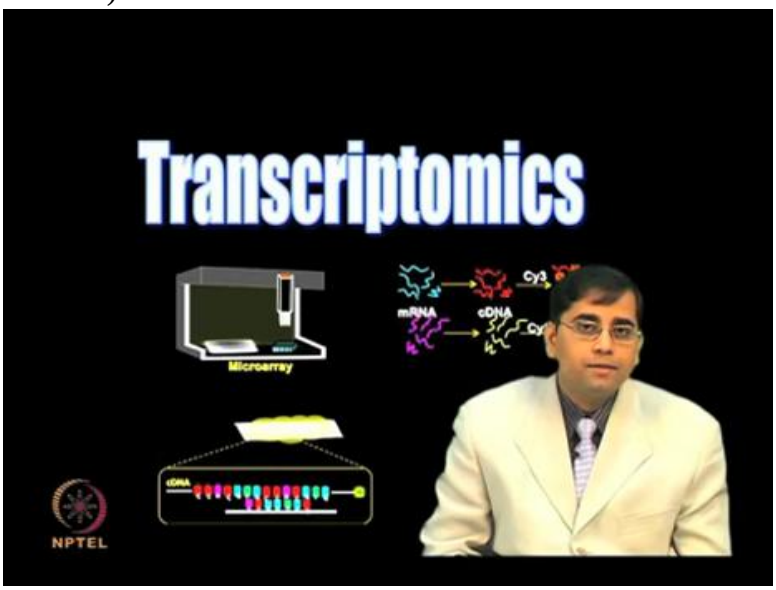
Further individual clones were sequenced by Shot Gun sequencing to provide the complete genome sequence.

(Refer Slide Time: 14:19)



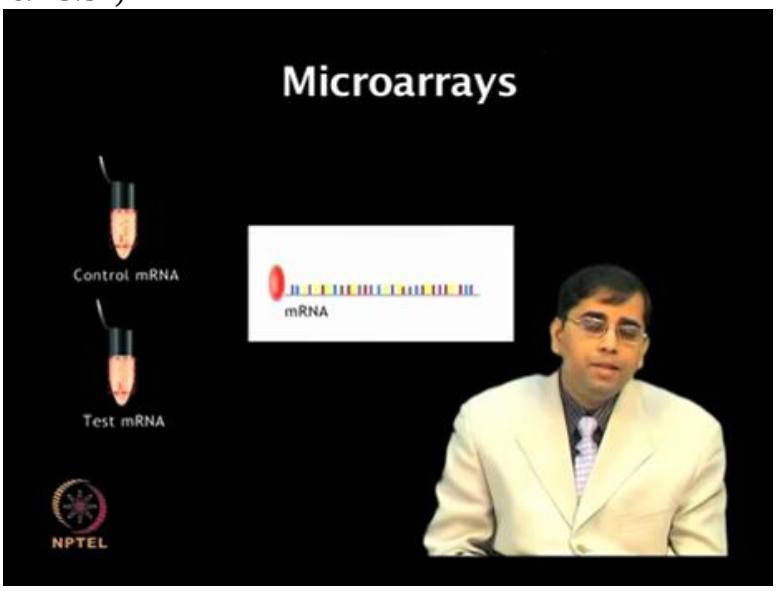
Recently Next Generation Sequencing NGS strategies have dramatically increased the pace of sequencing by several order of magnitudes. Next Generation Sequencing based on nano-pore structures is known as nano-pore sequencing.

(Refer Slide Time: 14:43)



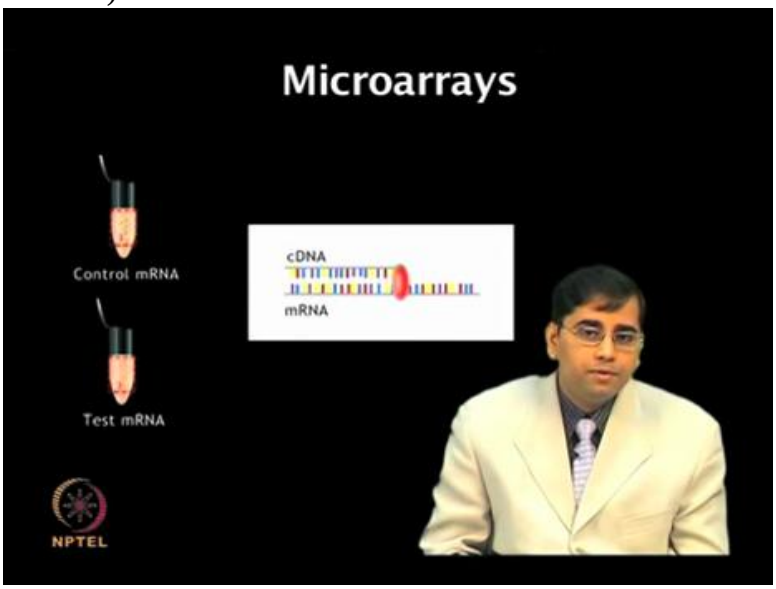
Transcriptomics: Study of all the mRNA molecules expressed by a particular cell type of an organism is known as transcriptomics. The transcriptomic analysis measures the genes that are being actively expressed at any given time and varies significantly with external environmental conditions. Various techniques such as microarrays, QRT PCR etc, have been widely used for transcriptional analysis.

(Refer Slide Time: 15:32)



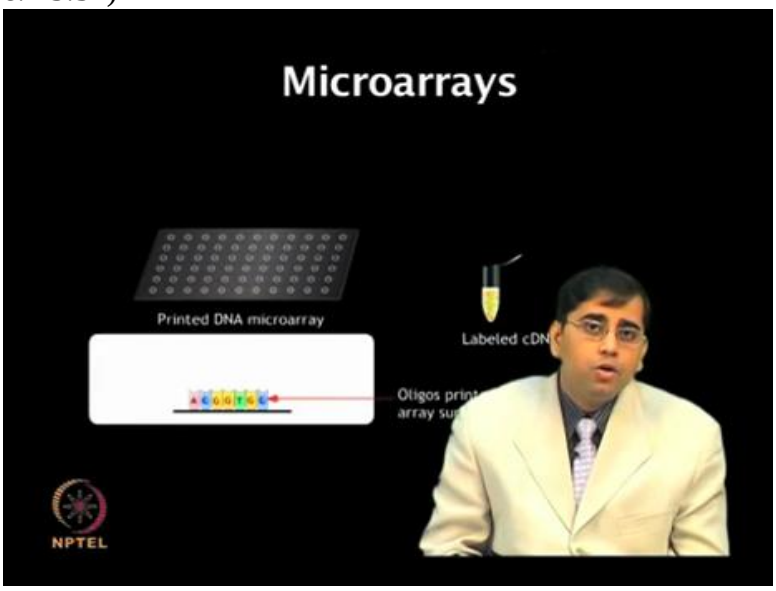
In a microarray experiment the mRNA from control and test samples are extracted….

(Refer Slide Time: 15:34)



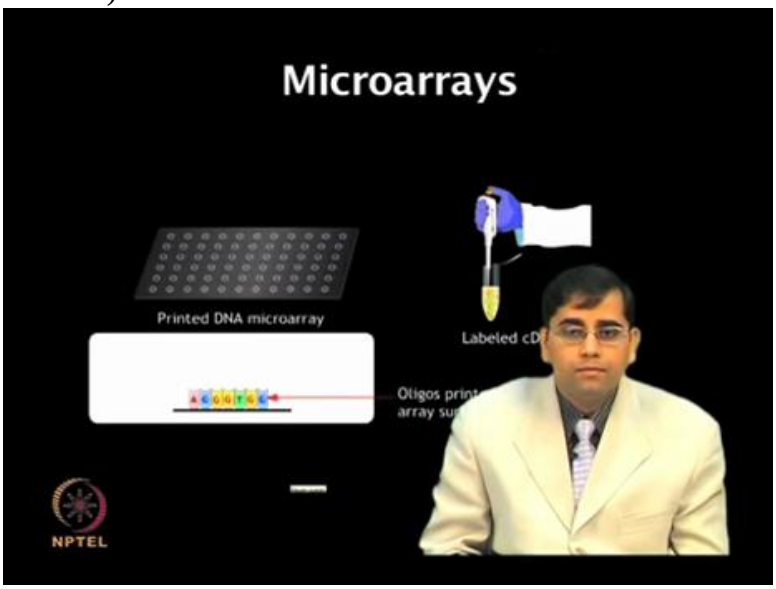
…and reverse transcribed …

(Refer Slide Time: 15:37)



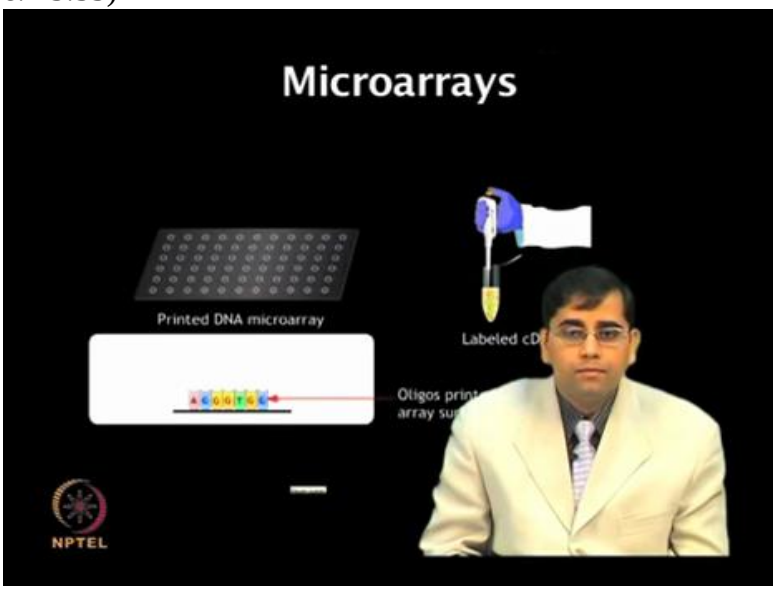
…into its corresponding cDNA.

(Refer Slide Time: 15:42)



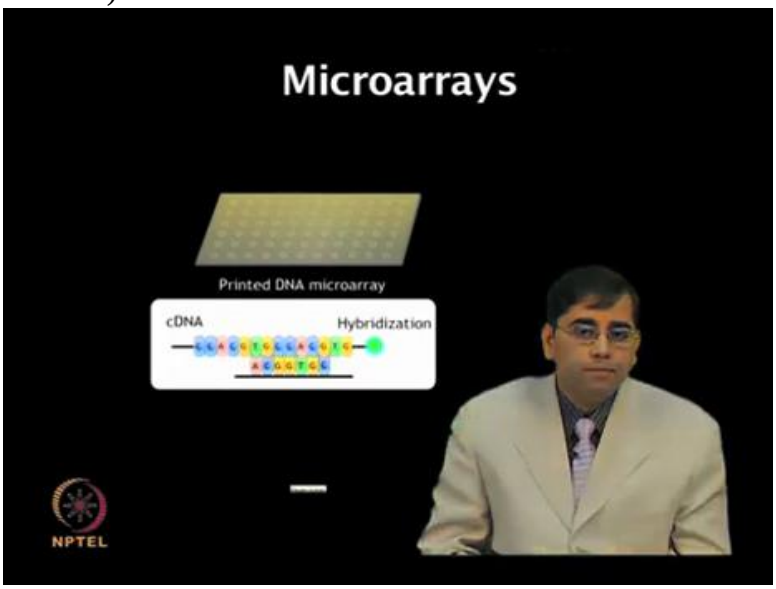
The cDNA samples are labeled with Cy5 and Cy3 dyes and

(Refer Slide Time: 15:55)



mixed cDNA sample is incubated on printed DNA microarray.

(Refer Slide Time: 16:16)

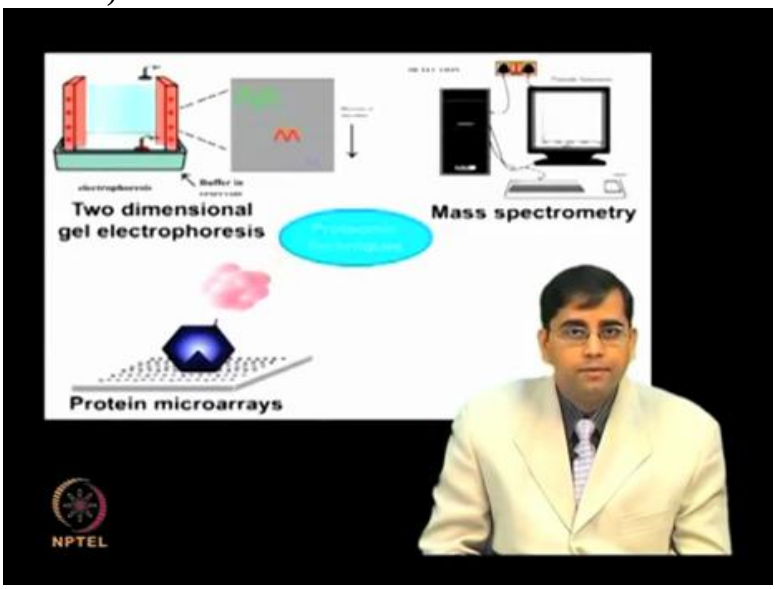


This allows hybridization to occur between the probed polynucleotide on array surface and the labeled cDNA sample of interest. In this manner, expression level of thousands of genes can be measured and analyzed simultaneously.

(Refer Slide Time: 16:23)

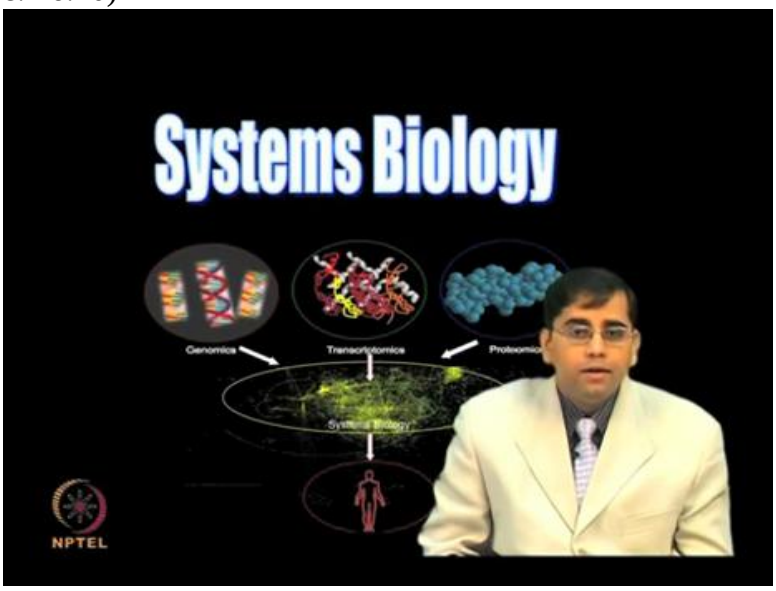
Different types of proteomics technologies such as…

(Refer Slide Time: 16:30)



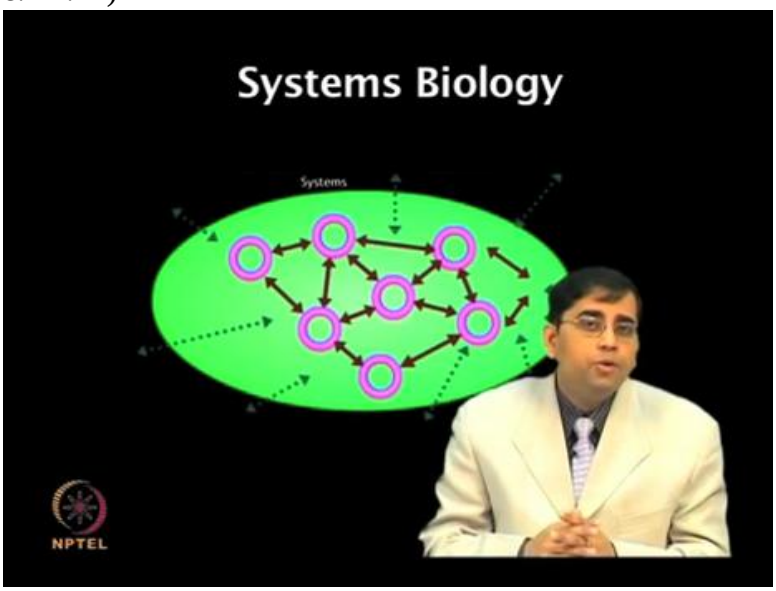
Two-dimensional electrophoresis, mass spectrometry, Microarrays and label-free techniques will be discussed in more detail later.

(Refer Slide Time: 16:40)



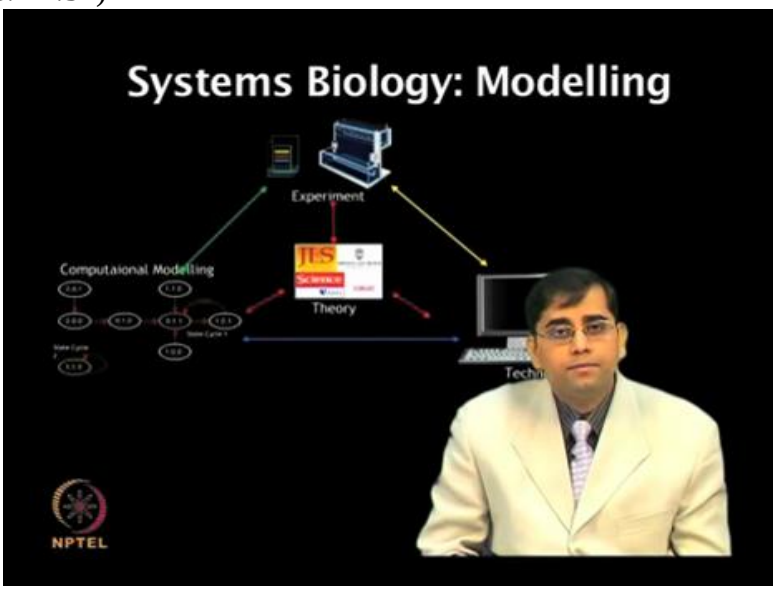
In Omics era, technological advancement in genomics, proteomics and metabolomics have generated large scale datasets in all the aspects of biology. These large datasets have motivated the computational biology and Systems approaches with objective of understanding the biological system as a whole.

(Refer Slide Time: 17:17)



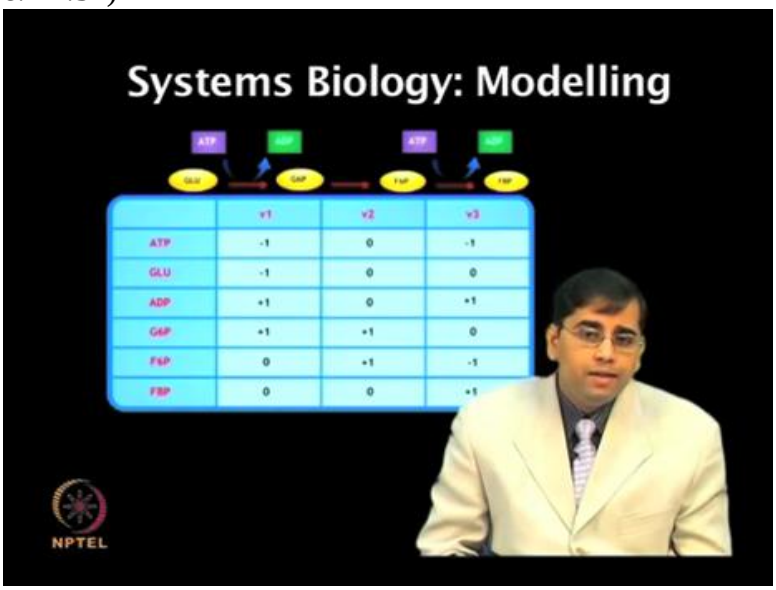
The system biology and biological network modeling aims to understand the biological processes…

Refer Slide Time: 17:34)



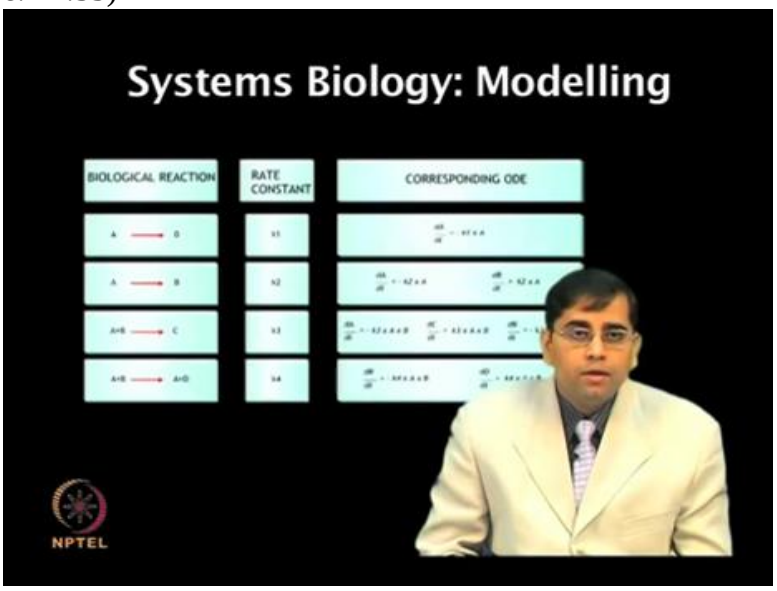
… as whole system rather than the isolated parts while synergistic application of experiment, theory, technology and modeling.

(Refer Slide Time: 17:37)



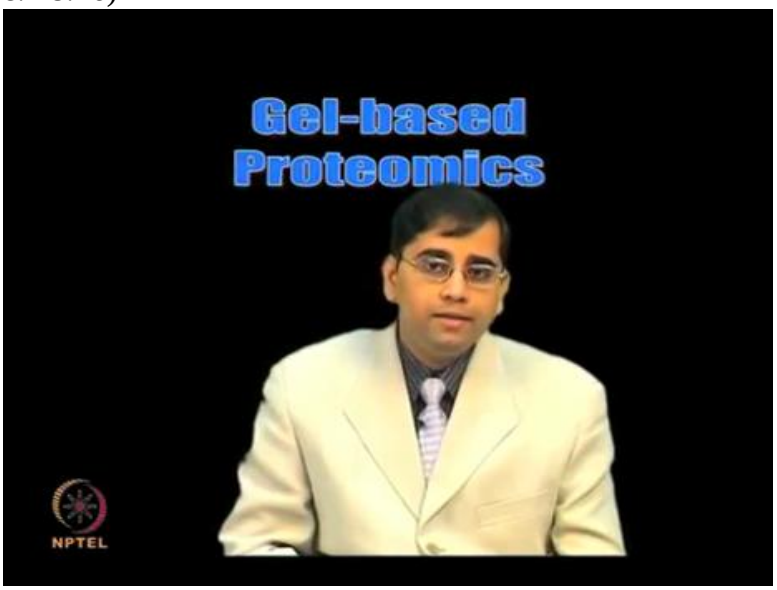
The systems level studies aim to develop computationally efficient and reliable models of underlying gene regulatory networks.

(Refer Slide Time: 17:53)



The quantitative analysis measures and aims to make models for precise kinetic parameters of system's network component. It also uses properties of network connectivity.

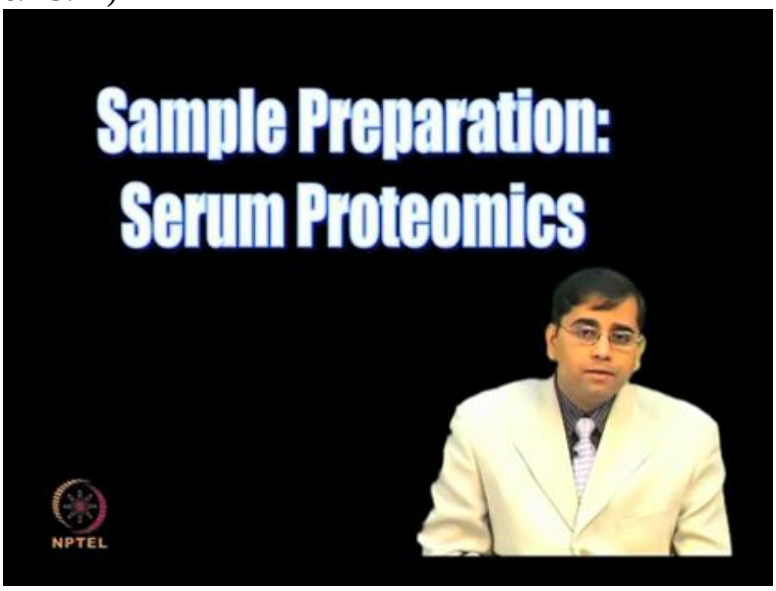
(Refer Slide Time: 18:16)



Several techniques used in proteomics typically aim to elucidate the expression, localization, interaction and cellular function of proteins.SDS PAGE, two-dimensional electrophoresis, difference gel electrophoresis are various commonly used gel-based proteomics. Protein extraction is the first step for the proteomic analysis.

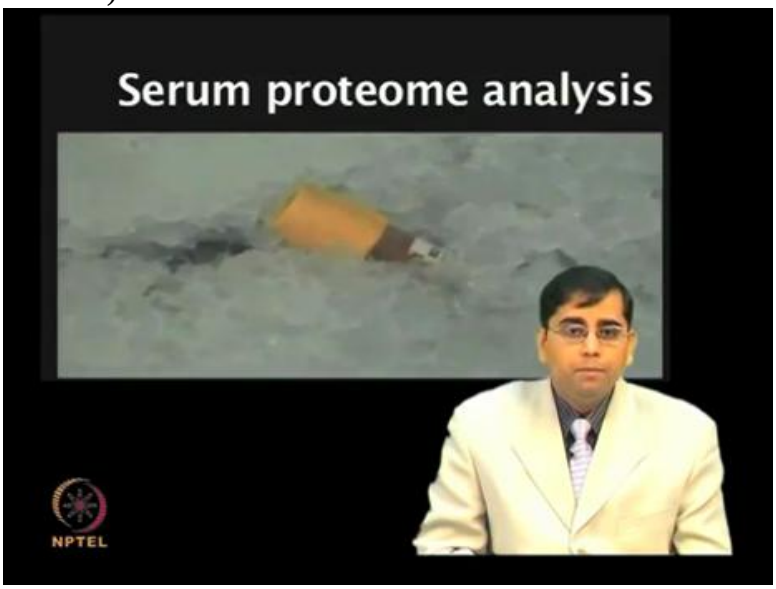
The protein extraction methods aim that most if not all the proteins in a cell or its organelle are extracted by the procedure and the presence of interfering components are reduced or minimized.

(Refer Slide Time: 19:12)



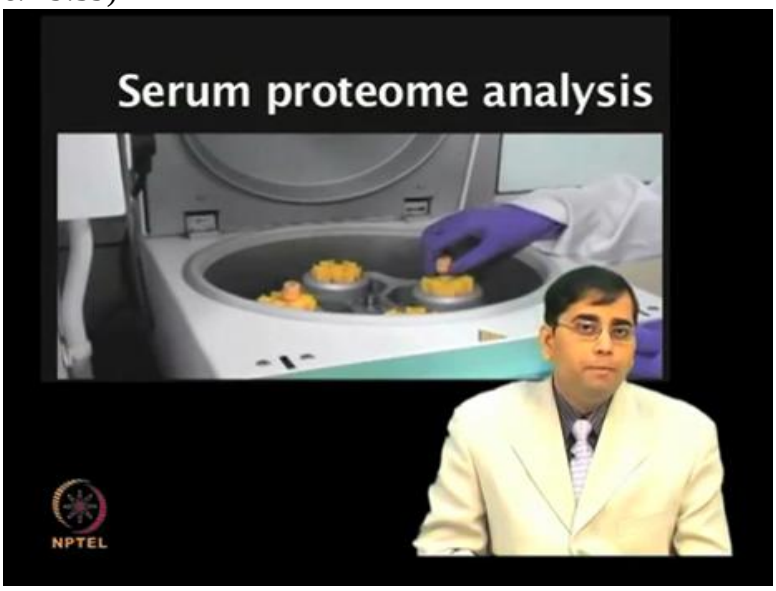
Different biological systems, different biological samples pose different types of challenges.

(Refer Slide Time: 19:36)



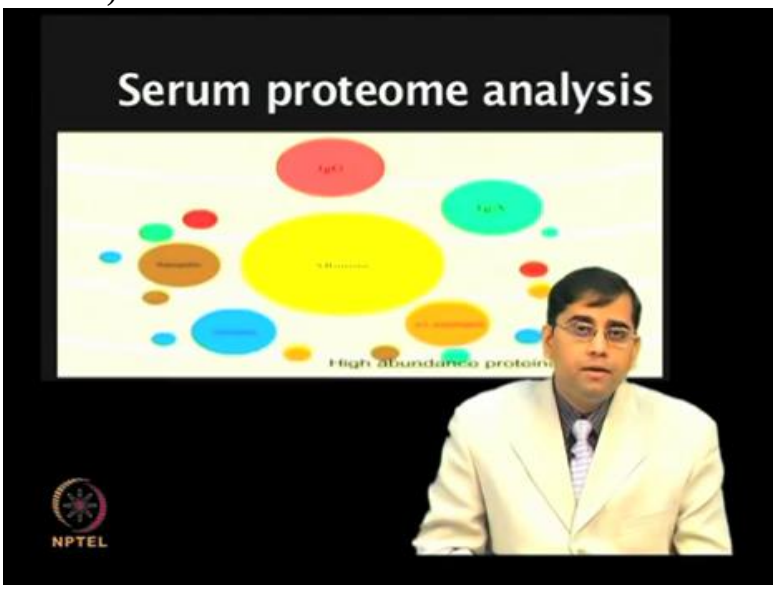
For example serum proteome analysis shown here illustrates that proteins in biological systems such as serum may have difference

(Refer Slide Time: 19:39)



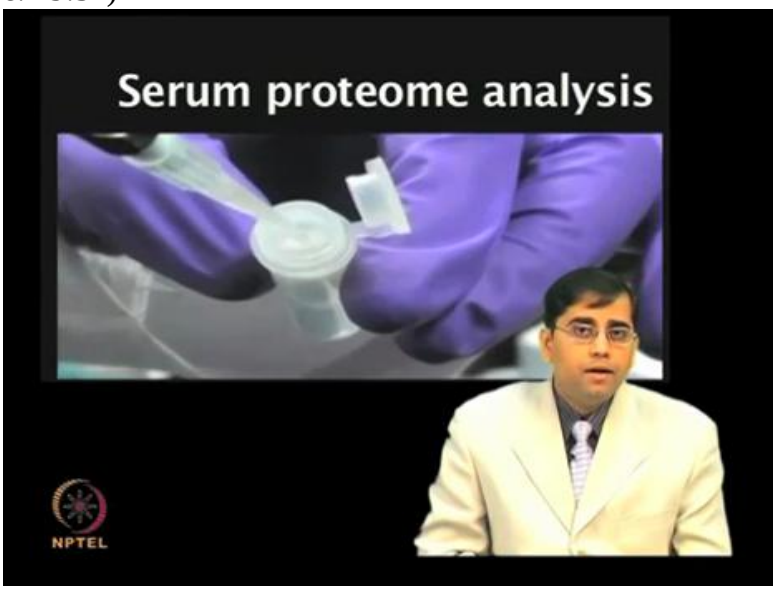
of several order of magnitude.

(Refer Slide Time: 19:42)



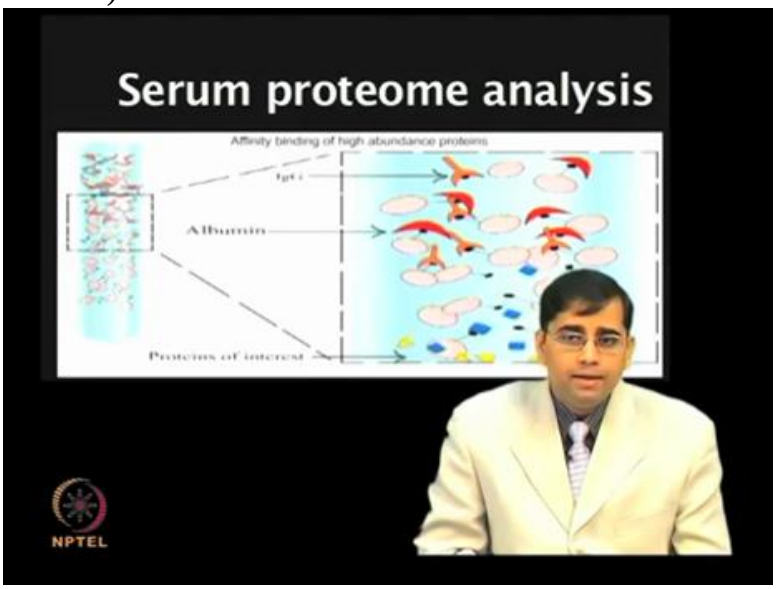
Albumin and immunoglobulin are the most abundant proteins in serum…

(Refer Slide Time: 19:52)



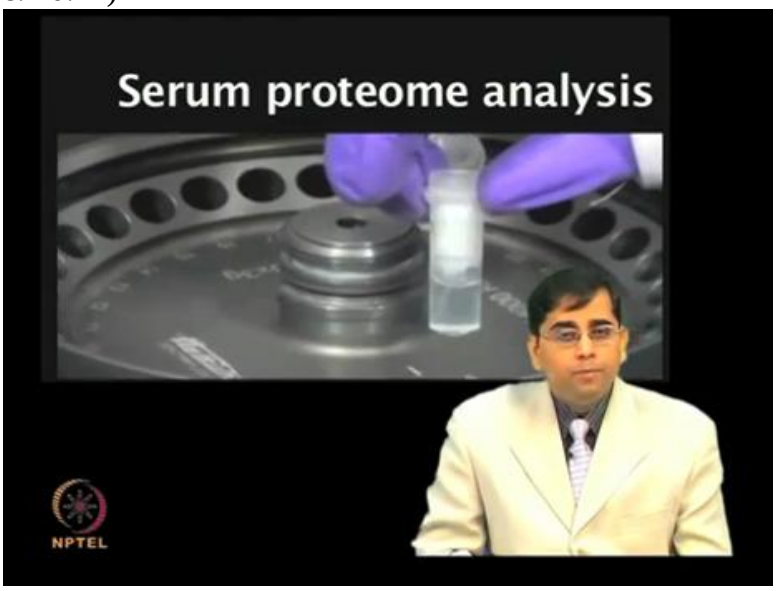
… which mask other low abundant proteins which are present in the lower concentration

(Refer Slide Time: 20:00)



It is therefore preferred to remove these high abundant proteins by using affinity chromatography methods.

(Refer Slide Time: 20:11)

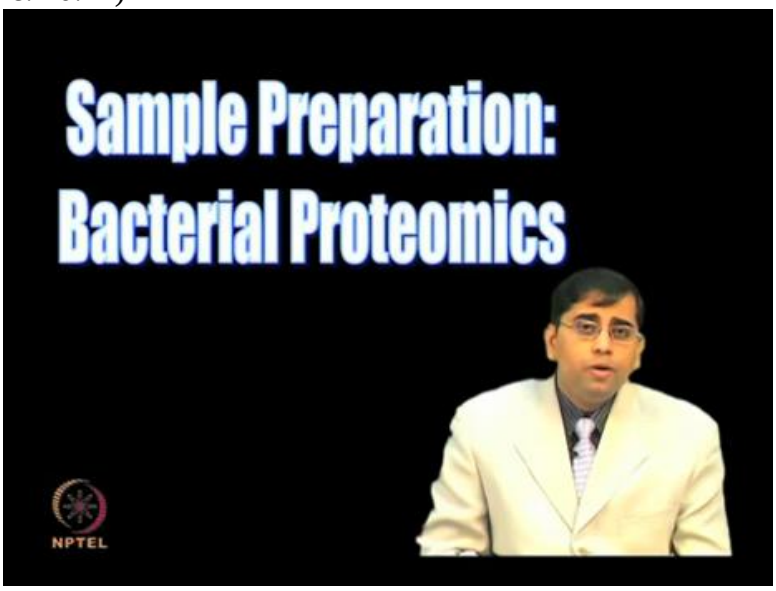


Once the serum has been processed using these chromatography methods…

(Refer Slide Time: 20:19)

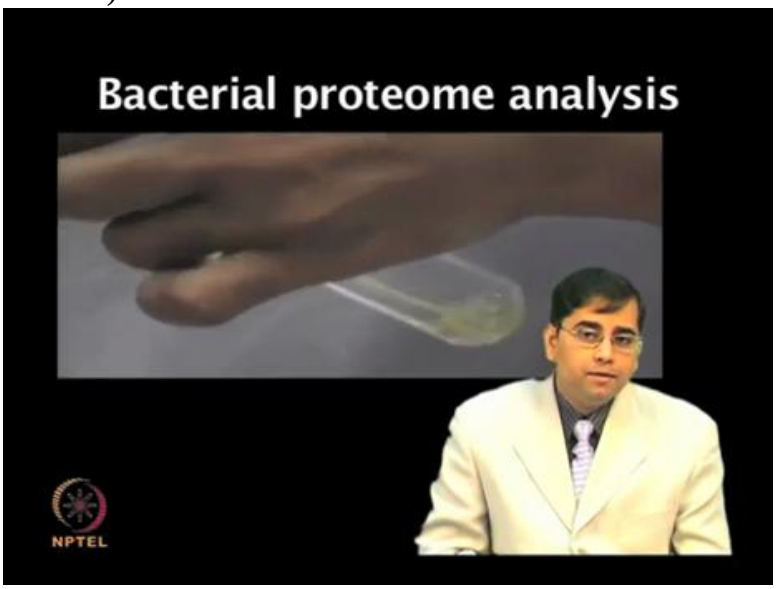
these proteins can be extracted.

(Refer Slide Time: 20:22)



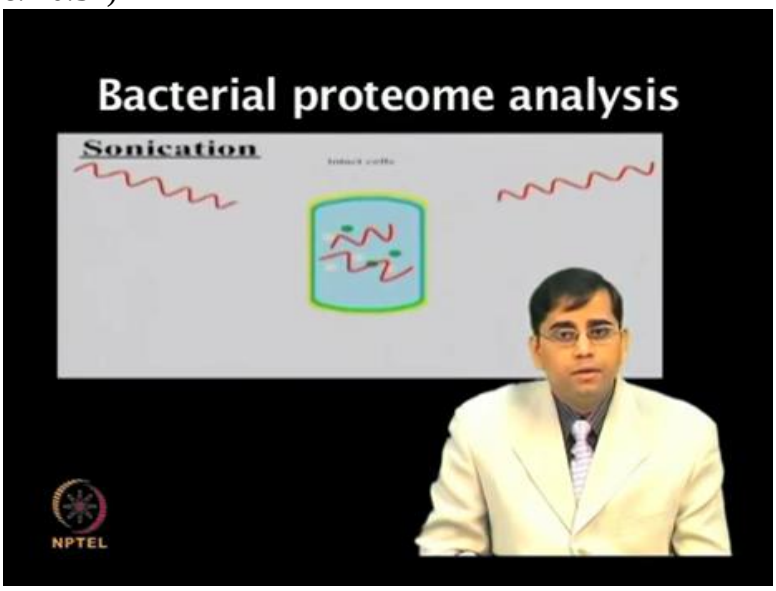
In bacterial protein sample preparation, sonication is an important step

(Refer Slide Time: 20:27)



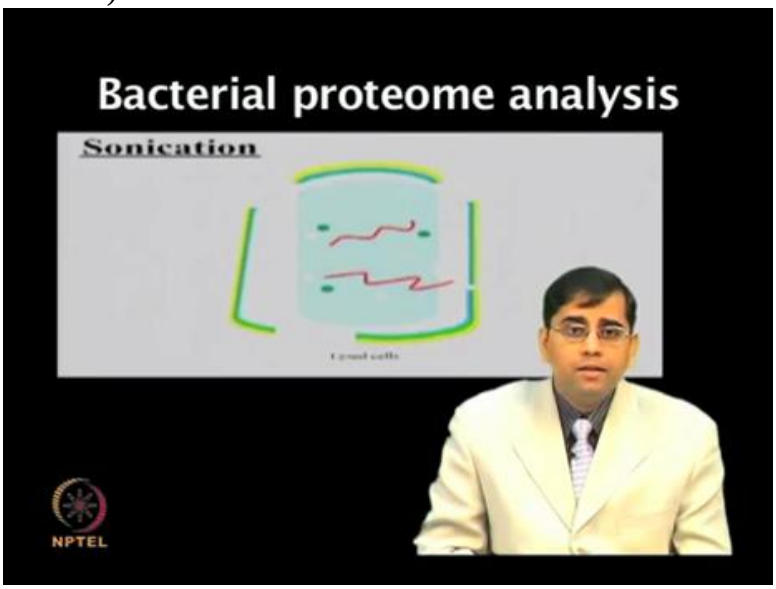
…to disrupt the bacterial membrane.

(Refer Slide Time: 20:31)



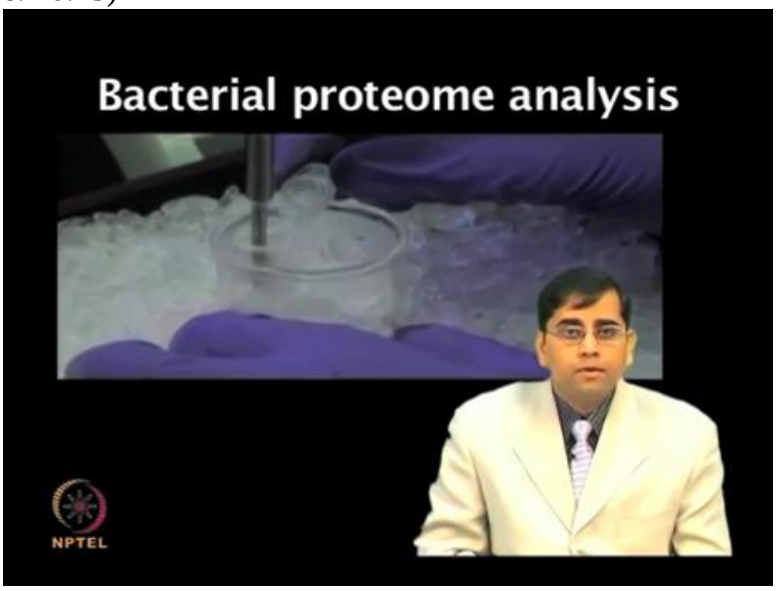
Sonciation breaks open the cellular contents…

(Refer Slide Time: 20:33)



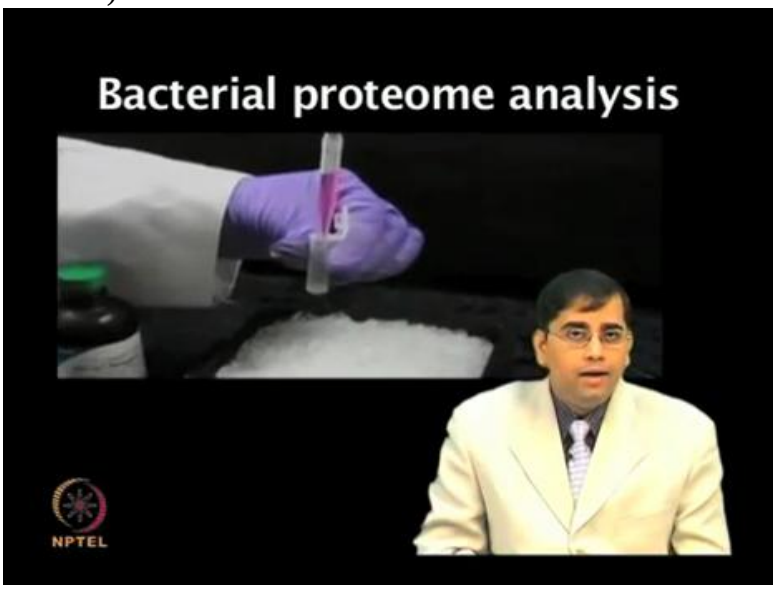
the cellular membrane to release the intracellular contents.

(Refer Slide Time: 20:43)



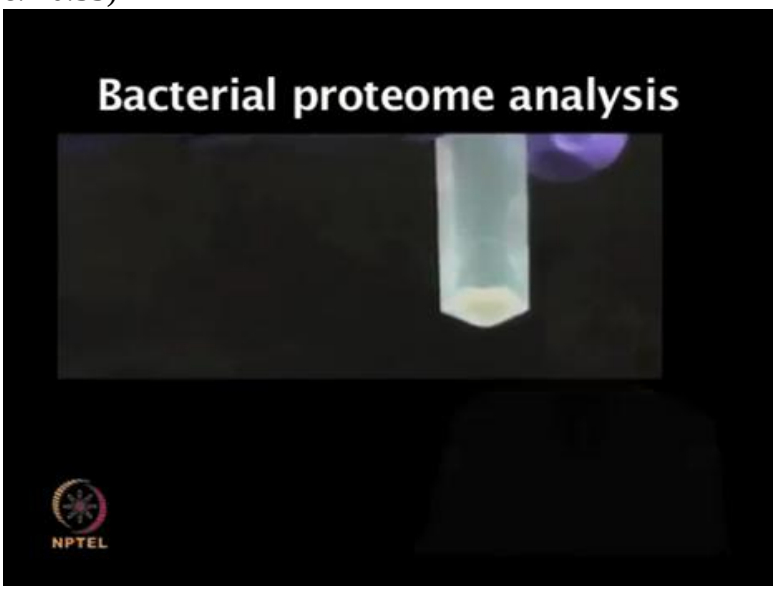
Protein extraction can be performed by using different methods…

(Refer Slide Time: 20:50)



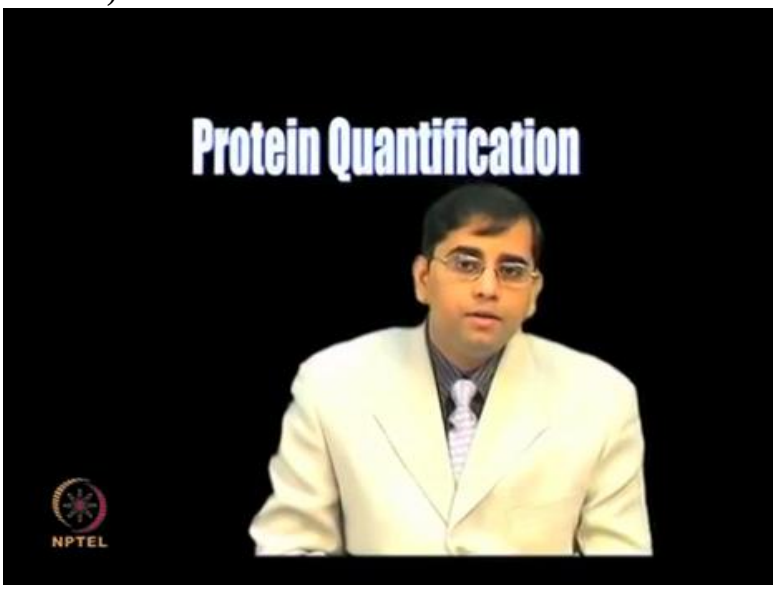
…and protein palettes are reconstituted in lysis buffer

(Refer Slide Time: 20:55)



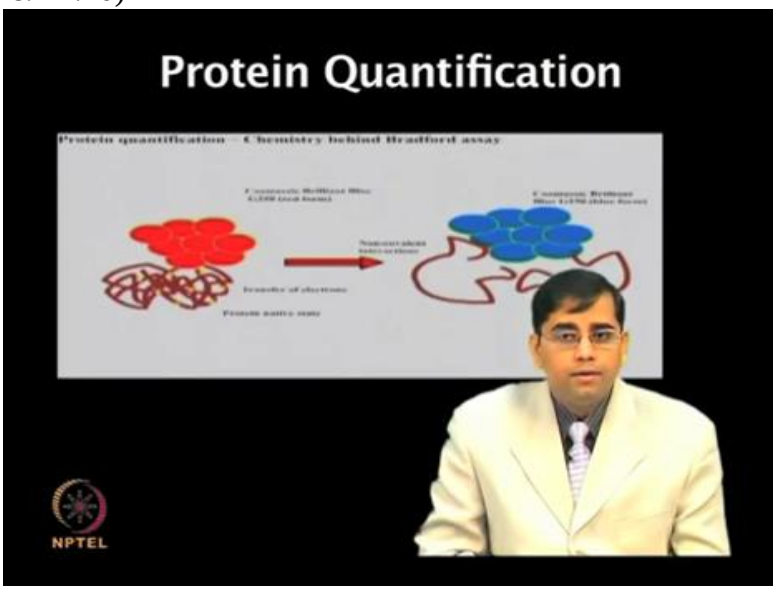
…for proteomic analysis.

(Refer Slide Time: 20:59)



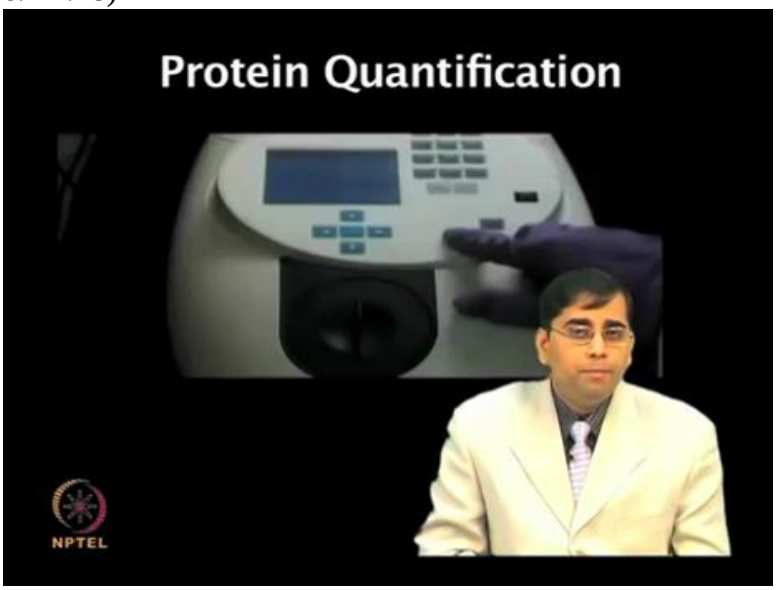
Protein quantification is sensitive to detergents or certain ions. Therefore it is crucial to select the correct quantification method.

(Refer Slide Time: 21:20)



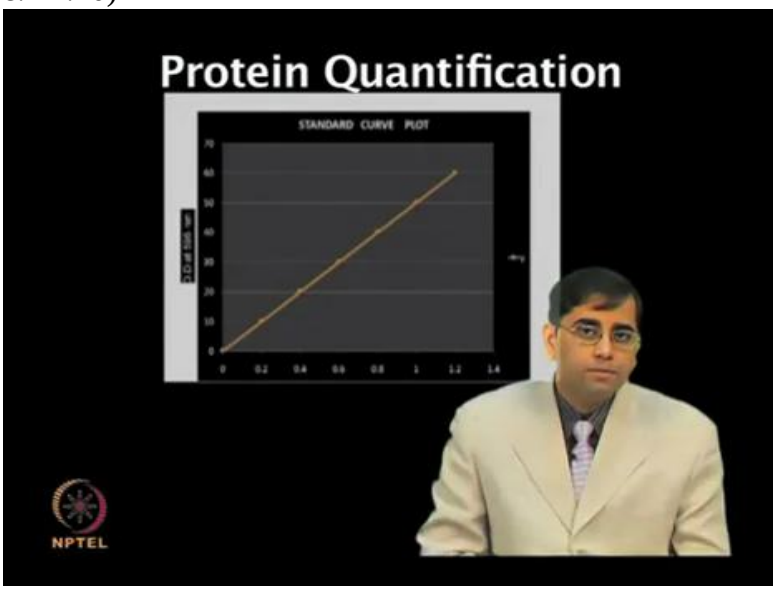
In Bradford color reagent, transfer of electrons converts the dye to its blue form thereby giving the solution blue color.

(Refer Slide Time: 21:28)



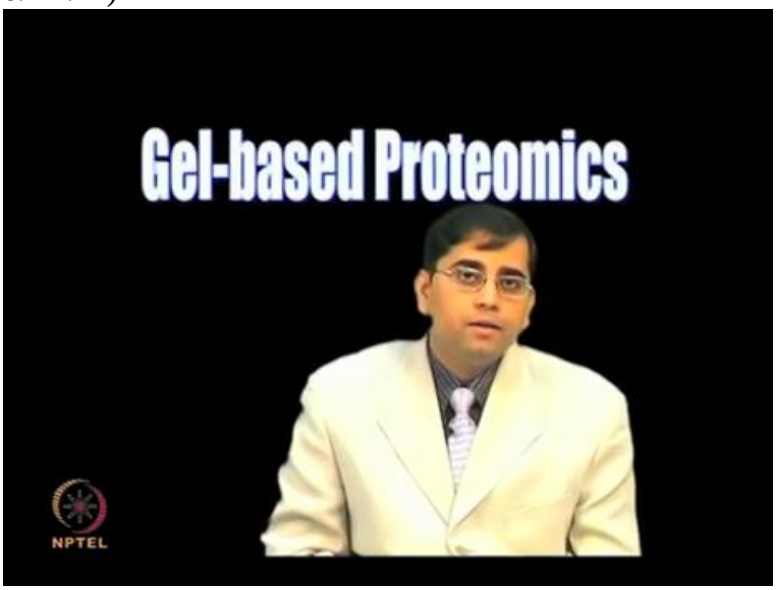
Absorbance of standard and unknown protein samples can be measured at 595 nanometers…

(Refer Slide Time: 21:40)



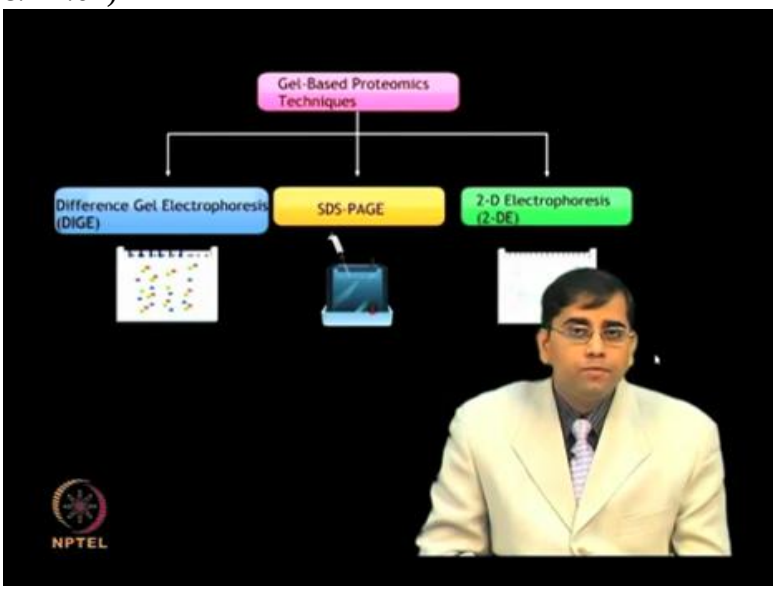
…and protein concentration can be determined from the standard plot of the absorbance values.

(Refer Slide Time: 21:47)



In gel-based proteomics, proteins are commonly analyzed using SDS PAGE and two-dimensional gel electrophoresis.

(Refer Slide Time: 22:07)

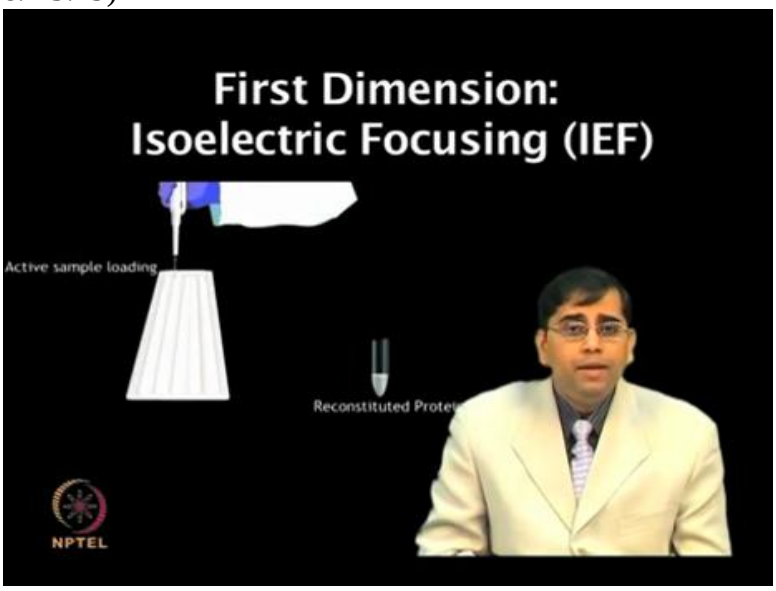


Separation in SDS PAGE occurs almost excessively on the basis of molecular weight whereas in 2DE, the first dimension separation is based on isoelectric point and second dimension separation based on molecular weight. Some of the limitations of two dimensional electrophoresis can be overcome by difference gel electrophoresis or DIGE technique.

2DE or DIGE in combination with mass spectrometry has been the standard technique for proteome analysis.

The two-dimensional electrophoresis involves protein separation on a pH gradient based on their isoelectric point using isoelectric focusing followed by separation in second dimension using SDS PAGE

(Refer Slide Time: 23:19)

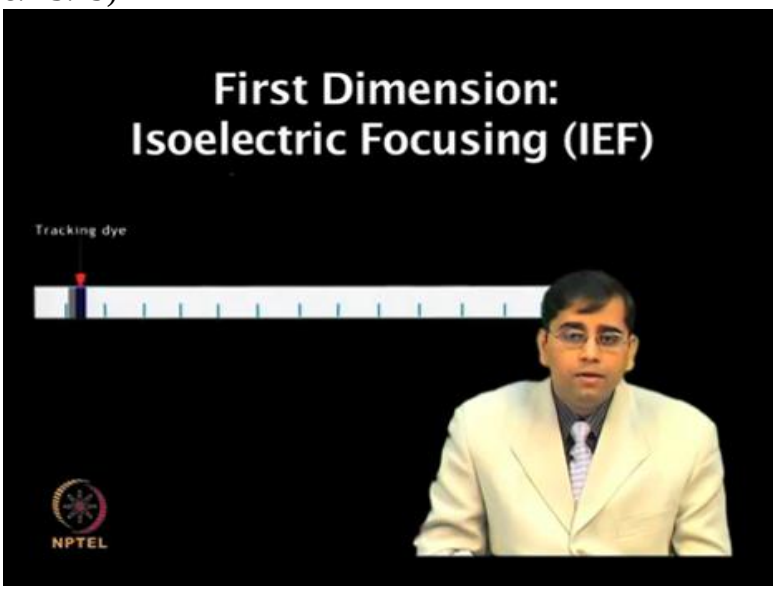


To perform 2DE add the reconstituted protein sample to the rehydration tray and place IPG strip for rehydration.

(Refer Slide Time: 23:23)

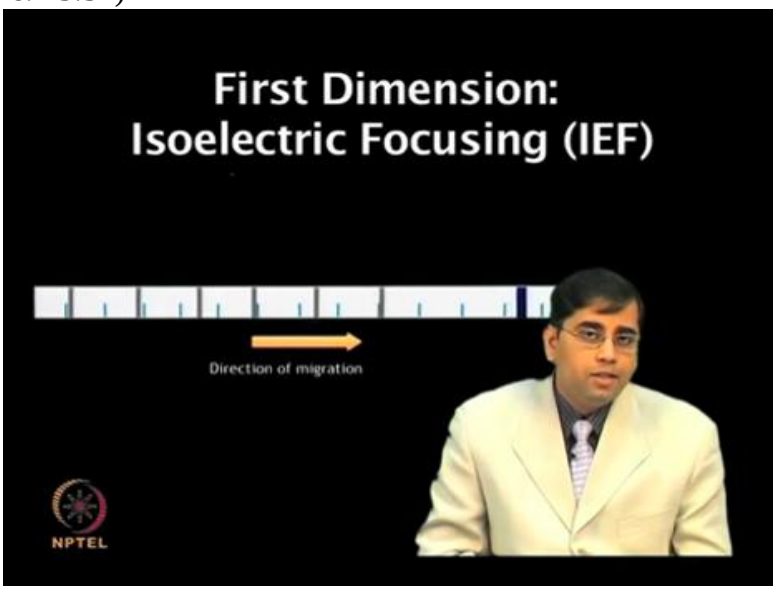
Isoelectric focusing involves the application of electric field which causes the proteins to migrate to the position on the pH gradient strips…

(Refer Slide Time: 23:49)



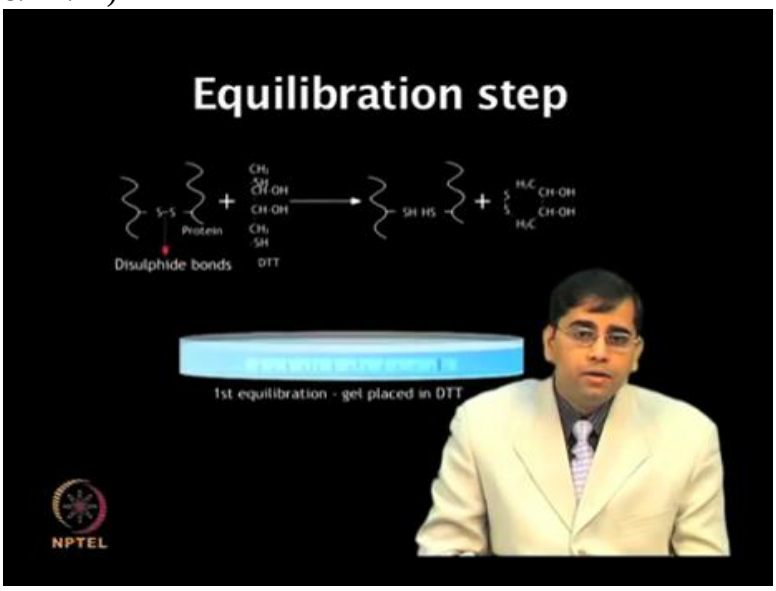
that matches the pI of the specific protein after which it does not move in the electric field owing to the lack of charge.

(Refer Slide Time: 23:57)



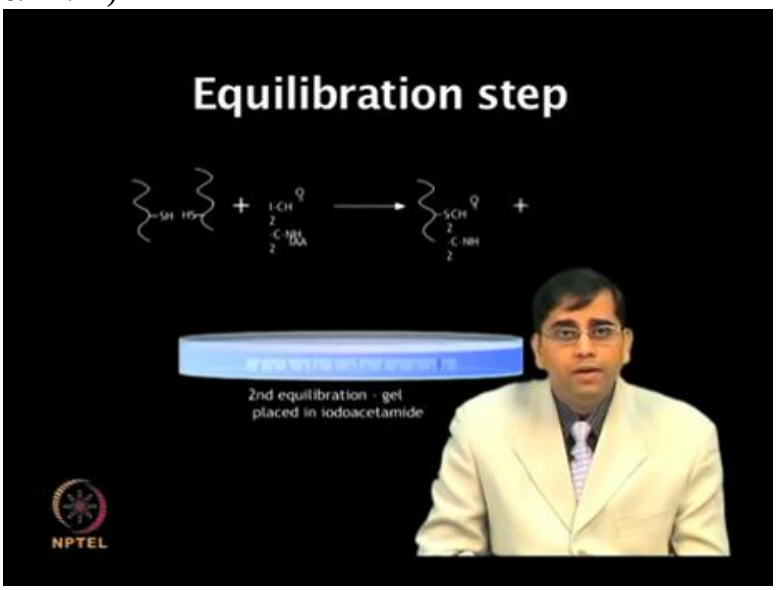
The proteins migrate along the strip and come to rest at a point where their net charge becomes zero known as isoelectric point.

(Refer Slide Time: 24:11)



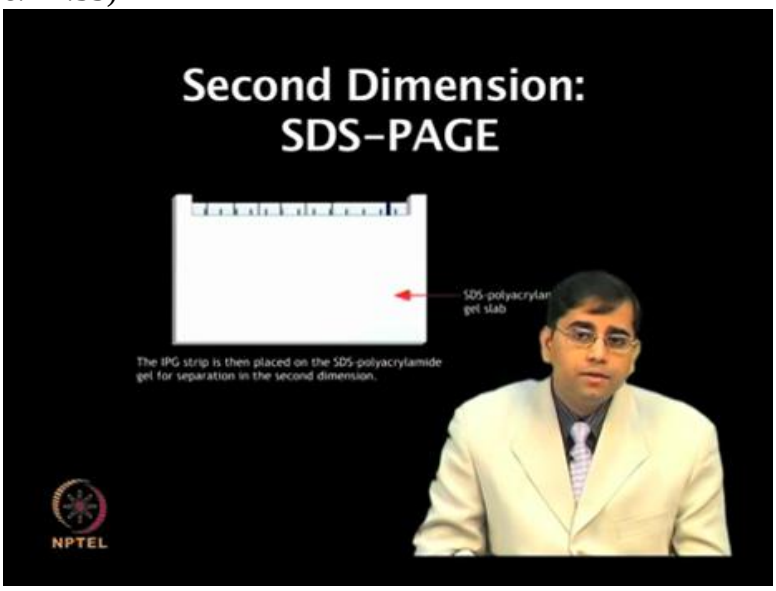
Prior to the second dimension separation, an equilibration step is required. In equilibration, dithiothreitol brings about cleavage of the protein disulfide bonds…

(Refer Slide Time: 24:24)



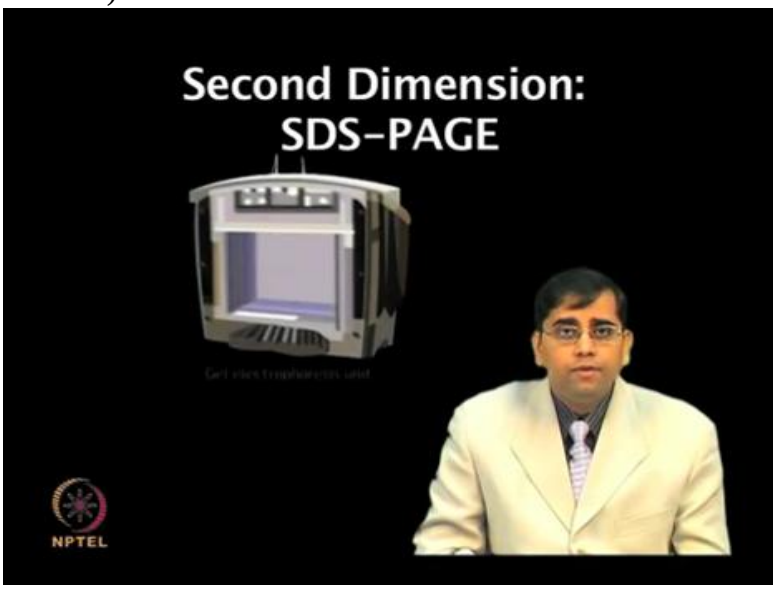
…while iodoacetamide prevents reformation of these bonds by binding to the sulfhydryl groups

(Refer Slide Time: 24:35)



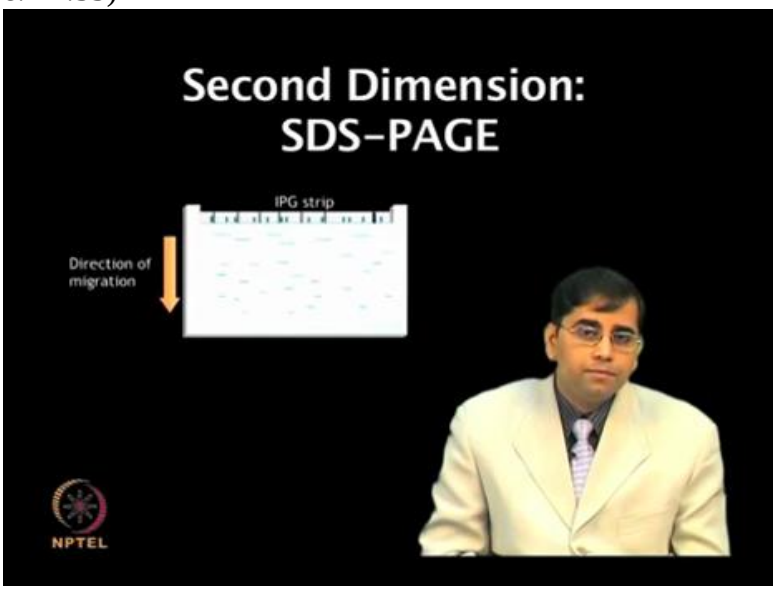
On SDS PAGE gel, proteins get separated.

(Refer Slide Time: 24:41)



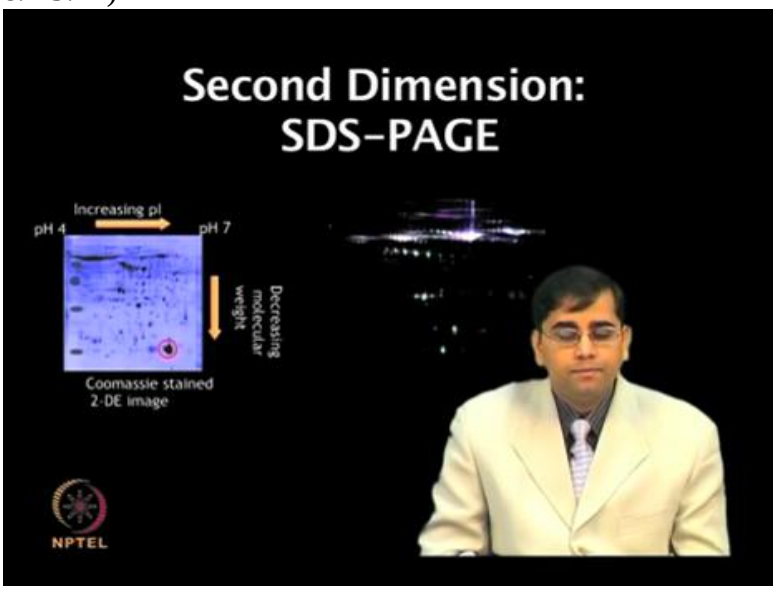
On the basis of their molecular weight, with the low molecular weight proteins having high mobility …

(Refer Slide Time: 24:55)



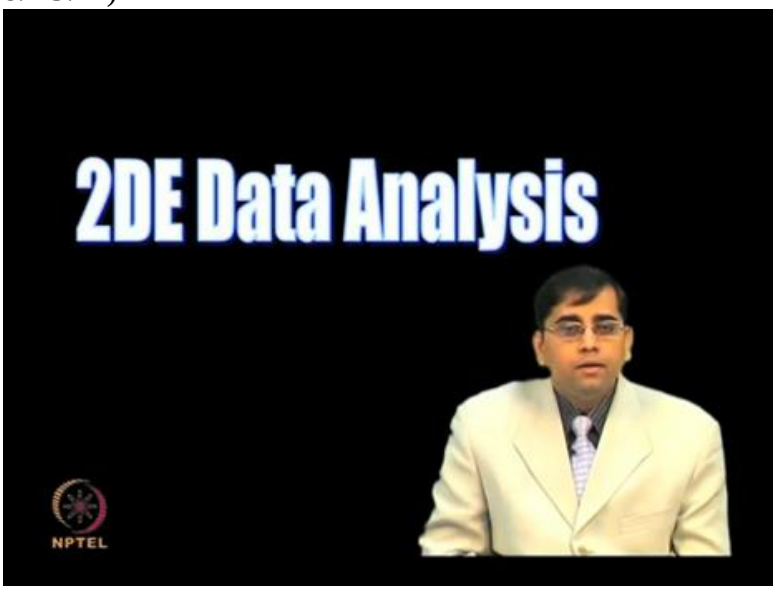
…are migrated further through the gel and the high molecular weight proteins remain close to the point of sample application.

(Refer Slide Time: 25:11)



Gel can be visualized by different staining methods such as Coomassie staining, silver staining and Cyanine dyes.

(Refer Slide Time: 25:12)

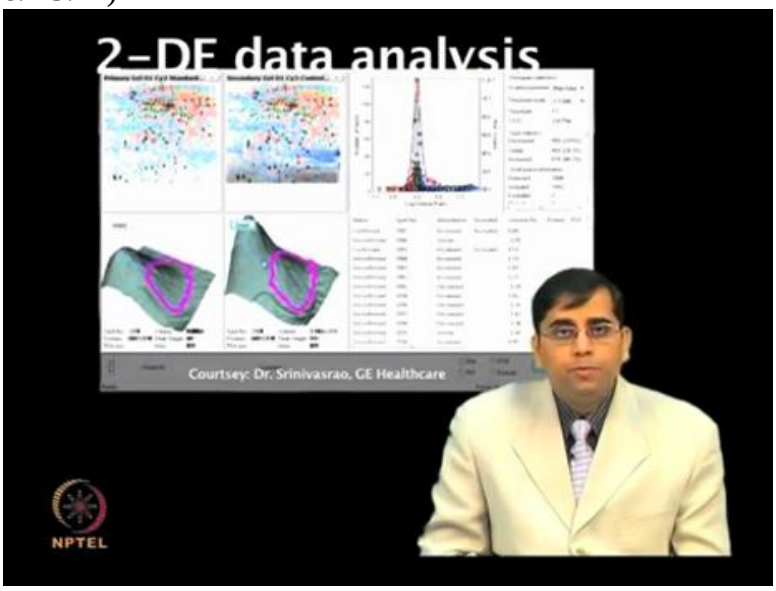


The gel data analysis will be discussed…

(Refer Slide Time: 25:16)

…with an application expert of GE Healthcare

(Refer Slide Time: 25:21)

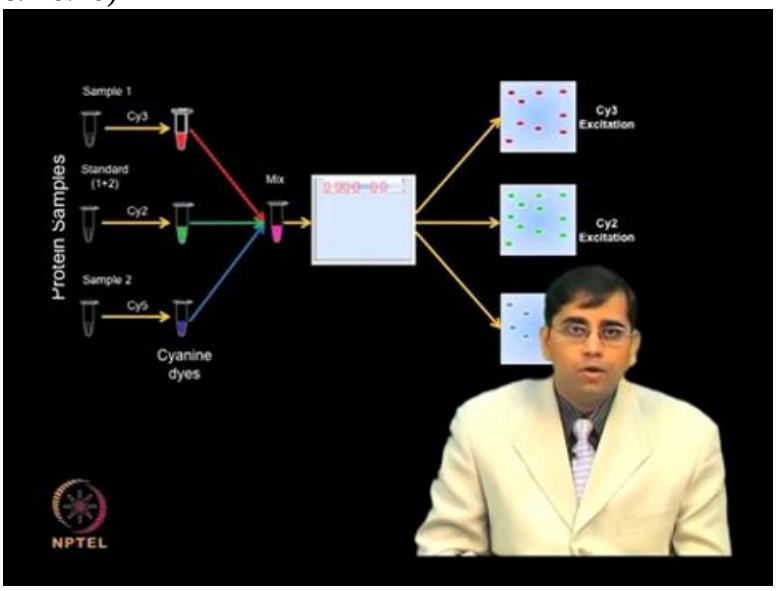


The gel analysis involves image processing, detection of spots, making matched sets, landmarking, viewing histograms etc. Various information regarding the spots such as their area, volume, intensity; and statistical parameters such as standard deviations can also be calculated.

(Refer Slide Time: 25:54)

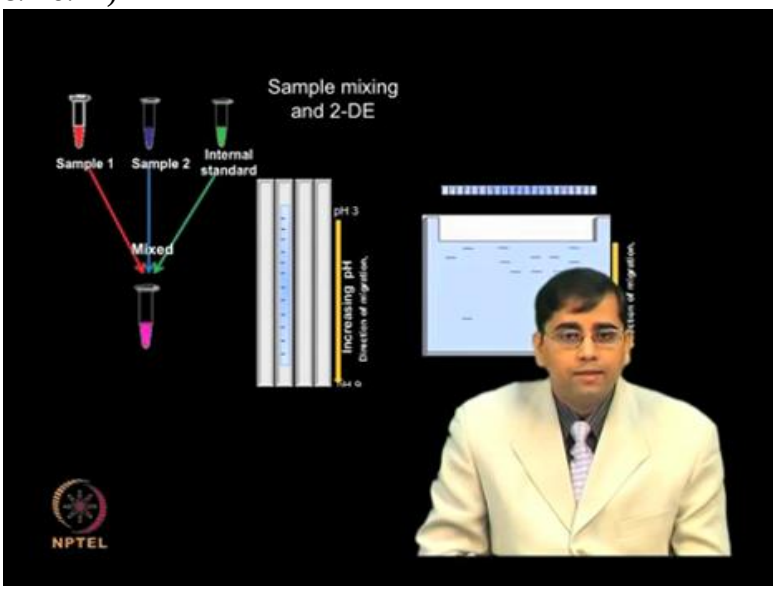
Two-dimensional electrophoresis has high resolving power but it has several limitations such as staining artifacts and reproducibility in gel to gel.

(Refer Slide Time: 26:10)



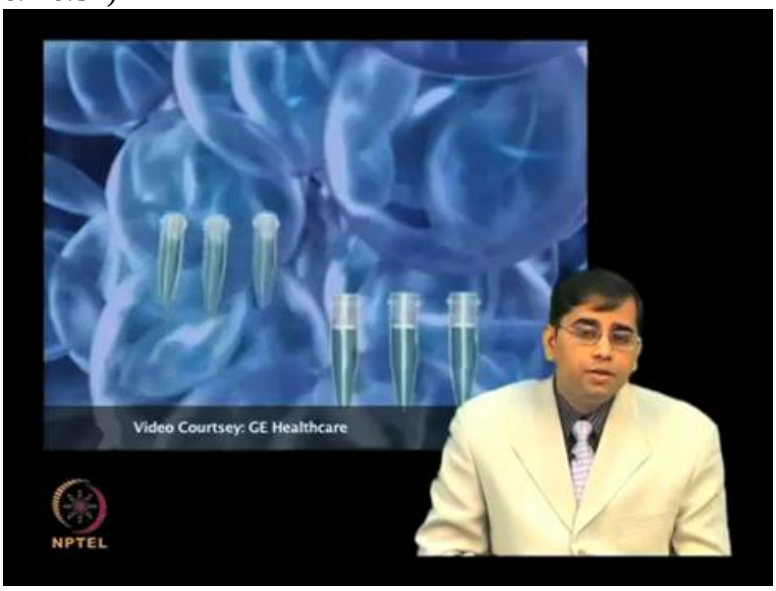
Fluorescence two-dimensional difference in gel electrophoresis or 2D DIGE is an advanced 2D technique that allows for…

(Refer Slide Time: 26:22)



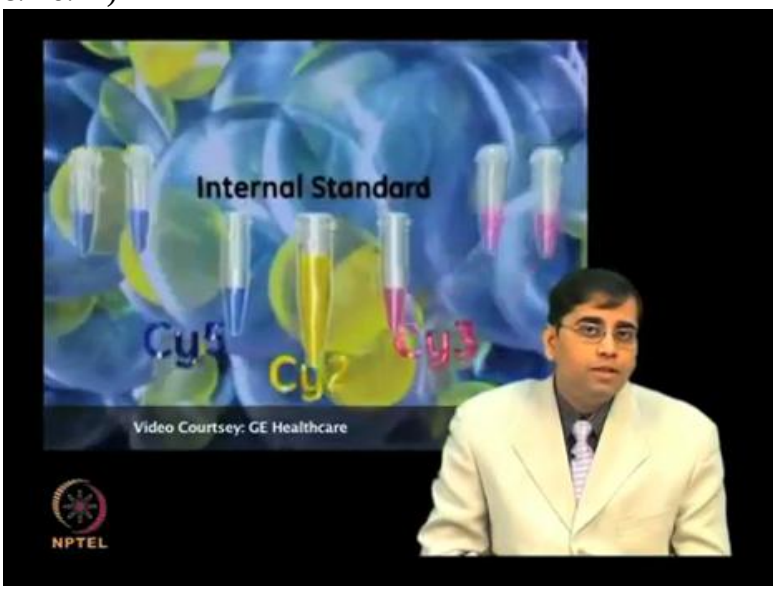
…accurate quantitation with statistical confidence while controlling non-biological variations

(Refer Slide Time: 26:32)



In DIGE, proteins extracted from different types of cells or tissue samples are labeled with different fluorescent reagents…

(Refer Slide Time: 26:42)

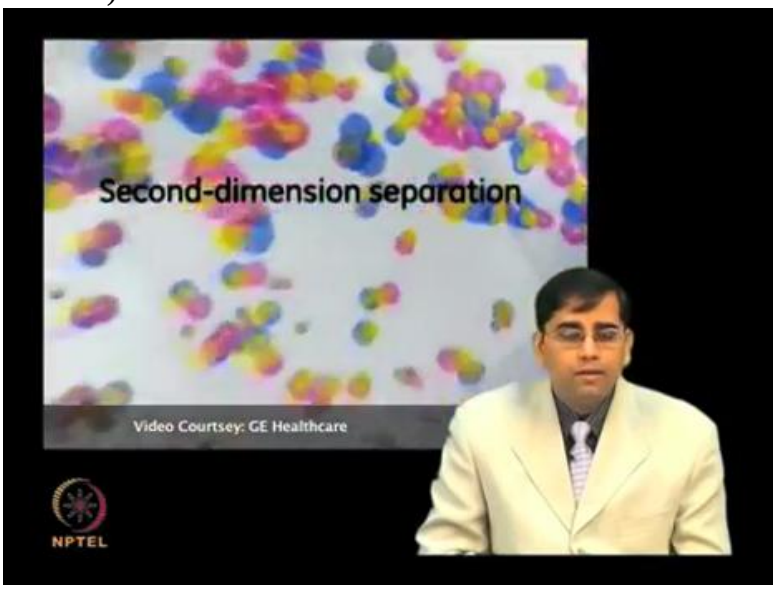


… such as Cy2, Cy3 and Cy5 …

(Refer Slide Time: 26:49)

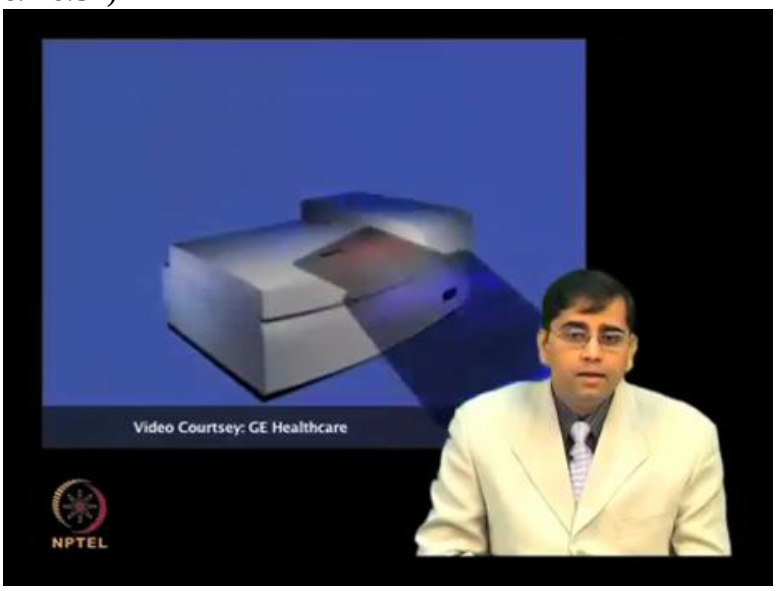
…mixed and then…

(Refer Slide Time: 26:53)



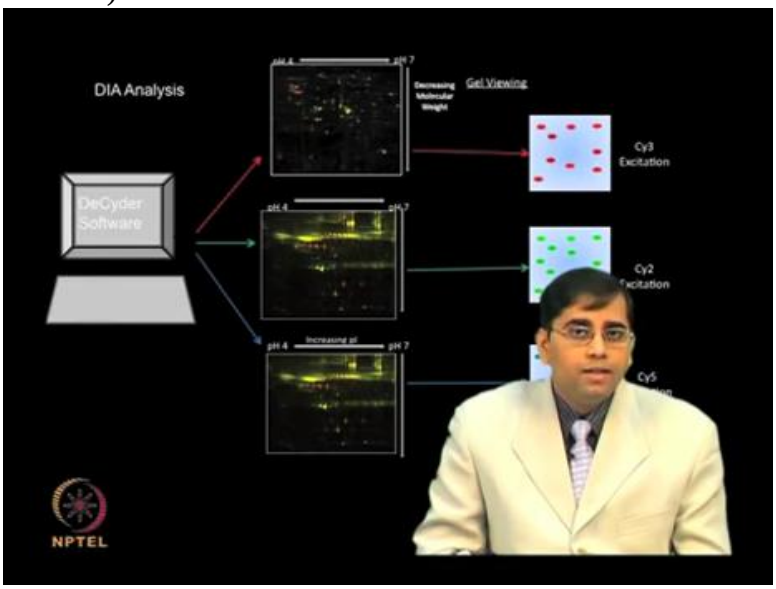
separated by two dimensional electrophoresis on a single gel.

(Refer Slide Time: 26:57)



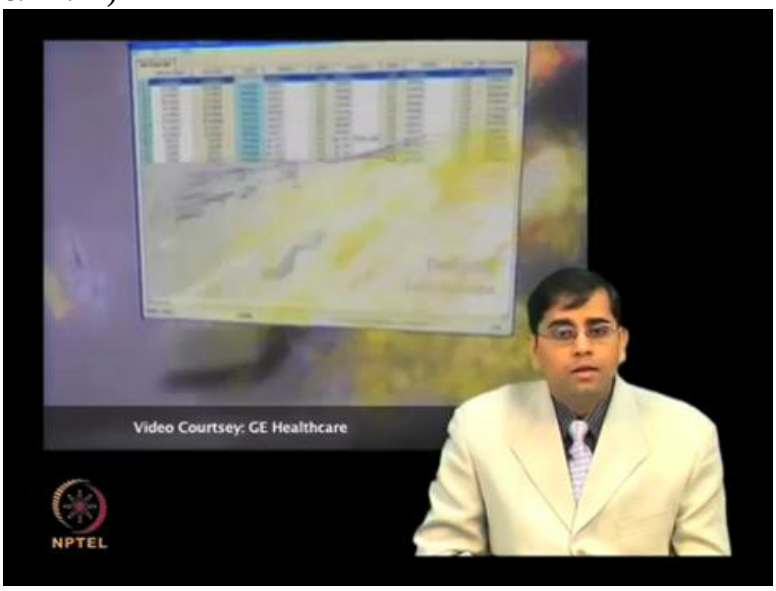
The proteins are detected separately using the excitation wavelength…

(Refer Slide Time: 27:08)



… specific to the different fluorescent reagents Cy2, Cy3 and Cy5.

(Refer Slide Time: 27:12)

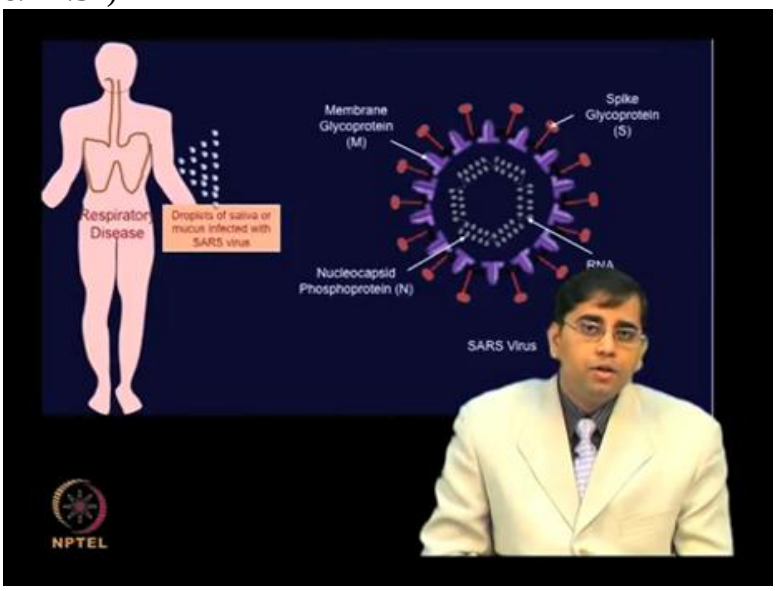


The commercial software such as DeCyder facilitate the automated analysis of DIGE gels and provide Differential Expression Analysis, Principal Component Analysis, Pattern and Discriminant Analysis.

(Refer Slide Time: 27:34)

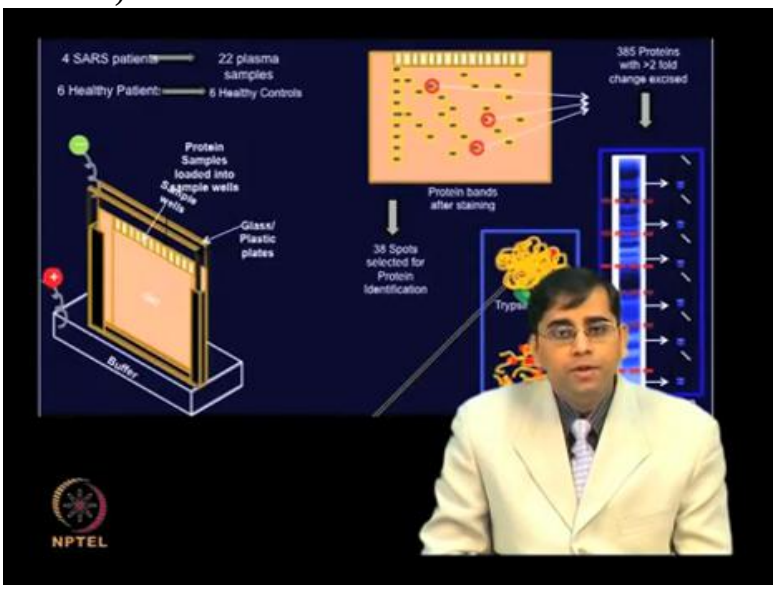
Two-dimensional electrophoresis, DIGE followed by mass spectrometry technique has been applied for…

(Refer Slide Time: 27:37)



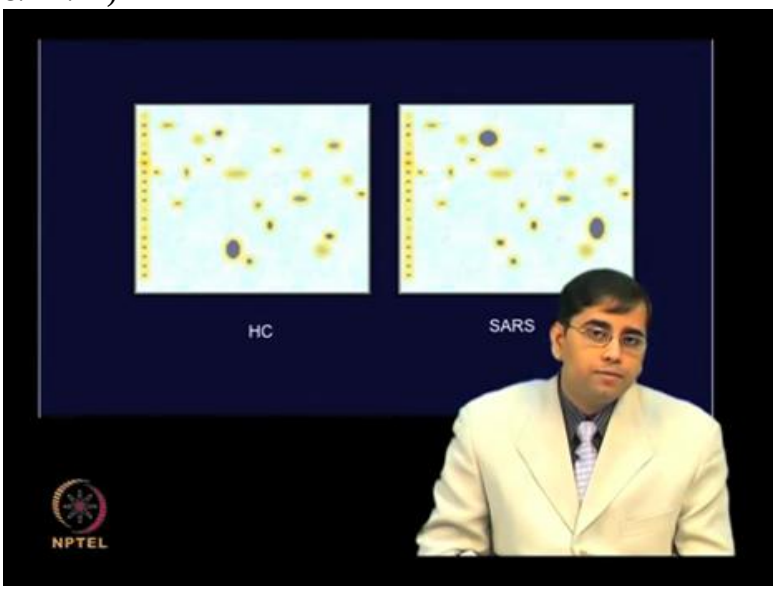
many…

(Refer Slide Time: 27:40)



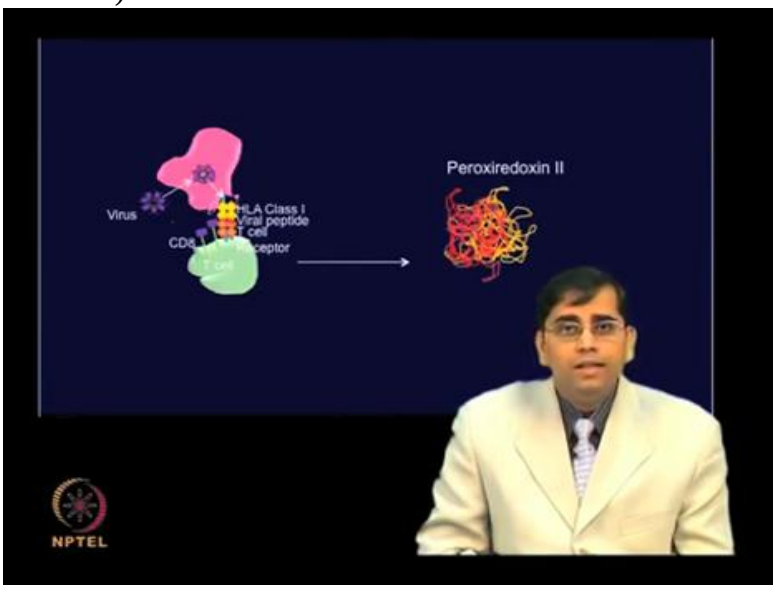
…applications. Some of these applications…

(Refer Slide Time: 27:42)



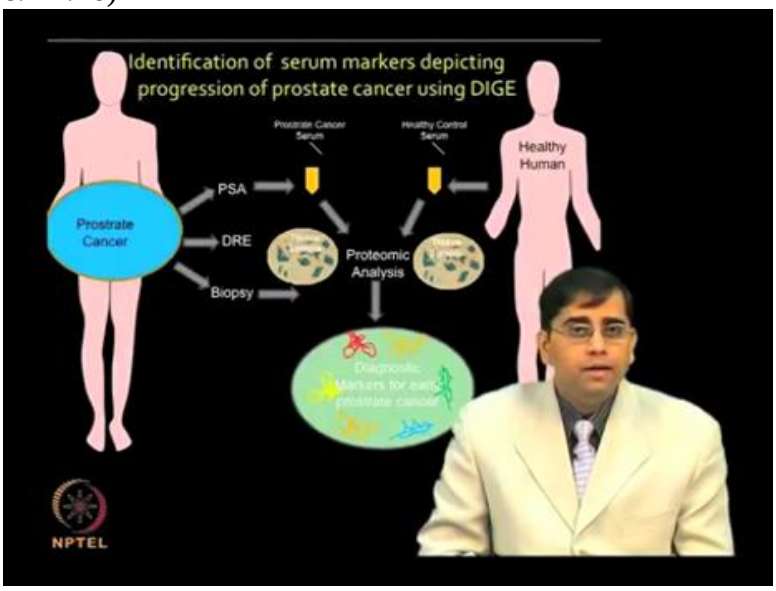
will be discussed…

(Refer Slide Time: 27:46)



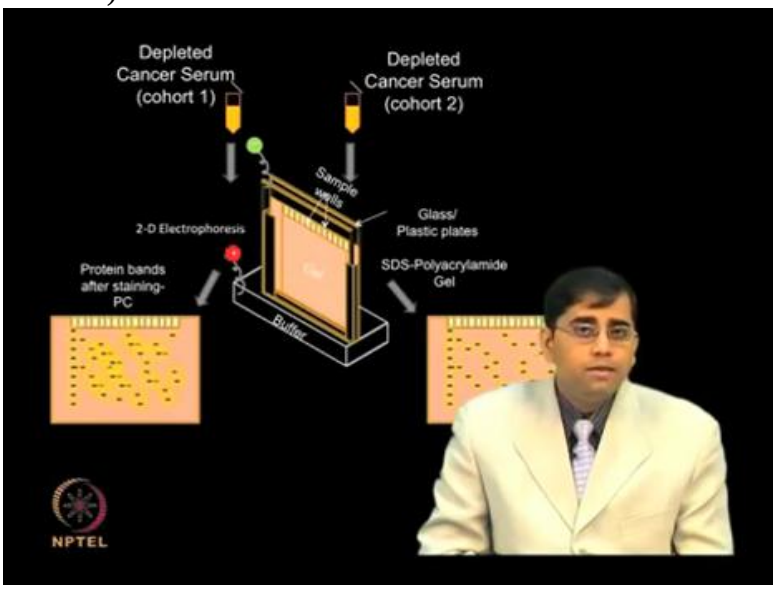
in this module…

(Refer Slide Time: 27:46)



of…

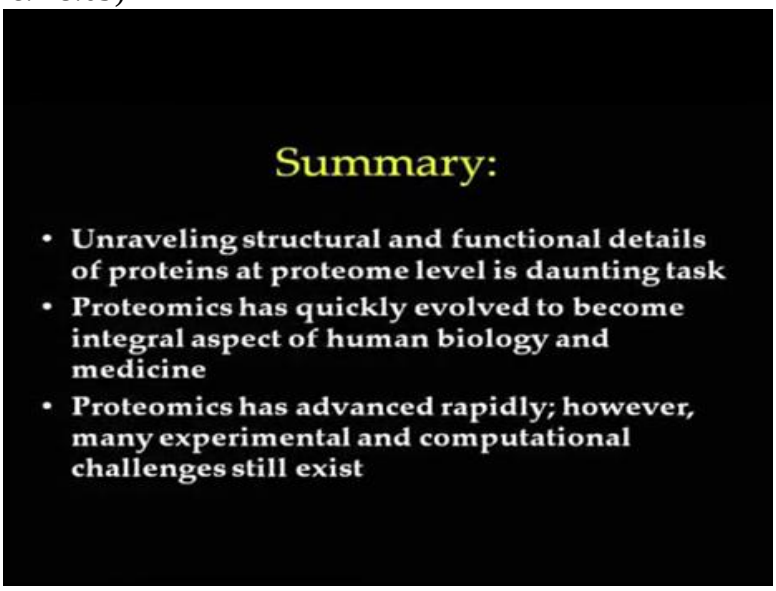
(Refer Slide Time: 27:50)



gel-based proteomics.

This century is considered as century of biology in which life science research is undergoing a profound transformation by employing various Omics technologies.

(Refer Slide Time: 28:09)



In summary unraveling structural and functional details of proteins at the proteomic scale is very daunting task. However proteomics has come to mean virtually everything in protein research and it has quickly evolved to become an integral aspect of human biology and medicine.

During the subsequent lectures I will take you to a journey of protein and proteomics research by providing basic concepts and details of proteomic techniques. I hope it will enthuse you to learn about proteomic techniques and proteomic concepts. Thank you for your attention.