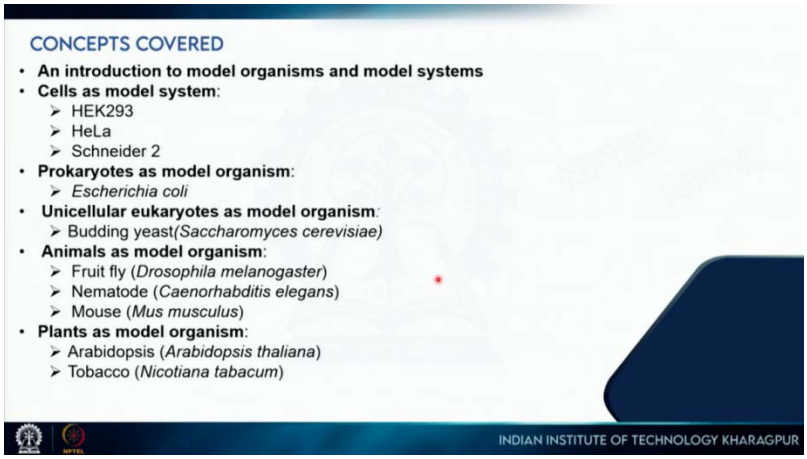
Introduction to Complex Biological Systems
Professor Dibyendu Samanta and Professor Soumya De
Department of Bioscience and Biotechnology
Indian Institute of Technology, Kharagpur

Lecture 40
Model organisms

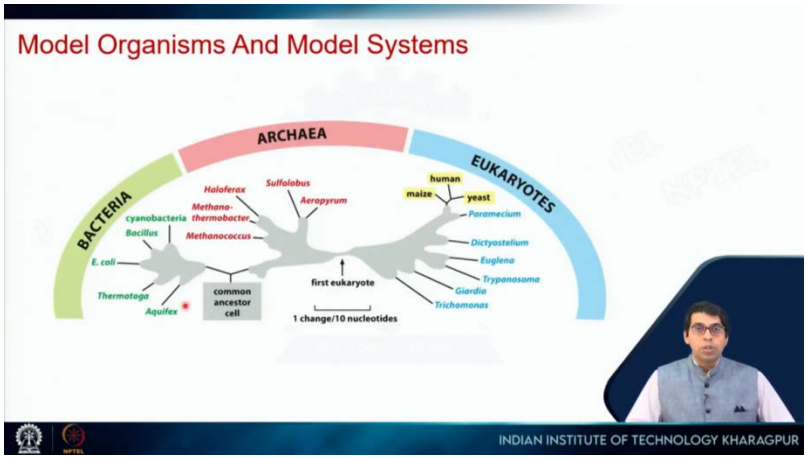
Welcome to the course Introduction to Complex Biological Systems. In this week 8, I will discuss the evolutionary history of life. So far, we have seen how evolution shapes the different life forms on this planet. Today, I am going to talk about certain model organisms. To do biological science, scientists have to use certain organisms. It turns out that some organisms are very amenable to scientific experiments conducted in laboratory settings. These types of organisms are typically referred to as model organisms. Today, I will discuss several model organisms. I will introduce their advantages and disadvantages.



But one thing I should point out is that this is not an exhaustive list. There are other organisms that scientists also use, but these are among the most common ones used in labs. For cellular experiments with eukaryotic cells, HEK293, HeLa cells, and Schneider 2 are used. I will discuss these cells in more detail in the coming slides. For prokaryotic systems, E. coli has been a model organism where many experiments have been conducted.

For unicellular eukaryotes budding yeast that is Saccharomyces cerevisiae has been used. For animals as model systems, fruit fly which is Drosophila melanogaster then a nematode which is C-elegans and mouse has been the workhorse of biological experiments. So these are model animals. For plants, these two plants Arabidopsis and Tobacco have been used

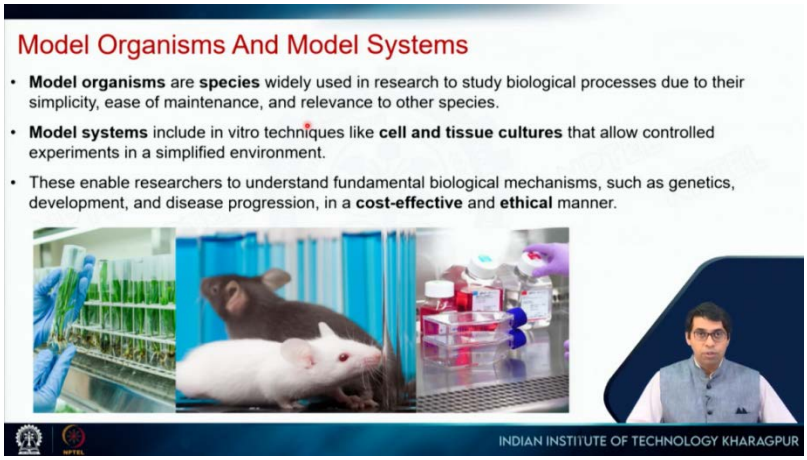
as model plants for scientific experiments. So this is the tree of life. So these are the different trees of life. You will also have viruses, maybe as the fourth kingdom, but these are the major three kingdoms of life.



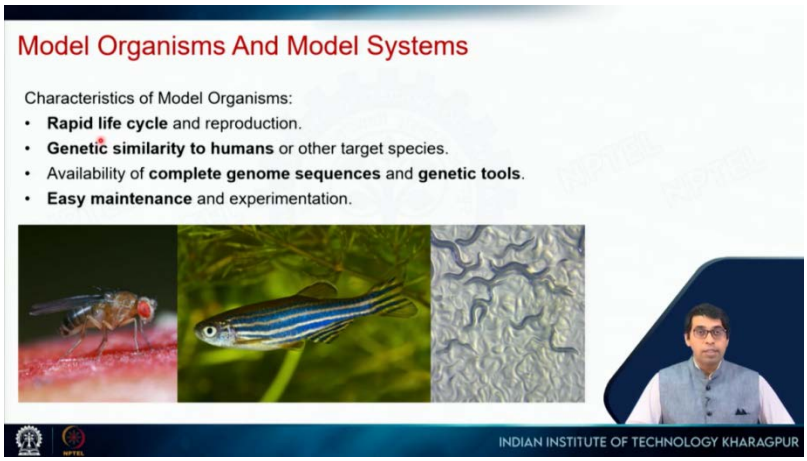
To study each of these kingdoms, different model organisms have been used, as we saw in the previous slide. So what are the characteristics? Model organisms are species widely used in research to study biological processes due to their simplicity, ease of maintenance, and relevance to other species. Simplicity means that they can be handled easily. Ease of maintenance means that they can reproduce easily.

Their genetics and other characteristics are already well known. Information obtained from experiments on these organisms can be translated to other organisms. So they are relevant to other species. Model systems include in vitro techniques like cell and tissue culture, which allow controlled experiments in a simplified environment. I will discuss this in more detail when I talk about model cells.

So what these types of systems do is enable researchers to understand fundamental biological mechanisms such as genetics. So, if you are doing a genetic experiment, you can use this. If you want to study development of how from a single cell a whole life form emerges you can use these model systems. For disease progression, which is very important that these model systems are used and since multiple labs worldwide use these organisms, there is consistency in experiments, and reagents have been developed to make them cost-effective.



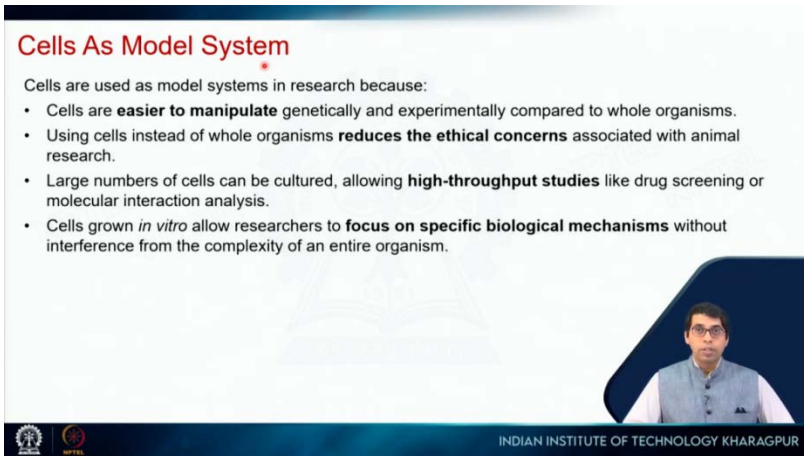
Also, guidelines have been designed so that experiments can be conducted ethically. So using model systems enables all of this. So what are the characteristics of these model systems? Why has a certain organism been chosen as a model organism? One important characteristic is that they should have a rapid life cycle, meaning they reproduce quickly so experiments can be done relatively fast.



Their genetics should be similar to humans. If we are using animal models, their genetics should resemble humans so that if we conduct an experiment in Drosophila or mice, we can translate that information to humans. If this particular protein works like this, most probably it will work in a similar manner for humans also. Availability of a complete genome sequence is very important.

So we should have the complete genome sequence of this model organism and we should know how many proteins are there and things like that and there should be genetic tools that are available because we want to make mutations, we want to delete certain proteins,

we want to introduce some other proteins. So all of these things can be done relatively easy and normally they can be done relatively easy. That is why these organisms are chosen as model organisms and also easy maintenance and experimentation. So cells as model organisms. So before we get into organisms, it is easier to do experiments in isolated cells. There are several reasons for that because cells can be easily manipulated compared to a whole organism. Cells are easier to manipulate genetically compared to a whole organism.

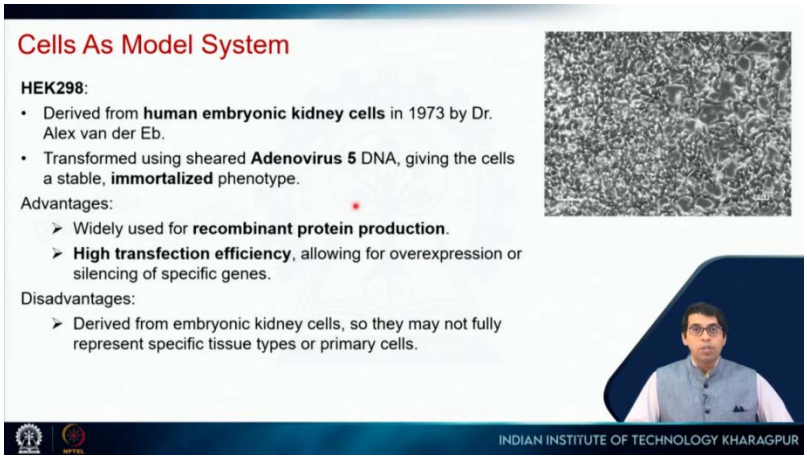


Using cells instead of whole organisms reduces the ethical concerns associated with animal research. So if we are doing some experiment in mice, the ethical considerations for doing experiments in mice versus ethical considerations for doing an experiment on an isolated cell, they are completely different. In cell, the ethical considerations are much simpler so one can do rapid experiments without worrying too much about all the ethical concerns.

Large number of cells can be cultured. So there is a limitation in the number of mice that you can use in a particular experiment. But when we are doing experiments in a cell, we can grow a large number of cells in a cell culture and this type of setting can be used for high throughput screening, for example, drug screening. So if you are trying to inhibit some protein or an enzyme, like we saw in the previous lectures, there is a library of 10,000 molecules. So a cell culture or isolated cells will be much better to start with to see whether this particular drug has the desired effect on the enzyme inside the cell rather than looking at the whole organism. If it works at the cellular level then it makes sense to go for the organism level. So this is how the hierarchy of experiments are typically designed. Cells

grown in vitro allow researchers to focus on specific biological mechanisms without interference from the complexity of an entire organism.

For example, if someone is looking at the effect of a particular molecule on liver cells. Now, if we are doing with cultured cells that molecule can be directly delivered to the cultured cells. However, if it is ingested, so if it is given in food and the mice digest that food, then it goes through its digestive system. So it will go through a lot of other environments before it reaches the liver. So, if we are doing cell culture, we do not have to worry about all these processes that occur before the molecule reaches the liver. However, you will have to worry about all those things. Once your cellular experiments are successful, then you will have to move to the organism level.

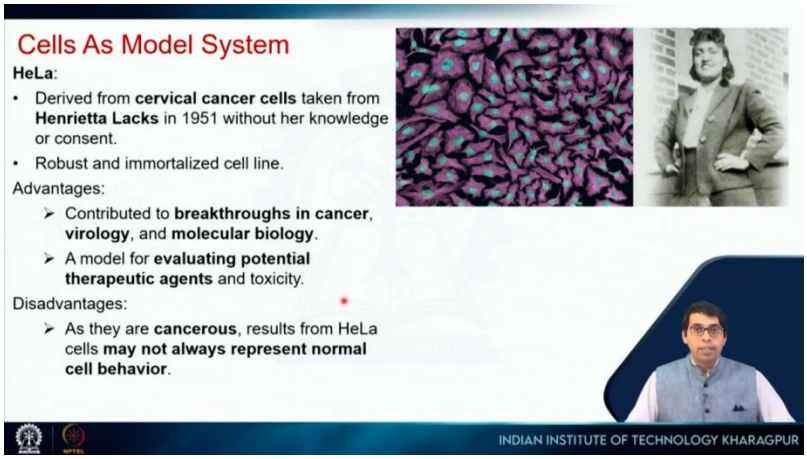


So, let us look at some model cellular systems. HEK298, so, HEK stands for human embryonic kidney. These human embryonic kidney 298 cells were derived in 1973 by Dr. Alexander Vanderem. These cells are very useful and are routinely used in different cell culture experiments throughout the world. They can be easily transformed using sheared adenovirus, giving cells that are stable and immortalized, meaning this cell line can be used repeatedly in different labs. What are the advantages? They are widely used for recombinant protein production in genetics. In recombinant DNA technology, you will learn that we typically use bacteria to produce proteins recombinantly. So, that is the first choice. However, many proteins do not express well in bacterial systems, especially those requiring post-translational modifications. In such cases, cells like HEK298 can be used

for overexpression of these target proteins. They also have high transcription efficiency, meaning silencing experiments can be performed relatively easily.

Disadvantage, so it is derived from embryonic kidney cells, so they are not fully representative of the specific tissue or primary cells. So since these are embryonic kidney cells, and these are isolated cells, plus they have been immortalized, they do not resemble the primary cells from the kidney. Also, they will not resemble the actual cells inside the organ and the third point is that they will be different from other cells.

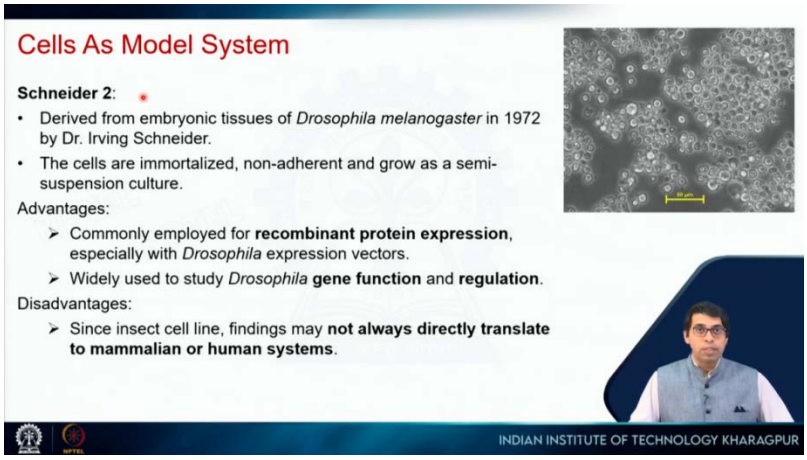
So if you want to do experiments in lung cells, then HEK298 cells will not be good enough. So you will actually have to take cells which are derived from lung cells so one has to be aware about the tissue specificity when they are using systems like this. Another extremely commonly used cell line is the HeLa cell line. So this cell line was derived from the cervical cancer cells from this woman named Henrietta Lacks in 1951 and since it was such a long time back, you can imagine that not many ethical guidelines were there. So it was actually taken from her without her knowledge or consent. However, it turns out that these cell lines are highly robust. These cells are highly robust and they are an immortalized cell line. So they can be used again and again and again.



So HeLa cells have been used in much breakthrough research over the past 70 years. So they have contributed to breakthroughs in cancer research, virology, and molecular biology. So they are a model for evaluating potential therapeutics, therapeutic agents, and toxicity. So you will see that in many labs, one of the first experiments people do is

performed on the HeLa cell line. A disadvantage is that they are cancerous cells, which means they do not represent normal cell behavior.

But many experiments can be done using these cell lines as long as we are aware of these differences. Another very commonly used model cell system or cell line is Schneider 2. So this is derived from embryonic tissues of Drosophila melanogaster, the fruit fly. So it was isolated in 1972 by Dr. Irving Schneider and is named after its inventor. These cells are also immortalized.

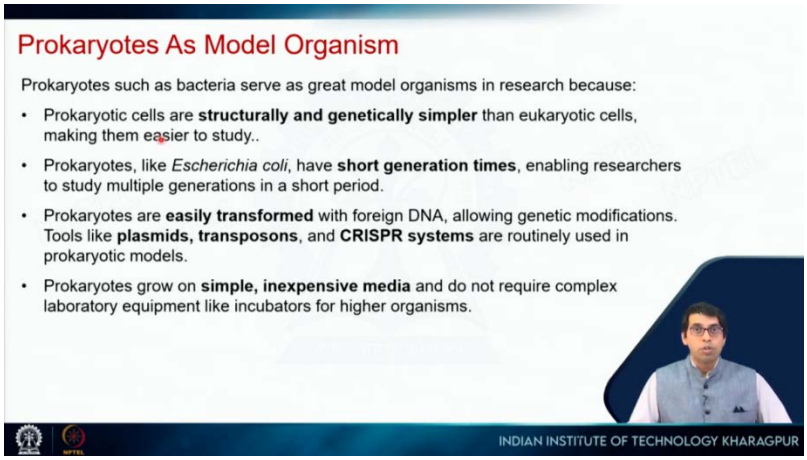


They are non-adherent and grow in semi-suspension in culture, which means you can use them in a shaker-like environment. So these cells are also a very robust system for expressing recombinant proteins, especially those from Drosophila or similar organisms. So these are widely used to study Drosophila gene function and regulation. Since it is an insect cell line, findings may not always directly translate to mammalian or human systems so one must be aware of the limitations of this particular cell line.

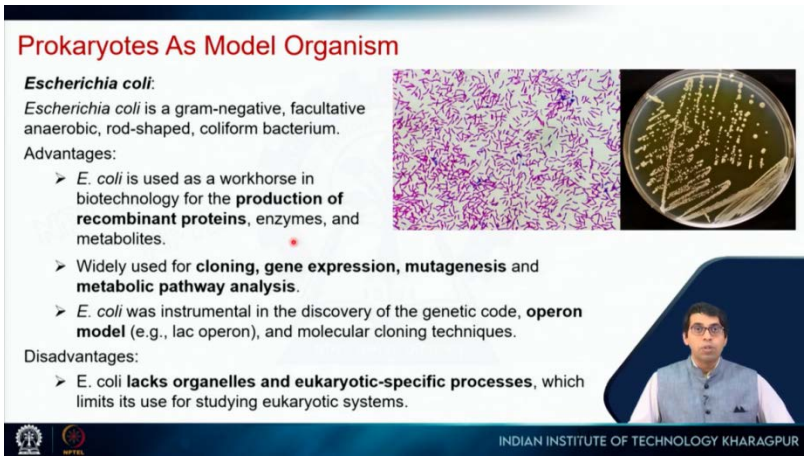
So these are the three cell lines that are derived from eukaryotes. Prokaryotes, also there are many studies that we do in prokaryotes. So we also have to have a model system from prokaryotes. So prokaryotes, such as bacteria, serve as great model organisms in basic research.

Prokaryotic cells are structurally and genetically simpler. So it is much easier to do experiments with them. E. coli is a model organism that has been used to study prokaryotes. They have a very short generation time, which means they can multiply very fast in a short

period of time. They can be easily transformed with foreign DNA, allowing genetic modifications.



Tools like plasmids, transposons, and CRISPR systems are routinely used in prokaryotic models. They grow on simple, inexpensive media. Growing bacteria is much less expensive compared to growing eukaryotic cells like HeLa or Schneider 2. So E. coli is a model prokaryotic organism. It is a gram-negative bacterium.

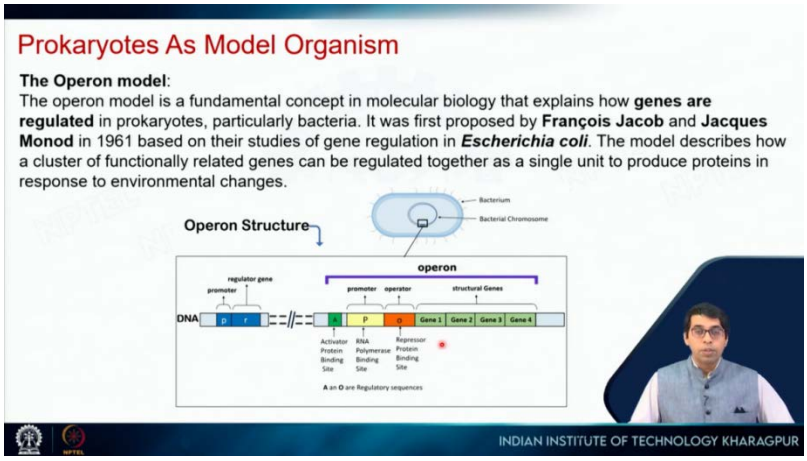


It is facultative anaerobic, which means that in the presence of oxygen, it will use oxygen to produce ATP. So it will go through the aerobic process. However, in the absence of oxygen, it will switch to anaerobic metabolism, the anaerobic process. It is rod-shaped. So you can see it's a rod-shaped structure here.

E. coli has been a workhorse in biotechnology for the production of recombinant proteins, enzymes, and metabolites. It is widely used for cloning, gene expression, mutagenesis, and

metabolic pathway analysis. It was instrumental in the discovery of the genetic code. So we have looked at the genetic code. We have looked at which codon codes for which amino acid.

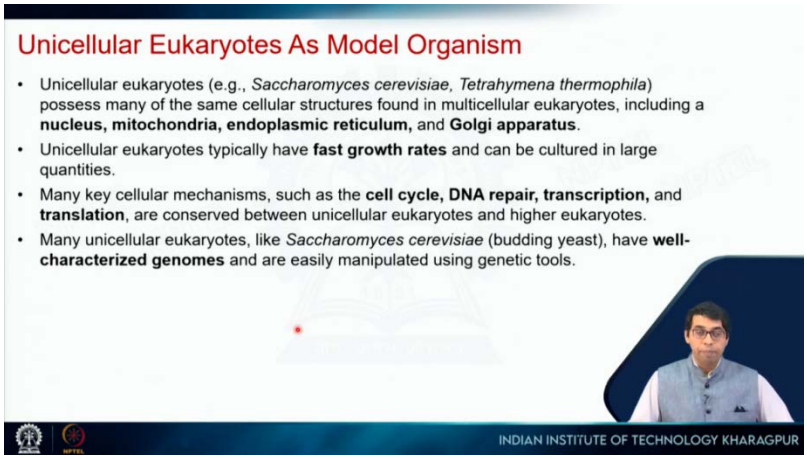
Those discoveries were made in E. coli. Another important discovery that was made was the operon model. The lac repression model was also discovered in E. coli. A disadvantage, of course, is that it lacks organelles and eukaryotic-specific processes. So it can serve as a model system for basic research or research that is suitable for prokaryotes but not for eukaryotic systems.



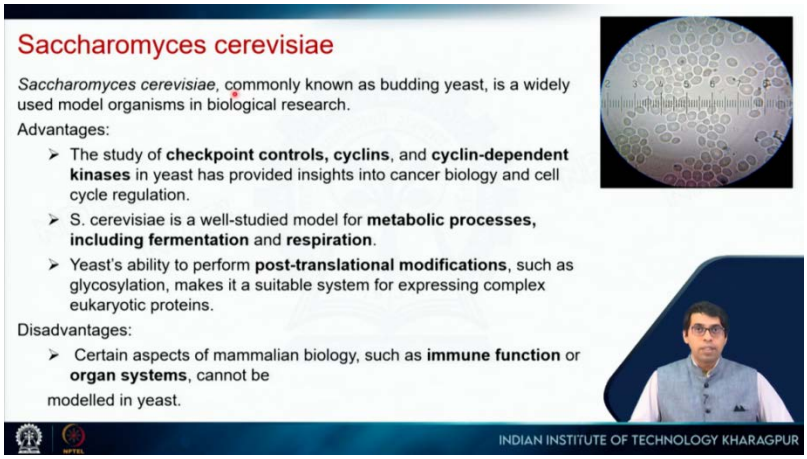
So, this is a brief description of the lac operon. So, you have already seen that. So, this lac operon system was discovered in E. coli by Jacob and Monod in 1961. So, briefly, in this repressor, the lac repressor comes and binds, and once there is lactose in the system, then this repressor gets converted to allolactose, falls off, and these genes get translated. So, this particular system, where consecutive genes are regulated by a single promoter, is called an operon, and it is found only in prokaryotes. So, this system was discovered in E. coli. Now, there are also eukaryotes that are unicellular. A model system for unicellular eukaryotes is Saccharomyces cerevisiae, that is, Baker's yeast.

Another model system is Tetrahymena thermophila. Again, these model systems have been used to make many groundbreaking discoveries. For example, Tetrahymena was used for the discovery of RNA splicing. So, RNA splicing was actually discovered using Tetrahymena. Saccharomyces cerevisiae, I will show one example in the next slide, which is the yeast two-hybrid assay, which is done using Saccharomyces cerevisiae. So these

types of eukaryotes have processes that are very similar to multicellular eukaryotes because they have all these internal organelles. So they can serve as a model system for studying eukaryotic systems. Since these are unicellular, they grow fast and can be cultured in large quantities in a relatively inexpensive manner.



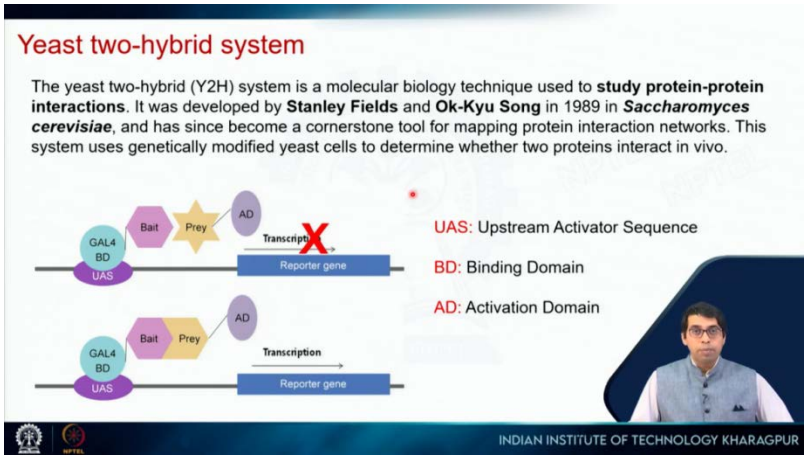
So many key cellular mechanisms, such as the cell cycle, DNA repair, transcription, translation, etc., are conserved in these unicellular eukaryotes as they are in higher eukaryotes. So Saccharomyces cerevisiae, commonly known as budding yeast, is a widely used model organism. It has been used to study checkpoint controls, cyclins, and cyclin-dependent kinases. So these things that we have seen in previous lectures were widely studied using this model system.



It is also a well-established model for studying metabolic processes, including fermentation and respiration. Its ability to perform post-translational modifications, such as glycosylation, also makes it a suitable system for expressing complex secretory proteins.

So proteins that need glycosylation cannot be expressed in bacteria, but they can be expressed in yeast. So what are the disadvantages? Since it is a unicellular organism, it lacks organ systems and more complex biological systems, such as immune function.

So you cannot study the immune system using this type of unicellular organism. Another important experiment that was first done in yeast is the yeast to hybrid system. So this is a very useful experiment that is used to study whether two proteins interact with each other. So this particular protein, the GAL4 binding domain and the activation domain, are normally linked together. So what was done is that they have been separated, and now, using recombinant DNA technology.

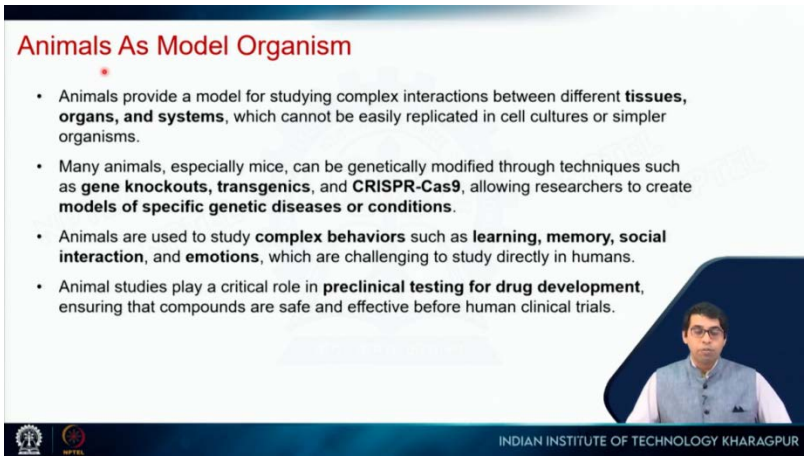


So you want to study the interaction between this protein, which is called the Bait, and this protein, which is called the Prey. So you want to screen whether these two proteins interact with each other or not. So you genetically fuse these two genes together and fuse these two genes together, and then you express them. If they interact, then the Bait and Prey come together. So they are binding tightly. GAL4, the binding domain, will bind to this upstream system. So when it binds to this, this is linked to it; this binds to this, this is linked to it. So it will bring in this activation domain, and it will activate the transcription of this reporter gene. This reporter gene can be anything. It can be a marker for antibiotic resistance, which means that when this interaction happens, those cells where this interaction occurs will survive because this protein is expressed and if you have given the antibiotic and the media, it will be able to survive. If there is no interaction, then there will be no survival, or it can be a fluorescent molecule. So, if there is interaction, there will be fluorescence; if there is

no interaction, there will be no fluorescence. Using this type of system, we can screen for molecules that interact with this bait. So, if you expect 20 different proteins that may or may not interact with this protein, you can clone them separately. So, you can have them separately, then transform them, and do this experiment to see which proteins are interacting. If you want to identify which residue, suppose I have these two proteins, I know they interact, but I don't know what the binding interface is.

So, you can mutate residues at this binding interface and screen them using this. So, you can screen all the different mutations using this type of system. So, anything that survives, you know that is needed. So, if you mutate some residue and it does not survive, you know that particular mutation is needed for this interaction. So, there are many ways these experiments can be done.

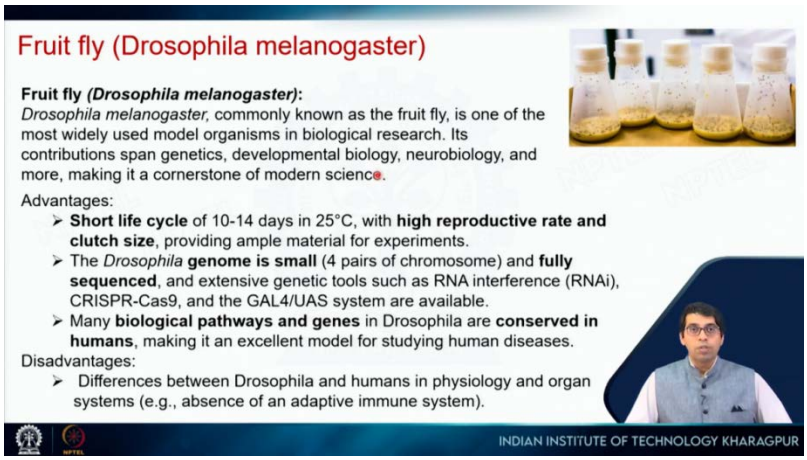
So, yeast actually contributed to the development of this type of experiment. Now animals as model organisms so we have looked at single-celled organisms, we have looked at prokaryotes. We have looked at eukaryotes.



Now, animals as model organisms so once you have done your experiments in the cellular environment and found interesting results, you may now move on to animal models. Animal models provide a system for studying complex interactions between different tissues, organs, and systems, which cannot be easily replicated in cell cultures or simpler organisms like yeast. Many animals, such as mice, can be genetically modified through techniques such as gene knockout, transgenics, CRISPR-Cas9, etc.

These types of techniques can be used to model certain systems. For example, if you want to study Alzheimer's disease, you can introduce certain mutations that will make the mice predisposed to Alzheimer's disease. Now you have a mouse model, which is an Alzheimer's mouse model. Then you can screen for different molecules or therapies to see whether they can reduce the effects of dementia in this particular mouse model. All these types of genetic manipulations can be done in these model organisms. Animals are used to study complex behaviors, such as learning, memory, social interaction, and emotions, which are challenging to study directly in humans. If you want to study what proteins are important in short-term or long-term memory then these types of animal models can also be used. Animal studies play a critical role in preclinical testing for drug development.

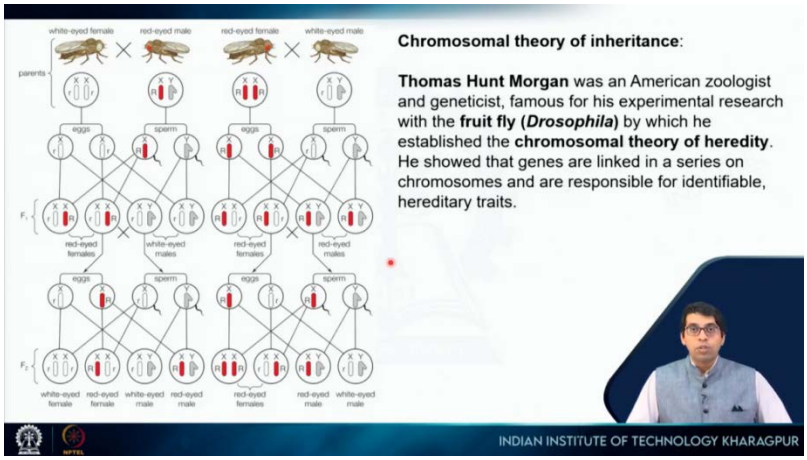
So, of course, after the cell, you go to these different animal models to see whether these drugs are actually effective in an organism setting. If they work, then, of course, you go to higher organisms and finally to human clinical trials. So, let's look at the fruit fly or Drosophila melanogaster. Drosophila melanogaster, or the fruit fly, is one of the most widely used model organisms in biological research. Its contributions span genetics, developmental biology, neurobiology, and many more.



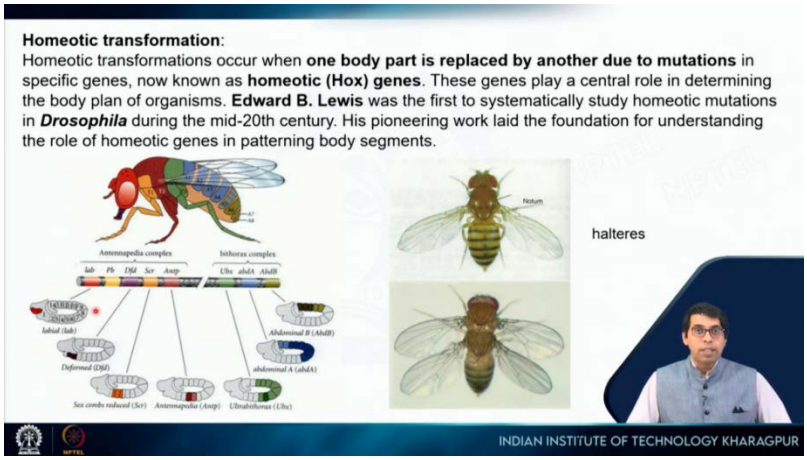
So, a lot of early research on developmental biology was done in fruit flies because of their ease of genetic manipulation and very fast life cycle. Drosophila has a short life cycle of 10 to 14 days at 25 degrees centigrade. They have a very high reproductive rate and clutch size, which means there is enough material to do experiments. The Drosophila genome is small.

It has only four pairs of chromosomes, which are fully sequenced, and there are extensive genetic tools available that can be used to manipulate the Drosophila genome. Many biological pathways and genes in Drosophila are conserved in humans, making it an excellent model for studying human diseases. For example, if you want to study development, the same set of genes is present in Drosophila as well as in humans. So, if we think of the Hox genes, there are 8 Hox genes in Drosophila and 39 Hox genes in humans. Even though the number of genes is different, the basic principles remain the same.

Disadvantages, difference between drosophila and humans in physiology and organ systems, for example, absence of adaptive immune system. So these things make it difficult to study certain phenomena in drosophila, which are important for humans. So in that case, you go to a different model, which will be a mice model. So Drosophila has contributed to many discoveries. One of the important discoveries that happened in Drosophila is the chromosomal theory of inheritance.

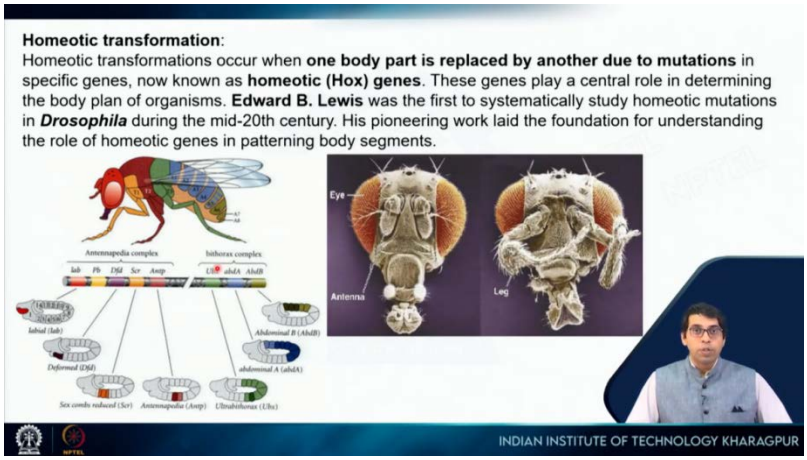


So Thomas Hunt Morgan, he was an American zoologist and geneticist. His famous experiment on Drosophila established this chromosomal theory of heredity. So he showed that genes are linked in a series on chromosomes and are responsible for identifiable hereditary traits. So another very important discovery that happened in Drosophila is the homeotic transformation. So it turns out that there are these master regulated genes called the Hox genes.

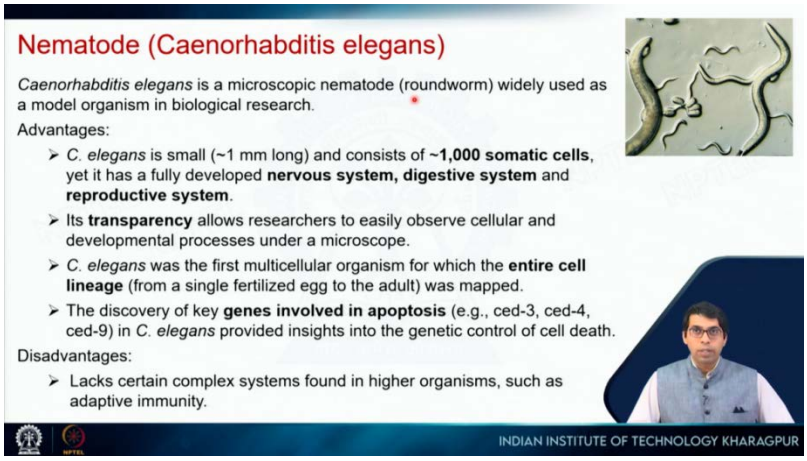


So in Drosophila, there are these eight Hox genes, labial, deformed, sex-combs reduced, antennapedia, ultrabithorax, abdominal A and abdominal B and they are arranged in this linear order on a chromosome. It turns out that each of these genes regulate the development of each of these segments in the body of this Drosophila. Now these genes are also present in us and they do exactly the same in us also. So this part, this gene is responsible for the development of the head, so on and so forth. Now there is also another very interesting discovery that was made, which is that if you copy one of these segments and add another copy; then, the same thing will be duplicated. So, for example, what happens here is that the Drosophila has two wings, and there are these small structures called halteres. They help in balancing when the Drosophila is flying. Now, a particular set of genes was copied, which resulted in the formation of two pairs of wings. So, instead of these halteres, there was another pair of wings found in this mutant Drosophila.

In fact, the position of these genes is also important, which is called the collinearity of these Hox genes. So, in this case, antennae, so from the head of the Drosophila, these antennae come out, and they are normally determined by the Antennapedia gene. Ultrabithorax is responsible for the development of the leg. Now, in one experiment, this gene was deleted from here and inserted before the Antennapedia. So, in those mutant flies, what happened was that instead of antennae, legs came out from the head. This means that these genes are highly modular in nature, and by shifting their position, the body plan can be changed.



This was a very important discovery in the development of vertebrates. Another important organism is a nematode called C. elegans. C. elegans is a microscopic nematode, or roundworm. It is very small. It is much smaller than mice and Drosophila and is widely used in biological research. It is one millimeter long.

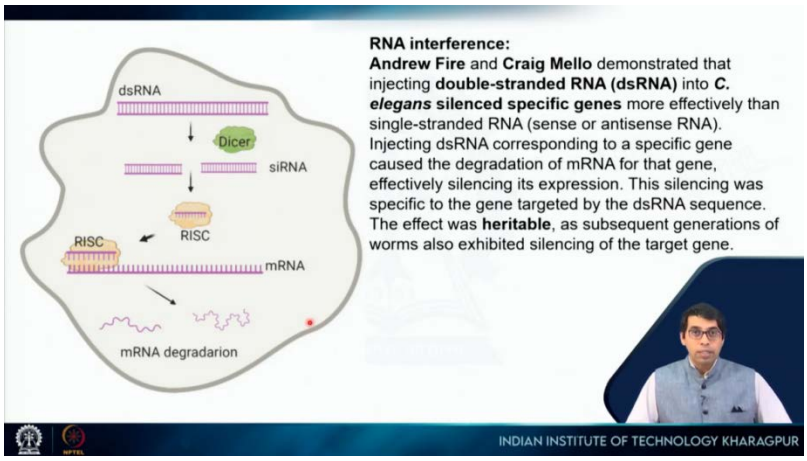


It contains roughly 1000 somatic cells. Yet it has a fully developed nervous system, digestive system, and a reproductive system. Now, since it is such a small organism, the lineage of all these cells is completely known. So, this organism is also transparent. So if you put it under a light microscope, you can see all the cells; you can actually see through this organism.

So its transparency allows researchers to easily observe cellular and developmental processes under a microscope. C. elegans was the first multicellular organism for which the entire cellular lineage was mapped. So from a single fertilized egg, one cell becomes two cells, then it becomes four, and so forth, where each cell goes, which cell is responsible

for forming the anterior side, and which cell is responsible for the posterior side, all of this was mapped for the first time in C. elegans. The discovery of key apoptosis genes, for example, Ced-3, 4, and 9, was done in C. elegans, and it provided important insights into the genetic control of cell death, that is, apoptosis.

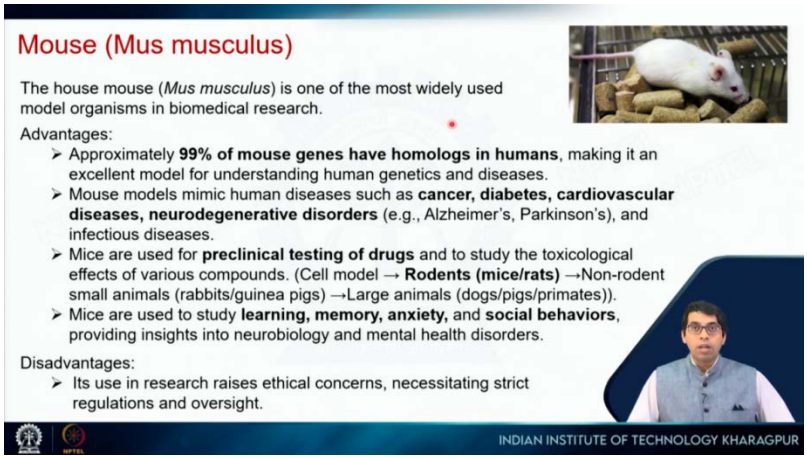
So what are the disadvantages? It lacks complete complex systems, which are found in higher organisms, such as adaptive immunity. So again, if you want to study something like that, then C. elegans is not the model organism. You have to go for something like a mouse. So another important discovery that was made using C. elegans was the phenomenon called RNA interference.



Andrew Fire and Craig Mello demonstrated that injecting double-stranded RNA, not single-stranded, but double-stranded RNA was very effective in silencing specific genes. This was done for the first time in C. elegans. It was also shown that this effect was heritable; meaning subsequent generations of the worm also exhibited silencing of the target gene, so a very important discovery. This is another important discovery made using C. elegans.

Now, the most widely used organism is the mouse model. The house mouse, or Mus musculus, is one of the most widely used model organisms in biomedical research. It has many advantages. Almost 99% of mouse genes are homologous to human genes. For example, if I talk about Hox genes, there are eight in Drosophila but 39 in humans.

Mice also have 39 Hox genes. These genes are very similar to human genes. This means that studying development in mice can provide detailed insights into human development. Mouse models mimic human diseases such as cancer, diabetes, cardiovascular diseases, neurodegenerative disorders like Alzheimer's and Parkinson's, and infectious diseases.

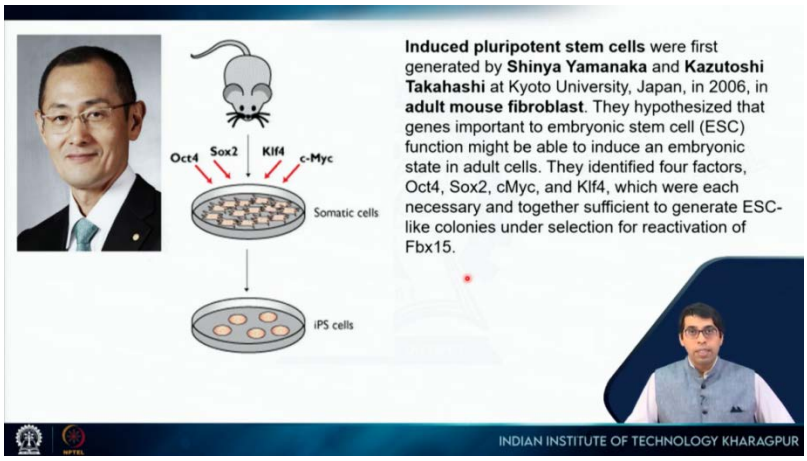


Mice are typically used for preclinical testing of drugs and to study toxicological effects of various compounds. So once you have screened a set of molecules using a cellular system and if you have come out with, let us say, four or five target molecules, then those will be tested in mice to see the effect on an organism level. Once that is successful, then typically we will go for higher animals such as rabbits or guinea pigs, and then dogs, pigs, or primates. Mice are also used to study complex behaviors like learning, memory, anxiety, and also social behaviors.

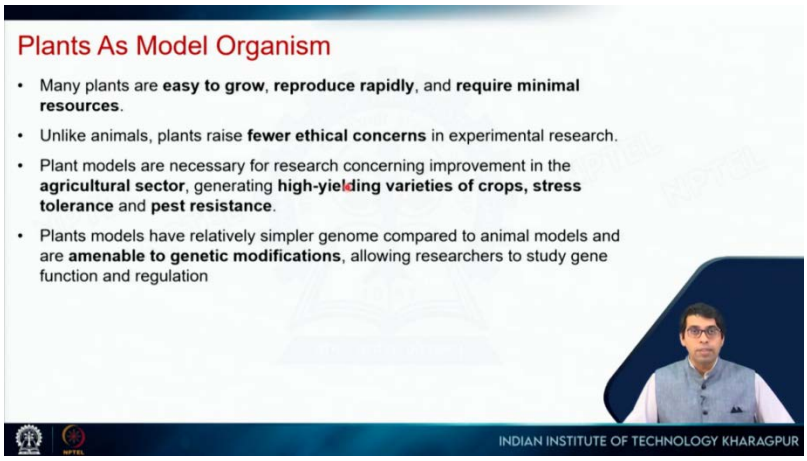
So in neuroscience, to study neuroscience, mice are also a very good model system. Disadvantages, well, it has, of course, these ethical issues. However, there are many things that are studied in mice that cannot be directly translated to humans. So that also poses a big disadvantage in using mice as a model system for studying really complex systems. So one very important discovery, there are many discoveries that have been done using mice but one important discovery that needs mention is the discovery of induced pluripotent stem cells.

So I have already discussed this earlier. So IPS cells, or induced pluripotent stem cells, were discovered by Yamanaka and Takahashi, and they did those using adult mouse embryos. So they showed that these four transcription factors are Oct4, Sox2, cMyc, and

Klf4. When you add them, you can actually turn somatic cells into these induced pluripotent stem cells.

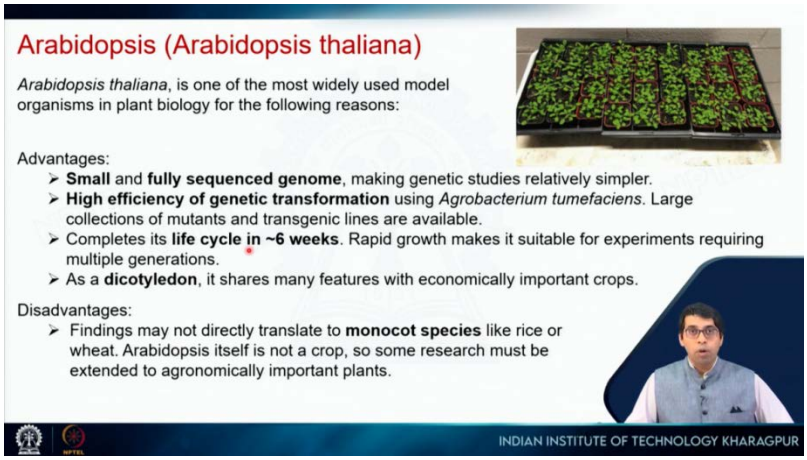


So mice are one of the model organisms in animals, but there are also model organisms for plants. In the case of plants, many are easy to grow, reproduce rapidly, and require minimal resources. Therefore, many plants can be used as model organisms. Unlike animals, plants raise fewer ethical concerns in experimental research. Plant models are necessary for research concerning improvements in the agricultural sector, which is a very important concern for us.

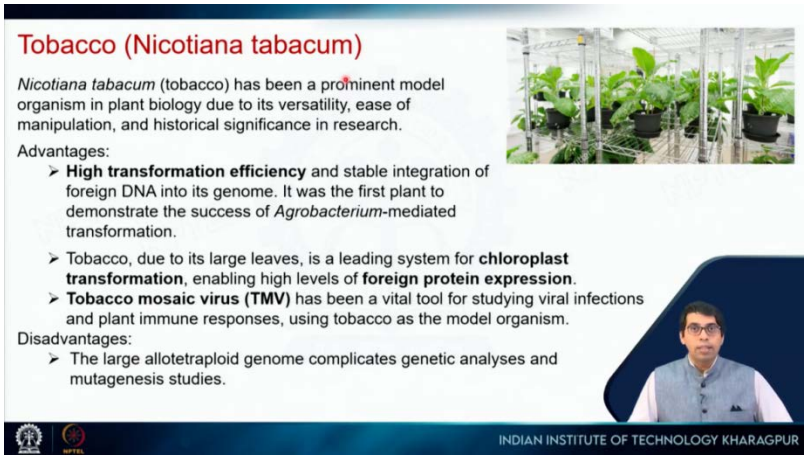


This allows us to generate high-yielding crops, crops that can tolerate different stress conditions and are pest-resistant. Plant models have relatively simpler genomes compared to animal models and are amenable to genetic modifications. One of the model plants used in many labs is Arabidopsis. It is one of the most widely used model organisms in plant

biology for the following reasons. It is small, has a fully sequenced genome, and has high efficiency for genetic transformation using Agrobacterium.



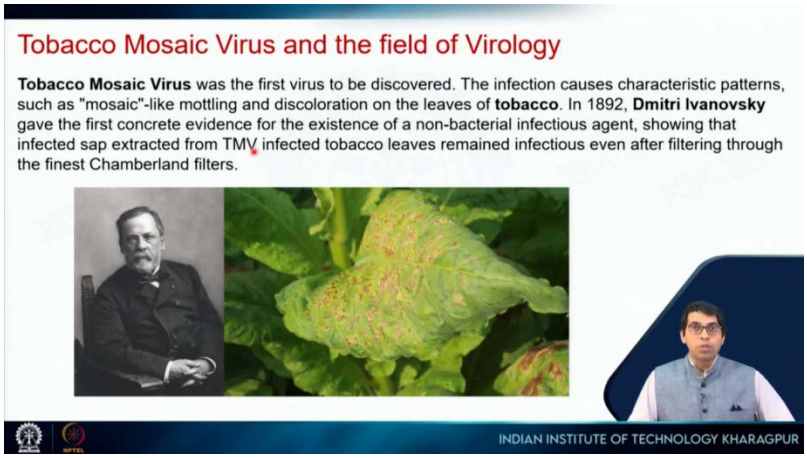
A large collection of mutants and transgenic lines can be easily produced for this particular plant. Its complete life cycle is six weeks, making it a rapidly growing plant, which is very useful for experimentation. It is a dicotyledon, or dicot, meaning it shares many features with economically important crops. However, findings in this dicot plant may not directly translate to monocot species like rice or wheat, meaning that for studying rice or wheat, we must use model organisms or model plants from these particular species. Another commonly used plant is tobacco.



So, it has been a prominent model organism in plant biology due to its versatility, ease of manipulation, and historical significance. So, it also has very high transformation efficiency in tobacco due to its large leaves. So, you can see that the leaves are very large. It is a leading system for chloroplast transformation, enabling high levels of protein

expression in these organelles. Tobacco mosaic virus, or TMV, has been a vital tool for studying viral infections and plant immune responses using tobacco as the model organism.

So, immunity in plants has been studied using this tobacco plant system, where tobacco mosaic virus has been used to induce these viral infections and there are, of course, disadvantages. For example, it's an allotetraploid, which means that it has a complicated genome, making its genetic analysis complex. So, the Tobacco Mosaic Virus turns out to have been the first virus to be discovered. So, the infection causes characteristic patterns, such as mosaic-like mottling and discoloration on the leaves of tobacco.

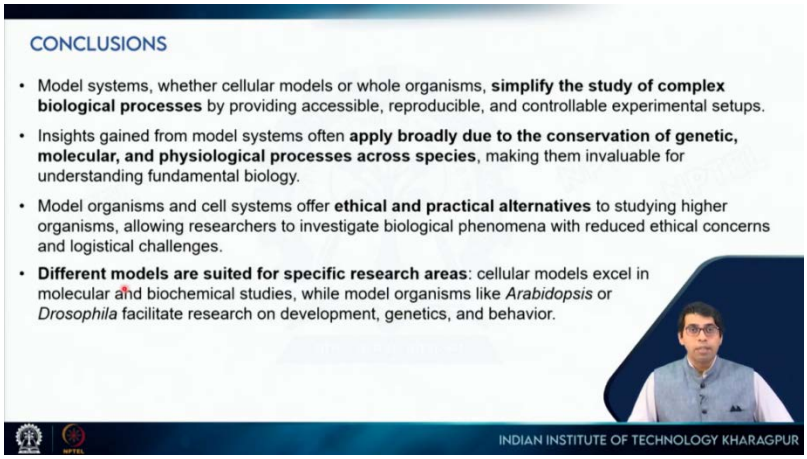


This is a picture. So, this is a big tobacco leaf, and infection makes it look like this. That is why it was called the tobacco mosaic virus. So, in 1982, Dmitri Ivanovsky provided the first concrete evidence for the existence of a non-bacterial infectious agent, showing that infected sap which was extracted from these leaves that infected tobacco leaves but even though you filter this sap through a filtering agent, it can still infect these tobacco leaves which means that there is something which is even smaller than known bacteria that can infect and that is how the first virus was discovered.

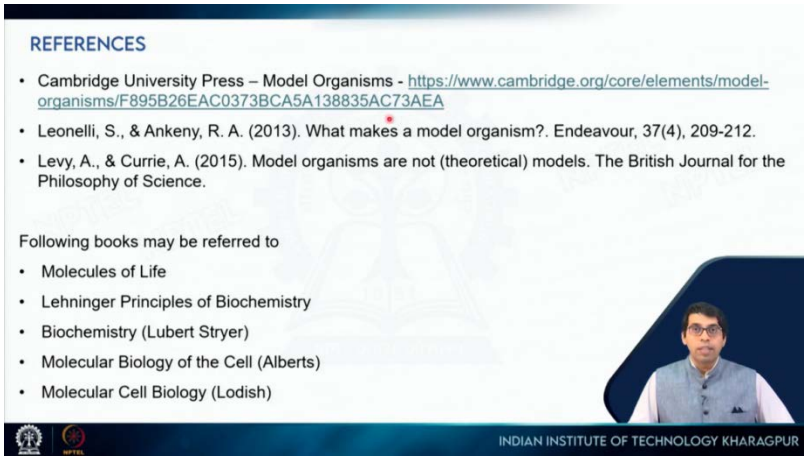
So here are the conclusion model systems where cellular models or whole organisms are used to simplify the study of complex biological systems. They provide insights that are applicable to other organisms. So they apply broadly due to their conserved genetics, molecular, and physiological processes across species. So whatever we learn from model systems, we can translate that information to other systems, and they can also contribute to

basic or fundamental research. Now, these model organisms or systems offer ethical and practical alternatives to studying higher organisms.

So, there are many ethical issues that come with actual organisms like mice or Drosophila. However, if we are using isolated cellular systems, we can circumvent some of these ethical issues. Different models are suited for specific research, as we have seen in the previous slide.



So here are some of the important references, and you can also refer to these books for this particular lecture.



Thank you.