Metabolic Engineering
Prof. Amit Ghosh
School of Energy Science and Engineering
Indian Institute of Technology – Kharagpur

Lecture – 13
Introduction to Systems Biology

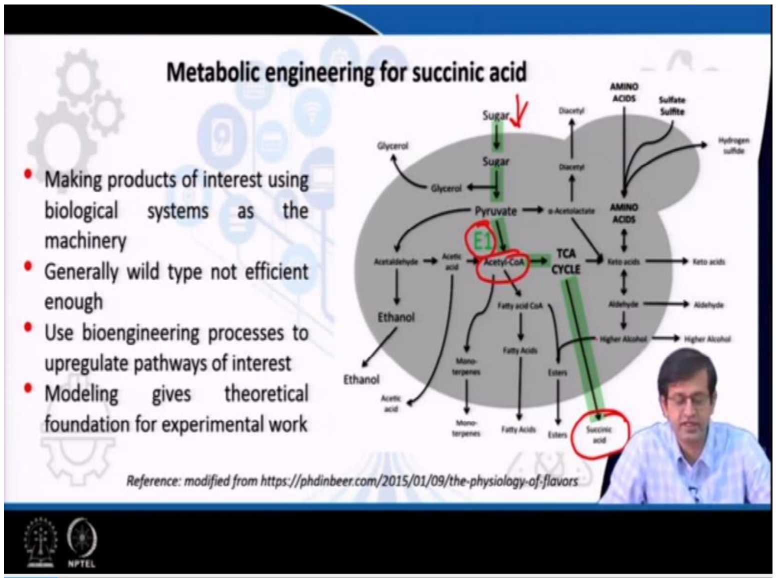
Welcome to the metabolic engineering course, today we will start a new topic that is the metabolic engineering where we will be learning about introduction to systems biology. Systems biology is another branch of study where you would like to know the systems perspective of the metabolism because this metabolism is well interconnected previous class; I told that it is a complex network.

So, metabolites are well connected is interlinked and when you try to change one of the metabolite or one of the gene then what happened? The entire network gets disturbed. So, it is a better that you know how they are interconnected. So, in this lecture, we will be learning the systems biology part, how they metabolite, how the reactions are interconnected in a network?

(Refer Slide Time: 01:28)

So, the concept covered in this lecture is basically challenges which I have already told the challenges of metabolic engineering, the need for systems biology, the genomics to genetic circuit and then application of systems biology. So, the challenges of metabolic engineering actually lead to systems biology because metabolic engineering involves regulating the fluxes and estimating the fluxes improving the production that actually lead to systems biology, systems biology problem you are addressing.

(Refer Slide Time: 02:02)



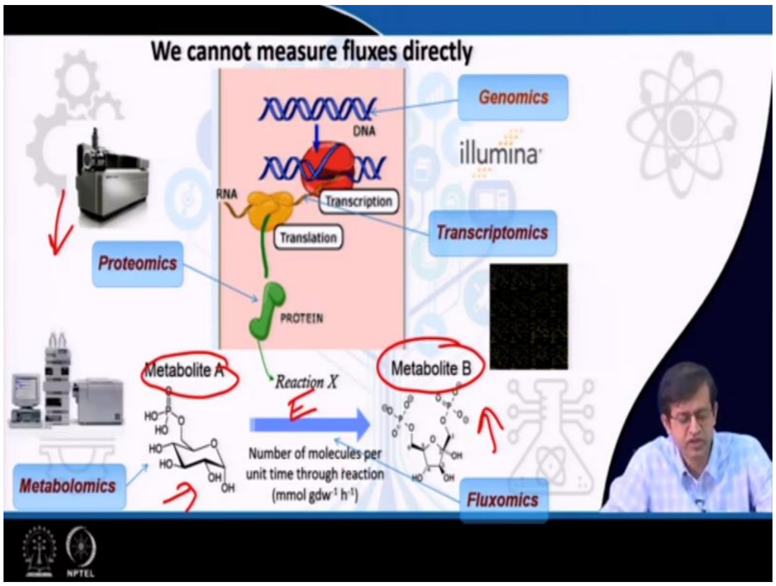
So, in the previous lecture, I told about how you can improve the succinic acid production. So, succinic acid production you want to improve the cell is growing on glucose. The glucose is entering from here, the carbon which is entering from the cell and ultimately you want to channelize the entire carbon, so that you get maximum succinic acid from the cell. So, to do that, you try to up regulate some of the genes which are on the pathway.

So, the green shaded arrow you can see that this is the pathway you are interested in. So, out of many metabolic pathways you are only interested in that pathway, such that the maximum carbon can flow through that pathway. But, earlier I told that you suppose that you want to increase the enzyme inside the cell, how do you increase the enzyme inside the cell? There is a way you can up-regulate the gene.

So, the gene which is associated with enzyme 1 can be upregulated and you can produce more enzymes in the cell and more enzymes is more carbon going through the cell or the rate of the reaction become higher, the more of acetyl coenzyme-A is formed inside the cell. So, the acetyl coenzyme-A is producing more inside the cell but not necessarily that acetyl coenzyme-A can go into this succinic acid pathway.

It can go into another pathway like fatty acid coenzyme-A and then monoterpenes, acetic acid, there are other pathways which are also involved in the TCA cycle are taking up acetyl coenzyme-A. So, this way is a systems biology problem. The system biology problem means you have to understand how the carbons are moving or going for each of the reaction. So, to understand the metabolic network that the; interconnectedness between these metabolites reaction, you need to understand in a systems perspective.

(Refer Slide Time: 04:11)



So, ultimately you would be you cannot measure flux directly and the metabolic fluxes that is the carbon flowing through the reaction. So, here I have shown that given 2 metabolite, metabolite A and metabolite B, the conversion rate of metabolite A to metabolite B is through a reaction. So, there is an enzyme, this enzyme is here and the enzyme is converting this metabolite A to metabolite B and you know already that the rate of reaction.

That is the number of molecule going through this reaction that is generally denoted by millimoles per gram dry cell weight per hour. So, this is the unit used for measuring the fluxes. So, the fluxes are actually not obtained directly in the sense that most of the data like genomic sequence and then transcriptomics data, proteomics data, metabolomics data they come directly from the equipment.

Suppose you are doing genomics, then you send the sequence to Illumina is a company who actually take your sequence and give you the DNA sequences in a high throughput way. So, you directly get the data from some equipment. Similarly, for transcriptomics also you get the expression level from the sample you are providing. And similarly, for proteomics also you have the spectroscopic setup those LC mass or mass spectrometers which are available.

We can measure the protein level, so, we can directly measure and the data can be directly interpreted to get the level of the protein. So, the protein level can be obtained from directly from the equipment. So, even in the metabolomics also you measure the metabolites through equipment and you get the level of the metabolites. So, you can measure this metabolite, each of the components you can measure but the flux is that is the rate of reaction that is not obtained directly from the equipment.

So, for that you have to actually do some metabolic modeling. So, we will come to it in details but just to give you an idea that why we want to do construct this metabolic network and the construction of the metabolic network, the primary purpose of the constructing a metabolic network is actually want to estimate the flux that is the rate of a molecule going through a reaction. So, the rate at which the; molecules are converted from metabolite A to B.

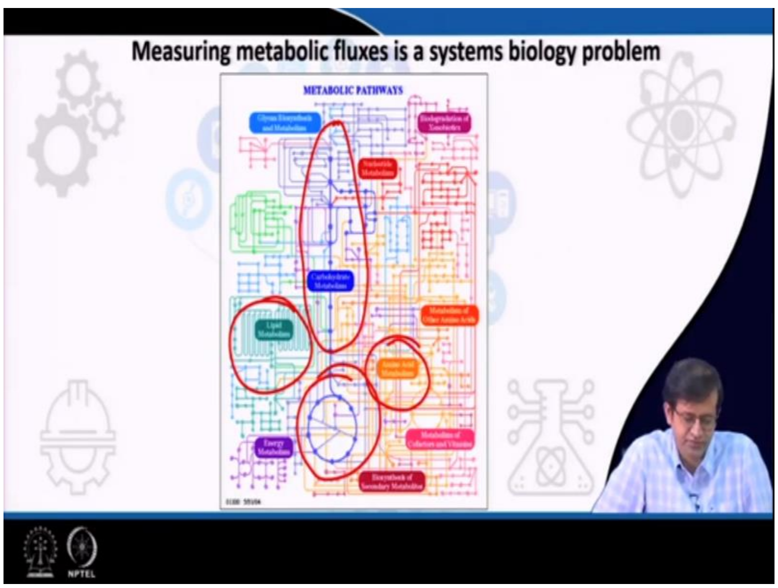
That you need more metabolic modeling and then followed by optimization, you get the metabolic fluxes. The fluxes are the most relevant phenotype of a cell that you should remember, you can do genomics, transcriptomics, proteomics, metabolomics but ultimately, how much flux is going that is the ultimate phenotype of a cell, you should remember that all these things you are doing genomic, transcriptomics, proteomics ultimately you want to know whether that reaction is active or not.

By just determining the protein level you may not know that the reaction flux is there or not, the protein maybe inhibited, a protein maybe to allosteric mechanism you know that protein get inhibited that even if you have sufficient protein inside the cell, it may not be sufficient for a reaction. So, similarly with the genes also genome sequence also we know the gene is exist but how much mRNA formed and then the level of mRNA is not always directly correlated with the protein level.

So, ultimately how much protein is from that protein catalyze that reaction and this way, the enter mechanism ultimately give rise to a new phenotype that is a flux through a reaction. So, here we are talking about the metabolic fluxes for all the reactions inside the cell because you know that and just by measuring the intracellular fluxes inside the cell, just for one reaction is not sufficient. So, in this lecture will focus on understanding how the intercellular fluxes that are the fluxes inside the cells are changing.

How you can measure those fluxes. So, experimentally you cannot measure the fluxes but you can you can do some metabolic modeling and try to estimate the fluxes. So that is the scheme used by many researchers. So, for metabolomics you know that you have LC mass where you can measure the metabolites level.

(Refer Slide Time: 09:08)

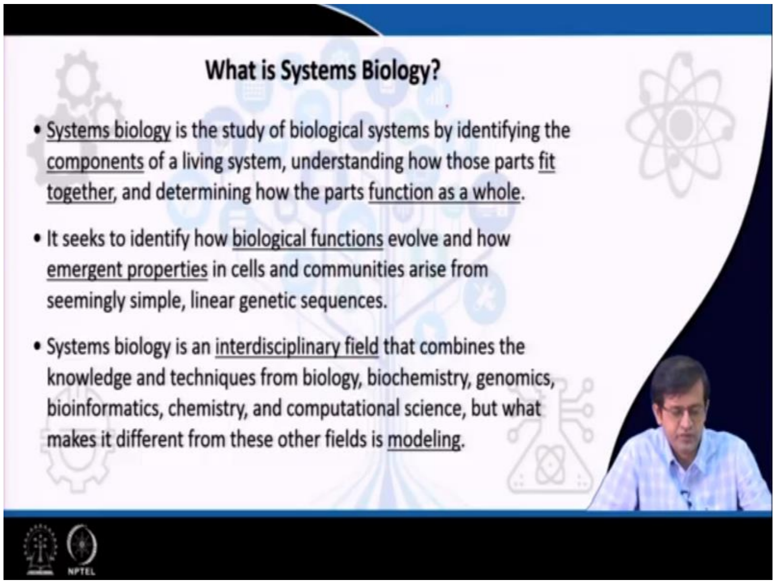


So, here you have measuring the metabolic fluxes are a systems biology problem. You can see that the various metabolic modules are there like you have the carbohydrate metabolic module. So, these module carbohydrate metabolic modules then you have TCA cycle and then we have the amino acid metabolism, lipid metabolism. So, these are the different module present in the metabolic network each of the modules has metabolic fluxes flowing and it is changing with time.

So, the basis of the flux estimation is the metabolic network that is constant but how much fluxes are flowing is actually changing with time that depends on the regulation and other important thing inside the cell. So, to understand the metabolic fluxes because metabolic engineering deal with metabolic fluxes that is the phenotype of the cell. The metabolic engineer you how much you ever you engineer ultimately you would be interested in how much fluxes you are actually producing.

Suppose, you are producing succinate production that is how much carbons are going in succinic acid production. So, this way you will try to estimate the different carbons flowing in different reaction and try to estimate how much succinic acid is produced.

(Refer Slide Time: 10:36)



So, what is systems biology? So, this systems biology problem we will start with basic like why? What is the definition of systems biology? How it came? We start with basic understanding why systems biology is a study of biological system by identifying a component of a living system and understanding how this fit together and determine the parts as a whole function. So, the driving force of systems biology is basically the high throughput technology.

Today, we have seen that a lot of high throughput data available for analysis the driving force which actually lead to enormous amount of biological data. And from this data you try to estimate what are the components. So, the component when I say; what the components are in a living system, it can be protein, DNA, RNA. So, this component individually identifies and gives some function to them.

So, the latter half of the 20th century actually, scientists were strongly influenced by reductionist approach like the generation of information about individual component. The individual component how this component but actually systems biology actually deals with interaction of this component which give rise to a biochemical reaction network. The biochemical reaction network or the metabolic network I just told about you is basically deal with when you think cell as a systems.

So, this system perspective has to come into picture where you try to connect the component and try to link each other and so that you can determine the component function as a whole. So, try to connect and summarize each and every component first you identify the function of each and every component and try to link each other, so that you can give a function as a whole. So, this is the objective of systems biology.

The chemical interaction of this component giving rise to the biochemical network and the high throughput technology; forcing biologists to view as systems because of the high throughput technology and like genomics, proteomics transcriptomics, metabolomics. So, this high throughput technology is generating huge amount of biological data and that actually forcing the biologist to understand the cell as a system.

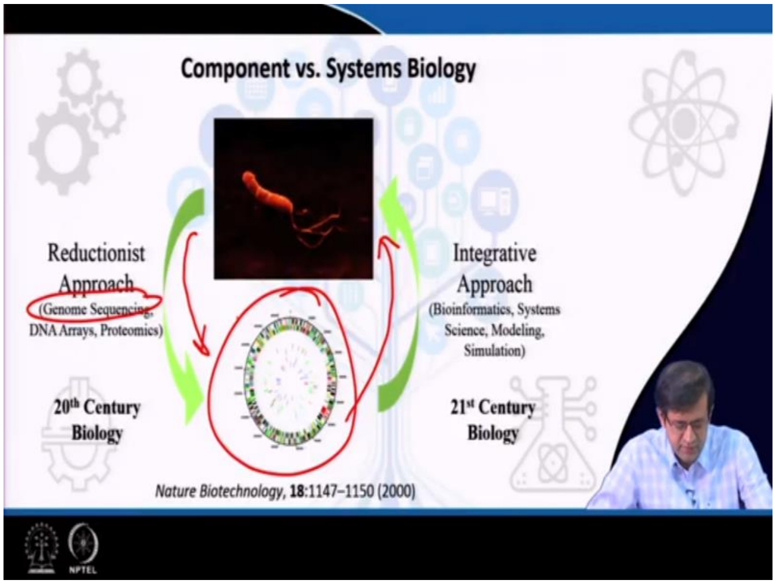
So, this it seeks to identify how biological function evolve and how emergent properties in cells and communities arise from seemingly simple and linear genetic sequence, you try to once you have the system, then you try to evolve the emergent properties inside the cell, what kind of function the biology, biological system can perform because of this interaction of different components. And this you are directly getting from the genome sequence the genetic sequence.

So, emergent properties because of the database because of you know already known protein. So, whenever a new protein comes from the genome sequence, you try to give a function was based on homology search. You try to search try to give a function to an unknown protein based on the based on the database you have for the known protein. The system biology is an interdisciplinary field that combines the knowledge and techniques from biology.

So, it is very interdisciplinary you need knowledge from biology, biochemistry, genomics, bioinformatics, chemistry and computational science but what makes it different from other field is the modeling. So, here you will be doing an extra part that is the modeling you have tried to combine these data and give a mathematical model. So, as you know, the metabolic networks need to be modeled, when you when you want to calculate the metabolic fluxes, you have to actually build a model first and the mathematical model.

So, in this course, we will be learning how to build a mathematical model and use the techniques which are available to build a metabolic network. So, the modeling is the essential part of this systems biology which includes knowledge from various disciplines biology, biochemistry, genomics, bioinformatics, chemistry and computational science.

(Refer Slide Time: 15:06)



So, here you can see there are 2 approaches whether that is the reductionist approach and the integrative approach. The reductionist approach where you can see that is the 20th century biology in 20th century, people will use to look for different components to identify 1 protein. Suppose, some genomic sequence identifies the number of gene and from, DNA array and proteomics identify different protein.

And try to give some function you try to allot some function and this is how we used to deal 20th century biology, where you identify different components like protein, DNA from the cell but 21st century biology is actually deal with it the other way around. So, it is in a way 20th century, we used to actually go in this way you identify the component from the cell and try to give some function.

But, 21st century deal with other way where you actually from the sequence you identify and build through bioinformatics, system science modeling, you try to give some function the cellular function. So, it is not the other way around, in other words, you identify the component but you are not giving any function to the cell but in this case, in the integrative approach, you are actually going from the sequence and finally, you are identifying the phenotype of the cell that is the emergent properties of the cell.

So, this is the only difference when you do systems biology. So, the component versus systems biology is now very crucial in the sense that we have many components available because of the high throughput technology. Because these sequences, what you see now are known for many organism and many are in pipeline because of the high throughput technology, you have a lot of genome sequence right now available, we have metagenomic sequence and so on.

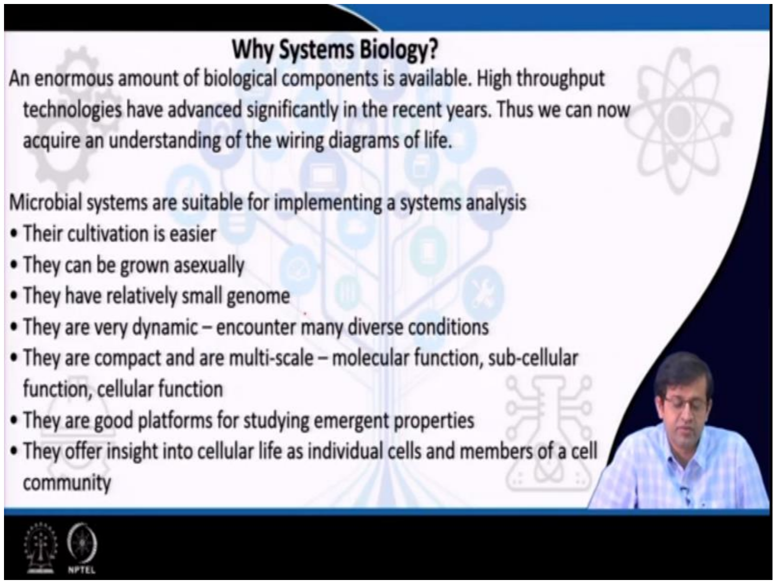
So, this sequence, you can actually build organisms based on the sequence and the systems biologists are working on that the job of the systems biologist is actually identifying different component from the sequence and then build the organism. So that you can predict new phenotype inside the cell.

(Refer Slide Time: 17:27)

So, this is in mechanical engineering and so you can build the car from this component. So, you can make tires, tubes and engine and try to assemble everything together and you build the new car but the biology is not done this way. In other engineering disciplines, you can see that each of the components they make and then they test it whether they are doing it good or not. And then finally, they assemble the component and to build the car.

And this technique is not followed in biology. But the systems biology technique is actually following the conventional way, where you the way the cars are built is exactly similar how systems biologist are building or making the component and try to assemble this component and build it metabolic model or metabolic network to identify different phenotypes of the cell.

(Refer Slide Time: 18:26)



So, why systems biology? The system biology an enormous amount of biological data is available. High throughput technologies have advanced significantly in recent years that is how we know and acquire an understanding of the wiring diagram of life. So, how these components are interacting each other the biological component inside the cell and that becomes the wiring diagram of life.

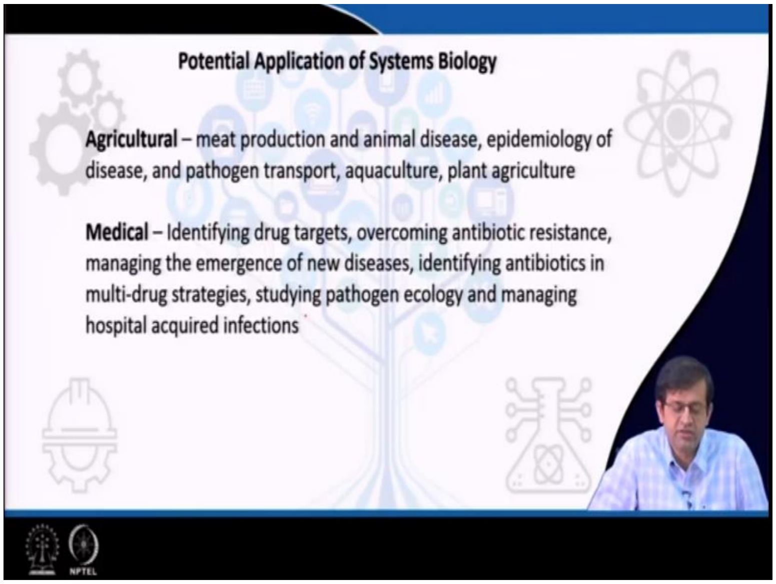
And then because of the advancement of high throughput technology, in the recent years, we have enormous amount of biological component that is a protein, DNA that have been identified and that you can use for building the wiring diagram of life. So, the microbial systems are suitable for systems analysis where we do metabolic engineering mostly we will be dealing with microbes.

So, systems biology of microbes are much comfortably because they can grow very easily the microbes are easy to grow, unlike any other mammalian cell which are very difficult to grow the microbial cells are easy they can grow asexually and then they have relatively small genome and then they have very dynamic encounter many diverse conditions. So, they can adapt in different condition you can grow them very easily.

They are compatible. They are multi scale molecular functional. Sub-cellular function and cellular, they are good platform for studying emergent properties. Ultimately, when you want to start systems biology of any organism, you start with the microbes because the microbes are actually easy to cultivate and they are good platform for studying a new function. Suppose you want to understand this cell and their function then it is easier using microbial system.

So, they offer insight into cellular life as individual cell and member of a cell and community. So, then through microbe you can learn back better than any other complex system. So that is why when you start with systems biology, you should start with the microbial system.

(Refer Slide Time: 20:48)

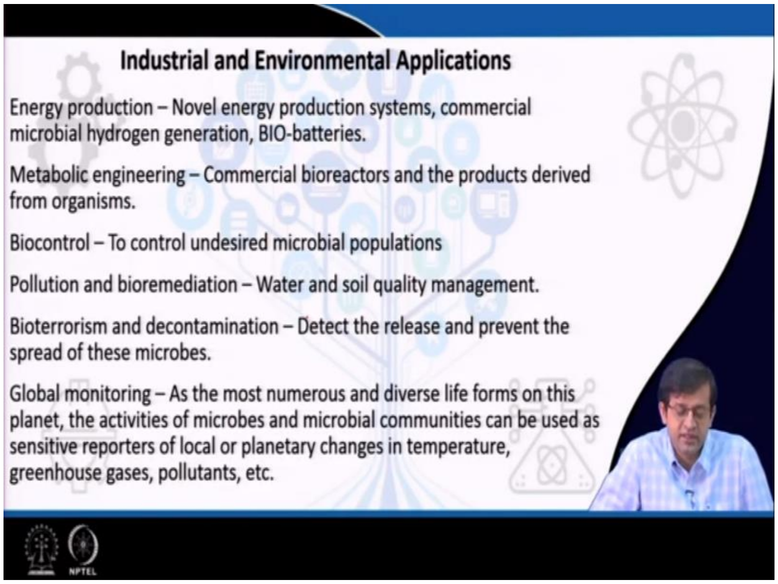


So, the application of systems biology; for example, the agriculture the systems biology very wide application. For example, the meat production animal disease, epidemiology of disease, pathogen transport, agriculture, plant agriculture. Systems biology is used extensively, where you understand how you can improve the production by understanding their metabolic gene regulation in a systems perspective.

Through systems perspective, you can handle the animal disease much better. And then in the medicine, also in medicine, you identify new target because you have observed that when you take a new drug, for suppose you have a disease and you are frequently taking that drug, then what happened, you may have worked on the disease but at the end of a couple of months, you end up with another disease.

So that that means the drugs you are taking is actually curing your problem but at the same time is you are actually affecting another protein or it is influencing another network and then that end up into another disease. So, this if you take it the systems perspective, then you can overcome these diseases, we are overcoming antibiotic resistance, managing the emergence of new diseases and then identifying antibiotics in multi-drug strategies studying pathogen ecology, managing a hospital acquired infection. So, those can be studied using systems biology.

(Refer Slide Time: 22:25)



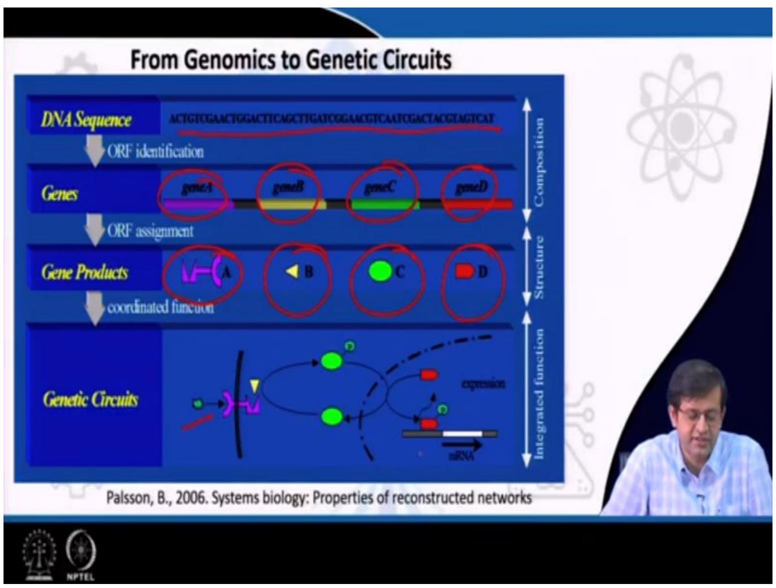
Then also the systems biology application in industry and environmental application for example, the energy production, the noble energy system, commercial microbial hydrogen production, the different energy production also require systems biology, understanding. And the metabolic engineering also requires understanding of the metabolic fluxes, how the metabolites the fluxes are changing with time you have to understand through systems biology.

And then bio-control to control undesired microbial population, then population, bioremediation water and soil quality management. Those can also be handled using systems biology. Bioterrorism, then decontamination detect the release and prevent the spread of the microbe and global monitoring as the most numerous and diverse life form on this planet, the activity of microbes and the microbial community can be used as a sensitive reporter of local or planetary changes in temperature, greenhouse gases, pollutant and etcetera.

So, the microbial communities you know which is present everywhere from your flower pot from your gut intestine and they are doing in the gut intestine and they are helping you in metabolism. But if you want to understand how these microbes are interacting and how they are actually influencing your health, it can be understood through systems biology, the metagenomic sequences are now available where you have many organisms in a same sample.

And you want to identify those microbes through their sequences are also feasible nowadays, through systems biology where systems biology plays an important role to understand how these microbes are interacting each other. So, the application of systems knowledge is very wide where you can apply the knowledge of systems biology to understand the system better. Systems perspective is required because you want to understand the system better so that much more emergent properties can be understood is from the cell.

(Refer Slide Time: 24:31)



So, this is the genetic circuit from genomics to genetic cell data. Suppose you have a DNA sequences and then from the DNA sequences, you identify how many open reading frames are there from the DNA sequence, supposing I give you a DNA sequence that is a genomic sequence the entire genome of the cell and try to identify how many open reading frames are there? Open reading frames are basically the start codon and the stop codon.

So that is where the gene starting and where the gene is ending. And based on ORF assignment, you identify how many genes are present inside the cell on the genome, Gene A, B, C like that you have 100s of genes inside the genome based on the order of assignment. And once you have the gene-A and gene-B identified in the chromosome then you identify how many gene products are there, how many gene products you are making out of this gene.

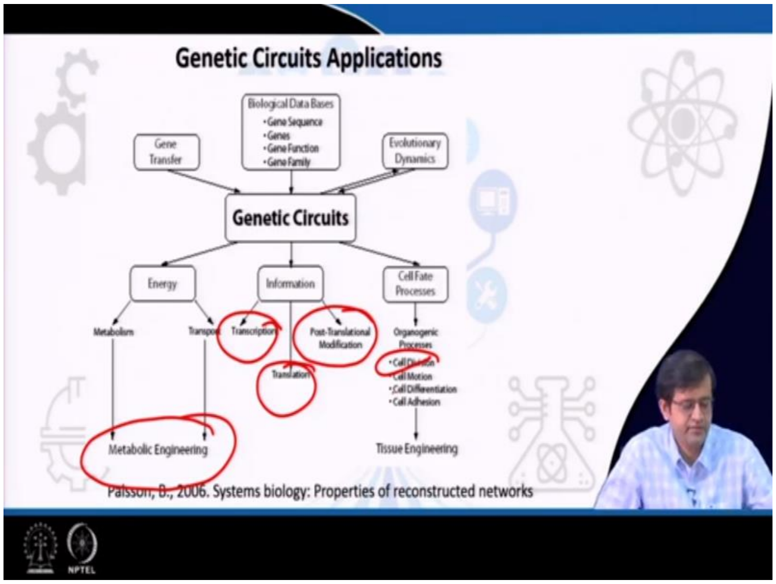
So, gene-A give rise to gene product A the gene products are basically RNA or protein, you identify your gene product gene. Gene-A gave rise to gene product A, gene-B gave rise to gene product B and then gene-C give rise to gene product C and gene-D gives rise to gene product D. And then this gene product; how they are interacting that is most crucial the interaction of this gene product.

The gene product which engage give rise to a cell new cellular function, the cellular function relies on the coordinated action of the gene product. So, how these gene products are interacting here you can see some kinds of signals is generated from external stimuli which is actually interacting with the protein and this protein is actually activating different transcription factor and this transcription factor is binding the DNA and where the gene is expressed.

So, this way the coordinated action of the various gene product giving rise to the expression of some gene and this is the integrated function, this genetic circuit is very crucial to understand the cell physiology or to understand the systems perspective of the cell you need to understand the gene genetic circuit present inside the cell. The function of genetic circuits is diverse that is DNA replication, translation glucose to pyruvate formation all requires different kinds of genetic circuit.

The information processing or this or any kind of cellular processes require the understanding of the genetic study. So, here from the genomics from the sequence that you identify how the genetic circuits are actually connected to genetic circuit remain very crucial in systems biology. And that you can get directly from the DNA sequence, from the DNA sequence you identify how many genes are there and the gene products are there and then how many genetic services are present?

(Refer Slide Time: 27:34)

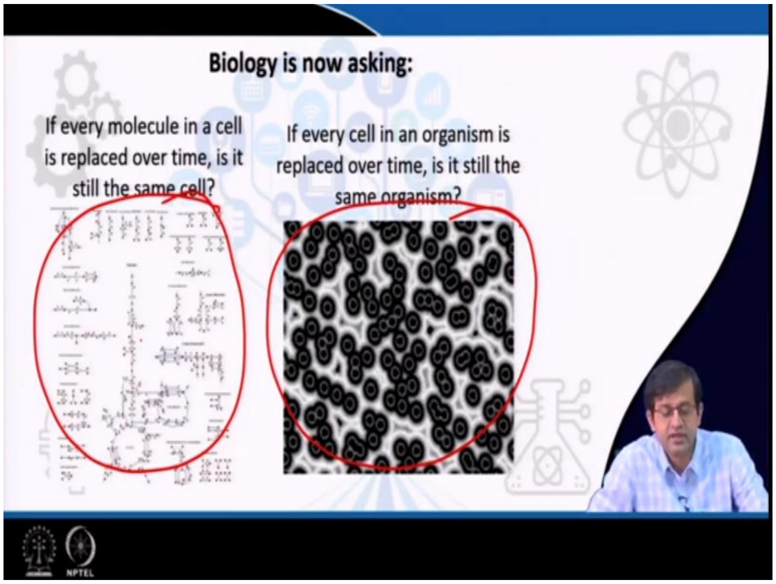


So, the application of genetic circuits you can see in energy production and then information processing or the cells process which I already told you energy, in the sense the metabolism where the currency metabolites are formed and how genetic circuits are present which are involved in metabolism, how the metabolite form because of different genetic circuits that you need to identify manually, it is helping in metabolic engineering.

Because metabolic engineering deal with removal or adding on new genes but, if you keep on adding and new genes or new things inside the cell, what will happen that your genetic circuit is getting disturb because the genetic circuit is unique for a given cell. So, understanding the genetic circuit will help better in metabolic engineering, it can also help genetic circuits are used for transcription and translation and post translation modification.

Also even it is used for cell division and cell adhesion, cell differentiation that gives rise to tissue engineering. So, all these dynamics inside the cells are actually given by the genetic circuit. The genetic circuit that you get from the; sequence the genome sequence and try to identify how many genetic circuit the cell may have.

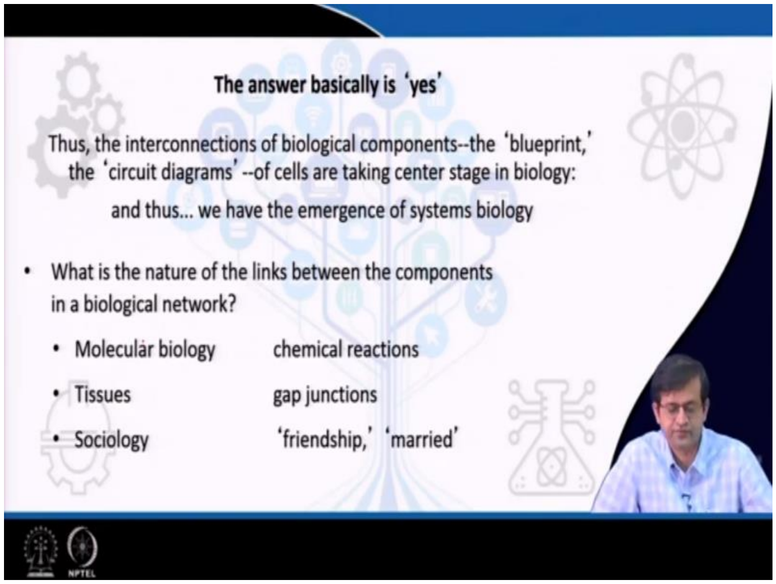
(Refer Slide Time: 28:59)



So, now, I asked a question if every molecule inside the cell is replaced over time, is it still the same cell or not? Suppose, you are actually replacing 1 cell at a time; removing replacing 1 cell at a time. Now, the question is that is it the same cell or it will be a different cell. Similarly, if every cell in every molecule inside the cell is replaced, first I told that this diagram is a cell where you have many cells and you are replacing 1 cell at a time with another organism.

And you want to see that whether the organisms remain the same or not. And in the other case, you are actually having the each of the molecule inside the cell you are replacing. So, each of the molecule which are basically metabolite replacing with and try to say that whether the cell remain the same or not because you are changing the cell component and also in the organism you are replacing the cell. Now, the question is that whether it will remain the same organism or it will become a different organism and what is the answer for you? Let me let us guess what could be the answer for this?

(Refer Slide Time: 30:12)



And the answer is that yes. So, the interconnection of the biological network, if every molecule is replaced over time is it remain the still same cell or it will become a different cell or if you replace the cell from the given organism will it remain the same cell or it will be a different organism, so, it will remain the same cell even if we change the component of the cell like metabolite it will remain the same even if you replace the cell of the organism.

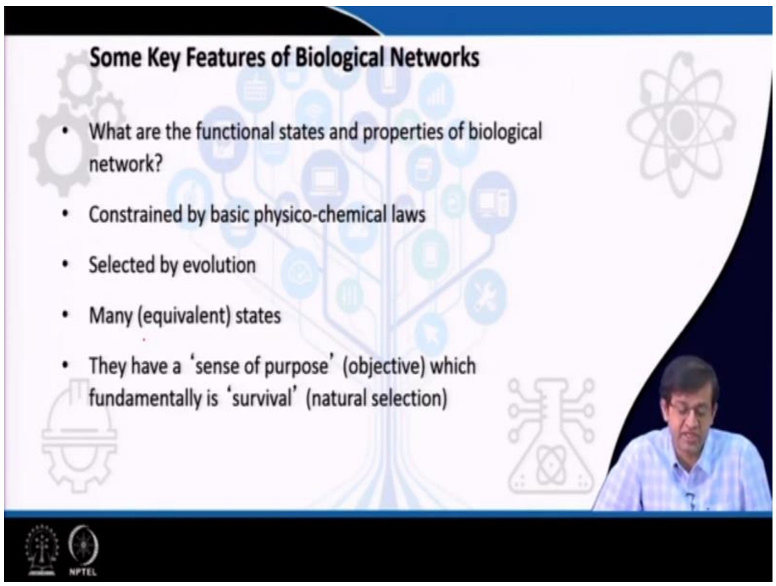
It will remain the same organism because the interconnection of the biological component that is the blueprint or the circuit diagrams of the cells are taking the center stage in biology. So, if you replace a component, you are not changing the interconnection. So that is the circuit diagram that is the genetic circuit present inside the cell it remains the same that is the blueprint of the biology and that is genetic circuits are more important that is what I want to say.

That the genetic circuits are more important and that gave rise to the emergence property of the cell that is the emergence of systems biology. If you know the genetic circuit that actually gives the wiring diagram of the life, so even if you change the component or the cell, this is not going to change the organism. So, then what is the nature of link between the component in a biological network, what kind of link do you have in a biological network.

The molecule in molecular biology the chemical reactions are basically the link between the metabolites and then in tissue we have the gap junction and then in sociology, we have friend how the 2 persons are connected to each other that is through friendship or they are married. So, this way, you have a network when you say the network any network has node and edges.

The connection between 2 nodes that is a connection between 2 persons in sociology is connected through that kind of relationship, you have. Similar in the biological network, you have the biochemical reaction network, the nodes are basically the metabolites and the link between the nodes are the reaction. So, this way, the networks are present in every field and you have to understand how they are connected. The connection between the; metabolites are important, rather than the metabolite itself that is what the circuit diagram means.

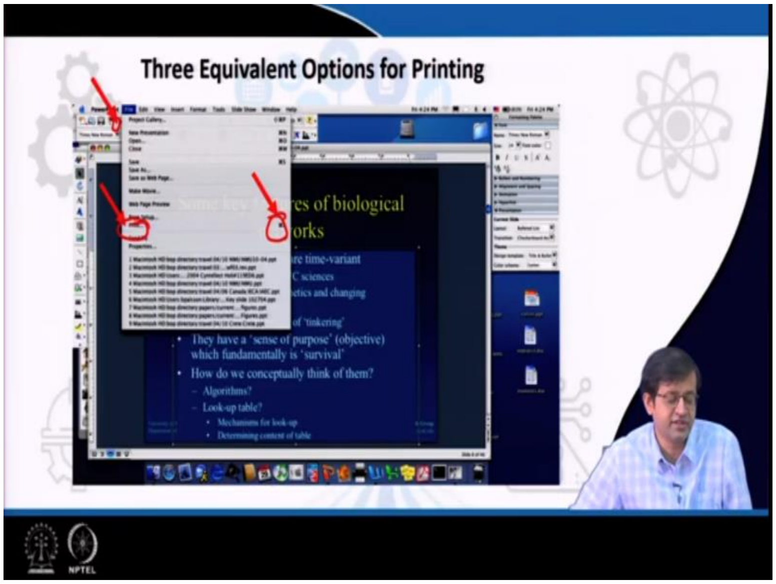
(Refer Slide Time: 32:47)



Some key feature of biological network, what are the functional state and the properties of the biological network, they are constrained by physiochemical laws selected by evolution, they have many states and they have a sense of purpose that is, they have an objective function. The objective function is basically the survival. So, every microbial cell have a sense of purpose, they want to grow more and they want to survive.

And these are the way you can define the objective function of the microbial cell. And the biological networks have in many states, many equivalent states and they are selected by evolution and they are also obeying laws of physics and chemistry as well. And the functional state and the properties of the biological network are defined by the different functional state present inside the cell.

(Refer Slide Time: 33:39)

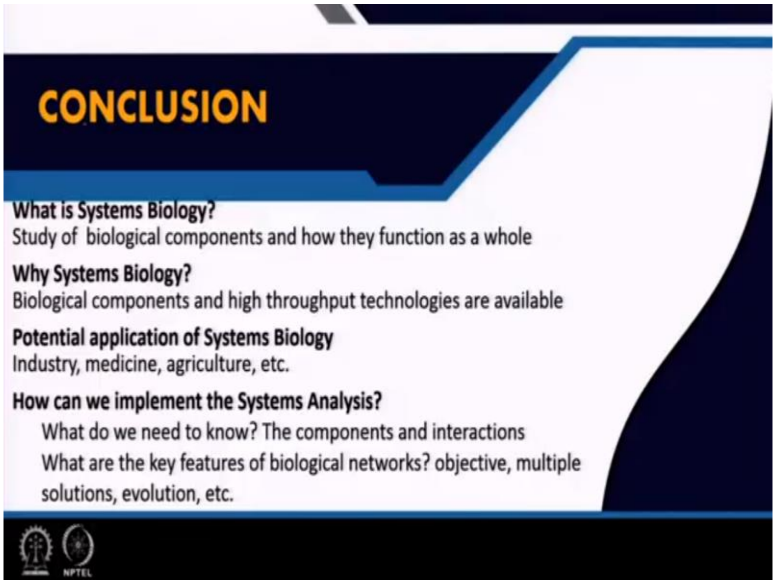


So, as I told the biological networks have many functional state, what does it mean? So, to give an analogy, I took an example. This is an example to understand how many equivalent state I told that biological networks have many equivalent states and that actually help the biological cell to be more robust, they are more they can survive in many condition because they have many equivalent state.

That is why the microbial cells are very robust compared to any other multicellular cell. And the biology of microbial cells are easy to grow because they have many, in overall like biological network, they do not have many state and that can be understood by using these example where you can see suppose you want to print this document. For printing the document, you have many options and either you go to the print option and print it, or you can use a command or you can actually have a button here on the tab.

So, 3 ways you can print these documents. So, you have multiple state equivalent options and these equivalent options are present inside the cell also, the cell can choose many ways to perform the same function. It can choose many ways in this cell, it can form many network circuit diagram that is generic circuit to perform the same function. So, this makes the biological network much more robust and they can survive in many conditions because the networks are many equivalent states are present.

(Refer Slide Time: 35:16)



So, in conclusion, we learned about what is systems biology? The systems biology basically study a biological component and how they function as a whole. So and then why there is systems biology, the systems biology actually deal with biological component and high throughput technology because of huge amount of biological data through high throughput technology which make system biology feasible.

And the potential applications of systems biology are there in industry, medicine, agriculture and etcetera. A huge amount of application, you can see through systems biology, how can you implement this system analysis, what do you need to know? System biology deal with systems analysis and for that you require the component that is the nodes and how they are interacting and that there is a connection between this node so that it can form a network.

What are the key feature of a biological network which as I discuss, it has an objective function that every network has an objective function and the multiple solution state exists and this is achieved through evolution. So, every biological network, the microbial cell, or you can consider any biological network, they evolve over time. So that the best interconnectedness or the best wiring diagram are taken up by the cell through evolution.

So, that gave rise to maximum survival, the cell can survive more by considering that particular network. So, this lecture, I gave you some introduction of about the systems biology which will help you to understand more about the biological network.

(Refer Slide Time: 36:58)

So, these are the books that you can follow the systems biology the properties are reconstructed network written by Bernard Palsson and also you can refer to some of the paper from Bernard Palsson that is the challenges in-Silico Biology. 2 dimensional annotation of genome. So, in the next lecture, we will learn more about the metabolic network, how we are going to use this system biology to construct the metabolic network. Thank you. Thank you for listening. Hope you understand systems biology in this lecture and will see you in the next lecture.