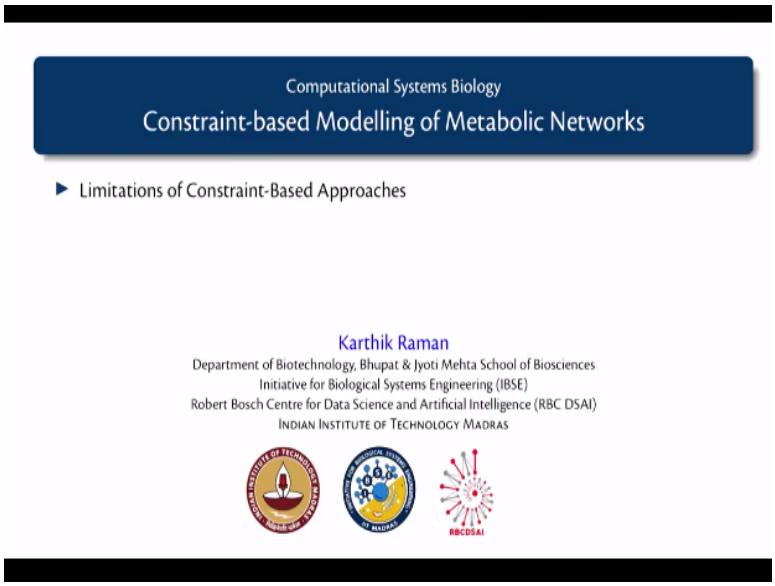
Computational Systems Biology
Karthik Raman
Department of Biotechnology
Indian Institute of Technology – Madras

Lecture - 70
Constraint-based Modelling of Metabolic Networks

In today's video we will look at some of the limitations of constraint-based approaches and we will see that most of these limitations arise from the model itself and not the modeling approach as such and we will see what they are.

(Refer Slide Time: 00:20)

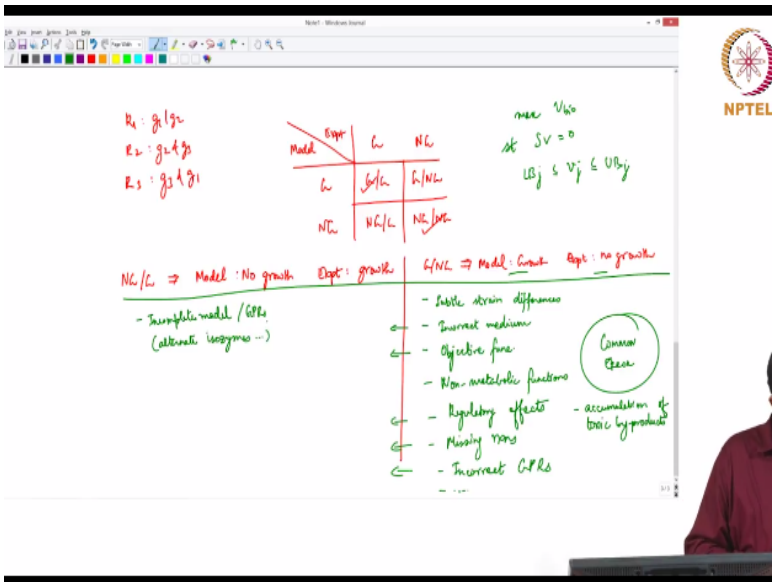


So let us study the limitation of constraint-based methods and then you know in the next classes we will go on to some applications, examples and you know how do we model compartmentalized models and so on. So what are the limitations of constraint-based methods, can you think of some already? So the first obvious set of limitations are what plaques every type of modeling approach is that you are limited by the data and so on, right.

You are limited by your own biochemical knowledge and so on but the other important aspect here is that you are somehow restricting yourself only to the metabolic lens of viewing the cell. You only view the cell through a metabolic lens. You cannot predict anything regulatory or something like that and a case in point is that even if you consider fba, I mean the lac operon, fba will make a wrong prediction on lac operon.

So it will assume that e-coli grows at a higher growth rate utilizing both lactose and glucose simultaneously instead of correctly figuring out that fba will first you know utilize all the glucose and then switch over to lactose because there is no regulatory logic that has been put in here, but that is something we will see in a future class as to how we can implement regulatory logic in flux-balance methods.

(Refer Slide Time: 01:41)



So there are 2 types of there are, you know, different kinds of prediction errors that can happen. So you can let us say this is the model and this is experiment, right. You can have growth, no growth, growth, no growth, right or rather you can say G/G. These 2 are good, right, true positives and true negatives, the other 2 are problem, false positives and false negatives. Let us first consider the NG/G scenario, which means model, no growth; experiment, there is growth.

What could be the reasons underlying this? and let us have the other thing in parallel as well so, what could be the reasons for this, **"Professor - student conversation starts"** wrong identification of lethals, **"Professor - student conversation ends"** yeah so that is the issue right. So why is the identification of lethal is wrong. So some common reasons could be you know missing data in the model and objective function and so on.

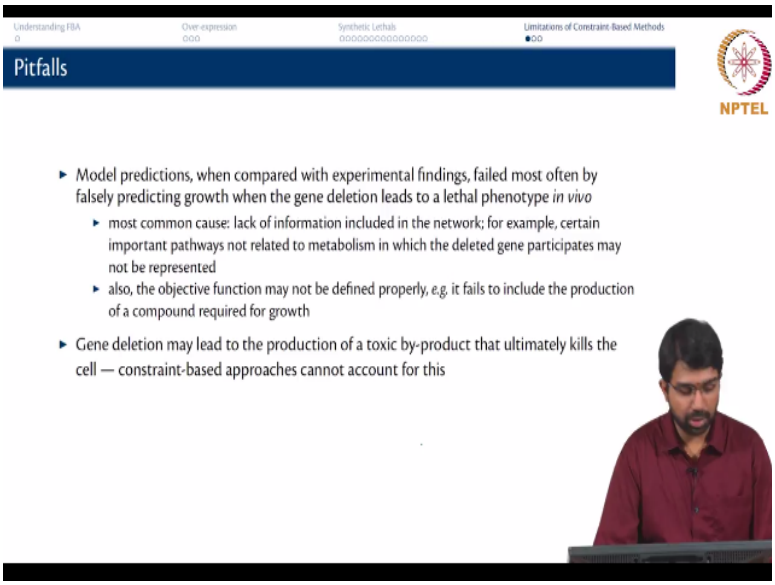
But beyond that what else could be the reasons, when would your model show growth, but the experiment shows no growth? **"Professor - student conversation starts"** there could be genetic difference with the strain they are working on and the model, so that could be an important thing, right, so subtle strain differences **"Professor - student conversation ends"**

There is still like a gap in the model, but it is really not that obvious, right you might be working with mg1655 in the lab and k12 in the model and so on, e-coli and that could give rise to some of these issues, very good, what else, yeah so incorrect medium. Some of these could also go here, right. So they do not need to be only one way errors they could cause errors in both places, incorrect medium.

See maybe you can just think of it in terms of the model. What is the model? max v/o such that Sv = 0 and each of these could be wrong, right. So you could have some problem with the objective function. Maybe it does not account for an important metabolite that needs to be produced. So the model seems to grow whereas in real life it does not grow.

Nonmetabolic functions, and commonly regulatory effects, missing reactions, incorrect gene protein reaction associations, and so on and so forth. So obviously this is the more common error. This error is a little less common, but that could again arise out of many of these things. So this can go here, this can go here, this can also come here, this can also come here, this can also come here, nearly all of the are, many of these are applicable.

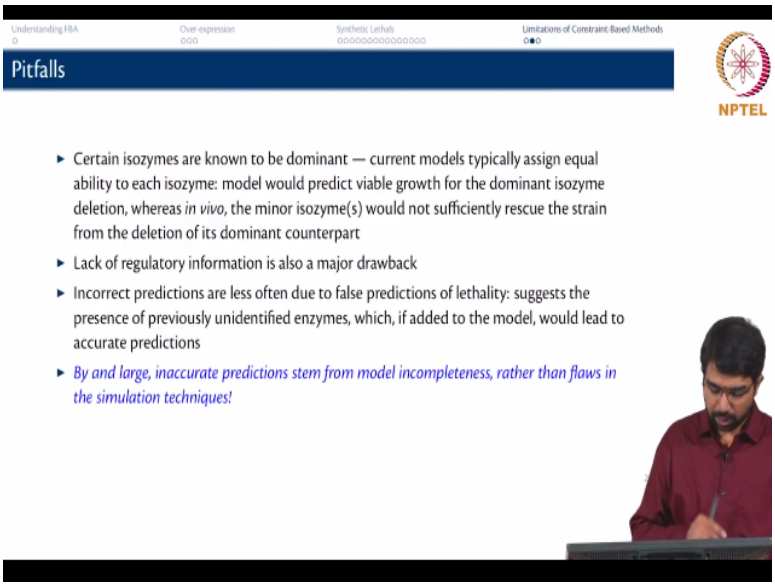
(Refer Slide Time: 06:46)



So the most common causes lack of information includes in the network for example certain important pathways not related to metabolism in which the deleted gene participates may have been ignored. Of course the objective function may not be defined properly, it does not include an important compound necessary for growth and gene deletion may lead to the production of a toxic byproduct which accumulates.

So this is something that we do not account for right, so here we can add that, right. **"Professor - student conversation starts"** yeah, so if the enzyme also has some structural role, right, so when you delete the enzyme it is metabolic reaction is not important but it destroy cell structure so the cell dies.

(Refer Slide Time: 08:01)



And the other issue is certain isozymes are known to be dominant right. Currently we basically assign equal weightage, right, in the gpr we just say A or B. So if A is there or B is there we assume the reaction goes on, but what happens if for A, the reaction upper bound is 100% whereas for B the reaction upper bound is 40%. If one gene is like weaker or doing the same job, so this may not sufficiently rescue the growth.

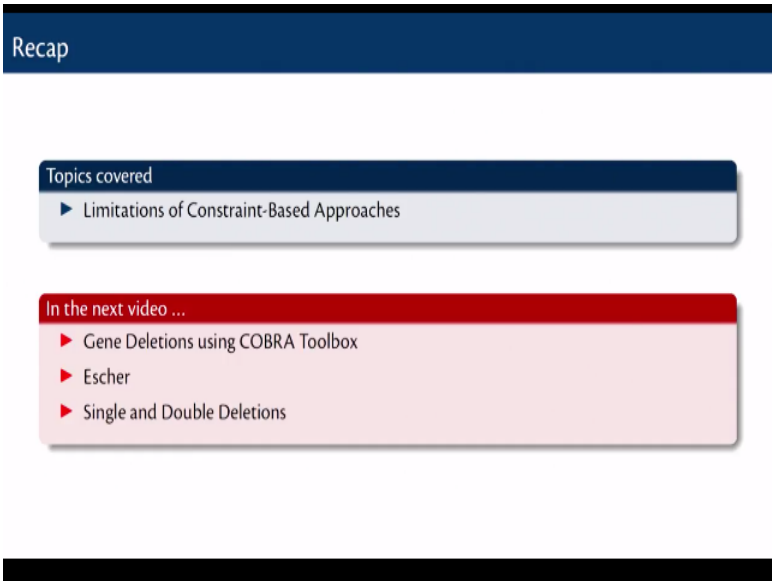
So when you delete A there is no effect or like when you delete B there is not much effect, but when you delete A the cell can only grow at 40% of it is maximum or something like that and the lack of regulatory information is also a major drawback right. And usually incorrect predictions are due to false predictions of lethality, this can also be because of other unidentified enzymes again.

You know some we did not exactly say incomplete model, so we could say incomplete model right you could have incomplete models or cprs like you know alternate isozymes et cetera. So in by and large the important thing to note of course is that the in accuracy stem from problems with the model not the simulation or the modeling approach itself, right, so fba as

such is very reasonable except for the fact that it ignores regulation or you know in some sense ignores non-metabolic activities of the cell.

Although some of it is somewhat accounted for by the ATP maintenance flux and so on, but the flux balance constraint-based modeling method by itself is not very fled there are you know very good agreements with what you find in literature and so on.

(Refer Slide Time: 09:50)



In today's video I hope I introduced you to some of the limitations of constraint-based approaches. As we have discussed even in a previous lecture, I mean all models are wrong so are constraint-based models, but they are very useful to give very interesting insights so and we also, I hope you are able to appreciate the fact that you know the modeling methodology is actually quite good, but you can have a lot of knowledge gaps which affect the accuracy of your predictions and so on.

In the next video we will do a lab session wherein we perform gene deletions using the COBRA toolbox, I will also show you this interesting tool called Escher and we will also look at how you do single or double gene deletions or synthetic lethals using the COBRA toolbox.