Results for presentation on 31-08-2015

Tamar Huygen

February 28, 2016

1 Algorithms for genome-scale models

1.1 Genetic instability

In biotechnology products are often produced by organisms which have been genetically modified in such a way that they produce a substance and excrete it out of the cell.

Often this is done by inserting a construct into its genome. One of the problems, which according to us has been given too little attention, is the genetic instability of these constructs. In most bioreactors there is a continuous selection pressure for fitness (chance of surviving until reproduction times the amount of offspring). If a new strain in a bioreactor has a slightly higher fitness it will soon take over the entire culture. Producing a substance has, almost always, a negative effect on an organisms fitness. Simply because it now has to divide its available resources over the production of biomass and the production of the said substance. Sometimes a single mutation in the promoter region is able to silence a whole gene construct. These 'cheaters' will then become more and more abundant, because of their elevated fitness, and will eventually take over the whole culture. This is undesirable in industry, since the culture has to be removed and re-initiated every time this happens, thus decreasing the yields.

One of the solutions of this problem is to produce a product an organism already can produce by deleting genes that would be responsible for the consumption of the product. Recreating a whole gene takes a lot more evolutionary time than a single mutation to silence one, and this would thus be more stable. One still has to be careful, because maybe the genes responsible for the production of the substance of interest can still be silenced. Therefore it is also desirable to make the production of the product also growth coupled. This means that the substance of interest is produced in a pathway that is used during growth.

An example of this is the acetate producing strain we use. In one of the reactions that is needed to produce biomass, acetate is produced as a side product. Normally acetate gets recycled in the organism by another reaction. The gene responsible for this recycling is knocked out. Now acetate can no longer be recycled into the system and gets excreted into the medium. This way of producing acetate is more stable than the insertion of a gene. In order to grow the organism has to produce it and it cannot simply silence the production by a mutation. We wanted to see if we could find more products which could be produced in this way. In order to do this we have created an algorithm that looks for genes that can be knocked out in order to produce a carbon product in a genetically stable way. For iGEM it is a useful tool to find other ways to create a stable consortium in which one of the organisms produces a substance which the others consume.

The algorithm is written to be used on genome-scale FBA models. The general outline of the algorithm looks like the following:

- Find all carbon products associated to the formation of biomass.
- For each of these products: find all reactions that are associated to the product.
- Find all different combinations of reactions that are associated to the product.
- For each reaction in each of these combinations, find all genes that are associated to the reaction.
- 'Knock out' these genes by turning off all reactions that have one of these genes also associated to them.
- 'Turning off' means that the boundaries of the steady state flux through these reactions are set to zero.
- Run the the model with the adjusted flux boundaries and optimize for the formation of biomass.
- Check if the organism still grows (biomass is still formed).
- Check if the product of interest is produced.

- Print out all successful combinations of reactions and associated genes.

There are still some issues with the algorithm:

- Some reactions in the models are not thermodynamically correct, this may yield some wrong results.
- Then there is still a possibility that in real life the organism doesn't necessarily need the products to form biomass and then a mutation in a gene responsible for the formation of the product can still yield a cheater.
- As of now the algorithm still only looks for reactions directly involved with the product, while more potential genes can be found if we would also look at reactions further away in a pathway. In other words, the algorithm doesn't find all possible genes. It is more of a heuristic algorithm than an exploitative one.
- Maybe the minimal production of the substance equals zero and the organism would still be able to grow without producing the substance. In order to check this, a flux variability analysis should be performed to see if the minimal production is higher than zero.

1.2 Auxotrophies

For our iGEM project we wanted to create an auxotrophic Synechocystis strain. Creating an auxotrophy can be useful in a consortium to make organisms dependent on each other. But in a much broader sense an auxotrophic organism has more boundaries as to where and when it grows, which can be used in a lot of ways. We created an algorithm which looks for genes that can be knocked out in order to make an auxotrophic organism. The algorithm again works on genome-scale FBA models. It also takes a list of substances for which it looks for ways to make an organism dependent on it. The general outline of the algorithm looks like the following:

- For each substance on the list, look for source reactions which create the substance in the extra cellular medium.
- Find reactions associated to the substance (which produce or consume the substance)
- Find all combinations of these reactions
- For each reaction in a combination, find all genes associated to it
- 'Knock out' these genes in the same way as the previous algorithm
- Run the model and optimize for biomass (At this stage there should be no growth possible)
- Create a source reaction of the substance in model.
- Run the model again and optimize again for biomass.
- If biomass is formed and the source reaction is used, knocking out the combination of genes is a possible way to create an auxotrophic organism.

With this algorithm there are also some issues:

- It still doesn't find all genes that can be knocked out in order to make an auxotroph since it only looks for reactions directly involved with the substance.
- It is possible that an organism is actually not able to transport a substance into the cell. If an organism lacks a transport mechanism, it will not be able to grow if the genes are knocked out even with the substance in the medium. The models are still not complete enough to be able to tell if an organism is able to transport a substance into the cell, so this has to be experimentally validated before the genes are knocked out.

1.3 Sinks

(I think that this is too much detail for the presentation, but I will say something about it in here anyway, in case there are questions.) In both algorithms there are reactions turned of by setting the flux boundaries to zero. In an FBA model the build-up of a species is not possible, since the system has to be in steady state. If this happens, the only steady-state solution that is left is no growth at all. In real life an organism might be able to get rid of a substance by excreting it to the extracellular space. In order to model this, we create sink reactions, which make a substance disappear into nothingness. We do this for all metabolites associated to the reactions which are turned off. For most substances these sinks won't be used, because the model is

optimized for growth. But for example for the substance we want the organism to produce in the new products algorithm, it has to use either a transport reaction or a sink. We check after the model has ran which sinks are used and how big the flux through the sink is. However if there is already an transport mechanism in place, it can sometimes be better to use the existing transport mechanism instead of the sink. These transport reactions usually cost energy in terms of ATP and if an in the model we can make the organism cheat by creating the sink. This may yield wrong results if we want to do a flux variability analysis.

2 Mathematical models

There are several questions about our consortium we want to answer with mathematical modeling. One of these questions is what the effect cell densities will have on the light conditions and how this will have an effect on the yields. Other questions are is about what initial conditions will be optimal to create the consortium? does the initial ratio of the biomass of the different organism have an effect? Are there other possible ways to create a consortium than cross-feeding? In order find out we are creating mathematical models base on first order differential equations in mathematica. These models are still under construction. The first model we created was created to look at the effect of shading of two species in a flask. One organism is phototrophic and the other is dependent on a limiting substrate the phototrophic produces. We started out really simple:

$$\frac{da}{dt} = \mu_a a \tag{1}$$

$$\frac{db}{dt} = \mu_b b \tag{2}$$

$$\frac{ds}{dt} = k_{syn}\frac{da}{dt} - k_{con}\frac{db}{dt}s \tag{3}$$

Herein is a the phototrophic organism which produces substance s and b is the organism growing on s. Organisms a and b have growth rates of μ_a and μ_b respectively. Substrate s is in this case only produced as a grows and only consumed when b grows, with k_{syn} the synthesis rate of s and K_{con} the uptake rate of s. is produced as To model the substrate dependence of b we used a simple Monod equation:

$$\mu_b([S]) = \mu_{max,b}\frac{[S]}{[K_S] + [S]} \tag{4}$$

Herein is μ the specific growth rate, $[S]$ the concentration of a limiting substrate and k_S the half velocity constant and $\mu_{max,b}$ the maximal growth rate. It looks very similar to a Michaelis-Menten equation for enzymes, but rather than theoretical arguments this one is mainly based on experimental results. Light intensity has also an effect on the growth rate of a. This is modeled using the following equation as given by Franco et al. (2006)[1]:

$$\mu(I) = \mu_{max,a}\frac{I}{K_{S,I} + I + \frac{I^2}{K_1}} \tag{5}$$

Herein is I the light intensity. There is however a bit of discussion as to whether the growth rate should be dependent on light intensity but another parameter to measure the amount of photons available. In the most recent model we made the light intensity linearly dependent on the amount of biomass of both species in the consortium, while in real life this not the case, but it is modeled this way just to keep it simple. Now in a flask you would expect that first species a will grow, which yields substance s, then b can grow. But the growth of a will go down, because of the shading. Eventually it will stop growing because there is no light left and then b will stop growing because there is no more s left after a while (see figure 1). The parameters are still not correct in these models, so the scales can be safely ignored. The behavior of the different species however may be explored.

Indeed we can see from figure 2, that indeed the growth rates drop to zero after a while. In figure 3 we can see the formation of s and we can see that it runs out again.

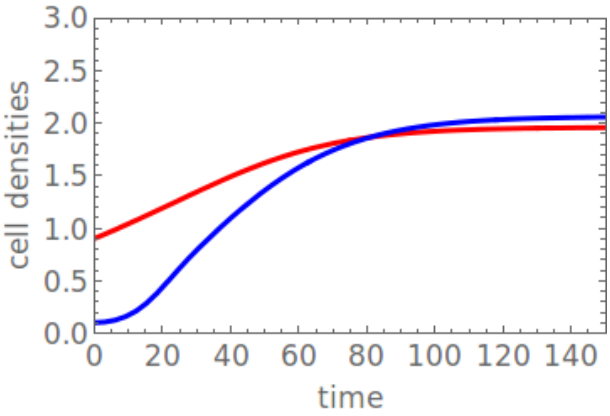

Figure 1: Cell densities of a(red) and b(blue) parameters are not yet correct, but we can see their biomass.

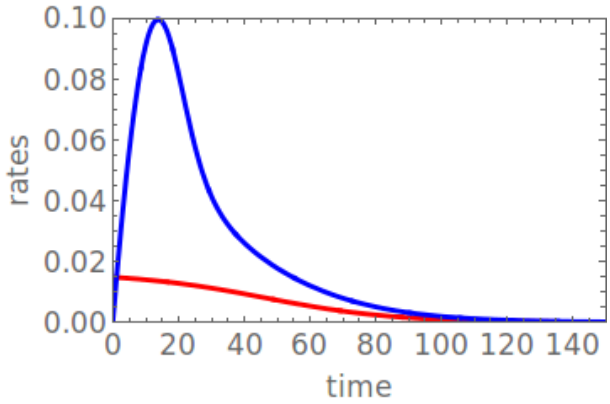

Figure 2: Growth rates of a(red) and b(blue). Parameters are not yet correct, but we can see the behavior of the growth rates.

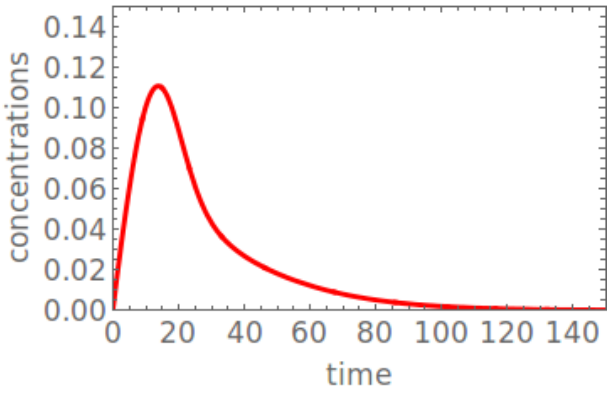

Figure 3: Growth rates of a(red) and b(blue) parameters are not yet correct, but we can see the s being formed and consumed.

References

[1] Ezequiel Franco-Lara, Jan Havel, Frank Peterat, and Dirk Weuster-Botz. Model-supported optimization of phototrophic growth in a stirred-tank photobioreactor. *Biotechnology and bioengineering*, 95(6):1177–1187, 2006.