Semana da Escola de Engenharia
October 24 - 27, 2011

Metabolic Engineering Approaches for Strain Optimization using Evolutionary Computation Techniques

Pedro Evangelista[1], Miguel Rocha[2] and Isabel Rocha[1]

1 - IBB-Institute for Biotechnology and Bioengineering, Centre of Biological Engineering, Address Universidade do Minho, 4710-057 Campus de Gualtar, Braga, Portugal

2 - CCTC - Computer Science and Technology Center, Address Universidade do Minho, 4710-057 Campus de Gualtar, Braga, Portugal.

E-mail: ptiago@deb.uminho.pt, mrocha@di.uminho.pt, irocha@deb.uminho.pt

KEYWORDS
Metabolic Engineering, Systems Biology.

ABSTRACT

One of the main purposes of Metabolic Engineering is the quantitative prediction of cell behavior to provide for selected genetic modifications. The purpose of the present study is to explore methods in which dynamical models provide for phenotype simulation such that strain optimization algorithms can be used to indicate enzyme under/over expression and deletion of a few reactions (knockouts) that maximize the production of target compounds with industrial interest. It is not apparent how to find the best set of genetic modifications in order to enhance the production of a target compound due to the high number of interacting reactions.

This work details on the developed optimization algorithms, based on Evolutionary Computation approaches, to enhance the production of a target metabolite by: (i) finding a (near) optimal set of reaction knockouts or (ii) changing the levels of expression of a set of enzymes.

As a basis for phenotype simulation, two metabolic dynamical models of selected pathways of *Eschericia coli* were used based on ordinary differential equations, namely:

(i) The full mechanistic model of the central carbon metabolism (Chassagnole 2002) consisting of mass balance equations for glycolysis and for the pentose-phosphate pathway;

(ii) The Lin-log version of the previous model extended with the TCA cycle, glyoxylate bypass and acetate metabolism.

Here dynamic models are used to generate a single steady-state solution without the need of specifying further assumptions like in the cases of Flux Balance Analysis (Kauffman 2003), Minimization of Metabolic Adjustment (Segre 2002), or Regulatory on/off minimization of metabolic flux changes (Shlomi 2005) for stoichiometric models.

In order to properly guide the optimization algorithms and evaluate the *in silico* strains, the following fitness function was employed: the mutant target metabolite drain flux value at steady-state was divided by the wild-type counterpart.

Three tasks were used to test the devised optimization techniques whose purpose was to maximize the production of a metabolite at steady-state: (i) Reaction Deletion: The objective of this task is to discover the best set of reaction deletions. The ideal number of reactions to remove is also determined in parallel; (ii) Reaction up/down regulation: The main goal is to find the best set of enzymes to tweak and the respective level of expression concerning the base values present in the original model; (iii) Reaction up/down regulation utilizing discrete expression values and a maximum predefined number of reaction modifications.

The aforementioned tasks were used in two case studies. In the first case study dihydroxyacetone phosphate was selected as target metabolite due to the fact of being present in both models, allowing to compare the obtained results. In the second case study only the Lin-log model was utilized with succinate as target metabolite. Nowadays these compounds play a major role in several industrial applications. Dihydroxyacetone is used in synthetic chemistry using the enzymatic aldol Syntheses and succinate is produced by petrochemical processes in order to manufacture bulk chemicals (Mckinlay 2007).

Semana da Escola de Engenharia
October 24 - 27, 2011

In the dihydroxyacetone knockout case study the best solutions obtained using the mechanistic and the lin-log model are similar except for one reaction as it can be observed in Table 1. Changing enzyme expression levels produces in this case study better results than the reaction deletions in the Lin-log model due to presence of the tricarboxylic acid cycle (results not shown).

Table 1: Best Reaction deletion set found for each target compound and model

Model	Target Compound	Reaction Deletion Set
Mechanistic	Dihydroxyacetone	G6PDH, PEPC, PK, TKA
Lin-log	Dihydroxyacetone	G6PDH, PEPC, PK
Lin-log	Succinate	G6PDH, PEPC, SDH, FUM, ACKA

Nonetheless, it is possible to envision differences between the lin-log and the mechanistic model. For instance the inclusion of the reaction transketolase, reaction A (TKA) in the best solution obtained using the lin-log model leads to a lower value of the objective function contrary to what happens in the mechanistic model. It is important to note that the lin-log model is a linearization around a reference state. These reaction knockouts can put the model ”far” from its reference state giving rise to erroneous extrapolations

In the succinate production case study, the results tend to indicate that the best set of genetic modifications is attainable by reaction knockouts without the need to tweak the enzyme expression levels. It is important to bear in mind that the glyxolate shunt is implicitly active and the pox-b reaction knocked-out in (Lin et al. 2005) is not present in the model. Nonetheless, the algorithm has the capacity to find a similar solution to the penta mutant described in (Lin et al. 2005), although there are some differences. On one hand the knockout of the Pepcarboxylase (PEPC) leads to an increase in pep metabolite concentration in both employed models. This overexpression has been previously described in literature to enhance succinate production *in vivo* (Gokarn et al. 1998). Another dissimilarity is the knockout of Glucose-6-phosphate dehydrogenase (G6PDH) suggested by the algorithm not described in the literature. The isocitrate dehydrogenase reaction (ICD) is not deleted because it influences negatively the succinate production contrary to what was done in (Lin

et al. 2005). The presented solution maintains both branches of the tricarboxylic acid cycle that lead to the production of succinate by knocking out succinate dehydrogenase (SDH). The fumarase (FUM) knockout effect is neutral since it is the only reaction after the SDH reaction.

In future work, the main issue to be tackled is the validation of the results in a *in vivo* setting.

REFERENCES

Chassagnole, C., Noisommit-Rizzi, N., Schmid, J., Mauch, K., and Reuss, M. (2002). Dynamic modeling of the central carbon. Biotechnol Bioeng, 79, 53–73.

Kauffman, K., Prakash, P., and Edwards, J. (2003). Advances in flux balance analysis. Current opinion in biotechnology, 14(5), 491–496.

Segre, D., Vitkup, D., and Church, G. (2002). Analysis of optimality in natural and perturbed metabolic networks. Proceedings of the National Academy of Sciences of the United States of America, 99(23), 15112.

Shlomi, T., Berkman, O., and Ruppin, E. (2005). Regulatory on/off minimization of metabolic flux changes after genetic perturbations. Proceedings of the National Academy of Sciences of the United States of America, 102(21), 7695.

McKinlay, J., Vieille, C., and Zeikus, J. (2007). Prospects for a bio-based succinate industry. Applied microbiology and biotechnology, 76(4), 727–740.

Gokarn, R., Eiteman, M., and Altman, E. (1998). Expression of pyruvate carboxylase enhances succinate production in escherichia coli without affecting glucose uptake. Biotechnology letters, 20(8), 795–798.

Lin, H., Bennett, G., and San, K. (2005). Metabolic engineering of aerobic succinate production systems in escherichia coli to improve process productivity and achieve the maximum theoretical succinate yield. Metabolic engineering, 7(2), 116–127.