Universidade do Minho
Escola de Engenharia

Semana da Escola de Engenharia
October 24 - 27, 2011

Methods for Strain Optimization Analysis Using Metabolic Network Comparison

José Pedro Basto Gouveia Pereira Pinto
CCTC and CEB/IBB
Campus Gualtar, Univ. Minho
E-mail: josepedr@di.uminho.pt or josepedr@gmail.com

KEYWORDS
Metabolic Engineering, Strain optimization, Metabolic networks, Network visualization.

ABSTRACT

Recent efforts in Bioinformatics and Systems Biology allowed the development of genome-scale metabolic models for several microorganisms (Adam M Feist et al. 2008). These models have been used to guide biological discovery promoting the comparison between predicted and experimental data, to foster Metabolic Engineering (ME) efforts in finding appropriate genetic modifications to synthesize desired compounds, to analyze global network properties and to study bacterial evolution (Adam M Feist and Bernhard Ø Palsson. 2008).

The most popular approach to phenotype simulation considers the cell to be in steady-state and takes reaction stoichiometry/reversibility in a constraint-based framework to restrict the set of possible values for the reaction uses. Cellular behaviour is thus predicted using for instance Flux Balance Analysis (FBA), based on the premise that microorganisms have maximized their growth along natural evolution (R.U. Ibarra et al. 2002). Using FBA, it is possible to predict the behaviour of microbes under distinct environmental/genetic conditions.

The combination of reliable models with efficient simulation methods has been the basis for different strain optimization algorithms. Their goal is to find the set of genetic modifications to apply to a given strain, to achieve an aim, typically related with the industrial production of a metabolite of interest.

In previous work, an approach based in the use of metaheuristics, such as Evolutionary Algorithms (EAs) and Simulated Annealing (SA), has been proposed to solve the optimization task of reaching an optimal (or near optimal) subset of reaction deletions to optimize an objective function related with the production of a given compound (M. Rocha et al. 2008). The idea is to force the microbe to synthesize a desired product, while keeping it viable.

The next logical step is to validate these results in the lab, a task that given its associated costs should be preceded by a thorough analysis of the solutions provided using computational methods. This screening process could identify more promising approaches and, thus, save resources in wet lab experiments.

In a first stage, the validation of the results involves the understanding of the strategies followed by these mutant strains to achieve the desired phenotype, by studying the different use of reactions/pathways to achieve the desired metabolite and still keep the strain viable. This becomes quite complex given the size of the networks involved in genome-scale models and the interactions between (sometimes distant) components. The manual verification and comparison of the phenotypes of different mutants is typically impossible.

The goal of this work is the development of automatic methods to analyze large sets of mutant strains for specific ME problems using network analysis tools. These methods take the phenotypes of a large number of possible solutions obtained by running strain optimization algorithms and attempt to identify shared patterns, taking advantage of methods from network topology analysis. The topological comparison between the networks provided by the wild type and mutant strains highlights the major changes, thus highly contributing to elucidate the strategies that lead to successful mutants.

Semana da Escola de Engenharia
October 24 - 27, 2011

The methodology developed can be divided into four distinct parts:

1. The processes necessary to create metabolic networks specific to certain strains of the organism by using the results of an *in silico* simulation to filter an original network with all available metabolic information related with the organism, the result is a smaller metabolic network containing only information related with parts of the metabolism which the specific strain uses according to the simulation.

2. Metrics for comparing the topology of different metabolic networks, these were developed because during the early stages of this project it was noticed that viable mutants have very similar metabolic networks to the wild type organism, this limits the application of many traditional network comparison metrics which lead to the necessity of developing new ones.

3. A generalized comparison algorithm which takes the results of multiple wild type-mutant network comparisons and creates a network, called variation network, which contains the more frequent differences.

4. A methodology for the analysis of the variation network in order to determine the biological meaning of the differences found.

Currently in order to validate this methodology it was applied to a case study considering *Escherichia coli* as the host and aiming at the production of succinate, by optimizing the set of gene knockouts to apply to the wild type. Large sets of solutions (mutants) provided by the use of SA and EAs are analyzed. To provide for large sets of possible solutions, the strain optimization algorithms were modified to keep all interesting solutions found during their execution. The initial results are positive but some aspects of the methodology are still undergoing refinement.

REFERENCES

Adam M Feist, Markus J Herrgard, Ines Thiele, Jennie L Reed, and Bernhard Ø Palsson. 2008. *Reconstruction of biochemical networks in microorganisms*. Nature Reviews Microbiology, 7(2):129.

Adam M Feist and Bernhard Ø Palsson. 2008. *Nature Biotechnology*, 26(6):659-67.

R.U. Ibarra, J.S. Edwards, and B.G. Palsson. 2002. *Escherichia coli k-12 undergoes adaptive evolution to achieve in silico predicted optimal growth*. Nature, 420:186-189.

M. Rocha, P. Maia, R. Mendes, J.P. Pinto, E.C. Ferreira, J. Nielsen, K.R. Patil, and I. Rocha. 2008. *Natural computation meta-heuristics for the in silico optimization of microbial strains*. BMC Bioinformatics, 9.

AUTHORS' BIOGRAPHIES

JOSÉ PEDRO PINTO was born in Porto, Portgal and went to the Univerity of Minho, where obtained a degree in systems and informatics
engineering.
Later he took a post graduation in Data Analysis Systems in the same a University.
Curretly he is an informatics PhD student still in the Univerity of Minho.