AN IMPROVED EVOLUTIONARY ALGORITHM-BASED FRAMEWORK FOR IDENTIFYING *IN SILICO* METABOLIC ENGINEERING TARGETS

Isabel Rocha, Centro de Engenharia Biológica, Universidade do Minho
Campus de Gualtar, Braga, 4710-057, Portugal
T: +351 253 604400, F: +351 253 678986, irocha@deb.uminho.pt
Miguel Rocha, Departamento de Informática, Universidade do Minho, Braga, Portugal
Eugénio C. Ferreira, Centro de Engenharia Biológica, Universidade do Minho, Braga, Portugal
Kiran Raosaheb Patil, Jens Nielsen, Center for Microbial Biotechnology, BioCentrum-DTU, Lyngby, Denmark

In metabolic engineering problems, due to the complexity of metabolic networks, it is often difficult to identify *a priori* which genetic manipulations will originate a given desired phenotype. Genome-scale metabolic models, available for several microorganisms, can be used to simulate the metabolic phenotype and therefore help the tasks of metabolic engineering. This simulation can be performed by calculating the fluxes through all metabolic reactions using techniques like the Flux Balance Analysis (FBA) or the MOMA approaches, among others.

Several algorithms have been developed that use genome-scale metabolic models to enable the identification of gene knockout strategies for obtaining improved phenotypes. However, the problem of finding optimal gene deletion strategy is combinatorial and consequently the computational time increases exponentially with the size of the problem.

In a previous study we reported that evolutionary algorithms (EAs) enable solving large gene knockout problems in relatively short computational time. The proposed algorithm – OptGene - also allows the optimization of non-linear objective functions and additionally provides a family of close to optimal solutions. Given the promising results obtained, this algorithm was modified with two main objectives: improve the predictions obtained and increase the flexibility. For these purposes, a new program was built by the authors using the Java programming language.

Regarding the optimization algorithms, in OptGene two distinct encoding schemes had been taken into account, binary and integer representations. The latter is more compact but potentially reduces the search space to a limited number of knockouts. In order to overcome this limitation, in this work a new feature was implemented by allowing the evolution of solutions with variable size. This allows maintaining the potential solutions with a relatively small number of genes while not defining *a priori* the exact number of knockouts. Furthermore, the EA's performance was boosted by the introduction of local search operators that look for improved solutions in the neighbourhood of the individual under consideration.

The quality of the solutions obtained by the EAs was compared to the ones obtained using a simpler algorithm, the well-known hill-climbing algorithm adapted for the present situation. The local search operator, in this case, considers all neighbours that imply the addition of a single knockout to the present solution and selects the best. The wild-type is considered as the starting solution and the local search operator is applied, until no improvement is possible.

Finally, a graphical user interface was developed that allows an easy utilization of any genome-scale metabolic model in SBML or other format, the manual modification of flux bound values, the selection of the appropriate simulation technique (FBA or MOMA) and the corresponding flux to be optimized for FBA. Additionally, the program allows the utilization of any of the optimization algorithms described above and the selection of a suitable (linear or non-linear) objective function, like yield or biomass coupled yield. A tool for the visualization of the flux distribution in the metabolic network is being developed.
This modified algorithm was validated using succinate production in both *Escherichia coli* and *Saccharomyces cerevisiae* as case studies. Potential metabolic engineering targets were identified and the results suggest that non-intuitive genetic modifications spanning several different pathways may be necessary for solving challenging metabolic engineering problems.

Poster Session II - 107 Experimental Techniques on X-omics: Isabel Rocha