

# Automatic termination of parallel optimization runs of global stochastic optimization methods by upper limit criterion

Natalja Bulipopa\*, Jurijs Sulins

Biosystems Group, Department of Computer Systems, Latvia University of Agriculture, Liela iela 2, LV-3001, Jelgava, Latvia

\*Corresponding author

[cvetkova.natalja@gmail.com](mailto:cvetkova.natalja@gmail.com)

Received: 28 May 2013; accepted: 31 May 2013; published online: 31 May 2013.  
This paper has no supplementary material.

**Abstract:** Many combinations of adjustable parameters should be tested in optimization experiments of biochemical networks to find the smallest subset of parameters enabling the best improvements of objective function both in case of design task and parameter estimation task. In case of optimization with global stochastic optimization methods one of the problems is the termination of the optimization run looking for a good compromise between spent computational resources and probability that the best found value of objective function will be the global optimum. Longer runs increase the possibility to reach the global optimum. Automatic termination criteria in case of consensus or stagnation of parallel optimization runs have been proposed as criteria for automatic termination. Varying the consensus and delay time settings different probability of reaching global optimum and duration of optimization can be reached. It is proposed to modify automatic optimization termination criteria of parallel optimization runs applying upper limit agreement of a number of parallel optimization runs. Automatic application of upper limit agreement would reduce the duration of scanning of the whole space of combination of adjustable parameters. This approach is tested on the yeast glycolysis model with six adjustable parameters using COPASI, CoRunner and ConvAn software for five parallel optimization runs per combination of adjustable parameters.

**Keywords:** optimization, termination criteria, global stochastic optimization method, parallel.

## 1. Introduction

Modelling becomes more and more important part of engineering cycle of biochemical processes (Banga, 2008; Hübner et al., 2011; Mauch et al., 2001). Optimization is one of the application fields of modelling. Improving the performance of a biochemical network for industrial purposes the goal is to use as few as possible alterations to the system to reach industrially interesting strain (Nikolaev, 2010; Pentjuss et al., 2013; Rodríguez-Acosta et al., 1999; Trinh and Srienc, 2009; Unrean et al., 2010).

The necessity to find the best combination per number of adjustable parameters leads to combinatorial explosion of combinations of adjustable parameters which have to be optimized (Stalidzans et al., 2012). Therefore, automatic screening of all the possible combinations becomes important. Unfortunately the systems of differential equations describing dynamics of biochemical networks can not be solved analytically and global computationally expensive stochastic optimization methods are used due to variety of reasons (Banga, 2008; Mendes and Kell, 1998). One of the problems of global stochastic optimization methods is the decision about the termination of optimization run because this kind of methods can not guarantee global optimality (Banga, 2008). That can be compensated by longer optimization runs. The optimization can be terminated when there are no changes of the best value of the objective function for a longer time. Due to stochastic nature the duration of optimization procedure becomes hardly predictable (Mozga et al., 2011; Nikolaev, 2010) even for the same model and constant number of adjustable parameters in combination (Mozga and Stalidzans,

2011a). Application of parallel optimization runs (Sulins and Stalidzans, 2012) with consensus and stagnation criteria is proposed to automate the termination of optimization runs (Sulins and Mednis, 2012).

It is proposed to modify automatic optimization termination criteria (Sulins and Mednis, 2012) of parallel optimization runs (Sulins and Stalidzans, 2012) and use upper limit agreement of a number of parallel optimization runs to reduce the duration of scanning of the whole space of combination of adjustable parameters. This approach is tested on the yeast glycolysis model of Galazzo and Bailey (Galazzo and Bailey, 1990) with six adjustable parameters.

## 2. Materials and methods

### 2.1. Model and optimization task setting

The yeast glycolysis model (Galazzo and Bailey, 1990) is used as an optimization task example. The optimization task is set according to the *in silico* optimization experiments of ethanol production performed by Rodríguez-Acosta on the same model (Rodríguez-Acosta et al., 1999). Concentrations of six enzymes catalyzing reactions ATPase, GAPD, Glucose in (Glu), Hexokinase (HK), Phosphofructokinase (PFK) and Pyruvate kinase (PK) are chosen as adjustable parameters. 63 combinations of six adjustable parameters (up to six out of six) are optimized. The range of changes of adjustable parameters is set within range from -99% to +900% (from 100-fold decrease to 10-fold increase) from their initial values. Maximization of the flow through reaction Pyruvate kinase (PK) (proportional to the ethanol production) is set as the objective function. Generally the proposed approach upper limit agreement can be used also for minimization tasks.

## 2.2. Software tools and optimization settings

The experiments are performed using *COPASI* (Hoops et al., 2006) Build 35 as the optimization tool. Particle swarm optimization method was used with following settings: iteration limit 30000, Swarm Size 50, Std. Deviation 1-e-06, Random Number Generator 1 and Seed 0. Steady state subtask is selected.

*CoRunner* software (Sulins and Stalidzans, 2012) is used for management of parallel optimization runs of *COPASI* files. Parallel optimization runs are stopped when all the five parallel runs have reached consensus with consensus corridor 1 % and delay time 15 minutes. The data about the dynamics of the objective function values of these optimization runs are used as test case to examine the reliability of proposed optimization termination criteria which are less demanding than consensus.

*ConvAn* software (Kostromins et al., 2012) is used for discretization and analysis of convergence dynamics of parallel optimization runs. The discretization step is set at 60 seconds.

## 2.3. Determination of upper limit

Consensus of two and more optimization runs at the best value of objective function is tested as termination criterion of parallel optimization runs of global stochastic optimization methods (Fig. 1). The best value of global objective function for each combination of adjustable parameters after reaching the consensus (as described in section 2.2) is set as global 100% value to assess the effect of less demanding consensus criterion on the reduction of objective function value. The necessary time to reach for consensus (as described in section 2.2) is set as 100% of optimization time to assess the reduction of optimization duration using upper limit agreement criteria. Only combinations of adjustable parameters which reach consensus (Sulins and Mednis, 2012) are used in this study.

Upper limit agreement of two (three, four...) runs termination criterion is satisfied when within 1% consensus corridor of the best objective function value among all the parallel runs are two (three, four...) best values of objective function of parallel runs. There is no delay time applied (delay time=0).

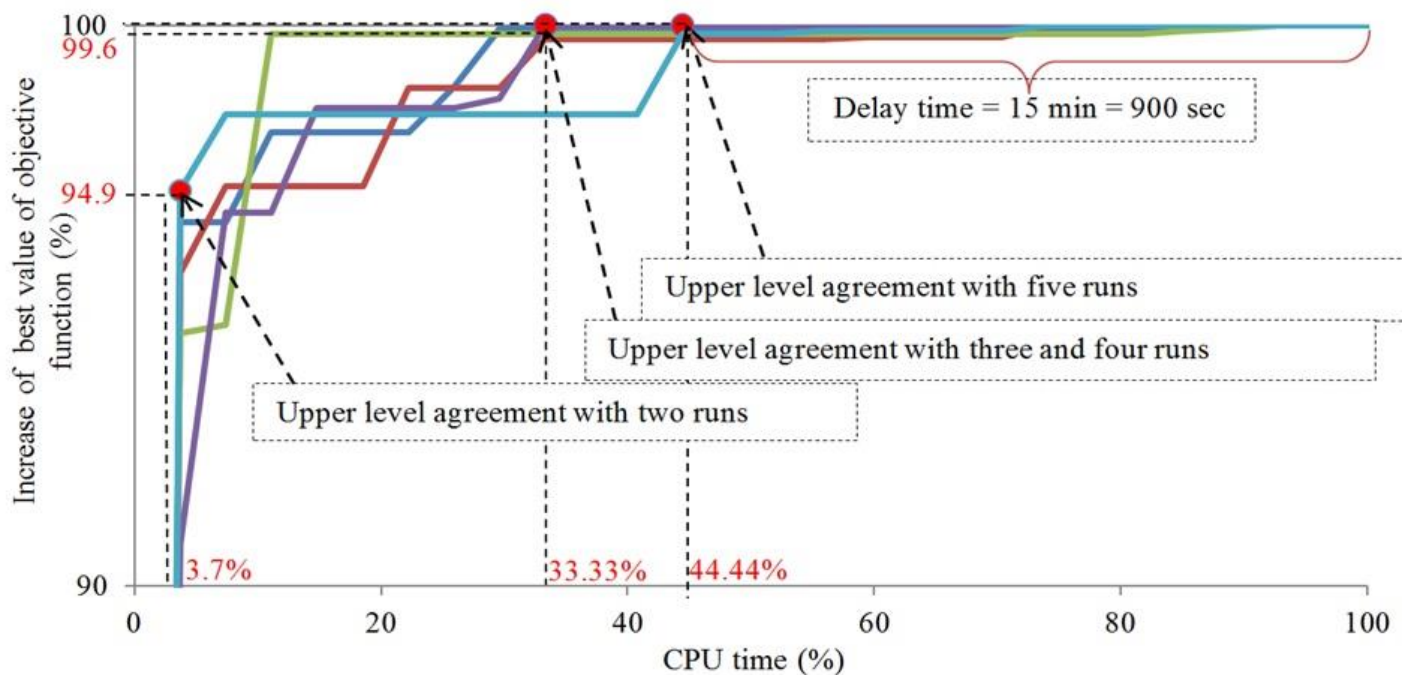


Fig. 1. Determination of upper level agreement for two and more parallel optimization runs and corresponding losses of best value and savings of computational time. 0% and 100% of increase of objective function best value correspond to the value of model before optimization and after delay time of consensus correspondingly. 100% of optimization time correspond to 1620 seconds.

## 3. Results

Earlier termination of parallel optimization runs reduces the duration of optimization and the best value of objective function (Fig. 2). Totally 63 combinations consist of correspondingly 6, 15, 20, 15, 6 and 1 combinations of 1, 2, 3, 4, 5 and 6 adjustable parameters in combination. The number of combinations that reached consensus/stagnation state are correspondingly 6/0 for one adjustable parameter in combination, 15/0 for two adjustable parameters in

combination, 17/3 for three adjustable parameters in combination, 19/1 for four adjustable parameters in combination, 5/1 for five adjustable parameters in combination and 1/0 for six adjustable parameters in combination. Totally five out of 63 combinations stagnated and are not included in the calculations for Fig. 2. There is no statistics about the case of six adjustable parameters as there is only one combination possible.

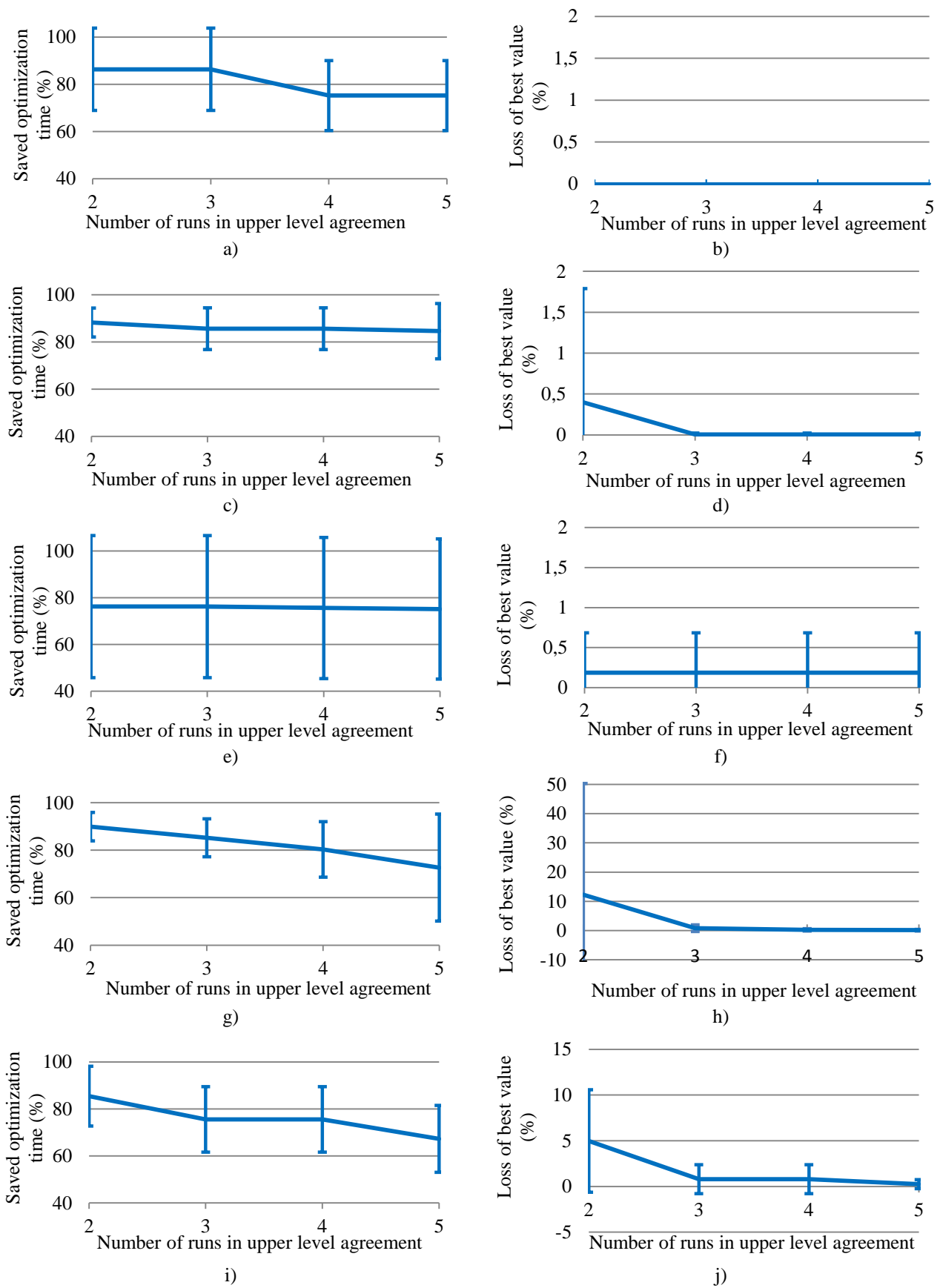


Fig. 2. Reduction of objective function value and on the increase of optimization time for one (a and b), two (c and d), three (e and f), four (g and h) and five (i and j) parameters correspondingly. Error bars demonstrate the standard deviation at n=6 for one, n=15 for two, n=20 for three, n=15 for four and n=6 for five parameters.

#### 4. Discussion

The reduction of the optimization time compared with full consensus within 1% corridor and 15 minutes delay time is about 80% independent on the number of adjustable parameters per combination. The curves are decreasing but still the average values always remain above 60% indicating that the saving of optimization time can be expected about 60-90%. In absolute numbers the optimization duration savings increase with the number of parameters in combination.

The average loss of best value can be significant in case of upper level agreement of two runs while it is below 1% starting from upper level agreement of three runs. Thus experiments demonstrate that consensus criterion (Sulins and Mednis, 2012) can be applied in a less strict way saving about 2/3 of optimization time and losing less than 1% of objective function improvement if upper level agreement of at least three parallel runs is reached.

More experiments are needed to generalize the conclusions about the efficiency of upper limit agreement criterion. In case bigger models and larger set of adjustable parameters the length of optimization duration would increase (Mozga and Stalidzans, 2011b) and time savings in percents would reduce while the time savings in absolute numbers would increase.

Looking at the savings of computational resources it can be calculated that faster result is reached by n-fold increase of computational resources where n is the number of parallel optimization runs. Thus in case of n-fold reduction of computational time would mean equal use of computational time (processor hours). More than n-fold reduction of computational time would mean additional benefit: savings of computational time in addition to the savings of optimization duration.

The 100% of optimization time in this study is determined experimentally and means the moment of automatic termination at consensus within 1% corridor of all parallel optimization runs. The time scale would therefore change if the corridor or delay time would be changed. The time scale would change even more if 100% time would be determined voluntarily by an expert.

The algorithm for upper level agreement criteria implementation can be executed automatically in optimization software.

#### 5. Conclusion

Significant time can be saved in case of approximate estimation of the best value of objective function for a particular combination of adjustable parameters using upper limit agreement criterion. Consensus of all parallel runs generally is a special case of upper level agreement criterion when all the parallel runs come to upper level agreement.

Computational experiments demonstrate that upper level agreement of at least three parallel optimization runs reduces the necessary optimization time by 60-90% and reduces the best value of objective function just by up to 1% compared to consensus within 1% of five parallel runs with delay time of 15 minutes. More extensive experiments would be needed to generalize this statement.

The optimization termination algorithm can be executed automatically.

#### References

Banga, J.R. (2008), "Optimization in computational systems biology.," *BMC systems biology*, Vol. 2, p. 47. <http://dx.doi.org/10.1186/1752-0509-2-47>

- Galazzo, J. and Bailey, J.E. (1990), "Fermentation pathway kinetics and metabolic flux control in suspended and immobilized *Saccharomyces cerevisiae*," *Enzyme and microbial technology*, Vol. 12 No. 3, pp. 162–172. [http://dx.doi.org/10.1016/0141-0229\(90\)90033-M](http://dx.doi.org/10.1016/0141-0229(90)90033-M)
- Hoops, S., Sahle, S., Gauges, R., Lee, C., Pahle, J., Simus, N., Singhal, M., et al. (2006), "COPASI—a Complex Pathway Simulator.," *Bioinformatics (Oxford, England)*, Vol. 22 No. 24, pp. 3067–74. <http://dx.doi.org/10.1093/bioinformatics/btl485>
- Hübner, K., Sahle, S. and Kummer, U. (2011), "Applications and trends in systems biology in biochemistry.," *The FEBS journal*, Vol. 278 No. 16, pp. 2767–857. <http://dx.doi.org/10.1111/j.1742-4658.2011.08217.x>
- Kostromins, A., Mozga, I. and Stalidzans, E. (2012), "ConvAn: a convergence analyzing tool for optimization of biochemical networks," *Biosystems*, Vol. 108 No. 1-3, pp. 73–77. <http://dx.doi.org/10.1016/j.biosystems.2011.12.004>
- Mauch, K., Buziol, S., Schmid, J. and Reuss, M. (2001), "Computer-Aided Design of Metabolic Networks," *AICHE Symposium Series*, pp. 82–91.
- Mendes, P. and Kell, D. (1998), "Non-linear optimization of biochemical pathways: applications to metabolic engineering and parameter estimation," *Bioinformatics*, Vol. 14 No. 10, pp. 869–883. <http://dx.doi.org/10.1093/bioinformatics/14.10.869>
- Mozga, I., Kostromins, A. and Stalidzans, E. (2011), "Forecast of Numerical Optimization Progress of Biochemical Networks," *Proceedings of the International Conference Engineering for Rural Development, 26-27 May 2011*, Jelgava, pp. 103–108.
- Mozga, I. and Stalidzans, E. (2011a), "Convergence dynamics of biochemical pathway steady state stochastic global optimization," *Proceedings of IEEE 12th International Symposium on Computational Intelligence and Informatics (CINTI), 21-22 November 2011*, Budapest, IEEE, pp. 231–235. <http://dx.doi.org/10.1109/CINTI.2011.6108504>
- Mozga, I. and Stalidzans, E. (2011b), "Convergence Dynamics of Biochemical Models to the Global Optimum," *Proceedings of the 3rd International Conference on E-Health and Bioengineering, 24-26 November 2011*, Iasi, pp. 227–230.
- Nikolaev, E.V. (2010), "The elucidation of metabolic pathways and their improvements using stable optimization of large-scale kinetic models of cellular systems.," *Metabolic engineering*, Elsevier, Vol. 12 No. 1, pp. 26–38. <http://dx.doi.org/10.1016/j.jmben.2009.08.010>
- Pentjuss, A., Odzina, I., Kostromins, A., Fell, D., Stalidzans, E. and Kalnenieks, U. (2013), "Biotechnological potential of respiring *Zymomonas mobilis*: a stoichiometric analysis of its central metabolism," *Journal of Biotechnology*, Vol. 165 No. 1, pp. 1–10. <http://dx.doi.org/10.1016/j.jbiotec.2013.02.014>
- Rodriguez-Acosta, F., Regalado, C.M. and Torres, N.V. (1999), "Non-linear optimization of biotechnological processes by stochastic algorithms: Application to the maximization of the production rate of ethanol, glycerol and carbohydrates by *Saccharomyces cerevisiae*," *Journal of Biotechnology*, Vol. 68 No. 1, pp. 15–28. [http://dx.doi.org/10.1016/S0168-1656\(98\)00178-3](http://dx.doi.org/10.1016/S0168-1656(98)00178-3)
- Stalidzans, E., Kostromins, A. and Sulins, J. (2012), "Two stage optimization of biochemical pathways using parallel runs of global stochastic optimization methods," *2012 IEEE 13th International Symposium on Computational Intelligence and Informatics (CINTI)*, IEEE, pp. 365–369. <http://dx.doi.org/10.1109/CINTI.2012.6496792>
- Sulins, J. and Mednis, M. (2012), "Automatic termination of parallel optimization runs of stochastic global optimization methods in consensus or stagnation cases," *Biosystems and Information technology*, Vol. 1 No. 1, pp. 1–5. <http://dx.doi.org/10.1159/bit.120501>
- Sulins, J. and Stalidzans, E. (2012), "Corunner: multiple optimization run manager for Copasi software," *Proceedings of 5th International Scientific Conference on Applied Information and Communication Technologies, 26-27 April 2012*, Jelgava, pp. 312–316.
- Trinh, C.T. and Srienc, F. (2009), "Metabolic engineering of *Escherichia coli* for efficient conversion of glycerol to ethanol.," *Applied and environmental microbiology*, Vol. 75 No. 21, pp. 6696–705. <http://dx.doi.org/10.1128/AEM.00670-09>
- Unrean, P., Trinh, C.T. and Srienc, F. (2010), "Rational design and construction of an efficient *E. coli* for production of diacylglycerol.," *Metabolic engineering*, Elsevier, Vol. 12 No. 2, pp. 112–22. <http://dx.doi.org/10.1016/j.jmben.2009.11.002>