A compartmented in silico model of *Rapeseed* central metabolism

Eleftherios Pilalis, Aristotelis Chatziioannou, Fragiskos Kolisis

Metabolic Engineering and Bioinformatics Group, Institute of Biological Research and Biotechnology, National Hellenic Research Foundation, 48 Vas. Konstantinou Ave, 11635, Athens, Greece

Introduction

Systems level modeling of cellular metabolism has proven to be indispensable for the design of rational genetic modification strategies for the redistribution of the metabolic flux network towards desired end-products. The exploitation of whole-genome pathway databases in combination with the appropriate mathematical techniques and modern high-throughput measurement methods, gives new perspectives in terms of understanding and controlling the pleiotropic functionality of complex biological reaction networks. In this work a large-scale in silico model is constructed for the simulation of the central metabolism of *Rapeseed* (*Brassica napus)* embryos. *Rapeseed* is an organism of particular interest in oil industry and the in silico reconstruction of its metabolism can help our understanding regarding the regulation of lipid biosynthesis. The model comprises 307 reactions and 249 metabolites extracted from the Aracyc and BRENDA databases. In order to validate this model, we performed constraint-based Flux Balance Analysis, through application of linear optimization methods, and by incorporating relevant experimental data from literature. Further exploiting the derived model, an in-silico gene deletion analysis and a systemic regulatory analysis were performed in order to evaluate and comprehend the plasticity of the real network and infer conclusions regarding its robustness as well as predict target for oil biomass overproduction.

Model construction

All reactions of the model plant *Arabidopsis thaliana* (1286 in total), which is phylogenetically close to *Brassica napus*, were retrieved from the Aracyc [1] database. The reactions were attributed to compartments based on the BRENDA enzyme database [2] and on literature. The stoichiometric table of the reactions was submitted to Singular Value Decomposition (fig. 1) in order to obtain the systemic reactions and metabolites [3], in regard to biomass production. The initial set of 1286 reactions was hence reduced to a final non-redundant model of 307 reactions and 249 metabolites (fig. 2) which represent the main metabolic pathways (Glycolysis, Pentoses-Phosphate, Citric Acid Cycle, Lipid Biosynthesis, Aminoacids Metabolism, biomass production).

Flux Balance Analysis by linear optimization

Flux balance analysis (FBA) is a constraint-based modeling approach that allows the prediction of metabolic steady-state fluxes by applying mass balance constraints to a stoichiometric model [4]. FBA uses linear optimization to determine the steady-state reaction flux distribution in a metabolic network by maximizing an objective function, such as ATP production or growth rate. In this study, we used as objective function the biomass production (triglycerides, proteins, starch). The composition of *Rapeseed* biomass during embryonic development was extracted from literature [5]. The optimization was performed using the Cobra Toolbox for Matlab [6].

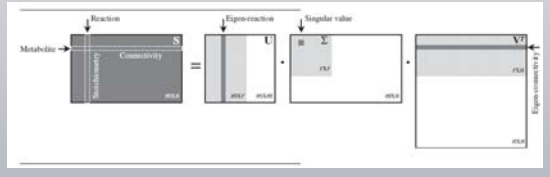

Figure 1. Obtaining the systemic reactions and metabolites by applying Singular Value Decomposition on the stoichiometric matrix

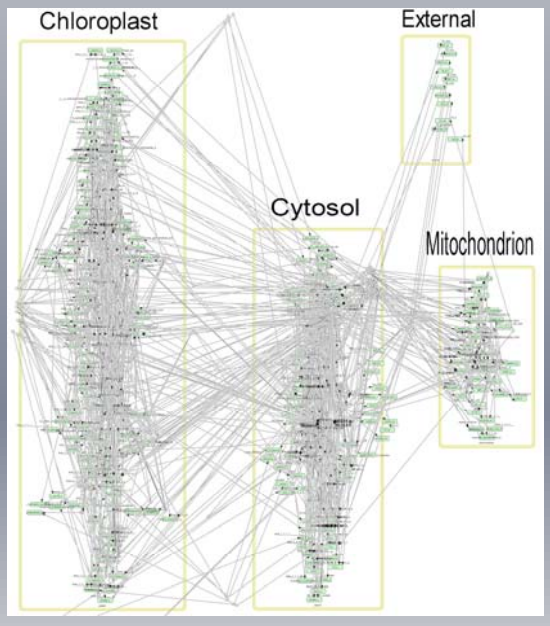

Figure 2. Graphical representation of the model

In silico Gene Deletion Analysis

The impact of single and double gene deletions on seed growth was simulated by a series of linear optimization runs. In each run, flux through single enzymes or all combinations of two enzymes were constrained to zero. By this means we tested the robustness of the model, which is the ability of adaptation to genetic perturbations by redistribution of the fluxes.

Systemic regulatory analysis by Singular Value Decomposition of the steady-state flux space

The application of the constraint-based modeling results to only one optimal steady-state flux distribution. In order to explore the set of all possible solutions (solution space), we performed uniform random sampling of the steady-state solution space. The obtained sample matrix was submitted to principal component analysis by Singular Value Decomposition [7]. By this approach we obtain a set of eigenvectors which represent the principal components of network regulation. Each reaction has a corresponding loading vector, which is the vector of the reaction's contribution to each of the eigenvectors. Specifically, if vij is the contribution of reaction j to eigenvector i, and there are m eigenvectors, then the loading vector for reaction j is lj = (v1j, v2j, ... , vmj). The score for reaction j is then given by the l2-norm of the loading vector: | lj | = (v1j 2 + v2j 2 + ... + vmj 2)1/2. The score of each reaction measures its relative importance in the regulation of the network, in regard to the specified objective function (table 1). The sampling was performed using the Cobra toolbox [6]. The principal component analysis was performed using Matlab 7.4.

Results and Conclusions

- We developed a large-scale in silico model of *Rapeseed* embryos central metabolism during development.
- The model is characterized by high robustness, as inferred by the gene deletion analysis. The compartmentation of metabolism enhances the ability of the plant to redistribute the fluxes in a way to preserve the optimal growth. The only lethal deletions were those of the external fluxes, i.e. the fluxes assuring the import of nutrients and the biomass production (data not shown).
- The regulatory role of all reactions was assessed by uniformly sampling the steady-state flux space and by performing principal component analysis on the sample matrix. The highest-scoring reactions were related to important cell functions such as respiration, carbon conversion efficiency (RUBISCO shunt [8]) and photosynthesis (table 1).
- The RUBISCO shunt was recently identified in *Rapeseed* , playing an important role to oil production [8]. Comparing to glycolysis, the Rubisco shunt results in an increase of carbon conversion efficiency, yielding 20% more acetyl-CoA with 40% less carbon loss. It is responsible for the production of 37% -75% of phosphoglycerate in the developing embryo.
- Overall, FBA on large-scale metabolic networks in combination with uniform flux space sampling is a powerful method for systemic regulatory analysis, in the absence of kinetic data.

Score	Enzyme name	Reaction	Function
1	glyceraldehyde-3-phosphate dehydrogenase (plastid)	D-glyceraldehyde-3-phosphate + phosphate + NAD+ <-> 1,3-diphosphateglycerate + NADH	RUBISCO shunt, glycolysis
1	RUBISCO	H2O + CO2 + D-ribulose-1,5-bisphosphate <->2 3-phosphoglycerate + 2 H+	RUBISCO shunt
0.804	cytochrome c oxidase	0.5 O2 (m) + Cytochrome C-H2 (m) -> 2 H (m) + Cytochrome C (m)	aerobic respiration
0.7	Carbon dioxide transporter	CO2(c) -> CO2 (e)	CO2 export
0.675	Acetate transporter	acetate(c) -> acetate (e)	fermentation
0.669	-	H2O (c) <-> H2O (e)	water diffusion
0.656	Phosphate transporter	phosphate (e) -> phosphate (c)	phosphate import
0.569	-	O2 (e) -> O2 (c)	aerobic respiration
0.561	-	NAD/NADP import (pseudoreaction)	oxidoreductive balance
0.541	Ethanol transporter	ethanol (c) -> ethanol (e)	fermentation
0.539	alcohol dehydrogenase	acetaldehyde (c)+ NADH (c) => NAD (c) + ethanol (c)	
0.472	-	O2 (c) <-> O2 (p)	Photosynthesis
0.442	NMN nucleosidase	nicotinamide_mononucleotide (p) + H2O => D_ribose_5_phosphate (p)	pentoses phosphate pathway, RUBISCO
0.44	NAD(+) diphosphatase	NAD(p) + H2O <-> nicotinamide_mononucleotide (p) + AMP (p)	shunt
0.433	glycolate transporter	glycolate (p) <-> glycolate (m)	
0.432	phosphoglycolate phosphatase	_2_phosphoglycolate (p) + H2O <-> phosphate (p) + glycolate (p)	glyoxylate cycle, RUBISCO shunt
0.432	Glycolate oxidoreductase	NAD (m) + glycolate (m) <-> glyoxylate (m) + NADH (m)	
0.432	Photosystem II	2 H2O + light -> 4 H (p) + 4 e_ + O2(p); v;	Photosynthesis
0.432	Photosystem I	2 H (p) + 4 e_ + 2 NADP (p) -> 2 NADPH (p)	
0.432	ribulose-bisphosphate carboxylase	O2 (p) + D_ribulose_1_5_bisphosphate (p) -> H (p) + _2_phosphoglycolate (p) + _3_phosphoglycerate (p)	RUBISCO shunt
0.402	oxygen diffusion to mitochondrion	O2 (c) -> O2 (m)	aerobic respiration
0.317	Glucose transporter	glucose (e) -> glucose (c)	glucose import
0.224	nitrate reductase	NO3 (c) + NADH (c) => NAD (c) + nitrite (c) + H2O	
0.224	nitrite reductase	nitrite (c) => ammonia (c) + H2O	nitrate assimilation
0.224	nitrate transporter	nitrate (e) -> nitrate (c)	
0.22	carbonate dehydratase	CO2 (p) + H2O <-> H (p) + HCO3 (p)	
0.218	acetyl-coa carboxylase	H (p) + acetyl_CoA (p) + HCO3 (p) + ATP (p) <-> phosphate (p) + ADP (p) + malonyl_CoA (p)	fatty acids biosynthesis

Table 1. The highest-scoring regulatory reactions contributing to the eigenvectors . Abbreviations: (c) : cytoplasmic, (m): mitochondrial, (p): plastidial, (e): external.

References

1. Mueller, L.A., P. Zhang, and S.Y. Rhee, *AraCyc: a biochemical pathway database for Arabidopsis.* Plant Physiol, 2003. **132**(2): p. 453-60.
2. Schomburg, I., et al., *BRENDA, the enzyme database: updates and major new developments.* Nucleic Acids Res, 2004. **32**(Database issue): p. D431-3.
3. Famili, I. and B.O. Palsson, *Systemic metabolic reactions are obtained by singular value decomposition of genome-scale stoichiometric matrices.* J Theor Biol, 2003. **224**(1): p. 87-96.
4. Edwards JS, Ramakrishna R, Schilling CH, Palsson BØ (1999) *Metabolic flux balance analysis.* In SY Lee, ET Papoutsakis, eds, Metabolic Engineering. Marcel Dekker, New York, pp 13–57
5. Schwender, J., Y. Shachar-Hill, and J.B. Ohlrogge, *Mitochondrial metabolism in developing embryos of Brassica napus.* J Biol Chem, 2006. **281**(45): p. 34040-7.
6. Becker, S.A., et al., *Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox.* Nat Protoc, 2007. **2**(3): p. 727-38.
7. Barrett, C.L., N.D. Price, and B.O. Palsson, *Network-level analysis of metabolic regulation in the human red blood cell using random sampling and singular value decomposition.* BMC Bioinformatics, 2006. **7**: p. 132.
8. Schwender, J., et al., *Rubisco without the Calvin cycle improves the carbon efficiency of developing green seeds.* Nature, 2004. **432**(7018): p. 779-82.