



Received for publication, December, 15, 2018

Accepted, August, 5, 2019

Original paper

Using genome-scale model to predict the metabolic engineering impact on *Escherichia coli* metabolism during succinic acid production optimization

ZSOLT BODOR^{1,*}, ANDREA (IUHÁSZ) FAZAKAS², KATALIN BODOR³,
ERIKA KOVÁCS¹, ILDIKÓ MIKLÓSSY¹, BEÁTA ALBERT¹

¹Sapientia Hungarian University of Transylvania, Department of Bioengineering, Libertatii square, No. 1, 530104, Miercurea Ciuc, Romania

²“Politehnica” University of Bucharest, Department of Inorganic Substances Technology and Environment Protection, Polizu street, No 1-7, 011061, Bucharest, Romania

³University of Pécs, Faculty of Natural Sciences, Doctoral School of Chemistry, Ifjúság 6, 7624 Pécs, Hungary

Abstract

In this study, we investigated the production possibility of succinic acid from renewable resources, such as glucose and glycerol (by-products resulted from different industries), and the impact of metabolic engineering on cellular metabolism was estimated using systems biology approaches. The cellular metabolic flux was estimated for aerobic, microaerobic and anaerobic conditions. The *in silico* simulation of metabolic processes in *Escherichia coli* has been concluded the possibility of increasing the yield of succinic acid from 0.008 to 0.9 (mol mol⁻¹ glucose) and from 0.0003 to 0.6 (mol mol⁻¹ glycerol) by removing three specific genes from the genome (pyruvate formate lyase – pflB, lactate dehydrogenase – ldhA and alcohol dehydrogenase – adhE) (anaerobic conditions). Comparing the microaerobic and anaerobic conditions was clear that a very small amount of oxygen has a vital contribution on cell survival. We have determined the deletion impact (p) on growth rate of each of the 1363 genes one by one under aerobic, microaerobic and anaerobic conditions in wild-type and mutant strains. Wet experiments were carried out and highest molar yield was obtained for glycerol, 0.51 mol mol⁻¹. This study informs other studies that environmental conditions have a stronger influence on the genes with reduced maximal growth than the elimination of different genes from redox metabolism. The knowledge gained in this study would help in microbial cell factories design for succinic acid production from renewable resources under various environmental conditions

Keywords *In silico*, *Escherichia coli*, succinic acid, metabolic engineering, glucose, glycerol.

To cite this article: BODOR Z, FAZAKAS AI, BODOR K, KOVÁCS E, MIKLÓSSY I, ALBERT B. Using genome-scale model to predict the metabolic engineering impact on *Escherichia coli* metabolism during succinic acid production optimization. *Rom Biotechnol Lett.* 2020; 25(3): 1666-1676. DOI: 10.25083/rbl/25.3/1666.1676

✉ *Corresponding author: ZSOLT BODOR, Sapientia Hungarian University of Transylvania, Department of Bioengineering, Libertatii square, No. 1, 530104, Miercurea Ciuc, Romania
E-mail: bodorzsolt@uni.sapientia.ro

Introduction

The role of *in silico* platforms is essential for the rational design of networks in order to improve the productivity of a microorganism, to increase the production of a target compound as well as to analyze the intracellular processes (LEWIS & al [1]; CHAN & al [2]). Systems biology has been shown to be successful in metabolic engineering (BODOR & al [3]; FONG [4]; FONG & al [5]; LEE & al [6], [7], [8]; MIKLÓSSY & al [9]; NA & al [10]; YIM & al [11]) especially in predicting the outcomes (e.g., production-, growth rates, etc.), while constraint-based modelling predicts the gene knockouts with reasonable accuracy (COOPER & DUFFIELD [12]; HUANG & al [13]; TEPPER & SHLOMI [14]). Different genome-scale models of metabolism (GSM) have been generated describing the cell and the processes inside (FEIST & PALSSON [15]; KIM & al [16]; ORTH & al [17]; SAHA, CHOWDHURY & MARANAS [18]). Recently, the generation of cell factories is directly linked to a holistic approach which involves several disciplines and techniques like genomics, proteomics, metabolomics, fluxomics, transcriptomics and phenomics (PEY & al [19]). Rational analysis, based on knowledge of cellular processes provides the opportunity for optimization using the highly curated genome-scale metabolic models. This widely used systems metabolic engineering enables target-oriented redesign of metabolism to increase the production of the target metabolite and to cope with huge complexities in biology (FAULON & CARBONELL, [20]; LEE & al [6]; LEE & al [21]; LEWIS & al [1]; STEPHANOPOULOS, [22]). Hence, *in silico* analysis is mandatory in deciphering metabolic changes under different environmental conditions and genetic manipulations (FEIST & al [23]; LEWIS & al [24]; BORDBAR & al [25]; ANDREOZZI & al [26]). *In silico* techniques were successfully used to design industrially important strains (HANSEN & al [27]; LEE & al [28]; MAIA & al [29]; YIM & al [11]), to analyse metabolic processes (CAMPODONICO & al [30]; LISHA & al [31]; RAMAN & al [32]). The relationship between genotype and phenotype can be analysed by using the well-known methods used in genetic engineering and phenotype analysis (CHOI & al [33]; NIELSEN & al [34]).

In vitro mutations, for example, can be performed by the λ -Red recombinase method for chromosomal gene disruption (DATSENKO & al [35]) and changes in metabolite profiles can be identified by using HPLC or GC-MS.

The production of biofuel generates large quantities of glycerol as waste (WENDISCH & al [36]) what could be an inexpensive carbon source (such as glucose) for many microorganisms (GARLAPATI & al [37]; YANG & al [38]; YIM & al [11]) and can be used as an appealing substrate for high-value added metabolites production like succinic acid. Succinic acid has received increasing attention during the last decades because it has an important role in the synthesis of high value added chemicals

(WERPY and PETERSEN, [39]; LIANG & al [40]; SALVACHÚA & al [41]). The chemical synthesis through catalytic hydrogenation of maleic acid is highly dependent on fossil resources, which gives rise to major concerns about sustainability. On the other hand, the environmental pollution even at regional levels should be taken into consideration (ILIE & al [42]; KERESZTESI & al [43]; SZÉP & al [44-50]). Taken together, biotechnology offers new sustainable alternatives, however, there are still many questions to be answered in order to better understand the interactions inside a cell, especially if different genetic modifications are carried out or the environmental conditions are modified.

The present work aimed at the analysis of succinic acid production impact on *E. coli* metabolism, under different genetic and environmental conditions. For this purpose systems biology tools were applied, namely the constraint-based approach of metabolic modeling (LEWIS & al [1]). The most important contributions are as follows: construction of an industrially important strain for succinic acid production from renewable resources such as glucose or glycerol, identification of the impact of gene knockouts and conditions on gene essentiality and finally the flux variations in wild-type and mutant strains were identified using flux variability analysis (FVA). To test the production potential basic experiments were carried out under anaerobic conditions.

Materials and Methods

A previously developed metabolic model of *E. coli* K12 MG1655 (ORTH & al [17]), available in SBML format at BioModels online database (<http://www.ebi.ac.uk/biomodels-main/>), was used for *in silico* studies. In order to avoid unrealistic behaviors and to improve the predictions quality the upper and lower bounds for the reactions CAT (catalase), DHPTDNR (dihydropteridine reductase), DHPTDNRN (dihydropteridine reductase (NADH)), FHL (Formate-hydrogen lyase), SPODM (superoxide dismutase), SPODMpp (superoxide dismutase – periplasm), SUCASPTpp (succinate:aspartate antiporter (periplasm)), SUCFUMtpp (succinate:fumarate antiporter (periplasm)), SUCMALtpp (succinate:malate antiporter (periplasm)), and SUCTARTtpp (succinate: D-tartrate antiporter (periplasm)) were constrained to zero. The steady-state flux distribution was calculated by using MATLAB (The MathWorks Inc., Natick, MA, USA) and COBRA Toolbox (BECKER & al [51]; SCHELLENBERGER & al [52]) software packages with TOMLAB CPLEX ((Tomlab Optimization Inc., San Diego, CA, USA) and GUROBI (Gurobi Optimizer version 6.0, Houston Texas) solvers.

Flux balance analysis (FBA)

Growth was simulated by maximizing flux through a defined biomass objective function. Steady-state balances

and stoichiometry of all metabolites are imposed as linear constraints as follows:

$$S \cdot v = 0 \quad (1)$$

- where S is an $m \cdot n$ stoichiometric matrix containing all the stoichiometric coefficients in the model of m metabolites and n reactions, v is the flux vector.

- constraints $vl \leq v \leq vu$; vl and vu are vectors with n elements each, which represent the lower and upper bounds on the fluxes, respectively. Because relationships are linear, the metabolic objective can be estimated by linear optimization:

$$\text{Maximize: } z = c^T v \quad (2)$$

In this equation, c is a row vector containing weighting factors for individual fluxes (v), while z is the objective function. Flux units are in mmol gDW⁻¹h⁻¹ (millimoles dry cell weight per hour), except biomass which has units of h⁻¹ (BECKER & al [51]; SCHELLENBERGER & al [52]).

***In silico* implementation and prediction of growth on glucose and glycerol under different environmental conditions**

The inputs were restricted to minimal media (M9) containing only inorganic salts and glucose or glycerol as carbon sources. Growth on glucose was simulated by constraining the lower bound of the glucose exchange to the experimentally determined 10 mmol gDW⁻¹h⁻¹ (VARMA & PALSSON, [53]). For growth on glycerol, the lower bound of the glucose exchange reaction was set to zero, while the lower bound of the glycerol exchange reaction was set to 10 mmol gDW⁻¹h⁻¹. Under aerobic growth, the oxygen uptake was set to 1000 mmol gDW⁻¹h⁻¹ (unlimited oxygen uptake), for microaerobic we used 5 mmol gDW⁻¹h⁻¹ and 0 to create anaerobic conditions.

The eliminated reactions were: $\Delta pflB$ (pyruvate formate lyase), $\Delta ldhA$ (lactate dehydrogenase), $\Delta adhE$ (alcohol dehydrogenase). Using glycerol as the sole carbon source the cell was able to grow aerobically and microaerobically, meanwhile, the $pflB$ gene elimination blocked the biomass formation under anaerobic conditions. Different co-substrates were tested to improve cell viability including glucose and amino acids (data not shown here).

Large-scale *in silico* evaluation of single gene deletions was conducted to determine the effect of gene deletion on cellular growth including the wild-type and designed mutant strains. Genes were categorized as non-lethal (unchanged maximal growth), with reduced-lethality (reducing maximal growth) and lethal/essential (no growth).

Flux variability analysis (FVA)

FVA is a method widely used to identify the maximum and minimum possible fluxes through each reaction of the genome-scale metabolic network for a given maximum objective value (GHOSH, & al [54]; MAHADEVAN & SCHILLING, [55]). Calculations were carried out by following the descriptions presented by (BECKER & al [51]). Briefly, the optimal predicted growth rate was constrained to 100% of the optimal value

under various environmental conditions and the minimum and maximum values were estimated to decipher the redundancy of reactions in the network. Two optimization problems need to be solved during FVA simulations for each flux v_j of interest,

$$\begin{aligned} & \max_{v_j} / \min_{v_j} \\ & \text{subject to } S_{ij} v_j = 0 \\ & c^T v = Z_{biomass} \\ & v_j^{min} \leq v_j \leq v_j^{max} \end{aligned} \quad (6)$$

where $Z_{biomass}$ is biomass optimal solution. Considering n as the number of reactions than FVA requires the $2n$ LP problems to be solved as mentioned previously.

Bacterial strain, culture conditions

The strain used in this study was *E. coli* K12 MG1655 from "Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH" (DSMZ 18039). All the work was done at 37°C, in minimal medium under anaerobic conditions with glucose or glycerol at a concentration of 5 g/L. Bacterial cells were grown in minimal medium in 5 mL medium to produce a starter culture and the seed culture was used to inoculate the fermentation medium (optical density 0.1). Cells were grown with shaking at 150 rpm (Certomat BS-1 Sartorius) for 24 h in serum bottles (50 mL) with 20 mL M9 medium. Samples were taken in every two hours for the analysis of cell growth. The optical density of the cell cultures was measured at 550 nm (OD₅₅₀) to quantify cell growth, using a Cary 50 Conc UV-Visible spectrophotometer.

Chromosomal gene deletion

For gene deletion strategy we used the λ -Red recombineering methods previously described by DATSENKO & al [35]. Plasmids (5 Strain Wanner Lambda Red Gene Disruption Kit) were obtained from the *E. coli* Genetic Stock Center (Yale University).

Analytical procedure

Metabolites were analysed with GC-MS (6890N/5975 Agilent) based on solid-phase micro extraction (SPME) with on-fiber silylation. Silylation was carried out using N, O bis trifluoroacetamide (BSTFA) following the procedures described by (LUAN & al [56]). For chromatograms and spectra analyses we used the MassLab software (ThermoQuest, Manchester, UK). Compounds were identified by comparing the mass spectra obtained with commercially available MS libraries (Wiley, NIST and LIBTX).

Results and Discussions

Under aerobic conditions, *E. coli* converted the substrates quantitatively to biomass and CO₂, without any by-product formation (highest biomass production). During microaerobic and anaerobic conditions in minimal medium, $pflB$ is highly active, catalyses the CoA-dependent cleavage of pyruvate to form acetyl-CoA and formic acid. Acetyl-CoA is then converted to acetic acid and ethanol (Fig. 1).

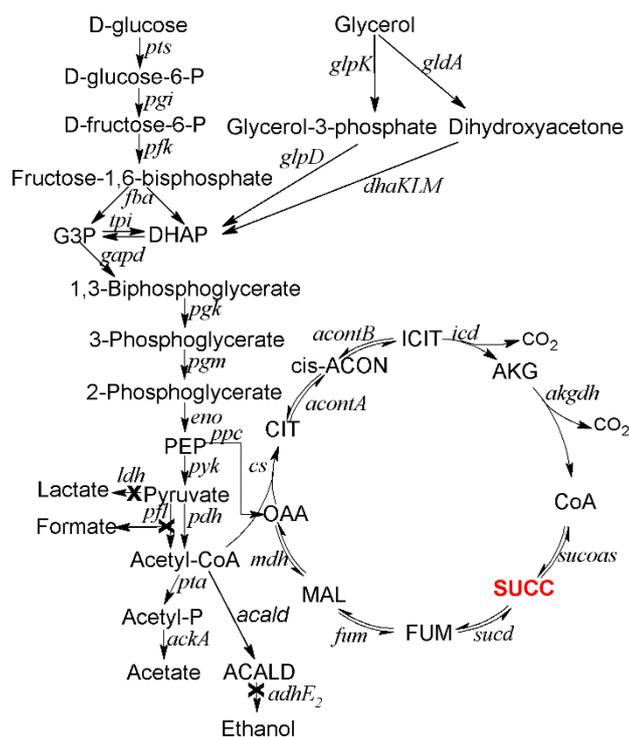


Figure 1. Routes for overproduction of succinic acid in the mutant strain. Crosses represent the reaction knockouts; the following pathways were eliminated (represented by a black x): pyruvate formate lyase- $\Delta pflB$ (formate), lactate dehydrogenase- $\Delta ldhA$ (lactate), alcohol dehydrogenase- $\Delta adhE$ (ethanol); Abbreviations: pts-D-glucose transport via PEP:Pyruvate PTS (periplasm); pgi – glucose-6-phosphate isomerase; pfk – phosphofruktokinase; fba – fructose-bisphosphate aldolase; tpi – triose-phosphate isomerase; gapd glycerinaldehyde-3-phosphate dehydrogenase; glpK – glycerol kinase; gldA – glycerol dehydrogenase; glpD – Sn-glycerol-3-phosphate dehydrogenase; dhaKLM – dihydroxyacetone kinase; pgk – phosphoglycerate kinase; pgm – phosphoglycerate mutase; eno – enolase; ppc – phosphoenolpyruvate carboxylase; pyk – pyruvate kinase; ldh-D – lactate dehydrogenase; pfl – pyruvate formate lyase; pdh – pyruvate dehydrogenase; pta – phosphotransacetylase; ackA – acetate kinase; acald – acetaldehyde dehydrogenase (acetylating); adhE2 – alcohol dehydrogenase (ethanol); cs – citrate synthase; acontA – aconitase (half-reaction A, Citrate hydro-lyase); acontB – aconitase (half-reaction B, Isocitrate hydro-lyase); icd – isocitrate dehydrogenase (NADP); akgdh – 2-Oxoglutarate dehydrogenase; sucoas – succinyl-CoA synthetase (ADP-forming); succD – succinate dehydrogenase (irreversible); fum – fumarase; mdh – malate dehydrogenase; cat – catalase; succCD – succinyl-CoA synthetase.

Deletion of *pflB* had little impact on cell growth in M9 mineral salts medium using glucose as the sole carbon source because the carbon flow is diverted to lactic acid (Table 1).

Table 1. Inactivating alternative NADH oxidizing pathways in M9 mineral salts medium (anaerobic conditions)

	Wild-type	$\Delta pflB$	$\Delta pflB, \Delta ldhA$	$\Delta pflB, \Delta ldhA, \Delta adhE$
μ (h^{-1})	0.24	0.19	0.19	0.11
Y_{Succ}^a ($mol\ mol^{-1}$)	0.008	0.006	0.006	0.910
Y_{For}^b ($mol\ mol^{-1}$)	1.730	0	0	0
Y_{Lac}^c ($mol\ mol^{-1}$)	0	1.720	0	0
Y_{EtOH}^d ($mol\ mol^{-1}$)	0.803	0	1.720	0
Y_{Ac}^e ($mol\ mol^{-1}$)	0.820	0	0	0.430
Y_{Ala}^f ($mol\ mol^{-1}$)	0	0	0	0.490
Y_{CO_2} ($mol\ mol^{-1}$)	0	0.040	1.760	0

Yield (Y); ^aSuccinic acid; ^bFormic acid; ^cLactic acid; ^dEthanol; ^eAcetic acid; ^fAlanine

Lactic acid production was blocked by blocking the lactate dehydrogenase enzyme (*ldhA*), however, the production yield of succinic acid did not change. The yield of succinic acid increased to 0.9 mol mol⁻¹ glucose with *adhE* deletion when both of the alternative NADH oxidation pathways were inactivated, but growth was reduced by more than 40% compared to the double mutant (*ΔpflB, ΔldhA*).

The main difference between microaerobic and anaerobic conditions is as follows: acetic acid production (higher during microaerobic) and succinic acid (higher during anaerobic conditions). On the other hand, alanine was present only under anaerobic conditions and in mutant strains.

Glycerol as carbon source

Glycerol is generated in huge quantities as a by-product during biofuel production, hence the value-added utilization in biotechnology is a promising future. The growth of wild-type *E. coli* K12 MG1655 was slowly on glycerol even under aerobic conditions. The reason is the difference between substrates molecular weights, but after calculations, it was clear that mass yield of cell-mass on glucose was 0.43 gDW g⁻¹ glucose while in case of glycerol 0.53 under aerobic conditions. With an uptake rate of 10 mmol glycerol the maximal growth rate was 0.55 under aerobic conditions, 0.33 under microaerobic and 0.08 (h⁻¹) under anaerobic conditions.

To block formic acid production the *pflB* reaction must be deleted as stated out for glucose, but our results show that the *ΔpflB* mutant fails to grow anaerobically. We decided to test different co-substrates including glucose and amino acids to improve cell viability. Three best candidates were detected, glutamic acid, glutamine and glucose. To modify the redox pathway and to estimate the effect on cell metabolism all 3 genes (*ΔpflB, ΔldhA, ΔadhE*) were eliminated as mentioned before. Using glucose as co-substrate the triple mutant succinic acid yield increased to 0.6 (mol mol⁻¹ glycerol) with 1 mmol gDW⁻¹ h⁻¹ glucose uptake rates, but the growth rate was significantly reduced (only 0.001 h⁻¹). Changing the environmental conditions- (limited oxygen availability/elimination of O₂) reduced biomass formation nearly by 58%-89% (wild-type – mutants) on glucose and 47%-99% on glycerol.

Figure 2 shows the deletion impact of the total genes (1363) on wild-type and mutant strains under different environmental conditions and substrate availability, and the distribution of relative growth rates for all the gene deletions in the model is presented. Using this method a good agreement with experimental deletion studies was reported earlier by (BECKER & al [51]).

With glucose as the sole source of carbon 200 of the 1363 model genes (Fig. 2. IA) were found to be lethal (critical genes), and 48 gene deletions resulting in reduced maximal growth rates under aerobic conditions in minimal medium. We were interested to investigate if the triple mutation has any effect. We found that the number of critical genes did not change and minimal changes were found in the list of genes with reduced maximal growth rates (Fig. 2. ID). Similar simulations were carried out for microaerobic and anaerobic conditions. Microaerobic predictions were in strong correlation with the aerobically obtained ones, 200 critical genes for wild-type and mutant

strains, 45 (Fig. 2. IB) respectively 54 genes (Fig. 2. IE) with reduced growth rates. Eliminating the oxygen from the system has increased the number of essential genes (Fig. 2. IC) to 204 and the growth rate was reduced significantly (~40%) in case of ten genes. The list of essential genes of the triple mutant strain (Fig. 2. IF) remained the same as in the case of wild-type (anaerobic conditions), suggesting that the modification of the redox metabolism does not change the number of essential genes. However, a significant difference was observed between genes with reduced growth rates (wild-type 25, mutant 47).

A similar tendency was observed after changing the carbon source to glycerol. Out of the 1363 genes in the model, 199 genes were considered lethal under aerobic (Fig. 2. IIA) and microaerobic conditions (Fig. 2. IIB) and 204 under anaerobic (Fig. 2. IIC) conditions. Genes with reduced maximal growths were as follows: 43 aerobic, 42 microaerobic and 21 for anaerobic conditions. Eliminating important genes from pyruvate metabolism has minimal impact on essential genes under aerobic (Fig. 2. IID) and microaerobic (Fig. 2. IIE) conditions (199) in wild-type. Interestingly under anaerobic conditions in the mutant strain with *singleGeneDeletion* we were unable to detect genes with reduced maximal growth rates (Fig. 2. IIF).

Changing the co-substrate to amino acids (glutamic acid or glutamine) under anaerobic conditions no significant differences were detected in essential genes compared to glucose (as co-substrate), and 8 genes were detected with reduced maximal growth.

The results indicate that *E. coli* metabolic network is robust to different changes and may activate other pathways (on different conditions) to strive alive. As we expected, minimal changes were observed in lethal genes after environmental and genetic modifications, however, genes with reduced maximal growths were mostly affected. Taking into account the environmental conditions (aerobic, microaerobic, anaerobic and the carbon sources) as well as the mutations (*ΔpflB, ΔldhA, ΔadhE*) we can conclude, that the profile of genes with reduced maximal growth rates was much more affected by substrates and the presence of oxygen than by genetic modifications.

Dynamic FBA of diauxic growth

Simulations were carried out by combining FBA with an iterative approach based on a quasi-steady-state assumption to identify diauxic growth. Briefly, the substrate concentration (S_c) is determined from the previous step (S_{co}) if it is the first time step from the initial substrate concentration. The amount of substrate available per unit of biomass per unit of time is estimated using the equation:

$$S_a = \frac{S_c}{X \Delta t} \quad (3)$$

where S_a is the amount of substrate available and X (gDW/L) represents cell density. Then the substrate uptake (S_u) and growth rate (μ) is calculated using FBA. The final concentrations for the next step are calculated from the differential equations:

$$\frac{dX}{dt} = \mu X \quad (4)$$

$$\frac{dS_c}{dt} = -S_u X \quad (5)$$

Considering the predictions, diauxic growth was not possible under anaerobic conditions, however under aerobic and microaerobic conditions the metabolites

(formic acid, acetic acid, ethanol or succinic acid) initially secreted were subsequently metabolized after glucose or glycerol exhaustion.

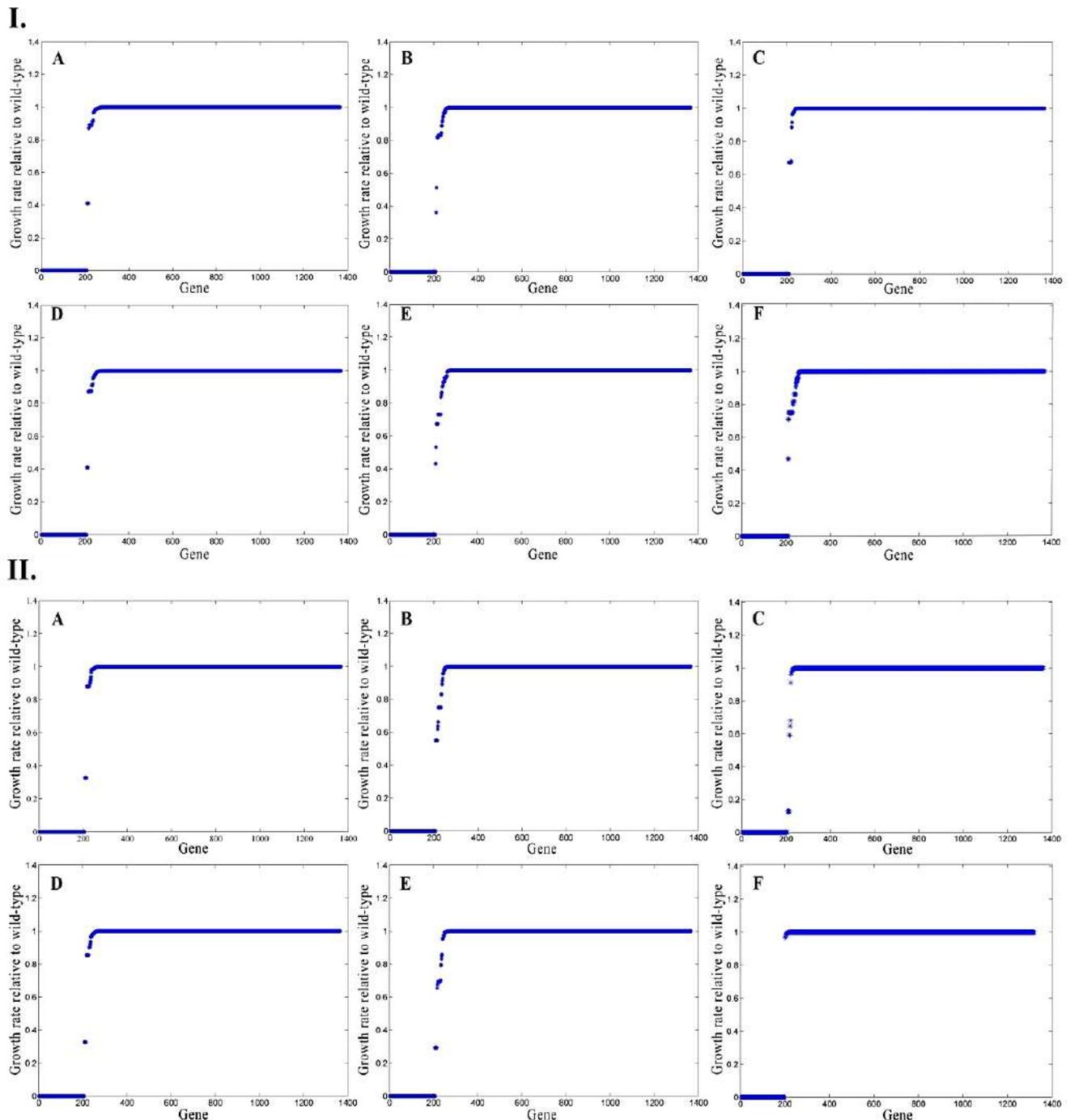


Figure 2. Single gene deletion results of 1363 genes (wild-type) and 1360 (mutant), distribution of relative growth rates; wild-type on glucose (I.) and glycerol (II.) under aerobic (A), microaerobic (B) and anaerobic (C) conditions; mutant strains ($\Delta pflB$, $\Delta ldhA$, $\Delta adhE$) under aerobic (D), microaerobic (E) and anaerobic (F) conditions.

Flux variability analysis within the metabolic network for wild-type and mutant strains

FVA was carried out by setting growth rates as the objective function and the maximum and minimum possible fluxes through each reaction was identified.

Briefly, the optimized objective value obtained with FBA was used as a constraint during calculations of feasible ranges of reaction fluxes (minimization and maximization of each flux value). Substrate consumption rates were set to 10 mmol $gDW^{-1}h^{-1}$, while the oxygen uptake rate

was modified to create anaerobic conditions. To analyse the impact of environmental conditions and the genetic modifications on flux ranges the changes were calculated for each condition. In Fig. 3 are presented reactions with a

normalized span >0.01 and sorted by magnitude. FVA is presented for wild-type and mutant strains on glucose and glycerol under anaerobic conditions.

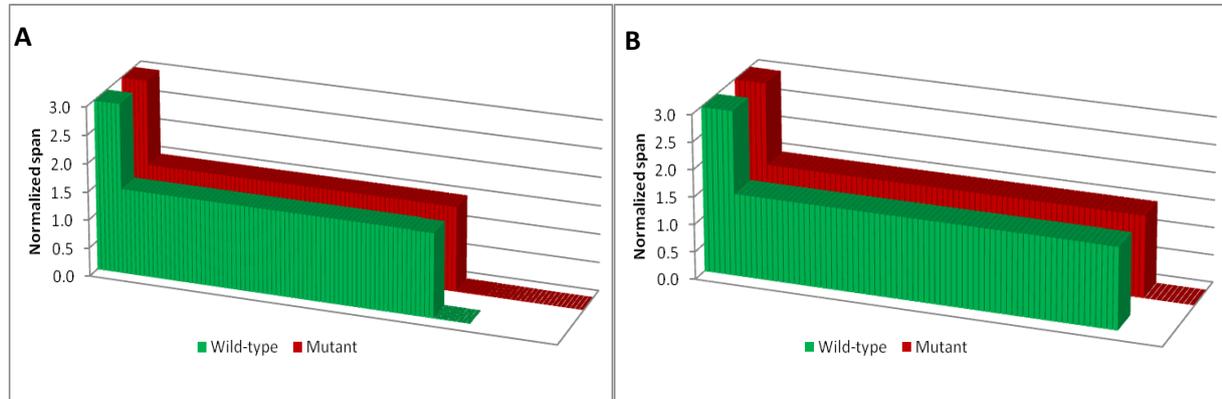


Figure 3. Range of flux values for wild-type and mutant strains analysed by flux variability analysis, **A.** for glucose and **B.** for glycerol. The span was normalized to the maximal flux span identified. Reactions with a normalized span >0.01 are shown for strains under anaerobic conditions.

In order to test the impact of knockouts FVA was implemented to calculate for each reaction in the model the flux variations under certain conditions. In this analysis two situations were compared: *i*) wild-type and mutant on glucose; *ii*) wild-type and mutant on glycerol. Fig. 3. shows the normalized values of the reactions span. Some reactions had an infinite span in both cases, (totally 6 reaction: ADK1 (adenylate kinase), ADK3 (adenylate kinase (GTP)), NDPK1 (nucleoside-diphosphate kinase (ATP:GDP)), ALATA_L (L-alanine transaminase), VALTA (valine transaminase), VPAMTr (Valine-pyruvate aminotransferase)), meaning that these reactions are not constrained. Many other reactions had a normalized span of 1.5 totally 70 and this number did not change during substrate availability. The remaining reactions are characterized by a normalized span <0.02. Important reactions with relevant biological meaning are as follows: TALA (transaldolase), FBA (fructose-bisphosphate aldolase), FBA3 (Sedoheptulose 1,7-bisphosphate D-glyceraldehyde-3-phosphate-lyase), PFK (phosphofructokinase), PFK_3 (phosphofructokinase (s7p)), F6PA (fructose 6-phosphate aldolase), DHAPT (dihydroxyacetone phosphotransferase), PYK (pyruvate kinase) for WT on glucose, and CITt3pp (citrate transport out via proton antiport (periplasm)), CITt7pp (citrate transport via succinate antiport (periplasm)), SUCt3pp (succinate transport out via proton antiport (periplasm)), FRD2 (fumarate reductase), FRD3 (fumarate reductase – 2-Demethylmenaquinol 8), NADH17pp (NADH dehydrogenase (menaquinone-8 & 3 protons) (periplasm)), NADH18pp (NADH dehydrogenase (demethylmenaquinone-8 & 3 protons) (periplasm)), FBA, TALA, FBA3, PFK, PFK_3, EX_h(e) (H⁺ exchange),

Htex (proton transport via diffusion (extracellular to periplasm)), EX_co2(e) (CO₂ exchange), CO2tex (CO₂ transport via diffusion (extracellular to periplasm)), CO2tpp (CO₂ transporter via diffusion (periplasm)), THD2pp (NAD(P) transhydrogenase (periplasm)), FADRx (FAD reductase), GLUDy (glutamate dehydrogenase), PDH (pyruvate dehydrogenase), EX_ala-L(e) (L-Alanine exchange), EX_nh4(e) (ammonia exchange), ALAtex (L-alanine transport via diffusion (extracellular to periplasm)), NH4tex (ammonia transport via diffusion (extracellular to periplasm)), NH4tpp (ammonia reversible transport (periplasm)), EX_ac(e) (acetate exchange), ACtex (acetate transport via diffusion (extracellular to periplasm)) for mutant, respectively. On the other hand, on glycerol the normalized flux span with at least 0.01 was predicted only for mutant strains and the reactions are: CITt3pp (citrate transport out via proton antiport (periplasm)), CITt7pp (Citrate transport via succinate antiport (periplasm)), FRD2, SUCt3pp (succinate transport out via proton antiport (periplasm)), FRD3, NADH17pp, NADH18pp, EX_ala-L(e), ALAtex. The result of the FVA analysis shows which fluxes are influenced by genetic engineering in a certain way; there are changes in reactions with large and small flux span in both substrates.

Time-course fermentation experiments were carried out for wild-type and mutant strains ($\Delta pf1B$, $\Delta ldhA$, $\Delta adhE$) to follow the changes of succinic acid. Results presented here (Fig. 4) have been obtained during anaerobic fermentations. Identification of metabolites was performed in a GC-MS experiment, data analysis was carried out with MassLab via comparison with mass spectra obtained from different libraries.

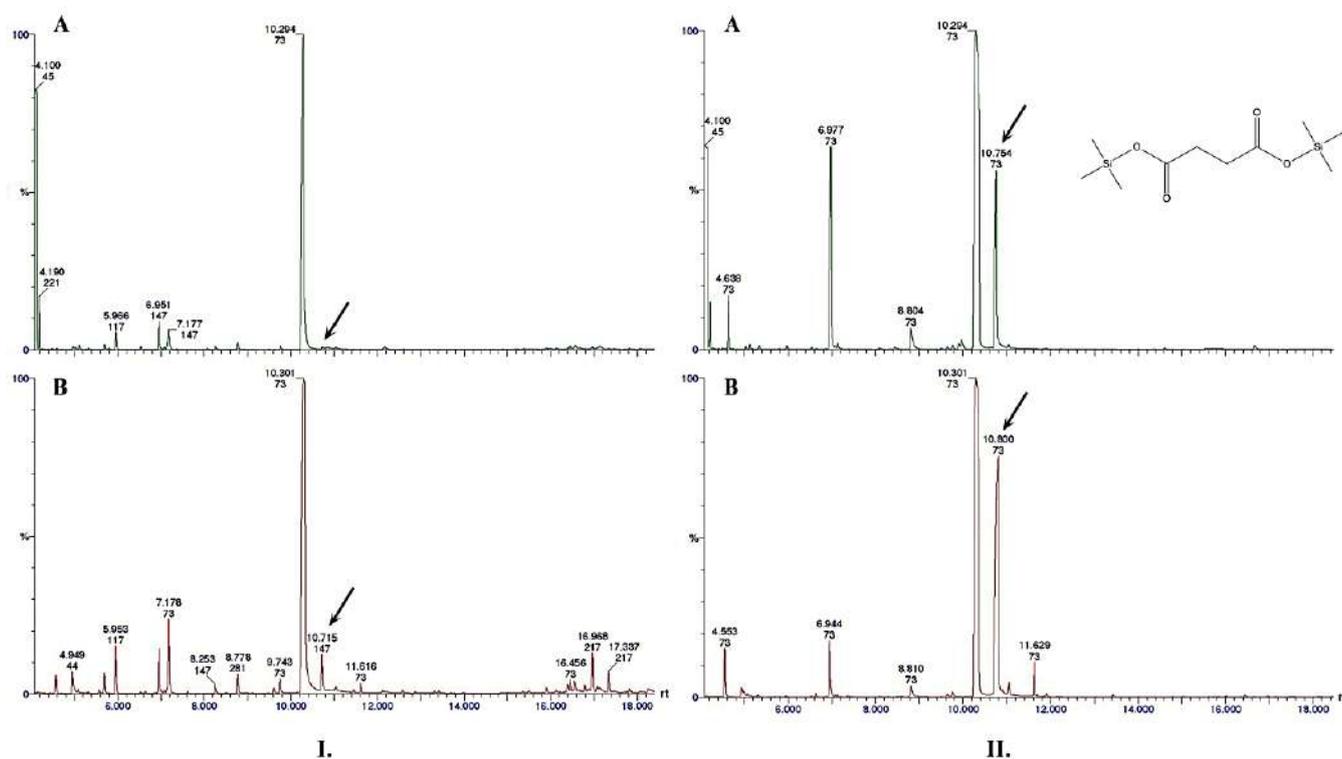


Figure 4. GC-MS analysis of determination of succinic acid from supernatants; I. glucose and II. glycerol (wild-type (A), mutant (B), (arrows indicate succinic acid peaks).

The compounds present in the fermentation mixtures were identified by GC-MS. The heights of the peak indicate the relative concentrations of the metabolites present in the fermentation mixtures. For example, in the fermentation mixtures the succinic acid was identified as trimethylsilylated succinic acid at 10^7 together with acetic acid, phosphoric acid, lactic acid, propanoic acid, butanoic acid, glucose and glycerol. The peak of succinic acid in mutant strains was significantly higher compared to wild-type strains (in both cases) and the molecular ion peak was at m/z 262 (succinic acid+2 trimethylsilyl) and the fragment ion peaks at m/z 247 (succinic acid+2 trimethylsilyl-methyl) and m/z 147 (2 trimethylsilyl+ H^+). Succinic acid peak area for wild-type and mutant strains on glucose was 0.65% and 12.14%, and on glycerol 3.77% and 29.69%, respectively.

As a proof of concept experiment, we demonstrate that using systems biology approach the production potential of succinic acid in *E. coli* under anaerobic conditions is a sustainable approach. Significant changes were observed between wild-type and mutant strains on both substrates. To our surprise, the presence of succinic acid in the fermented broth was significantly higher in case of glycerol compared to glucose. Succinic acid concentration after 6 days was as follows: wild-type 3.43 ± 0.15 mM and 13.5 ± 0.08 mM with a molar yield of $0.51 \text{ mol mol}^{-1}$. The production in mutant strain was 4x higher compared to wild-type. As stated out earlier (JEONG & al [57]) during adaptive laboratory evolution the production potential can be significantly improved in

order to reach a stable and industrially important strain with increased succinic acid yield and productivity on both substrates. This is an initial strategy for strain development and further analysis will be carried out using knowledge, genome-scale metabolic models and constraint-based modeling strategies to identify new gene targets to further improve the production potential.

Conclusions

Sustainable development is environment-connected and needs to be harmonized to sustain a continuous growth, and hence, the application of (bio-based) green technologies is crucial. A potential approach for conversion of biomass and renewable resources to fuels and chemicals is to combine biochemical and chemical processes.

Genetically engineered strains were designed and constructed using systems biology and metabolic engineering methods, which are capable of metabolizing glucose and glycerol (resulted from different industries as by-products) and producing succinic acid a key platform chemical. The present work presents *in silico* and experimental methods for strain simulation and genetic modification to design strains by deleting different genes and testing the substrate and environmental effects on cell metabolism. The λ -Red recombineering technology was successfully used for chromosomal modifications in *E. coli*.

The highest succinic acid production was under anaerobic conditions, using glucose the yield was 0.9 mol mol^{-1}

and in the case of glycerol with glucose as co-substrate 0.6 mol mol⁻¹. Blocking the NADH reoxidation results a complex redirection of flux from formic acid, lactic acid and ethanol to succinic acid. Using minimal medium and the elimination of tree genes was sufficient for succinic acid production from glucose or glycerol under micro-aerobic and anaerobic conditions.

We found that the number of essential genes is less influenced by factors, such as substrate availability, genetic engineering or environmental conditions. Meanwhile, important changes were predicted on genes which reduce maximal growth. These genes are much more influenced by environmental conditions or by substrate availability than by genetic modifications of redox metabolism. Diauxic growth was predicted to aerobic and microaerobic conditions when all the glucose or glycerol has been consumed the cell began to metabolise the metabolites secreted earlier. Using FVA simulations the flux span changes were identified for both substrates, as a conclusion, we can confirm that genetic engineering impact on cell metabolism can be analysed by redundancy approach.

At this stage at the proof-of concept level, we were interested to test *in vivo* the production potential, and after six days a 0.51 mol mol⁻¹ molar yield was achieved on glycerol, however with long-term adaptation the production potential can be significantly improved.

Bio-based production of different chemicals from renewable feedstocks could be a sustainable solution to fulfill the demands of organic chemicals.

Conflicts of Interest: none

Acknowledgements

This work was supported by the „Sectoral Operational Programme Human Resources Development 2007-2013 of the Romanian Ministry of Labour, Family and Social Protection through the Financial Agreement POSDRU/6/1.5/S/19.”, “BIOBILD-Synthesis of some C₄, C₅ carboxylic acid building block chemicals from renewable biomass resources” PN-II-PCCA-2011-3.2-1367.

Our work was supported by the Sapientia University 2017/2018 Research Program, grant nr. 227/2/17.05.2017 and by Collegium Talentum. The GC-MS analysis was carried out at Center for Organic Chemistry “Costin D. Nenitescu”.

References

- LEWIS, N.E., JAMSHIDI, N., THIELE, I., PALSSON, B.O. Metabolic systems biology: a constraint-based approach. *Encyclopedia of Complexity and Systems Science*, 5535 (2009).
- CHAN, H.W., MOHAMAD, S.M., DERIS, S., ILLIAS, M.R. A Review of computational approaches for *in silico* metabolic engineering for microbial fuel production. *Current Bioinformatics*, 8(2), 253-258 (2013).
- BODOR, ZS., FAZAKAS (IUHASZ), A., KOVÁCS, E., LÁNYI, S., ALBERT, B. Systems biology and metabolic engineering for obtaining *E. coli* mutants capable to produce succinate from renewable resources. *Romanian Biotechnological Letters*, 19(4), 9625-9636 (2014).
- FONG, S.S. Computational approaches to metabolic engineering utilizing systems biology and synthetic biology. *Computational and Structural Biotechnology Journal*, 11(18), 28-34 (2014).
- FONG, S.S., BURGARD, A.P., HERRING, C.D., KNIGHT, E.M., BLATTNER, F.R., MARANAS, C.D., PALSSON, B.O. *In silico* design and adaptive evolution of *Escherichia coli* for production of lactic acid. *Biotechnology and Bioengineering*, 91(5), 643-648 (2005).
- LEE, D.H., PALSSON, B.O. Adaptive evolution of *Escherichia coli* K-12 MG1655 during growth on a nonnative carbon source, L-1,2-propanediol. *Applied and Environmental Microbiology*, 76(13), 4158-4168 (2010).
- LEE, S.Y., KIM, H.U. Systems strategies for developing industrial microbial strains. *Nature Biotechnology*, 33(10), 1061-72 (2015).
- LEE, J.W., YI, J., KIM, T.Y., CHOI, S., AHN, J.H., SONG, H., LEE, S.Y. Homo-succinic acid production by metabolically engineered *Mannheimia succiniciproducens*. *Metabolic Engineering*, 38, 409-417 (2016).
- MIKLÓSSY, I., BODOR, ZS., SINKLER, R., ORBÁN, K.C., LÁNYI, S., ALBERT, B. *In silico* and *in vivo* stability analysis of a heterologous biosynthetic pathway for 1,4-butanediol production in metabolically engineered *E. coli*. *Journal of Biomolecular Structure and Dynamics*, 35(9), 1874-1889 (2016).
- NA, D., YOO, S.M., CHUNG, H., PARK, H., PARK, J.H., LEE, S.Y. Metabolic engineering of *Escherichia coli* using synthetic small regulatory RNAs. *Nature Biotechnology*, 31(2), 170-4 (2013).
- YIM, H., HASELBECK, R., NIU, W., PUJOL-BAXLEY, C., BURGARD, A., BOLDT, J., VAN DIEN, S. Metabolic engineering of *Escherichia coli* for direct production of 1,4-butanediol. *Nature Chemical Biology*, 7(7), 445-452 (2011).
- COOPER, I., DUFFIELD, M. The *in silico* prediction of bacterial essential genes. *Science Against Microbial Pathogens*, 551-559 (2011).
- HUANG, D., LI, S., XIA, M., WEN, J., JIA, X. Genome-scale metabolic network guided engineering of *Streptomyces tsukubaensis* for FK506 production improvement. *Microbial Cell Factories*, 12, 52 (2013).
- TEPPER, N., SHLOMI, T. Predicting metabolic engineering knockout strategies for chemical production: Accounting for competing pathways. *Bioinformatics*, 26(4), 536-543 (2009).
- FEIST, A.M., PALSSON, B.Ø. The growing scope of applications of genome-scale metabolic reconstructions using *Escherichia coli*. *Nature Biotechnology*, 26(6), 659-667 (2008).

16. KIM, T.Y., SOHN, S.B., KIM, Y.B, KIM, W.J., LEE, S.Y. Recent advances in reconstruction and applications of genome-scale metabolic models. *Current Opinion in Biotechnology*, 23(4):617-623 (2012).
17. ORTH, J.D., CONRAD, T.M., NA, J., LERMAN, J.A., NAM, H., FEIST, A.M., PALSSON, B.Ø. A comprehensive genome-scale reconstruction of *Escherichia coli* metabolism – 2011. *Molecular Systems Biology*, 7, 1-16 (2011).
18. SAHA, R., CHOWDHURY, A., MARANAS, C.D. Recent advances in the reconstruction of metabolic models and integration of omics data. *Current Opinion in Biotechnology*, 29(1) 39-45, (2014).
19. PEY, J., VALGEPEA, K., RUBIO, A., BEASLEY, J.E., PLANES, F.J. Integrating gene and protein expression data with genome-scale metabolic networks to infer functional pathways. *BMC Systems Biology*, 7, 134 (2013).
20. FAULON, J., CARBONELL, P. Systems Metabolic Engineering. *Systems Metabolic Engineering: Methods and Protocols, Methods in Molecular Biology*, 985(1), 149-173 (2013).
21. LEE, S.Y., KIM, H.U., CHAE, T.U., CHO, J.S., KIM, J.W., SHIN, J.H., JANG, Y.-S. A comprehensive metabolic map for production of bio-based chemicals. *Nature Catalysis*, 2(1), 18-33 (2019).
22. STEPHANOPOULOS, G. Synthetic Biology and Metabolic Engineering. *{ACS} Synth. Biol.*, 1(11), 514-525 (2012).
23. FEIST, A.M., ZIELINSKI, D.C., ORTH, J.D., SCHELLENBERGER, J., HERRGARD, M.J., PALSSON, B.O. Model-driven evaluation of the production potential for growth-coupled products of *Escherichia coli*. *Metabolic Engineering*, 12(3), 173-186 (2010).
24. LEWIS, N.E., NAGARAJAN, H., PALSSON, B.O. Constraining the metabolic genotype–phenotype relationship using a phylogeny of in silico methods. *Nature Reviews Microbiology*, 10(4), 291-305 (2012).
25. BORDBAR, A., MONK, J.M., KING, Z.A., PALSSON, B.O. Constraint-based models predict metabolic and associated cellular functions. *Nature Reviews Genetics*, 15(2), 107-120 (2014).
26. ANDREOZZI, S., CHAKRABARTI, A., SOH, K.C., BURGARD, A., YANG, T.H., VAN DIEN, S., HATZIMANIKATIS, V. Identification of metabolic engineering targets for the enhancement of 1,4-butanediol production in recombinant *E. coli* using large-scale kinetic models. *Metabolic Engineering* (2016).
27. HANSEN, A.S.L., LENNEN, R.M., SONNENSCHNEIN, N., HERRGÅRD, M.J. Systems biology solutions for biochemical production challenges. *Current Opinion in Biotechnology*, 45, 85-91 (2017).
28. LEE, J.W., NA, D., PARK, J.M., LEE, J., CHOI, S., LEE, S.Y. Systems metabolic engineering of microorganisms for natural and non-natural chemicals. *Nature Chemical Biology*, 8(6), 536-546 (2012).
29. MAIA, P., ROCHA, M., ROCHA, I. *In silico* constraint-based strain optimization methods: the quest for optimal cell factories. *Microbiology and Molecular Biology Reviews : MMBR*, 80(1), 45-67 (2016).
30. CAMPODONICO, M.A., ANDREWS, B.A., ASENJO, J.A., PALSSON, B.Ø., FEIST, A.M. Generation of an atlas for commodity chemical production in *Escherichia coli* and a novel pathway prediction algorithm, GEM-Path. *Metabolic Engineering*, 25, 140-158 (2014).
31. LISHA, K.P., SARKAR, D. Dynamic flux balance analysis of batch fermentation: Effect of genetic manipulations on ethanol production. *Bioprocess and Biosystems Engineering*, 37(4), 617-627, (2014).
32. RAMAN, K., CHANDRA, N. Flux balance analysis of biological systems: Applications and challenges. *Briefings in Bioinformatics*, 10(4), 435-449 (2009).
33. CHOI, K.R., SHIN, J.H., CHO, J.S., YANG, D., LEE, S.Y. Systems Metabolic Engineering of *Escherichia coli*. *EcoSal Plus*, 7(1) (2016).
34. NIELSEN, J., KEASLING, J.D. Engineering Cellular Metabolism. *Cell*, 164(6), 1185-1197, (2016).
35. DATSENKO, K.A., WANNER, B.L. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proceedings of the National Academy of Sciences of the United States of America*, 97(12), 6640-6650 (2000).
36. WENDISCH, V.F., LINDNER, S.N., MEISWINKEL, T.M. Use of Glycerol in Biotechnological Applications. *Biodiesel – Quality, Emissions and By-Products*, 305-340 (2011).
37. GARLAPATI, V.K., SHANKAR, U., BUDHIRAJA, A. Bioconversion technologies of crude glycerol to value added industrial products. *Biotechnology Reports*, 9, 9-14 (2016).
38. YANG, F., HANNA, M.A., SUN, R. Value-added uses for crude glycerol--a byproduct of biodiesel production. *Biotechnology for Biofuels*, 5(1), 13 (2012).
39. WERPY, T., PETERSEN, G. Top Value Added Chemicals from Biomass Volume I — Results of Screening for Potential Candidates from Sugars and Synthesis Gas Top Value Added Chemicals From Biomass Volume I : Results of Screening for Potential Candidates (2004).
40. LIANG, L.A., LIU, R.M., MA, J.F., CHEN, K.Q., JIANG, M., WEI, P. Increased production of succinic acid in *Escherichia coli* by overexpression of malate dehydrogenase. *Biotechnology Letters*, 33(12), 2439-2444 (2011).
41. SALVACHÚA, D., SMITH, H., ST. JOHN, P.C., MOHAGHEGHI, A., PETERSON, D.J., BLACK, B.A., BECKHAM, G.T. Succinic acid production from lignocellulosic hydrolysate by *Basfia succiniciproducens*. *Bioresource Technology*, 214, 558-566 (2016).
42. ILIE, M., MARINESCU, F., SZÉP, R., GHIȚA, G., DEAK, G., ANGHEL, A.M., URȚESCU, B.

- Ecological risk assessment of heavy metals in surface sediments from the Danube River. *Carpathian Journal of Earth and Environmental Sciences*, 12(2), 437-445 (2017).
43. KERESZTESI, Á., PETRES, S., GHITA, G., DUMITRU, F.D., MONCEA, M.A., OZUNU, A., SZÉP, R. Ammonium neutralization effect on rainwater chemistry in the basins of the Eastern Carpathians-Romania. *Revista de Chimie* (2018).
 44. SZÉP, R., BODOR, Z., MIKLÓSSY, I., NITA, I.A., OPREA, O.A., KERESZTESI, Á. Influence of peat fires on the rainwater chemistry in intra-mountain basins with specific atmospheric circulations (Eastern Carpathians, Romania). *Science of the Total Environment*, 647, 275-289 (2019).
 45. SZÉP, R., KERESZTES, R., CONSTANTIN, L. Multi-model assessment of tropospheric ozone pollution indices of risk to human health and crops, and ozone deposition in Ciuc Depression – Romania. *Revista de Chimie*, 67(3), 408-413 (2016).
 46. SZÉP, R., KERESZTES, R., DEAK, G., TOBA, F., GHIMPUSAN, M. The dry deposition of the PM₁₀ and PM_{2.5} to the vegetation and its health effect in the Ciuc basin. *Revista de Chimie (Bucharest, Romania)* (2016).
 47. SZÉP, R., KERESZTES, R., KORODI, A., TONK, S., NICULAE, A.G., BIRLOIU, A.M. Dew point – Indirect particulate matter pollution indicator in the Ciuc Basin – Harghita, Romania. *Revista de Chimie*, 67(10), 1914-1921 (2016).
 48. SZÉP, R., MATEESCU, E., NECHIFOR, A.C., KERESZTESI, Á. Chemical characteristics and source analysis on ionic composition of rainwater collected in the Carpathians “Cold Pole,” Ciuc basin, Eastern Carpathians, Romania. *Environmental Science and Pollution Research*, 24(35), 27288-27302 (2017).
 49. SZÉP, R., MATEESCU, E., NIȚĂ, I.A., BIRSAN, M.V., BODOR, Z., KERESZTESI, Á. Effects of the Eastern Carpathians on atmospheric circulations and precipitation chemistry from 2006 to 2016 at four monitoring stations (Eastern Carpathians, Romania). *Atmospheric Research*, 214(April), 311-328, (2018).
 50. SZÉP, R., MÁTYÁS, L. The role of regional atmospheric stability in high-PM₁₀ concentration episodes in Miercurea Ciuc (Harghita). *Carpathian Journal of Earth and Environmental Sciences*, 9(2), 241-250, (2014).
 51. BECKER, S.A., FEIST, A.M., MO, M.L., HANNUM, G., PALSSON, B.Ø., HERRGARD, M.J. Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox. *Nature Protocols*, 2(3), 727-738 (2007).
 52. SCHELLENBERGER, J., QUE, R., FLEMING, R.M.T., THIELE, I., ORTH, J.D., FEIST, A.M., PALSSON, B.Ø. Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox v2.0. *Nature Protocols*, 6(9), 1290-1307 (2011).
 53. VARMA, A., PALSSON, B.O. Stoichiometric flux balance models quantitatively predict growth and metabolic by-product secretion in wild-type *Escherichia coli* W3110. *Applied and Environmental Microbiology*, 60(10), 3724-3731 (1994).
 54. GHOSH, A., ZHAO, H., PRICE, N.D. Genome-scale consequences of cofactor balancing in engineered pentose utilization pathways in *Saccharomyces cerevisiae*. *PLoS ONE*, 6(11) (2011).
 55. MAHADEVAN, R., SCHILLING, C.H. The effects of alternate optimal solutions in constraint-based genome-scale metabolic models. *Metabolic Engineering*, 5(4), 264-276 (2003).
 56. LUAN, T.G., YU, K.S.H., ZHONG, Y., ZHOU, H.W., LAN, C.Y., TAM, N.F.Y. Study of metabolites from the degradation of polycyclic aromatic hydrocarbons (PAHs) by bacterial consortium enriched from mangrove sediments. *Chemosphere*, 65(11), 2289-2296 (2006).
 57. JEONG, H., LEE, S.J., KIM, P. Procedure for Adaptive Laboratory Evolution of Microorganisms Using a Chemostat, (June), 1-5 (2016).