DOI 10.18699/BGRS/SB-2020-106

Genome-scale metabolic modeling of 2,3-butanediol production by *Geobacillus icigianus*

Mikhail Kulyashov
BIOSOFT.RU, LLC
Institute of Computational Technologies SB RAS,
Novosibirsk, Russia
Novosibirsk State University,
Novosibirsk, Russia
m.kulyashov@mail.ru

Ilya R. Akberdin
BIOSOFT.RU, LLC
FRC Institute of Cytology and Genetics SB RAS,
Novosibirsk, Russia
Novosibirsk State University
Novosibirsk, Russia
akberdinir@gmail.com

***Abstract* — 2,3-butanediol (2,3-BTD) is an important substrate for chemical production and at the same time is highly promising bacterial-based platform substances. *Geobacillus icigianus* is a strain of thermophilic genus *Geobacillus*, which is currently considered as the potential bacterial chassis that can be used in biotechnology. A genome-scale metabolic model of the bacteria has been built using a computational pipeline for autogeneration with consequent manual curation. The current version of the model comprises 1678 reactions, 1590 metabolites and 1316 genes and it is the largest known model for genus *Geobacillus*. In this work we demonstrate that *Geobacillus icigianus* can be potentially used for the production of 2,3-butanediol from different carbon sources, one of which is glycerol – a byproduct of chemical production. Furthermore, this model can be used as a theoretical platform to gain insight into the metabolism of the thermophilic bacteria and to predict more favor pathways for genetic modifications of *Geobacillus icigianus* strain in biotechnological goals.**

Keywords* — *Genome-scale modeling, Geobacillus icigianus, 2,3-butanediol, thermophilic organisms, biotechnology

Introduction

Due to the gradual depletion of fossil fuel sources, rising oil prices, and the aggravating environmental situation, which lead to the tight control of the chemical industry, a question of creating biological factories for the production of chemical substances is becoming crucial. One of the important substances in the chemical industry is 2,3-butanediol (butadiene glycol-2,3). The potential of bacteria in production of 2,3-butanediol was shown in the early 20th century [1]. Moreover, the production of 2,3-BTD was shown for thermophilic bacteria of genus *Geobacillus* [2]. Based on published data we decided to identify this capability by *Geobacillus icigianus* - new strain of thermophilic bacteria [3]. To solve this issue we have been used a genome scale metabolic modeling approach, which could give an opportunity to investigate a bacterial metabolism [4].

Materials and Methods

Reconstruction of the mathematical model

The complete genome of *Geobacillus icigianus* strain was extracted from NCBI Refseq database [5]. Initially we re-annotated this genome by means of RAST [6]. The procedure was conducted using standard RAST annotation scheme. At the next step we used automatic generation pipeline which is presented in web-service Kbase [7] to generate a genome-scale metabolic model. Module from Kbase for building of genome-scale models was harnessed with standard parameters including gap-filling algorithm. Biomass equation was generated automatically and stoichiometrically equivalent to biomass equation of *Bacillus subtilis*. A quality of the draft model was checked out using Memote web service [8], which demonstrated that consistency of the developed model is 92%. To improve the model consistency, we manually curated the draft model in order to modify SEED [9] reaction names and ID's on their equals from BIGG database [10]. We wrote a script on Python 3.6 using Cobrapy package and replaced all IDs which compose information about BIGG ID. Afterwards, we added boundary conditions for drain reactions which are necessary to describe wildtype growth of the strain. D-Glucose (glucose) was used as the first carbon source for the growth and lower bound was set up equal to -17 mmol/gDCWl/h^{-1} according to the published data for closely related species [11]. The modified GSM model of the strain was uploaded and analyzed via Optflux tool [12]. Flux balance analysis conducted in this tool using pFBA approach showed that growth rate on glucose, as a single carbon source, compose 0.5 mmol/gDCWl/h^{-1}.

To consider the growth of the bacteria on other carbon sources we added exchange reactions for metabolites growth ability on which was shown in the published data: *glycerol, L-arabinose and D-xylose* [3]. Metabolic analysis of the model revealed that to consider the growth on xylose and arabinose some metabolic reactions, which were not presented in the model, are required. Analysis of the metabolic pathways in closely related species of genus *Geobacillus* and metabolic pathways of *Bacillus subtilis* in SEED, KEGG [13] and BIGG databases resulted in an addition of the next set of reactions for growth on xylose: D-Xylose exchange, D-xylose reversible transport, Xylulokinase (EC: 2.7.1.17) reactions; for growth on arabinose: L-ribulokinase (EC:2.7.1.16) and L-ribulose-phosphate 4-epimerase (EC: 5.1.3.4) reactions. The presence of all proteins encoded in the *Geobacillus icigianus* genome for this set of reactions was checked out using Blastp web-service. Flux balance analysis for growth of the strain on abovementioned carbon sources was perfomed in Optflux tool using pFBA [14] approach too. Visualization of pFBA outcomes was carried out through Escher web-service [15].

Model modification for 2,3-butanediol production

To identify a list of reactions required for 2,3-butanediol production we conducted a literature analysis. It turned out that metabolic reactions to synthesize the substance in the bacteria comprise: acetolactate synthase (EC:2.2.1.6), acetolactate decarboxylase (EC: 4.1.1.5) and (R,R)-butanediol dehydrogenase (EC:1.1.1.4). Acetolactate synthase was originally presented in the model, while other metabolic reactions were added using Cobrapy. Moreover, we needed to add (R,R)-butanediol transport и (R,R)-

butanediol exchange reactions. The final version of the GSM model was uploaded into Memote web-service which demonstrated that the model consistency did not change compared to the draft model and equals to 92%.

Model analysis for 2,3-butanediol production optimization

To identify genetic modifications in order to increase 2,3-butanediol production and simultaneously do not significantly reduce biomass value we used evolutionary optimization approach via Optflux. To conduct this type of the analysis we selected 5 basic simulation algorithms: pFBA, MiMBL [16], MOMA [17], LMOMA [18] and ROOM [19]. All algorithms were started with 5000 maximum evolutionary functions and with maximum number modifications equal to 2. Optimization algorithm was chosen considering specific options of simulation methods. LMOMA, MOMA and pFBA simulation methods were run with Strengh Pareto Evolutionary Algorithm for reaction under/over expression. MiMBL and ROOM methods were initialized with Strengh Pareto Evolutionary Algorithm for the gene under/over expression.

Results

Thus, we generated the first GSM model for *G.icigianus* using Kbase web-service and final version of the model includes 1678 reactions, 1590 metabolites and 1316 genes. Flux balance analysis of the model showed that flux distribution in *G.icigianus* differs from *B.subtilis*, a model microorganism metabolic pathways and biomass equation of which were employed as a template for our model at the building stage. For instance, there are changes in electron donor/acceptor reactions in the citric acid cycle (TCA); missed reaction PGL (EC: 3.1.1.31), which is catalyzed by not thermostable enzyme, and the model describes the metabolic feature. Moreover, simulations of the GSM model demonstrated that an electron donor for oxidative phosphorylation depends on the carbon source. Growth on all substrates excluding glycerol showed a presence of lactate and succinate as excreted compounds, and it is consistent with experimental data [3]. The developed model predicts that glucose and glycerol (0.5 mmol/gDCWl/h^{-1}) are the most effective substrates for the growth, but growth on glycerol needs more oxygen. Furthermore, we matched built GSM model with published one for *Bacillus subtilis* (iYO844) [19]. As a result, the growth rate of *G.icigianus* is higher than one of *B.subtilis* for analogous substrate uptake rate which corresponds to 1% glucose concentration in the media according to the published data [20].

Model optimization for 2,3 butanediol production

Optimization analysis of the model in Optflux tool indicated that all carbon substrates can be used for 2,3-BTD production. It worth to note that MOMA and ROOM algorithms were not able to calculate the model. However, LMOMA, MiMBL, pFBA algorithms predicted glycerol as the most promising substrate for 2,3-BTD production by *G.icigianus*. All reaction modifications to improve the 2,3-BTD production predicted by the algorithm somehow affect two metabolic aspects: 1) modifications of the TCA cycle that lead the reduction of succinate production and 2) modifications that result in anaerobic or microaerobic growth conditions. It is interesting that all predicted ways of the metabolic modifications were earlier experimentally verified for another biotechnological species [1] and for closely related *Bacillus subtilis* [21].

References

[1] Celińska, E., and W. Grajek. "Biotechnological production of 2, 3-butanediol—current state and prospects." Biotechnology advances 27.6 (2009).

[2] Xiao, Z., Wang, X., Huang, Y. et al. “Thermophilic fermentation of acetoin and 2,3-butanediol by a novel Geobacillus strain”. Biotechnol Biofuels 5, 88 (2012).

[3] Bryanskaya, A. V., Rozanov, A. S., Slynko, N. M., Shekhovtsov, S. V., & Peltek, S. E. Geobacillus icigianus sp. nov., a thermophilic bacterium isolated from a hot spring. International journal of systematic and evolutionary microbiology, 65(3), 864-869, (2015).

[4] Simeonidis, E., Price, N.D. “Genome-scale modeling for metabolic engineering”. J Ind Microbiol Biotechnol 42, 327–338 (2015).

[5] Bryanskaya A. V. et al. “Draft genome sequence of *Geobacillus icigianus* strain G1w1T isolated from hot springs in the Valley of Geysers, Kamchatka (Russian Federation)”, Genome Announc.. – T. 2. – №. 5. – C. e01098-14 (2014).

[6] Aziz, R.K., Bartels, D., Best, A.A. et al. “The RAST Server: Rapid Annotations using Subsystems Technology”, BMC Genomics 9, 75 (2008).

[7] Arkin, A., Cottingham, R., Henry, C. et al. “KBase: The United States Department of Energy Systems Biology Knowledgebase”, Nat Biotechnol 36, 566–569 (2018).

[8] Lieven, C., Beber, M.E., Olivier, B.G. et al. “MEMOTE for standardized genome-scale metabolic model testing”, Nat Biotechnol 38, 272–276 (2020).

[9] Devoid S., Overbeek R., DeJongh M., Vonstein V., Best A.A., Henry C. (2013) “Automated Genome Annotation and Metabolic Model Reconstruction in the SEED and Model SEED”. In: Alper H. (eds) Systems Metabolic Engineering. Methods in Molecular Biology (Methods and Protocols), vol 985. Humana Press, Totowa, NJ.

[10] Charles J Norsigian, Neha Pusarla, John Luke McConn, James T Yurkovich, Andreas Dräger, Bernhard O Palsson, Zachary King, “BiGG Models 2020: multi-strain genome-scale models and expansion across the phylogenetic tree”, Nucleic Acids Research, Volume 48, Issue D1, 08 January 2020, Pages D402–D406.

[11] Lisowska, B., “Genomic analysis and metabolic modelling of *Geobacillus thermoglucosidasius* NCIMB 11955”, Thesis, Department of Biology & Biochemistry of University of Bath, (2016).

[12] Rocha, I., Maia, P., Evangelista, P. et al. “OptFlux: an open-source software platform for in silico metabolic engineering”, BMC Syst Biol 4, 45 (2010).

[13] Kanehisa M. et al. “The KEGG database”, Novartis Foundation Symposium. – Chichester; New York; John Wiley; 1999, 2002. – C. 91-100.

[14] Lewis NE, Hixson KK, Conrad TM, et al. “Omic data from evolved E. coli are consistent with computed optimal growth from genome-scale models”. Mol Syst Biol. (2010).

[15] Zachary A. King, Andreas Dräger, Ali Ebrahim, Nikolaus Sonnenschein, Nathan E. Lewis, and Bernhard O. Palsson. “Escher: A web application for building, sharing, and embedding data-rich visualizations of biological pathways”, PLOS Computational Biology 11(8), (2015).

[16] Brochado AR, Andrejev S, Maranas CD, Patil KR. “Impact of stoichiometry representation on simulation of genotype-phenotype relationships in metabolic networks”, PLoS Comput Biol. 2012.

[17] Segrè D, Vitkup D, Church GM. “Analysis of optimality in natural and perturbed metabolic networks”, Proc Natl Acad Sci U S A. 2002.

[18] Becker, S., Feist, A., Mo, M. et al. “Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox”, Nat Protoc 2, 727–738 (2007).

[19] Shlomi, Tomer et al. “Regulatory on/off minimization of metabolic flux changes after genetic perturbations.” Proceedings of the National Academy of Sciences of the United States of America vol. 102,21 (2005).

[20] Oh, You-Kwan, et al. "Genome-scale reconstruction of metabolic network in *Bacillus subtilis* based on high-throughput phenotyping and gene essentiality data." Journal of Biological Chemistry 282.39 (2007).

[21] Li, L., Zhang, L., Li, K. et al. A newly isolated *Bacillus licheniformisstrain* thermophilically produces 2,3-butanediol, a platform and fuel bio-chemical. Biotechnol Biofuels 6, 123 (2013).