Towards the Reconstruction of the Genome-Scale Metabolic Model of *Lactobacillus acidophilus* La-14

Emanuel Cunha[1], Ahmad Zeidan[2], and Oscar Dias[1(✉)]

[1] Centre of Biological Engineering, University of Minho, 4710-057 Braga, Portugal
emanuel.r.cunha@gmail.com, odias@ceb.uminho.pt
[2] Computational Systems Biology Discovery, R&D, Chr. Hansen A/S, Hørsholm, Denmark
DKAHZE@chr-hansen.com

Abstract. *Lactobacillus acidophilus* is a probiotic lactic acid bacterium used in food and dietary supplements for many years. However, despite its importance for industrial development and recognized health-promoting effects, no genome-scale metabolic model has been reported. A GSM model for *L. acidophilus* La-14 was developed, accounting 494 genes and 783 reactions. A genome annotation was performed to identify the metabolic potential of the bacterium. The biomass composition was determined based on information available in literature and previously published models. The model was validated by comparing *in silico* simulations with experimental data, regarding the aerobic and anaerobic growth. The reconstruction of the metabolic model has confirmed the fastidious requirements of *L. acidophilus* for amino acids, fatty acids, and vitamins. This model can be used for a better understanding of the metabolism of this bacterium and identification of industrially desirable compounds.

Keywords: Genome-scale metabolic model · Lactic acid bacteria · Probiotic · *Lactobacillus acidophilus* La-14 · Metabolic reconstruction

1 Introduction

Genome-scale metabolic (GSM) models and constraint-based modeling are increasingly important tools in systems biology. These models contain all known metabolites, reactions, and pathways of a target organism, allowing to look at the cell from a global perspective. Hence, the metabolic mechanisms that lead to the final phenotype can be better understood, accelerating the industrial development of biological processes and decreasing not only the costs but also the time required for this kind of exploration [1]. The reconstruction of GSM models involves four major steps: genome annotation, metabolic network assembly and curation, conversion of the network into a stoichiometric model, and model validation [2]. Due to the complexity and time required for this process, it is highly recommended the utilization of a user-friendly tool designed for this purpose, like Metabolic Models Reconstruction Using Genome-Scale Information

G. Panuccio et al. (Eds.): PACBB 2020, AISC 1240, pp. 205–214, 2021.
https://doi.org/10.1007/978-3-030-54568-0_21

(*merlin*) [3], or others [4]. Moreover, biological online databases and primary literature provide valuable information regarding genome sequences, enzymes, transport proteins, and metabolic pathways.

Lactic acid bacteria (LAB) is a heterogeneous group of gram-positive bacteria that produce lactic acid as the main product of fermentation processes. These bacteria can be classified in homofermentative, facultative heterofermentative, and heterofermentative. In homofermentative LAB, lactic acid is the unique fermentation by-product, while heterofermentative organisms produce ethanol and/or acetate besides lactic acid. Facultative heterofermentative can present a homofermentative or heterofermentative metabolism depending on the environmental conditions [5].

Lactobacillus acidophilus is a homofermentative LAB used by the food industry in products like yogurt, sweet acidophilus milk, and cheese [6]. This organism uses glycolysis to ferment hexoses, producing L and D-lactic acid in a racemic mixture [7]. In 2013, the complete genome sequence of the *L. acidophilus* La-14 strain was published [8]. Several health-promoting properties have been reported, including oxalate-degradation capability [9], and beneficial contribution to the immune system [10], in addition to the ordinary cooperation to prevent and treat gastrointestinal disorders found in other *L. acidophilus* strains.

This work aimed at obtaining a high-quality GSM model for *L. acidophilus* La-14, allowing to improve the knowledge of the metabolism of this bacterium.

2 Results and Discussion

2.1 Biomass Composition

The biomass equation includes nine different entities (Table 1) representing the complex macromolecules found in *L. acidophilus*.

Table 1. Biomass composition of *L. acidophilus*.

Macromolecule	*L. acidophilus* La-14	*L. lactis* MG1363	*L. plantarum* WCFS1	Reference
Protein	0.288	0.460	0.261	[11]/*merlin*
DNA	0.019	0.023	0.019	[11]/*merlin*
RNA	0.090	0.107	0.090	[11]/*merlin*
Lipid	0.063	0.034	0.063	[11–13]
Peptidoglycan	0.145	0.118	0.145	[11, 14]
Exopolysaccharides	0.119	0.120	0.099	[11, 15]
Wall Teichoic acids	0.138	–	0.138	[11, 12]
Lipoteichoic acids	0.080	0.080	0.410	[16, 17]
Cofactors	0.058	0.058	–	[17, 18]/*merlin*

Lactobacilli present a low protein content (around 30% of the biomass) when comparing to other gram-positive bacteria [11, 12]. The 28.8% inferred from the *L. plantarum* model agrees with this. The DNA, RNA, and protein composition were determined with the BiomassX [19] tool, available in *merlin*. The remaining macromolecule contents were determined according to experimental data [20–26]. The growth-associated maintenance (GAM) energy requirement was defined as 27.4 mmol gDW^{-1}, whereas the non-growth associated maintenance (NGAM) energy requirement was adjusted to 1.50 mmol h^{-1} gDW^{-1}.

2.2 Carbohydrate Uptake

The pathways used by *L. acidophilus* for the degradation of carbohydrates (glucose, fructose, galactose, lactose, and sucrose) were analyzed accounting in silico simulations and available information. Figure 1 shows a reconstruction of these pathways.

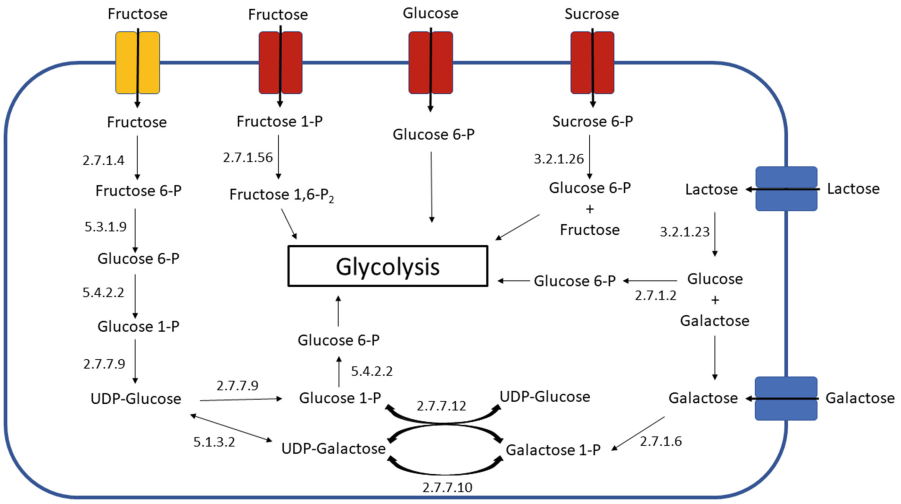

Fig. 1. Pathway reconstruction of the metabolism of carbohydrates in *L. acidophilus*. The uptake of sucrose, fructose, and glucose is made through the phosphotransferase system (PTS) (red). An ABC system for the uptake of fructose is presented in yellow. Galactose and lactose enter the cell through a permease (blue). The EC numbers have the following correspondence: 2.7.1.2 – glucokinase; 2.7.1.4 – fructokinase; 2.7.1.6 – galactokinase; 2.7.1.56 – 1-phosphofructokinase; 2.7.7.9 – UTP-glucose-1-phosphate uridylyltransferase; 2.7.7.10 – UTP-hexose-1-phosphate uridylyltransferase; 2.7.7.12 – UTP-hexose-1-phosphate uridylyltransferase; 3.2.1.23 – β-galactosidase; 3.2.1.26 – β-fructofuranosidase; 5.1.3.2 – UDP-glucose 4-epimerase; 5.3.1.9 – glucose-6-phosphate isomerase; 5.4.2.2 – phosphoglucomutase.

The glucose uptake through the PTS allows this sugar to enter the cell as glucose 6-phosphate, which is mainly directed to glycolysis. However, a minor amount is converted to UDP-glucose, required for the synthesis of exopolysaccharides and wall teichoic acids. The sucrose uptake is also performed through the PTS. The β-fructofuranosidase activity

allows the degradation of sucrose 6-phosphate to glucose 6-phosphate and fructose. The PTS is the main transport system responsible for the uptake of exogenous fructose, forming fructose 1-phosphate, which is phosphorylated to fructose 1, 6-bisphosphate. In *L. lactis*, part of the fructose 1, 6- bisphosphate is hydrolyzed to fructose 6-phosphate, which is used for the formation of UDP-glucose. However, fructose-bisphosphatase appears to be absent in the genome of *L. acidophilus*. Hence, fructose required for biomass may enter the cell through a mechanism other than the PTS. In *in silico* simulations, fructose required for biomass production enters the cell through an ABC system and is then phosphorylated to fructose 6-phosphate.

Lactose and galactose are both assimilated through symport with H^+. Lactose is degraded into glucose and galactose by a β-galactosidase. Glucose enters in the EMB pathway, while galactose follows the Leloir pathway.

2.3 Model Validation

The *in silico* growth rate (0.81 h^{-1}) and the uptake rate for carbon sources (16.4 mmol h^{-1} gDW^{-1}) was defined according to available experimental data [27]. *Soska* [28] obtained an identical growth rate (0.83 h^{-1}), using a similar Chemically Defined Media (CDM) (with the addition of L-asparagine), though not measuring, the glucose consumption rate. The *in silico* ratio between the lactic acid production and glucose consumption is in agreement with literature (1.8 $mol_{lactate}/mol_{glucose}$) [29].

L. acidophilus achieves optimal growth in anaerobic and microaerophilic conditions, even though most strains are aerotolerant. The lower growth rate in aerobic conditions has been associated with the damage caused by reactive oxygen species to biomolecules like proteins and lipids [30]. In anaerobic conditions, *L. acidophilus* requires acetate for growth [31], which was also observed in *L. lactis* and *L. johnsonii* [31, 32]. In these species, the activity of acetate kinase allows the phosphorylation of acetate, generating acetyl-phosphate, which can be converted to acetyl-CoA by phosphate acetyltransferase. The production of acetyl-CoA through this pathway is essential since genes encoding pyruvate dehydrogenase and pyruvate formate lyase are absent in the *L. acidophilus* genome. In *L. johnsonii*, the requirement for acetate can be replaced by oxygen, due to the activity of pyruvate oxidase (POX) [31].

The rerouting of pyruvate through this pathway allows the production of additional ATP, while NAD^+ is regenerated either by NADH oxidase and NADH peroxidase. *In silico* simulations in aerobic conditions predict the production of acetate through the POX pathway and a higher growth rate. Nevertheless, *L. acidophilus* is obligatory homofermentative and does not produce acetate, even if oxygen is available [33, 34]. To access the *in silico* requirement for acetate production, Flux Variability Analysis (FVA) were performed in anaerobic and aerobic conditions (Table 2).

In anaerobic conditions, acetate is consumed while lactate production is mandatory. In these conditions, lactate (originated from pyruvate) is the only by-product, and the lactate dehydrogenase activity assures the regeneration of NAD^+. In aerobic conditions, the production of lactate, acetate or ethanol is not required to achieve the defined growth rate. No information was found explaining why *L. acidophilus* does not produce acetate in these conditions. In *L. johnsonii*, the flux through pyruvate oxidase is limited but sufficient to eliminate the requirement for acetate, without producing this compound [31].

Table 2. Minimum and maximum fluxes of acetate, lactate, and ethanol determined through FVA simulations in aerobic (oxygen uptake unconstrained) and anaerobic conditions. In both conditions, the growth rate was limited to 0.81 h^{-1}.

	Aerobic		Anaerobic	
Compound	Minimum flux	Maximal flux	Minimum flux	Maximal flux
Acetate	−0.27	−0.26	0	47.68
Lactate	29.88	29.89	0	29.67
Ethanol	0	0.001	0	4.98

Hence, POX may be used by *L. acidophilus* and *L. johnsonii* just to provide acetyl-CoA, and not for acetate production.

2.4 Model Summary

GSM models for six LAB species are available at the moment, in which *L. plantarum*, *L. casei*, and *L. lactis* are the most closely related species to *L. acidophilus*. An overview of the available metabolic models for these species is presented in Table 3.

Table 3. Overview of the *GSM* models of five lactic acid bacteria.

	L. acidophilus La-14	*L. acidophilus* NCFM	*L. plantarum* WCFS1	*L. lactis* MG1363	*L. casei* LC2W
Genes	494	540	721	518	846
Gene coverage (%)	26.3	29.0	23.5	19.9	27.7
Total Reactions	783	1460	762	754	969
Internal reactions	541	923	413	530	604
Transport reactions	122	132	118	119	227
Exchange reactions	120	405	113	105	139
Metabolites	682	1120	658	650	785
Internal	562	715	549	551	—*
External	120	405	113	105	—*
Unique metabolites	572	1009	554	552	604
Compartments	(c, e)	(c, e)	(c, e)	(c, e)	(c, p, e)

*data not available; c: cytoplasm; e: extracellular space; p: periplasm.

The GSM model contains 494 genes, which corresponds to 26.3% of the total genes in the genome of *L. acidophilus* La-14. This percentage is slightly higher than in *L. plantarum* and *L. lactis*.

In general, the number of genes, reactions, and metabolites in the metabolic model is higher than the respective ones in the *L. plantarum* and *L. lactis* models. This may be associated with a more restrictive reconstruction approach in these GSM models. In the *L. casei* LC2W model, an additional compartment (periplasm) was included, which is

unusual in GSM models of gram-positive bacteria. The number of internal and exchange reactions available in the AGORA model [35] is surprisingly high, which might be a result of the semi-automatic reconstruction and lack of manual curation. In fact, this model includes the fatty acid biosynthetic pathway regardless of the absence of the genes required for this pathway.

3 Materials and Methods

The *L. acidophilus* La-14 genome sequence was retrieved from the GenBank [36] repository, with the accession number CP005926 (assembly ID ASM38967v2).

3.1 Tools and Online Resources

Merlin was used to support the reconstruction of the GSM model. This tool provides a user-friendly graphical interface, allowing to perform several steps of this process semi-automatically. Through all stages, the GSM model reconstruction process was supported by information available in different online databases.

3.2 Genome Annotation

Enzymes annotation was performed using *SamPler* [37], a tool included in *merlin* that allows obtaining a semi-automatic annotation. Similarity searches were performed using Basic Local Alignment Search Tool (BLAST) [38] against Swiss-Prot and UniProtKB [39] in March 2018, setting an e-value threshold of 10^{-30}.

3.3 Network Assembly and Curation

A draft network was assembled by integrating the enzymes annotation and loading KEGG's [40] metabolic. The reactions stoichiometry was corrected using biological databases to retrieve information on metabolites formula and reaction stoichiometry.

The correction of the reversibility of reactions was performed automatically using the tool available in merlin for this purpose, setting the Ma and Zeng [41] database as the data source. Through the reconstruction process, the reversibility and direction of some reactions were manually corrected, which was performed by analyzing information available in online databases. The sub-cellular protein location was obtained using PSORTb 3.0 [42]. Transport reactions were automatically obtained using TranSyT, a tool available in *merlin*. Nevertheless, additional transport reactions were added if necessary. In that case, the TCDB [43], BiGG [44], and TransportDB 2.0 [45] databases were used to retrieve information for substrates, mechanisms, and genes associated with each transport reaction.

3.4 Conversion of the Network into a Stoichiometric Model

The complex macromolecules included in the biomass equation were inferred from the *Lactobacillus plantarum* WCFS1 model. The content of each macromolecule and respective precursors was determined using three different sources: experimental data, GSM models of closely related organisms, and *merlin.*

The GAM energy requirement was determined using information retrieved from the GSM model of *L. plantarum* WCFS1. The NGAM energy requirement was adjusted to experimental data by plotting the growth rate vs ATP maintenance value.

3.5 Validation of the Metabolic Model

The model validation stage was performed using Optflux v3.4.0 [46] for simulation and analysis of the metabolic model, followed by comparison with experimental data and literature. A CDM was defined according with *Morishita et al.* [47]. The uptake rate of all compounds included in the CDM was left unconstrained, except for the carbon source (glucose or other) whose uptake was settled as 16.4 mmol h^{-1} gDW^{-1}. *In silico* simulations in anaerobiosis were performed by constraining the lower bound of the oxygen exchange reaction to zero, while aerobiosis was defined by unconstraining the oxygen uptake.

In silico simulations were performed using Parsimonious Flux Balance Analysis (pFBA) [48, 49] maximizing the biomass production, by setting the biomass reaction as the objective function. To assess the growth rate, rich medium was used for comparison and adjustment to experimental data, in anaerobic conditions.

Acknowledgements. This study was performed under the scope of the project "BIODATA.PT – Portuguese Biological Data Network" (ref. LISBOA-01-0145-FEDER-022231), funded by FCT/MCTES, through national funds of PIDDAC, Fundo Europeu de Desenvolvimento Regional (FEDER), Programa Operacional de Competitividade e Internacionalização (POCI) and Programa Operacional Regional de Lisboa (Lisboa 2020).

References

1. Russell, R., Paterson, M., Lima, N.: Molecular biology of food and water borne Mycotoxigenic and Mycotic Fungi. Food Microbiol. (2015). https://doi.org/10.1145/2110147.2110155
2. Thiele, I., Palsson, B.Ø.: A protocol for generating a high-quality genome-scale metabolic reconstruction. Nat. Protocols **5**, 93–121 (2010). https://doi.org/10.1038/nprot.2009.203
3. Dias, O., Rocha, M., Ferreira, E.C., Rocha, I.: Reconstructing genome-scale metabolic models with merlin. Nucleic Acids Res. **43**, 3899–3910 (2015). https://doi.org/10.1093/nar/gkv294
4. Faria, J.P., Rocha, M., Rocha, I., Henry, C.S.: Methods for automated genome-scale metabolic model reconstruction. Biochem. Soc. Trans. **46**, 931–936 (2018). https://doi.org/10.1042/BST20170246
5. Lahtinen, S., Ouwehand, A.C., Salminen, S., von Wright, A.: Lactic acid bacteria — microbiological and functional aspects. In: Beneficial Microbes, 4th edn, vol. 3, pp. 245–247 (2012). https://doi.org/10.3920/BM2012.x003
6. Gopal, P.K.: Lactic acid bacteria | Lactobacillus spp.: Lactobacillus acidophilus. In: Encyclopedia of Dairy Sciences, pp. 91–95. Elsevier (2011)

7. Anjum, N., Maqsood, S., Masud, T., et al.: Lactobacillus acidophilus: characterization of the species and application in food production. Critical Rev. Food Sci. Nutr. **54**, 1241–1251 (2014). https://doi.org/10.1111/j.2003.00350.x
8. Stahl, B., Barrangou, R.: Complete genome sequence of probiotic strain lactobacillus acidophilus La-14. Genome Announcements (2013). https://doi.org/10.1128/genomeA.00376-13
9. Han, H., Segal, A.M., Seifter, J.L., Dwyer, J.T.: Nutritional management of kidney stones (Nephrolithiasis). Clin. Nutr. Res. **4**, 137–152 (2015). https://doi.org/10.7762/cnr.2015.4.3.137
10. Paineau, D., Carcano, D., Leyer, G., et al.: Effects of seven potential probiotic strains on specific immune responses in healthy adults: a double-blind, randomized, controlled trial. FEMS Immunol. Med. Microbiol. (2008). https://doi.org/10.1111/j.1574-695X.2008.00413.x
11. Teusink, B., Wiersma, A., Molenaar, D., et al.: Analysis of growth of Lactobacillus plantarum WCFS1 on a complex medium using a genome-scale metabolic model. J. Biol. Chem. (2006). https://doi.org/10.1074/jbc.M606263200
12. Xu, N., Liu, J., Ai, L., Liu, L.: Reconstruction and analysis of the genome-scale metabolic model of Lactobacillus casei LC2W. Gene (2015). https://doi.org/10.1016/j.gene.2014.10.034
13. Exterkate, F.A., Otten, B.J., Wassenberg, H.W., Veerkamp, J.H.: Comparison of the phospholipid composition of Bifidobacterium and Lactobacillus strains. J. Bacteriol. **106**, 824–829 (1971)
14. Wu, Z., Pan, D., Zeng, X., et al.: Phosphorylation of peptidoglycan from Lactobacillus acidophilus and its immunoregulatory function. Int. J. Food Sci. Technol. (2016). https://doi.org/10.1111/ijfs.13028
15. Bergmaier, D., Champagne, C.P., Lacroix, C.: Growth and exopolysaccharide production during free and immobilized cell chemostat culture of Lactobacillus rhamnosus RW-9595M. J. Appl. Microbiol. (2005). https://doi.org/10.1111/j.1365-2672.2004.02462.x
16. Fischer, W., Koch, H.U., Rösel, P., Fiedler, F.: Alanine ester-containing native lipoteichoic acids do not act as lipoteichoic acid carrier. Isolation, structural and functional characterization. J. Biol. Chem. **255**, 4557–4562 (1980)
17. Oliveira, A.P., Nielsen, J., Förster, J.: Modeling Lactococcus lactis using a genome-scale flux model. BMC Microbiol. (2005). https://doi.org/10.1186/1471-2180-5-39
18. Hamana, K., Akiba, T., Uchino, F., Matsuzaki, S.: Distribution of spermine in bacilli and lactic acid bacteria. Can. J. Microbiol. (2009). https://doi.org/10.1139/m89-069
19. Santos, S., Rocha, I.: Estimation of biomass composition from genomic and transcriptomic information. J. Integr. Bioinform. **13**, 285 (2016). https://doi.org/10.2390/biecoll-jib-2016-285
20. Laws, A.P., Chadha, M.J., Chacon-Romero, M., et al.: Determination of the structure and molecular weights of the exopolysaccharide produced by Lactobacillus acidophilus 5e2 when grown on different carbon feeds. Carbohydr. Res. **343**, 301–307 (2008). https://doi.org/10.1016/j.carres.2007.10.028
21. Yamamoto, Y., Nunome, T., Yamauchi, R., et al.: Structure of an exocellular polysaccharide of Lactobacillus helveticus TN-4, a spontaneous mutant strain of Lactobacillus helveticus TY1-2. Carbohydr. Res. (1995). https://doi.org/10.1016/0008-6215(95)00077-7
22. Staaf, M., Yang, Z., Huttunen, E., Widmalm, G.: Structural elucidation of the viscous exopolysaccharide produced by Lactobacillus helveticus Lb161. Carbohydr. Res. **326**, 113–119 (2000). https://doi.org/10.1016/S0008-6215(00)00027-6
23. Yang, Z., Staaf, M., Huttunen, E., Widmalm, G.: Structure of a viscous exopolysaccharide produced by Lactobacillus helveticus K16. Carbohydr. Res. (2000). https://doi.org/10.1016/S0008-6215(00)00201-9

24. Staaf, M., Widmalm, G., Yang, Z., Huttunen, E.: Structural elucidation of an extracellular polysaccharide produced by Lactobacillus helveticus. Carbohydr. Res. (1996). https://doi.org/10.1016/S0008-6215(96)00166-8
25. Oh, Y.-K., Palsson, B.O., Park, S.M., et al.: Genome-scale reconstruction of metabolic network in Bacillus subtilis based on high-throughput Phenotyping and Gene essentiality data. J. Biol. Chem. **282**, 28791–28799 (2007). https://doi.org/10.1074/jbc.M703759200
26. Robijn, G.W., Gallego, R.G., van den Berg, D.J.C., et al.: Structural characterization of the exopolysaccharide produced by Lactobacillus acidophilus LMG9433. Carbohydr. Res. (1996). https://doi.org/10.1016/0008-6215(96)00097-3
27. Lv, X,, Liu, G., Sun, X., et al.: Short communication: nutrient consumption patterns of Lactobacillus acidophilus KLDS 1.0738 in controlled pH batch fermentations. J. Dairy Sci. https://doi.org/10.3168/jds.2017-12607 (2017)
28. Soska, J.: Growth of Lactobacillus acidophilus in the absence of folic acid. J. Bacteriol. **91**, 1840–1847 (1966)
29. Gomes, A.M.P., Malcata, F.X.: Bifidobacterium spp. and Lactobacillus acidophilus: biological, biochemical, technological and therapeutical properties relevant for use as probiotics. Trends Food Sci. Technol. (1999). https://doi.org/10.1016/S0924-2244(99)00033-3
30. Talwalkar, A., Kailasapathy, K., Peiris, P., Arumugaswamy, R.: Application of RBGR - a simple way for screening of oxygen tolerance in probiotic bacteria. Int. J. Food Microbiol. (2001). https://doi.org/10.1016/S0168-1605(01)00563-3
31. Hertzberger, R.Y., Pridmore, R.D., Gysler, C., et al.: Oxygen relieves the CO2 and Acetate dependency of Lactobacillus johnsonii NCC 533. PLoS ONE (2013). https://doi.org/10.1371/journal.pone.0057235
32. Collins, E.B., Bruhn, J.C.: Roles of acetate and pyruvate in the metabolism of Streptococcus diacetilactis. J. Bacteriol. **103**, 541–546 (1970)
33. Sakamoto, M., Komagata, K.: Aerobic growth of and activities of NADH oxidase and NADH peroxidase in lactic acid bacteria. J. Ferment. Bioeng. (1996). https://doi.org/10.1016/0922-338X(96)88810-6
34. Talwalkar, A., Kailasapathy, K., Hourigan, J., et al.: An improved method for the determination of NADH oxidase in the presence of NADH peroxidase in lactic acid bacteria. J. Microbiol. Methods (2003). https://doi.org/10.1016/S0167-7012(02)00189-6
35. Magnúsdóttir, S., Heinken, A., Kutt, L., et al.: Generation of genome-scale metabolic reconstructions for 773 members of the human gut microbiota. Nat. Biotechnol. (2017). https://doi.org/10.1038/nbt.3703
36. Sayers, E.W., Barrett, T., Benson, D.A., et al.: Database resources of the National Center for Biotechnology Information. Nucleic Acids Res. **37**, D5–D15 (2009). https://doi.org/10.1093/nar/gkn741
37. Cruz, F., Lagoa, D., Mendes, J., et al.: SamPler – a novel method for selecting parameters for gene functional annotation routines. BMC Bioinform. (2019). https://doi.org/10.1186/s12859-019-3038-4
38. Altschul, S.F., Gish, W., Miller, W., et al.: Basic local alignment search tool. J. Mol. Biol. **215**, 403–410 (1990). https://doi.org/10.1016/S0022-2836(05)80360-2
39. Apweiler, R., Martin, M.J., O'Donovan, C., et al.: Ongoing and future developments at the universal protein resource. Nucleic Acids Res. **39**, D214–D219 (2011). https://doi.org/10.1093/nar/gkq1020
40. Kanehisa, M., Goto, S.: KEGG: kyoto encyclopedia of genes and genomes. Nucleic Acids Res. **28**, 27–30 (2000). https://doi.org/10.1093/nar/27.1.29
41. Ma, H., Zeng, A.P.: Reconstruction of metabolic networks from genome data and analysis of their global structure for various organisms. Bioinformatics (2003). https://doi.org/10.1093/bioinformatics/19.2.270

42. Nakai, K., Horton, P.: PSORT: a program for detecting sorting signals in proteins and predicting their subcellular localization. Trends Biochem. Sci. **266**, 594–600 (1999)
43. Saier, M.H.: A functional-phylogenetic classification system for transmembrane solute transporters. Microbiol. Mol. Biol. Rev.: MMBR **64**, 354–411 (2000). https://doi.org/10.1128/MMBR.64.2.354-411.2000
44. King, Z.A., Lu, J., Dräger, A., et al.: BiGG models: a platform for integrating, standardizing and sharing genome-scale models. Nucleic Acids Res. (2016). https://doi.org/10.1093/nar/gkv1049
45. Ren, Q., Chen, K., Paulsen, I.T.: TransportDB: a comprehensive database resource for cytoplasmic membrane transport systems and outer membrane channels. Nucleic Acids Res. (2007). https://doi.org/10.1093/nar/gkl925
46. 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). https://doi.org/10.1186/1752-0509-4-45
47. Morishita, T., Deguchi, Y., Yajima, M., et al.: Multiple nutritional requirements of lactobacilli: genetic lesions affecting amino acid biosynthetic pathways. J. Bacteriol. **148**, 64–67 (1981)
48. Famili, I., Forster, J., Nielsen, J., Palsson, B.O.: Saccharomyces cerevisiae phenotypes can be predicted by using constraint-based analysis of a genome-scale reconstructed metabolic network. Proc. Nat. Acad. Sci. USA **100**, 13134–13139 (2003). https://doi.org/10.1073/pnas.2235812100
49. Varma, A., Palsson, B.O.: Metabolic capabilities of escherichia coli. II. Optimal growth patterns. J. Theor. Biol. (1993). https://doi.org/10.1006/jtbi.1993.1203