Genome-scale model of *Rhodotorula toruloides* metabolism

Ievgeniia A Tiukova[1,2], Sylvain Prigent[3], Jens Nielsen[1], Mats Sandgren[2], Eduard J Kerkhoven[1,*]

[1]Systems and Synthetic Biology, Department of Biology and Biological Engineering, Chalmers University of Technology, Göteborg, Sweden

[2]Department of Molecular Sciences, Swedish University of Agricultural Sciences, Uppsala, Sweden

[3]UMR 1332 BFP, INRA, Univ Bordeaux, F33883 Villenave d'Ornon, France

* corresponding author: EJK (eduardk@chalmers.se)

Abstract

The basidiomycete red yeast *Rhodotorula toruloides* is a promising platform organism for production of biooils. We present *rhto-GEM*, the first genome-scale model of *R. toruloides* metabolism, that was largely reconstructed using RAVEN toolbox. The model includes 4869 genes, 897 reactions, and 3334 metabolites. Lipid metabolism was described using the SLIMEr formalism, which allows direct integration of lipid class and acyl chain experimental distribution data. The simulation results confirmed that the *R. toruloides* model provides valid growth predictions on glucose, xylose and glycerol, while prediction of genetic engineering targets to increase production of linolenic acid and triacylglycerols highlighted genes that have previously been successfully used to increase lipid production. This renders *rtho-GEM* useful for future studies to improve the production of other industrially important oleochemicals including both value-added fatty acids and carotenoids, while it will also be valuable tool for system-wide omics-data analysis in *R. toruloides*. Expanding the portfolio of GEMs for lipid accumulating fungi contributes to both understanding of metabolic mechanisms of the oleaginous phenotype but also uncover particularities of lipid production machinery in *R. toruloides*.

Introduction

Rhodosporidium toruloides (syn. *Rhodotorula toruloides*) is a basidiomycetous yeast belonging to the subphylum Pucciniomycotina and occurs naturally in a wide range of habitats including surfaces of leaves, soil and sea water [1]. The broad substrate range of *R. toruloides* and its ability to accumulate lipids exceeding half of its cell dry weight has made this yeast a popular system for production of biological oils from inedible substrates (e.g. pentose sugars, crude glycerol) [2]. The ready scalability and fast rate of lipid production by oleaginous yeasts such as *R. toruloides* makes this approach a competitive alternative to emerging "third generation" biodiesel produced from carbon dioxide. A commercially viable process of biodiesel production using *R. toruloides* was recently patented in Brazil where conversion of a mixture of sugarcane juice and urea into lipids achieved a productivity of

0.44 g/L h [3]. The *R. toruloides* lipid fraction contains omega-3 linolenic acid and heptadecenoic acid, which also makes this yeast a promising organism for production of pharma- and nutraceuticals. *R. toruloides* also produces several carotenoid pigments including torularhodin, torulene, γ-carotene and β-carotene [4].

Genome-scale metabolic models (GEMs) are comprehensive summaries of metabolic reactions in an organism, which are derived from the available genome sequence. GEMs can be used for identification of potential metabolic engineering targets, or select the best performing metabolic network among alternatives [5]. Extending the portfolio of GEMs for various organisms can enhance the benefits of bioprospecting and aid in the design and improvement of performance of synthetic organisms. The generation of a genome scale metabolic model of *R. toruloides* lipid synthesis would provide a deeper understanding of oleaginous ability and facilitate genetic engineering strategies. A number of metabolic models of lipid production in *R. toruloides* have been reported to date [6,7].

One of the major determinants of the oleaginous phenotype of *R. toruloides* is its capacity for acetyl-CoA production [8]. Unlike non-oleaginous yeasts such as the baker's yeast Saccharomyces cerevisiae, *R. toruloides* possesses the enzyme ATP:citrate lyase, which is the main source of acetyl-CoA for lipid synthesis. In addition, a mitochondrial beta-oxidation pathway provides additional source of acetyl-CoA in this yeast [9]. Lipid biosynthetic reactions downstream of acetyl-CoA synthesis do not differ between oleaginous and non-oleaginous yeast species.

A better understanding of lipid metabolism in *R. toruloides* has applications beyond the production of lipids for fuel and nutraceuticals. *R. toruloides* has also been used in screens for novel anti-obesity drugs against candidates [10], while lipid synthesis has been shown to be a promising target for antifungal therapies [11]. Several species within the *Rhodotorula* genus have been identified as emerging pathogens in both animals and people [12], and intracellular glycerol accumulation caused by hydrolysis of storage triglycerides by pathogenic fungi has been shown to play a crucial role in turgor generation for penetration and invasion of tissues [13,14]. Modelling of *R. toruloides* lipid metabolism can therefore also aid in unravelling the pathobiology of this group of yeasts.

Methods

Model reconstruction and analysis

The genome-scale model, named *rhto-GEM*, is based on the genome sequence of *R. toruloides* strain NP11 [8]. The reconstruction of *rhto-GEM* was primarily performed using RAVEN 2, a MATLAB toolbox for genome-scale model reconstruction [15]. All steps of the reconstruction are documented in detail on the GitHub repository under the folder `ComplementaryScripts/reconstruction`. More information on the GitHub repository is provided below.

To reconstruct those parts of metabolism that are relatively conserved between fungal species, the well-curated GEM of *Saccharomyces cerevisiae* was taken as template model (yeast-GEM version 8.2.0, [16]), while orthologous genes were identified via bi-directional BLASTP against the *S. cerevisiae* S288c reference genome. All 1-1 orthologs were included, after cut-offs of E-value <1e-20; identity >35% and alignment length >150 bp. Additional orthologs between both species were identified as provided by MetaPhOrs [17], filtered for ortholog pairs with confidence scores of 1 and whose PhylomeDB tree contained at least 2 members.

Gap-filling was performed with meneco [18], by using the biomass components as targets and yeast-GEM version 8.2.0 as repair network. Meneco typically finds multiple alternative solutions that all fill the identified gaps. Not to favour one proposed reaction over the other, as we have no proof of what solution is more likely to be present, we added the union of proposed reactions. Gap-filling was complemented with additional manual curation to identify missing reactions and incorrect gene associations.

Flux balance analysis was performed with RAVEN toolbox, using constraints on exchange fluxes as specified in the text, while also detailed in the relevant scripts in the `ComplementaryScripts` folder. To predict genetic targets for metabolic engineering the FSEOF (flux scanning of enforced objective function) implementation of RAVEN was used. Exchange reactions were added for linolenic acid and triglyceride (1-18:0, 2-18:1, 3-18:0), which were optimized with glucose as carbon source. The slope identified by FSEOF was assumed to be indicative of which reactions are promising targets for overexpression to increase productivity.

Reconstructing lipid metabolism

Lipid metabolism was described using the SLIMEr formalism, which Splits Lipids Into Measurable Entities [19]. To somewhat reduce the number of unique lipid species to include, we included only representatives of each lipid with a unique chain composition. As example, when cardiolipin (1-18:0, 2-18:1, 3-16:0, 4-18:1) is included in the model, cardiolipin (1-16:0, 2-18:1, 3-18:0, 4-18:1) is not included (cf. the acyl chains on the 1- and 3-position). To facilitate adjusting the lipid composition in the model by considering measured lipid class and acyl chain distributions, we provide the functions `adjustRhtoBiomass` and `scaleAbundancesRhto` in the `ComplementaryScripts/experimental` folder.

Further model development and distribution via GitHub repository

The growth-associated and non-growth associated energy requirements were fitted to measured glucose uptake rates from continuous cultivations of *R. toruloides* as reported in literature [20]. The biomass composition was modified from yeast-GEM to include *R. toruloides* lipid class and acyl chain distributions, as provided in `ComplementaryData/data`. The *rhto-GEM* model is hosted on a

dedicated GitHub repository (http://github.com/SysBioChalmers/*rhto-GEM*). Here, all scripts for model reconstruction are provided, in addition to the model in various file formats, e.g. SBML, YAML, TXT and XLSX, and scripts for performing the simulations detailed in this manuscript. This environment allows for versioning and comparison of the model, reporting and tracking of issues, organization of development and continuous integration.

Results

Genome-scale models of metabolism are valuable tools in studies of metabolism, to e.g. identify metabolic engineering targets or integration of omics data. To support the development of *R. toruloides* as promising microbial biocatalyst for oleochemical production, we developed a genome-scale model for *R. toruloides* strain NP11 through a combination of semi-automated reconstruction and manual curation based on literature data (**Figure 1**). The reconstruction and curation process is tracked in a publicly available Git repository to support reproducibility and continuous curation of the model. Each step of the reconstruction is documented in the `reconstruction` subfolder.

Bidirectional protein-blast [21] with RAVEN [15] querying the *R. toruloides* and *S. cerevisiae* genomes identified 628 pairs of 1-1 orthologous genes that were annotated to the template yeast-GEM model, while querying MetaPhOrs [17] yielded 571 pairs of orthologs model-annotated genes. Combined, this first step of model reconstruction identified 791 *R. toruloides* genes, connecting to 2046 reactions that represent parts of metabolism that are well conserved between fungal species.

In the second step of model reconstruction, a large number of non-gene associated reactions, primarily pseudo-reactions, exchange and intracellular transport reactions, were transferred from the template to the draft reconstruction. While such reactions are required to obtain a functional model, not all the template-derived reactions will be required to support growth in the *R. toruloides* model. Therefore, unconnected non-gene associated reactions were removed from the model at a late stage of the reconstruction. In addition, through extensive manual curation of central metabolism a large set of additional genes and reactions could be identified, resulting in an extended draft model with 3259 reactions and 837 genes.

R. toruloides specific reactions and pathways, such as carotene and torulene biosynthesis, synthesis and degradation of C18:2 and C18:3 fatty acids, and mitochondrial beta-oxidation were subsequently manually curated in the third step of model reconstruction, to increase the reaction and gene count to 3373 and 899. As this model was still unable to support the production of biomass, gap-filling in step 4 identified four additional reactions and one gene that were required for obtaining a functional draft model that can produce biomass.

As particular interest on *R. toruloides* is focused on its oleaginous nature, attention was paid to accurately depict lipid metabolism. Recently, we have developed the SLIMEr formalism for describing lipids in genome-scale models, which Splits Lipids Into Measurable Entities [19]. This formalism represents the flexibility of lipid metabolism while allowing incorporation of measurements of lipid classes and acyl chain distributions. Therefore, in the fifth step of reconstruction we applied the SLIMEr formalism as previously described for *S. cerevisiae*. As the acyl chain distribution of *R. toruloides* is different from *S. cerevisiae*, e.g. the presence of C18:2 and C18:3, this required extensive manual curation of the SLIMEr reactions. The lipid-curated draft model contained 4889 reactions and 904 genes, and its lipid composition can be modified based on experimental measurements of lipid class and acyl chain distribution.

To transform this functional draft model to the first version of the *R. toruloides* GEM, additional manual curation was performed where remaining template-derived genes were replaced by their *R. toruloides* ortholog where possible and otherwise deprecated. In addition, in step 7 of the reconstruction the (non-)growth associated maintenance energy was fitted to experimentally determined growth and glucose uptake rates [20]. The resulting model, abbreviated to *rhto-GEM*, is the first curated genome-scale model of *R. toruloides* with a total of 4869 reactions, 3334 metabolites and 897 genes.

To validate *rhto-GEM* we compared its growth rate to experimental measured values. From literature we gathered glucose, glycerol and xylose uptake rates from *R. toruloides* bioreactor cultivations [6,20,22,23], constraint the uptake rates in *rhto-GEM* to the reported values and optimized for biomass production. A good correlation could be found during growth on glucose, glycerol and xylose (**Figure 2**).

Significant interest in *R. toruloides* is focused on its potential as biocatalyst for production of triacylglycerol for biofuels and other applications. Ideally, such a biocatalyst should be able to consume a variety of different carbon sources, to allow flexibility in feedstock used. We used *rhto-GEM* to predict metabolic engineering targets for increased production of triacylglycerol (TAG) and linolenic acid while alternatively using glucose or xylose as carbon source, two sugars that are major parts of hydrolysed plant biomass. We performed FSEOF analysis using triacylglyceride (1-18:0, 2-18:1, 3-18:0) as a representative TAG for biofuel production, while this TAG is furthermore also of interest as major component of cocoa butter. The slope parameter derived from FSEOF is indicative of how strong each reaction is contributing towards a shift from growth towards production of the target compound. Comparison of glucose and xylose as carbon sources, with TAG and linolenic acid as desired compound (**Table 1**) indicates significant overlap in predicted targets, which can be anticipated due to the chemical similarity between the two compounds and the similarity in the catabolic pathways of glucose and xylose. Regardless, increasing production of TAGs seems to

require more significant rewiring of metabolism, represented by the higher slopes predicted from FSEOF in comparison to those predicted for linolenic acid, while catabolism of xylose affects more gene associated reactions. Most significant targets in all simulations were acetyl-CoA carboxylase and acetyl-CoA synthetase, which are located at the start of the lipid biosynthetic pathway. Overexpression of these and other genes from this list have recently been shown as effective strategy to increase TAG production [24,25], validating that this analysis is able to yield valuable predictions.

Conclusion

Previous studies have presented GEMs of several oleaginous fungal species, including oleaginous ascomycete *Yarrowia lipolytica* [26], zygomycetes *Mortierella alpina* [27] and *Mucor circinelloides* [28]. Our study presents the first reconstruction of GEM of lipid-accumulating basidiomycete *R. toruloides.* This will allow for understanding of molecular mechanisms of the oleaginous phenotype across phylogenetically distant fungal species.

The model incorporates knowledge obtained from genomics and proteomics data generated for *R. toruloides* [8] and was validated using cultivation data [6,20,22,23]. *R. toruloides* has broad substrate range including inedible substrates, such as pentoses and crude glycerol. Simulations of *rhto-GEM* on various carbon sources demonstrated good agreement with experimentally reported growth rates. Analysis of model allowed to identify potential genetic engineering strategies for enhanced lipid production. Some of these genetic targets were found to agree with published experimental studies [24,25]. As such, *rhto-GEM* emerges as a valuable tool for future analysis of oleaginous and lipid metabolism, while its distribution through a Git repository allows for continuous improvement and tracking of model development.

Acknowledgements

Financial support was provided by grants from The Swedish Research Council for Environment, Agricultural Sciences and Spatial Planning (FORMAS) number 213-2013-80, FORMAS Mobility (2016-00767) and the Novo Nordisk Foundation (NNF10CC1016517). EJK acknowledges financial support from Åforsk Foundation.

References

1. Sampaio, J. P. Rhodosporidium Banno (1967). *The Yeasts* **3,** 1523–1539 (2011).
2. Park, Y.-K., Nicaud, J.-M. & Ledesma-Amaro, R. The Engineering Potential of Rhodosporidium toruloides as a Workhorse for Biotechnological Applications. *Trends Biotechnol.* **36,** 304–317 (2018).
3. Soccol, C. R. *et al.* Pilot scale biodiesel production from microbial oil of Rhodosporidium toruloides DEBB 5533 using sugarcane juice: Performance in diesel engine and preliminary economic study. *Bioresour. Technol.* **223,** 259–268 (2017).
4. Buzzini, P. *et al.* Carotenoid profiles of yeasts belonging to the genera *Rhodotorula* , *Rhodosporidium* , *Sporobolomyces* , and *Sporidiobolus*. *Can. J. Microbiol.* **53,** 1024–1031 (2007).

5. Kerkhoven, E. J., Lahtvee, P.-J. & Nielsen, J. Applications of computational modeling in metabolic engineering of yeast. *FEMS Yeast Res.* **15,** 1–13 (2015).

6. Bommareddy, R. R., Sabra, W., Maheshwari, G. & Zeng, A.-P. Metabolic network analysis and experimental study of lipid production in Rhodosporidium toruloides grown on single and mixed substrates. *Microb. Cell Fact.* **14,** 36 (2015).

7. Castañeda, M. T., Nuñez, S., Garelli, F., Voget, C. & De Battista, H. Comprehensive analysis of a metabolic model for lipid production in Rhodosporidium toruloides. *J. Biotechnol.* **280,** 11–18 (2018).

8. Zhu, Z. *et al.* A multi-omic map of the lipid-producing yeast Rhodosporidium toruloides. *Nat. Commun.* **3,** 1112 (2012).

9. Vorapreeda, T., Thammarongtham, C., Cheevadhanarak, S. & Laoteng, K. Alternative routes of acetyl-CoA synthesis identified by comparative genomic analysis: involvement in the lipid production of oleaginous yeast and fungi. *Microbiology* **158,** 217–228 (2012).

10. Lee, N. K., Cheon, C. J. & Rhee, J.-K. Anti-Obesity Effect of Red Radish Coral Sprout Extract by Inhibited Triglyceride Accumulation in a Microbial Evaluation System and in High-Fat Diet-Induced Obese Mice. *J. Microbiol. Biotechnol.* **28,** 397–400 (2018).

11. Pan, J., Hu, C. & Yu, J.-H. Lipid Biosynthesis as an Antifungal Target. *J. fungi (Basel, Switzerland)* **4,** 72 (2018).

12. Wirth, F. & Goldani, L. Z. Epidemiology of rhodotorula: An emerging pathogen. *Interdiscip. Perspect. Infect. Dis.* **2012,** (2012).

13. Wang, C. & St. Leger, R. J. The Metarhizium anisopliae perilipin homolog MPL1 regulates lipid metabolism, appressorial turgor pressure, and virulence. *J. Biol. Chem.* **282,** 21110–21115 (2007).

14. Nguyen, L. N. *et al.* Autophagy-related lipase FgATG15 of Fusarium graminearum is important for lipid turnover and plant infection. *Fungal Genet. Biol.* **48,** 217–224 (2011).

15. Wang, H. *et al.* RAVEN 2.0: A versatile toolbox for metabolic network reconstruction and a case study on Streptomyces coelicolor. *PLOS Comput. Biol.* **14,** e1006541 (2018).

16. Sánchez, B. J., Li, F., Lu, H., Kerkhoven, E. J. & Nielsen, J. SysBioChalmers/yeast-GEM: yeast 8.2.0. *Zenodo* (2018). doi:10.5281/zenodo.1495483

17. Pryszcz, L. P., Huerta-Cepas, J. & Gabaldón, T. MetaPhOrs: Orthology and paralogy predictions from multiple phylogenetic evidence using a consistency-based confidence score. *Nucleic Acids Res.* **39,** (2011).

18. Prigent, S. *et al.* Meneco, a Topology-Based Gap-Filling Tool Applicable to Degraded Genome-Wide Metabolic Networks. *PLoS Comput. Biol.* **13,** 1–32 (2017).

19. Sánchez, B. J., Li, F., Kerkhoven, E. J. & Nielsen, J. SLIMEr: probing flexibility of lipid metabolism in yeast with an improved constraint-based modeling framework. *BMC Syst. Biol.* **13,** 4 (2019).

20. Shen, H. *et al.* Kinetics of continuous cultivation of the oleaginous yeast Rhodosporidium toruloides. *J. Biotechnol.* **168,** 85–89 (2013).

21. Camacho, C. *et al.* BLAST+: architecture and applications. *BMC Bioinformatics* **10,** 421 (2009).

22. Azambuja, S. P. H., Bonturi, N., Miranda, E. A. & Gombert, A. K. Physiology and lipid accumulation capacity of different Yarrowia lipolytica and Rhodosporidium toruloides strains on glycerol. *bioRxiv* 1–18 (2018). doi:10.1101/278523

23. Bonturi, N., Crucello, A., Viana, A. J. C. & Miranda, E. A. Microbial oil production in sugarcane bagasse hemicellulosic hydrolysate without nutrient supplementation by a Rhodosporidium toruloides adapted strain. *Process Biochem.* **57,** 16–25 (2017).

24. Zhang, S. *et al.* Engineering Rhodosporidium toruloides for increased lipid production. *Biotechnol. Bioeng.* **113,** 1056–1066 (2016).

25. Díaz, T. *et al.* Combining evolutionary and metabolic engineering in Rhodosporidium toruloides for

lipid production with non-detoxified wheat straw hydrolysates. *Appl. Microbiol. Biotechnol.* **102,** 3287–3300 (2018).

26. Loira, N., Dulermo, T., Nicaud, J.-M. & Sherman, D. J. A genome-scale metabolic model of the lipid-accumulating yeast Yarrowia lipolytica. *BMC Syst. Biol.* **6,** 35 (2012).

27. Ye, C. *et al.* Reconstruction and analysis of a genome-scale metabolic model of the oleaginous fungus Mortierella alpina. *BMC Syst. Biol.* **9,** 1 (2015).

28. Vongsangnak, W. *et al.* Genome-scale metabolic modeling of Mucor circinelloides and comparative analysis with other oleaginous species. *Gene* **583,** 121–129 (2016).

Tables

Table 1. Comparison of TAG production targets predicted from FSEOF simulation on glucose or xylose as carbon source.

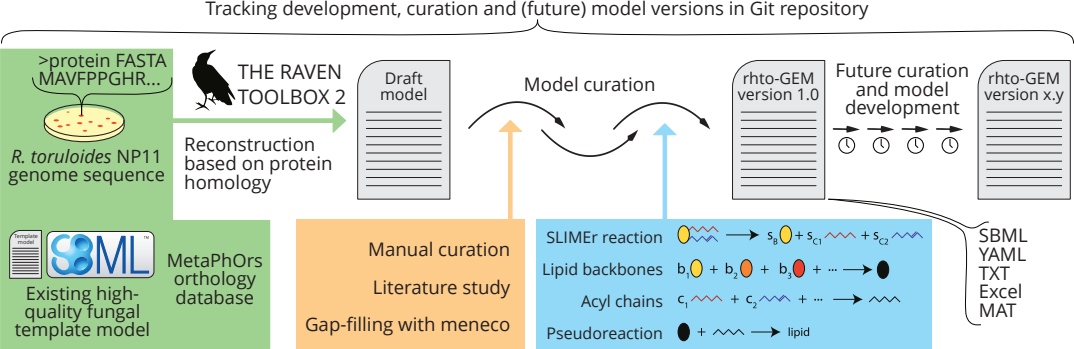
	triacylglycerol		linelonic acid			
rxn	**glucose**	**xylose**	**glucose**	**xylose**	**reaction name**	**gene association**
r_0112	30.1	25.1	10.0	8.4	acetyl-CoA synthetase	RHTO_08027
r_0148	23.7	18.7	7.0	5.3	adenylate kinase	RHTO_02749 or RHTO_03117 or RHTO_04701
r_0109	22.4	22.4	7.5	7.5	acetyl-CoA carboxylase, reaction	RHTO_02161 and RHTO_02004
r_0569	20.7	15.6	6.0	4.3	inorganic diphosphatase	RHTO_00177
r_1667	19.3	19.0	6.5	6.4	bicarbonate formation	RHTO_04980
r_0962	17.2	17.2	5.8	5.8	pyruvate kinase	RHTO_01610
r_0366	16.5	16.5	5.6	5.6	enolase	RHTO_00323
r_0893	16.5	16.5	5.6	5.6	phosphoglycerate mutase	RHTO_04793 or RHTO_07535 or RHTO_07773
r_0486	3.4	2.7	1.4	1.0	glyceraldehyde-3-phosphate dehydrogenase	RHTO_01292 or RHTO_03746
r_0892	3.4	2.7	1.4	1.0	phosphoglycerate kinase	RHTO_00033
r_0534	3.3		0.06		hexokinase (D-glucose:ATP)	RHTO_06072 or RHTO_06870
r_2141	2.9	2.9	1.0	1.0	fatty-acyl-CoA synthase (n-C18:0CoA)	RHTO_02032 or RHTO_02139
r_2337	1	1			1-acyl-sn-glycerol-3-phosphate acyltransferase (1-18:0, 2-18:1)	RHTO_05332 or RHTO_06718
r_2349	1	1			PA phosphatase (1-18:0, 2-18:1)	RHTO_04894
r_2389	1	1			diacylglycerol acyltransferase (1-18:0, 2-18:1, 3-18:0)	RHTO_00726 or RHTO_01962
r_2310	0.99	0.99			glycerol-3-phosphate acyltransferase (18:0)	RHTO_03058
r_2183	0.96	0.96	0.99	0.99	stearoyl-CoA desaturase (n-C18:0CoA -> n-C18:1CoA)	RHTO_03911
r_0491	0.91	0.91			glycerol-3-phosphate dehydrogenase (NAD)	RHTO_02273 or RHTO_08013
r_0173		28.4		8.5	aldehyde dehydrogenase (acetaldehyde, NADP)	RHTO_04310 or RHTO_04425 or RHTO_05838
r_0959		25.2		8.4	pyruvate decarboxylase	RHTO_00098
r_0091		9.8		3.8	6-phosphogluconolactonase	RHTO_07939
r_0466		9.8		3.8	glucose 6-phosphate dehydrogenase	RHTO_07853
r_0889		9.8		3.8	phosphogluconate dehydrogenase	RHTO_02788
r_0467		7.8		3.1	glucose-6-phosphate isomerase	RHTO_04058
r_1050		3.6		1.4	transketolase 2	RHTO_03248 or RHTO_03251
r_1049		3.3		1.3	transketolase 1	RHTO_03248 or RHTO_03251
r_1048		3.3		1.5	transaldolase	RHTO_06955
r_0982		2.9		1.1	ribose-5-phosphate isomerase	RHTO_06311
r_0887		0.02			phosphofructokinase (s7p)	RHTO_00494
r_0990		0.02			sedoheptulose 1,7-bisphosphate D-glyceraldehyde-3-phosphate-lyase	RHTO_03043
t_0042			1	1	oleoyl-CoA desaturase (n-C18:1CoA - n-C18:2CoA)	RHTO_03911
t_0943				1	fatty-acid--CoA ligase (linolenoyl-CoA)	RHTO_00058 or RHTO_04350

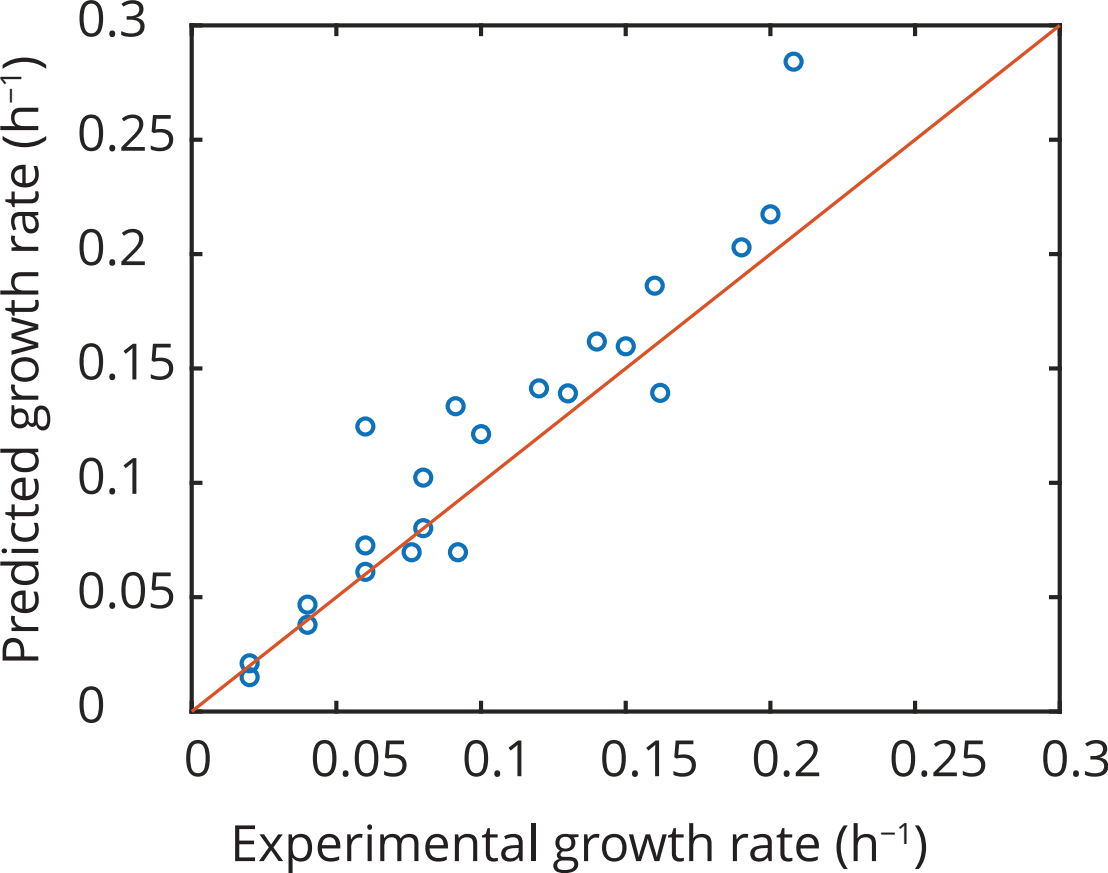
Indicated are slopes derived from FSEOF, indicative of whether gene expression should be increased or decreased to direct flux from growth towards production. Only reactions with gene associations are shown.

Figure legends

Fig. 1. Development and distribution of *rhto-GEM*. Schematic overview of the reconstruction pipeline, starting from generating a draft model using RAVEN, a protein FASTA of the *R. toruloides* genome and a template model. Through manual curation and inclusion of the SLIMEr formalism for lipid metabolism a functional model *rhto-GEM* is constructed, which is available in a number of different file formats. The whole development is tracked in a Git repository, which furthermore makes the model amenable to trackable future curation and development.

Fig. 2. Growth prediction from experimental data. Comparison of experimentally measured growth rates [6,20,22,23] and model predicted growth rates when constraint for the same carbon source uptake rates. Each data-point corresponds to one experimental growth measurement. The red line indicates perfect correlation.


Tracking development, curation and (future) model versions in Git repository
>protein FASTA
MAVFPPGHR...
R. toruloides NP11
genome sequence
THE RAVEN
TOOLBOX 2
Reconstruction
based on protein
homology
Draft
model
Model curation
rhto-GEM
version 1.0
Future curation
and model
development
rhto-GEM
version x.y
SBML
YAML
TXT
Excel
MAT
Template model
SBML
Existing high-
quality fungal
template model
MetaPhOrs
orthology
database
Manual curation
Literature study
Gap-filling with meneco
SLIMEr reaction
Lipid backbones
Acyl chains
Pseudoreaction
lipid


Predicted growth rate (h⁻¹)
0.3
0.25
0.2
0.15
0.1
0.05
0
0
0.05
0.1
0.15
0.2
0.25
0.3
Experimental growth rate (h⁻¹)