

# Xylose Metabolism and the Effect of Oxidative Stress on Lipid and Carotenoid Production in *Rhodotorula toruloides*: Insights for Future Biorefinery

Marina Julio Pinheiro<sup>1,2,‡</sup>, Nemailla Bonturi<sup>1,‡</sup>, Isma Belouah<sup>1</sup>, Everson Alves Miranda<sup>2</sup>, Petri-Jaan Lahtvee<sup>1,\*</sup>

<sup>1</sup>Institute of Technology, University of Tartu, Tartu, Estonia

<sup>2</sup>Department of Materials and Bioprocess Engineering, School of Chemical Engineering, State University of Campinas, Campinas, Brazil

<sup>‡</sup>Contributed equally

\* Correspondence:

Petri-Jaan Lahtvee  
petri.lahtvee@ut.ee

**Keywords:** Microbial oil, carotenoids, *Rhodotorula toruloides*, Genome-scale modelling, xylose, biorefinery, absolute proteomics

## Abstract

The use of cell factories to convert sugars from lignocellulosic biomass into chemicals in which oleochemicals and food additives, such as carotenoids, is essential for the shift towards sustainable processes. *Rhodotorula toruloides* is a yeast that naturally metabolises a wide range of substrates, including lignocellulosic hydrolysates, and converts them into lipids and carotenoids. In this study, xylose, the main component of hemicellulose, was used as the sole substrate for *R. toruloides*, and a detailed physiology characterisation combined with absolute proteomics and genome-scale metabolic models was carried out to understand the regulation of lipid and carotenoid production. To improve these productions, oxidative stress was induced by hydrogen peroxide and light irradiation and further enhanced by adaptive laboratory evolution. Based on the online measurements of growth and CO<sub>2</sub> excretion, three distinct growth phases were identified during batch cultivations. The intracellular flux estimations correlated well with the measured protein levels and demonstrated improved NADPH regeneration, phosphoketolase activity and reduced beta-oxidation, correlating with increasing lipid yields. Light irradiation resulted in 70% higher carotenoid and 40% higher lipid yields. The presence of hydrogen peroxide did not affect the carotenoid yield but culminated in the highest lipid yield of 0.65 g/g<sub>DCW</sub>. The adapted strain showed improved fitness and 2.3-fold higher carotenoid yield than the parental strain. This work presented a holistic view of xylose conversion into microbial oil and carotenoids by *R. toruloides* for further cost-effective and renewable production of these molecules.

## 1 Introduction

*Rhodotorula toruloides* is considered one of the most promising oleaginous yeasts for industrial applications. This microorganism is a natural producer of lipids (microbial oil) and high-value compounds, such as carotenoids and enzymes for pharma and chemical industries (L-phenylalanine ammonia-lyase and D-amino acid oxidase) (Park et al., 2017). The microbial oil, which is primarily composed of triacylglycerides (TAGs), is a potential raw material for oleochemicals that can be used

as biodiesel, cosmetics, and coatings as well as a replacement of vegetable oil in fish feed (Unrean et al., 2017; Blomqvist et al., 2018; Yang et al., 2018). Carotenoids are important molecules for different industries, such as the food, chemical, pharmaceutical and cosmetics industries. In addition to its colorants properties, carotenoids can be metabolised into vitamin A and have antioxidant activity that has been explored, for example, in the prevention of cancer, immune diseases and as skin protection against radiation (Stahl and Sies, 2007; Du et al., 2016; Kot et al., 2019). The global market for carotenoids should reach US\$2.0 billion by 2022 (BBC Research, 2018), while the global demand for fatty acids (FAs) and alcohols is expected to reach over 10 Mt in 2020 (Adrio, 2017).

In addition to the ability to produce a variety of relevant compounds, *R. toruloides* can consume a range of carbon and nitrogen sources (Park et al., 2017; Lopes et al., 2020a), including lignocellulosic hydrolysates (Bonturi et al., 2017; Lopes et al., 2020b). Following cellulose, hemicellulose is the second most abundant fraction of lignocellulose, and this fraction consists of polymerised five-carbon sugars, mainly xylose. The efficient utilisation of xylose by a microorganism is essential to improve the conversion of lignocellulosic materials into target compounds, thus increasing the economic viability of the biotechnological processes in biorefineries. Therefore, understanding of the metabolic mechanisms involved in the production of lipids and carotenoids from xylose by *R. toruloides* is crucial to further improve the titres, yields and rates of this bioprocess.

The cellular content of lipids and carotenoids is affected by several factors, including medium composition and cultivation conditions (Mata-Gómez et al., 2014). Previous studies have described the increase in carotenoid production in the presence of oxidative stress, such as hydrogen peroxide ( $H_2O_2$ ) and light irradiation. In the presence of  $H_2O_2$  (5 mmol/L), a 5-fold increase in carotenoid production by *Rhodotorula mucilaginosa* was observed (Irazusta et al., 2013), illustrating how optimisation of cultivation conditions can improve production yields of the desired metabolites. Under light irradiation (4,000 lux), the production of carotenoids and lipids by *Rhodotorula glutinis* increased 60% (Gong et al., 2020). The cellular response mechanism against oxidative stress is not clear for yeast, including *R. toruloides*. This condition is associated with the presence of reactive oxygen species (ROS), including  $H_2O_2$ , superoxide ( $O_2^-$ ) and hydroxyl radicals ( $HO\bullet$ ). ROS are potent oxidants that can damage all cellular components, including DNA, lipids and proteins. Microbial cells possess two defensive systems against oxidative damage: enzymatic and non-enzymatic. The former is mainly constituted by enzymes superoxide dismutase and catalase, and the latter is involved in direct scavenging of ROS or recycling of oxidised compounds, such as ascorbate, glutathione, alpha-tocopherol and carotenoids (Irazusta et al., 2013).

Integration of large-scale data sets is crucial for a better understanding of the genomic organisation and metabolic pathways in living cells. Complete genome sequences are available for several *R. toruloides* strains (Kumar et al., 2012; Zhu et al., 2012; Morin et al., 2014; Hu and Ji, 2016; Sambles et al., 2017; Coradetti et al., 2018). The lipid formation process during different growth phases of cultivation on glucose has been investigated through proteomic analysis (Liu et al., 2009) and compared with cells grown on xylose (Tiukova et al., 2019b). Multi-omics analyses have identified metabolism modification under nitrogen (Zhu et al., 2012) and phosphate limitation (Wang et al., 2018b). The latter studies have identified higher lipid accumulation under nitrogen or phosphate limitation, which has been correlated to higher activation in nitrogen recycling but also lipid degradation and autophagy. Carotenoid production from glycerol was investigated using global metabolomics, revealing reduced abundance of metabolites involved in TCA and amino acid biosynthesis (Lee et al., 2014).

Genome-scale metabolic models (GEMs) are another powerful tool to understand and provide a holistic view of metabolic fluxes, energy and redox metabolism or even suggest targets for metabolic engineering. GEMs are constructed based on the available genome sequence of a specific organism, thus providing a summary of the metabolic network (Kerkhoven et al., 2015). Regarding *R. toruloides*, a number of metabolic models are available to assess lipid production (Bommareddy et al., 2015; Castañeda et al., 2018), and two reports of GEMs are available (Dinh et al., 2019; Tiukova et al.,

2019a). Lopes et al. (2020a) reported the first study that combined data from cultivations of *R. toruloides* under different carbon sources with the GEM. The approach proved to be useful for understanding metabolic fluxes and identifying targets to improve lipid production either by metabolic engineering or process optimisation.

Therefore, the current study aimed at providing a holistic view of lipid and carotenoid production by *R. toruloides* using xylose as a sole carbon source by combining detailed physiological characterisation with the quantitative proteomics and GEM analysis. Oxidative stress ( $H_2O_2$  and light irradiation) and adaptive laboratory evolution were employed to improve lipid and carotenoids production. To our knowledge, this is the first work combining such approaches for this strain, and the data obtained here can be used to establish future bioprocesses in biorefineries.

## 2 Material and methods

### 2.1 Strain and inoculum

*Rhodotorula toruloides* CCT 7815 was obtained from “Coleção de Culturas Tropicais” (Fundação André Tosello, Campinas, Brazil) and stored at  $-80^{\circ}C$  in 10% (v/v) glycerol. This strain was derived from *R. toruloides* CCT 0783 (synonym IFO10076) after short-term adaptation in sugarcane bagasse hemicellulosic hydrolysate (Bonturi et al., 2017). The inoculum was prepared in YPD medium at 200 rpm and  $30^{\circ}C$  for 24 h. The cells were washed twice with 0.9% (v/v) NaCl before inoculation.

### 2.2 Adaptive laboratory evolution

Adaptive laboratory evolution (ALE) was performed by successive shake flask cultivations of *R. toruloides* in the presence of  $H_2O_2$  in rich medium (30.0 g xylose, 2.5 g glucose, 0.9 g yeast extract, 0.2 g  $(NH_4)_2SO_4$ , 1.5 g  $KH_2PO_4$  and 0.9 g  $MgSO_4 \cdot 7 H_2O$ ). ALE was performed in two cycles with the aim of improving the performance of this yeast under this oxidative environment. The initial  $H_2O_2$  concentration was 10 mmol/L (the highest concentration in which the cells grew in successive cultivations) in the first cycle and 20 mmol/L in the second cycle (the highest concentration that the parental strain tolerated). In each cycle, the cells were harvested at the exponential growth phase and transferred to a fresh medium with the same  $H_2O_2$  concentration. The initial OD (at 600 nm) of every passage was 0.5. The end of the cycle was determined by the stabilisation of the maximum specific growth rate and the length of lag phase.

### 2.3 Yeast cultivation

Batch cultivations were performed in 1-L bioreactors (Applikon Biotechnology, Delft, The Netherlands) with a working volume of 600 mL at pH 6.0 and controlled with the addition of 2 mol/L KOH. Dissolved oxygen was maintained at greater than 25% using 1-vvm airflow and stirring speeds between 400 and 600 rpm. The cultivation started with 1% (v/v) of overnight culture inoculum. Oxidative stress was induced by the addition of  $H_2O_2$  (20 mmol/L) at the start of the cultivation or by light irradiation (white LED light, 40,000 lux) throughout the experiment. The composition of  $CO_2$  and  $O_2$  in the gas outflow was measured using an online gas analyser (BlueSens GmbH, Herten, Germany), and biomass was monitored online using a Bug Lab BE3000 Biomass Monitor (Bug Lab, Concord, CA, USA) at 1,300 nm. Data were collected and processed with BioXpert V2 software v. 2.95 (Applikon Biotechnology, Delft, the Netherlands). All experiments were performed in triplicate.

The medium composition used in the bioreactor experiments was, per litre, 70.0 g xylose, 1.95 g  $(NH_4)_2SO_4$ , 3.0 g  $KH_2PO_4$ , 0.5 g  $MgSO_4 \cdot 7 H_2O$ , 1.0 mL vitamins solution and 1.0 mL trace metal solution (Lahtvee et al., 2017), supplemented with 100  $\mu$ l antifoam 204 (Sigma-Aldrich, MO, USA). Samples for biomass, carotenoid and extracellular metabolite analyses were collected every 24 h.

Samples were withdrawn from bioreactors, transferred into precooled 2-mL Eppendorf tubes and centrifuged for 20 s at 4°C and 18,000 g. The supernatant was collected and stored at -20°C for extracellular metabolite quantification. The biomass was snap-frozen in liquid nitrogen, stored at -80°C, and further used for proteomics analysis.

## 2.4 Quantification of biomass, extracellular metabolites, carotenoids, total lipids and proteins

Online turbidity measurements were calibrated by gravimetrically measuring the dry cellular weight (DCW) every 24 h. Extracellular metabolites in the broth were measured using HPLC (LC-2030C Plus, Shimadzu, Kyoto, Japan) equipped with a refractive index detector (RID-20A, Shimadzu, Kyoto, Japan). Concentrations of xylose, organic acids and glycerol were measured using an HPX-87H column (Bio-Rad, CA, USA) at 45°C, and 5 mmol/L H<sub>2</sub>SO<sub>4</sub> served as the mobile phase with isocratic elution at 0.6 mL/min. Xylitol and arabitol were quantified using Rezex RPM-Monosaccharide column (Phenomenex, CA, USA) at 85°C, and purified water (MilliQ Ultrapore Water System, Merck, Darmstadt, Germany) used as the mobile phase with isocratic elution at 0.6 mL/min.

For quantification of carotenoids (modified from Lee et al., 2014), 2 mL of cells were harvested by centrifugation, washed twice in distilled water and resuspended in 1.0 mL of acetone. The cells were lysed with acid-washed glass beads (400 – 650 µm) using the FastPrep homogeniser for three cycles (4 m/s for 20 s) (MP Biomedicals, CA, USA). After centrifugation at 15,000 g for 5 min, the acetone solution containing carotenoids was collected and stored at 4°C. These steps were repeated until the cell debris was colourless. Then, the solvent was evaporated in Concentrator Plus (Eppendorf, Hamburg, Germany), and the remaining extracts were resuspended in a known volume of acetone. Carotenoids were measured using Acquity UPLC (Waters, MA, USA) equipped with a TUV detector (Waters, MA, USA) and C18 column (BEH130, 1.7 µm, 2.1 x 100 mm, Waters, MA, USA). The mobile phase was a gradient from 80 to 100% of acetone in purified water at a flow rate of 0.2 mL/min. Detection was performed at 450 nm (modified from Weber et al., 2007). All identified peaks were quantified using the β-carotene standard (Alfa Aesar, MA, USA). Detected peaks were identified according to the known carotenoid retention time profile (Weber et al., 2007; Lee et al., 2014).

Lipids were extracted according to an adaptation of the Folch method (Folch et al., 1957) described by Bonturi et al. (2015). Briefly, a mixture of chloroform and methanol (2:1 v/v) was added to dried biomass. After 24 h, the solvent was evaporated in a rotary evaporator (Buchi, Switzerland), and the total lipid content was determined gravimetrically.

Total proteins were extracted from 600 µg of biomass resuspended in Y-PER solution (Thermo Fisher, MA, USA) in a 2-mL Eppendorf tube. This suspension was incubated at 30°C for 45 min. Then, glass beads were added in the tube, and cell lysis was performed in a FastPrep homogeniser during ten cycles (4 m/s for 20 s). Between the cycles, the tubes were placed on ice for 3 min. After centrifugation at 18,000 g and 4°C for 10 min, the supernatant was carefully removed and stored at 4°C for further protein quantification. Y-PER reagent was added to the remaining biomass in the tube, and the cell disruption process was repeated. This step was performed until no protein was detected. Total protein was quantified using a commercially available assay (Micro BCA™ Protein Assay Kit, Thermo Fisher), and samples were diluted in the linear range of BSA protein standard (0.5 to 20 µg/ml).

## 2.5 Proteome analysis

Fully labelled biomass to be used as the internal standard in absolute proteome analysis was produced by cultivating *R. toruloides* in minimal mineral medium containing labelled heavy 15N, 13C-lysine (Silantes, Munich, Germany). Heavy labelling of the proteinogenic lysine was measured at 96.6% (data not shown). Absolute proteomics and internal heavy-labelled standard preparation and analyses were performed similarly as described in Lahtvee et al. (2017) and Kumar and Lahtvee (2020). Briefly, cells

were resuspended in the lysis buffer (6 mol/L guanidine HCl, 100 mmol/L Tris-HCl pH 8.5 and 100 mmol/L dithiothreitol) and homogenised with glass beads using the FastPrep24 device. The supernatant was removed by centrifugation (17,000 g for 10 min at 4°C) and precipitated overnight with 10% trichloroacetic acid (TCA) at 4°C. Pellets from the previous precipitation step were spiked in 1:1 ratio with the heavy-labelled standard. This mixture was further precipitated with 10% TCA. The pellet was resuspended in a buffer containing 7 mol/L urea and 2 mol/L thiourea in 100 mmol/L ammonium bicarbonate (ABC) followed by reduction using 5 mmol/L DTT and alkylation with 10 mmol/L chloroacetamide. Peptides were digested at room temperature with *Achromobacter lyticus* Lys-C (Wako Pure Chemical Industries, Osaka, Japan) for 4 h at the ratio of 1:50 (enzyme:protein) followed by overnight digestion of the previous solution diluted 5 times in 100 mmol/L ABC buffer. Peptides were desalted using in-house prepared C18 (3M Empore, Maplewood, MO, USA) tips and were reconstituted in 0.5% trifluoroacetic acid (TFA). For separation, 2 µg of peptides was injected on an Ultimate 3000 RSLCnano system (Dionex, Sunnyvale, CA, USA) coupled to a C18 cartridge trap-column in a backflush configuration and an analytical 50 cm x 75 µm emitter-column (New Objective, MA, USA) in-house packed with 3 µm C18 particles (Dr. Maisch, Ammerbuch, Germany). Eluted peptides were sprayed to a quadrupole-orbitrap Q Exactive Plus (Thermo Fisher Scientific, MA, USA) tandem mass spectrometer. MaxQuant 1.4.0.8 software package was used for raw data identification and identification (Cox and Mann, 2008). *R. toruloides* NP11 served as a reference proteome database in UniProt ([www.uniprot.org](http://www.uniprot.org)). Protein quantification was performed following the total protein approach described in (Sánchez et al., 2020) and assuming 90% coverage from the total protein abundance.

LC-MS/MS proteomics data were deposited in the ProteomeXchange Consortium (<http://proteomecentral.proteomexchange.org>) via the PRIDE partner repository (Vizcaino et al., 2013) and can be retrieved using the dataset identifier PRIDE: PXD019305. Processed quantitative data are presented in Supplementary Table S7. Triplicated quantitative proteomics data were used for differential expression analysis. P-values were adjusted for multiple testing using the Benjamini-Hochberg procedure (Benjamini and Hockberg, 1995). Additional data analysis included gene set enrichment analysis (carried out using PIANO platform; Våremo et al., 2013) and gene enrichment analysis (g:Profiler; Raudvere et al., 2019).

## 2.6 Genome-scale modelling

The intracellular flux patterns were predicted using version 1.2.1. of the *R. toruloides* metabolic network at the genome scale (Tiukova et al., 2019a; Lopes et al., 2020a) and rhto-GEM (<https://github.com/SysBioChalmers/rhto-GEM/releases/>). The model was improved by adding carotenoids ( $\beta$ -carotene,  $\gamma$ -carotene, torulene and torularhodin) into the biomass reaction. Flux balance analysis was performed to calculate internal fluxes using RAVEN Toolbox (Wang et al., 2018a) on MATLAB (The MathWorks Inc., MA, USA), Gurobi solver (Gurobi Optimization Inc., TX, USA) and by optimising for non-growth related ATP maintenance (NGAM). The latter was followed by flux variability analysis (random sampling at  $n = 5000$ ) at 95% from the maximal NGAM value. Experimental data obtained from the yeast cultivations were used to constrain the model if not stated otherwise. Biomass composition was adjusted to the measured total protein, lipid and carotenoid content.

## 3 Results

### 3.1 Three distinct phases of *R. toruloides* growth on xylose

*R. toruloides* growth was characterised with xylose as a sole carbon source under aerobic batch conditions on a mineral medium. The initial xylose concentration of 70 g/L was chosen, and the amount

of nitrogen was adjusted to result in a C/N ratio of 80 mol/mol. On-line monitoring of culture turbidity, CO<sub>2</sub> production and O<sub>2</sub> consumption was complemented by off-line analysis of sugars, alcohols and biomass composition (total lipids, proteins and carotenoids; Figure 1). During the batch cultivation of *R. toruloides*, three distinct growth phases were observed based on growth dynamics and substrate consumption patterns. In the first growth phase (P1), cells were growing exponentially without any observable limitation, and xylose used as the sole carbon source. Arabitol, xylitol and CO<sub>2</sub> were the main fermentation by-products detected. Phase two (P2) started with a sudden decrease in the specific growth rate due to nitrogen limitation. At that point, approximately 23 g/L of xylose was consumed, indicating a critical C/N ratio of 26 for *R. toruloides* to reach nitrogen limitation. P2 lasted until the depletion of the primary carbon source, namely, xylose. Consumption of arabitol and xylitol under nitrogen limitation defined the third growth phase (P3). To our knowledge, no previous study has provided a characterisation of *R. toruloides* physiology on xylose in such detail. These phases were further analysed in this work, aiming to identify the cellular metabolic behaviour in response to the environment changes during the batch growth.

P1 comprised the highest specific growth rate ( $\mu$ ,  $0.060 \pm 0.001 \text{ h}^{-1}$ ) and the specific xylose uptake rate ( $r_{\text{XYL}}$ ,  $1.74 \pm 0.12 \text{ mmol/g}_{\text{DCW}}\cdot\text{h}$ ) (Supplementary Table S1), while no nutrient-level limitations was detected. During the exponential growth, approximately one-third of the consumed carbon was secreted as arabitol and xylitol. An additional one-third of the consumed carbon was secreted as CO<sub>2</sub>. Although the carotenoid yield in P1 was low ( $0.66 \pm 0.06 \text{ mg/g}_{\text{DCW}}$ ), the specific production rate ( $r_{\text{CAR}}$ ) was the highest ( $0.042 \pm 0.003 \text{ mg/g}_{\text{DCW}}\cdot\text{h}$ ) due to the higher  $\mu$  (Figure 2A, Supplementary Table S1). The biomass and lipid yields were  $0.25 \pm 0.02 \text{ g}_{\text{DCW}}/\text{g}$  and  $0.18 \pm 0.03 \text{ g/g}_{\text{DCW}}$ , respectively. Carbon balance in this phase was estimated at 93%, indicating a small amount of undetected by-products.

Nitrogen limitation under the xylose-excess conditions marked the start of P2. The limitation led to a decrease in  $\mu$  (average value  $0.020 \pm 0.002 \text{ h}^{-1}$ ) and  $r_{\text{XYL}}$  (average value  $0.41 \pm 0.06 \text{ mmol/g}_{\text{DCW}}\cdot\text{h}$ ). Compared to P1, the carotenoid yield did not show a significant difference; however,  $r_{\text{CAR}}$  decreased 4-fold compared with P1 (Supplementary Table S1). As expected, lipid accumulation doubled compared to P1 due to the positive influence of nitrogen limitation, reaching  $0.38 \pm 0.05 \text{ g/g}_{\text{DCW}}$  (Figure 2B). P3 started after the depletion of xylose. Here, xylitol and arabitol were simultaneously consumed by the cells. The average specific growth rate was the lowest of the three growth phases ( $0.005 \pm 0.0003 \text{ h}^{-1}$ ). The highest accumulation of carotenoids per biomass was detected in this phase, increasing substantially to  $1.87 \pm 0.21 \text{ mg/g}_{\text{DCW}}$ ; however, the  $r_{\text{CAR}}$  was the lowest of all phases (Figure 2 and Supplementary Table S1). The lipid yield remained at the same level as that noted in P2. In addition, 50% carbon loss (undetected carbon) was observed in P3, which can be partially explained by the technical uncertainty of measurements for off-gas at the very low growth rate conditions in the P3. Total carotenoid and lipid yields for the whole batch cultivation were  $0.85 \pm 0.01 \text{ mg/g}_{\text{DCW}}$  and  $0.33 \pm 0.07 \text{ g/g}_{\text{DCW}}$ , respectively (Table 1, column REF).

### 3.2 Understanding intracellular flux patterns among three growth phases

To understand the changes in intracellular flux patterns between the observed growth phases, simulations using *R. toruloides* GEM were performed (Supplementary Tables S2; Tiukova et al., 2019a; Lopes et al., 2020a). GEM is a mathematical reconstruction of the metabolic network based on genome annotations and information, such as gene-protein-reaction relationships. GEM in combination with flux balance analysis (FBA) allows simulation of intracellular flux patterns. In addition to intracellular rates, FBA can be used to predict potential by-products, which were not identified experimentally. Given that 7 and 5% of carbon was missing in P1 and P2, respectively, the model predicted that glyoxylate and acetaldehyde as the most likely undetected products secreted in P1 and P2, respectively. For the flux distribution comparison between the observed growth phases, fluxes were normalised by the total carbon uptake rate (Supplementary Tables S3).

NADPH supply is crucial for lipid production given that every elongation step of fatty acid (FA) synthesis requires the oxidation of two NADPH (Wasylenko et al., 2015). In addition to FA synthesis, xylose and L-arabitol utilization also require NADPH. For all three growth phases, the highest NADPH usage was observed for substrate uptake: xylose reductase (r\_1093) in P1 and P2 and L-xylulose reductase (t\_0882) in P3 (Supplementary Figure S1, Table S5). During xylose metabolism (P1 and P2), arabitol production via L-xylulose reductase partially regenerated the oxidised NADPH. Once xylose was exhausted (P3), NADPH was required for arabitol catabolism (Supplementary Figure S1). Our simulations noted that the oxidative branch of the pentose phosphate pathway (PPP), namely, glucose 6-phosphate dehydrogenase (r\_0466) and phosphogluconate dehydrogenase (r\_0889), was responsible for 83, 87 and 96% of NADPH regeneration in P1, P2 and P3, respectively. Although NADPH demand in the substrate consumption and amino acid biosynthesis pathways decreased in P2 and P3 compared to P1, fluxes in lipid biosynthesis increased 1.4-fold.

At the xylulose-5P branch point, 91% of carbon entered into the central carbon metabolism via traketolase (r\_1049, r\_1050) in P1 (Supplementary Figure S1). The remaining 9% was converted into glyceraldehyde-3-phosphate and acetyl-phosphate by the phosphoketolase reaction (t\_0081). Under nitrogen limitation, the activity of phosphoketolase was approximately tripled compared to P1. The phosphoketolase pathway further generates acetyl-CoA, a precursor of FA synthesis, by phosphate transacetylase (t\_0082) without losing a carbon compared to the pathway originating from pyruvate.

At the pyruvate branch point, on average, ca. 70% of the pyruvate produced in the cytosol was transported to mitochondria (r\_1138 and r\_2034) to be converted by pyruvate dehydrogenase (r\_0961) into acetyl-CoA, which is used in the tricarboxylic acid (TCA) cycle by citrate synthase (r\_0300). The remaining cytosolic pyruvate was either converted to cytosolic oxaloacetate (pyruvate carboxylase, r\_0958) or into acetyl-CoA by three enzymatic steps (pyruvate decarboxylase, r\_0959; acetaldehyde dehydrogenase, r\_2116; acetyl-CoA synthase, r\_0112). Acetyl-CoA can also be produced from citrate by ATP citrate lyase (y200003) or from acetyl-P by phosphate transacetylase. Under excess nitrogen conditions (P1), acetyl-CoA synthase was responsible for 60% of the flux. However, under nitrogen limitation (P2 and P3), ca. 71% of acetyl-CoA originated via phosphate transacetylase (Supplementary Figure S1).

Our physiological data showed that the lipid yield was higher under nitrogen-limitation phases (P2 and P3), and the carotenoids yield was higher in P3. These results can be explained by the higher predicted fluxes through reactions involving phosphoketolase, FA and carotenoids synthesis, and NADPH regeneration.

### 3.3 Impact of oxidative stress via light irradiation or the presence of H<sub>2</sub>O<sub>2</sub>

Further, we were interested in how oxidative stress created by either 20 mmol/L H<sub>2</sub>O<sub>2</sub> or light irradiation (40,000 lux white light) affects cellular growth and lipid and carotenoid accumulation. Cultivation of *R. toruloides* under these oxidative stress conditions presented the same three growth phases described for the reference condition (REF) and a similar growth and substrate consumption profile (Supplementary Figure S2 A, B). The specific growth rate differences compared to the reference condition were insignificant under the light irradiation (LIG) condition, but a significant 50% decrease was observed under the H<sub>2</sub>O<sub>2</sub> stress (PER) in P1 (Supplementary Table S1). The most significant difference was detected in the longer lag phase (approximately 90 h) shown in the PER condition (Supplementary Figure S2B). Moreover, the accumulation profiles of carotenoids and lipids showed altered behaviour compared to the reference (Figure 2). The stress caused by H<sub>2</sub>O<sub>2</sub> negatively affected carotenoid production in P1. However, in P3, the total carotenoid yield was the highest of all conditions ( $4.72 \pm 0.47$  mg/g<sub>DCW</sub>); thus, the highest  $r_{CAR}$  ( $0.026 \pm 0.012$  mg/g<sub>DCW</sub>.h) was reached in the third phase (Figure 2A, Supplementary Table S1). Lipid production exhibited a different behaviour, presenting a higher yield in P1 and P2 compared to the reference condition. The specific lipid production rate ( $r_{LIP}$ )

in P2 was the highest of all conditions in all phases ( $0.020 \pm 0.004$  g/g<sub>DCW</sub>.h), albeit the specific growth rate was not amongst the highest obtained in this study.

Regarding the overall results, for the entire cultivation under light irradiation, the cells presented 70% increased carotenoid yield ( $1.45 \pm 0.14$  mg/g<sub>DCW</sub>) and 40% increased lipid yield ( $0.46 \pm 0.12$  g/g<sub>DCW</sub>) compared to the reference condition. H<sub>2</sub>O<sub>2</sub> stress did not affect carotenoid production by the parental strain, and the production yield was maintained at  $0.89 \pm 0.04$  mg/g<sub>DCW</sub>. Surprisingly, this condition showed the highest lipid yield ( $0.65 \pm 0.06$  g/g<sub>DCW</sub>), which was increased by two-fold compared with the reference condition (Table 1). The achieved lipid yield was only slightly lower than the highest lipid yield reported in the literature for *R. toruloides*; specifically, 0.68 g/g<sub>DCW</sub> has been reported in fed-batch cultivation on a rich glucose-based medium (Li et al., 2007).

### 3.4 Increasing carotenoid accumulation via adaptive laboratory evolution

Adaptive laboratory evolution (ALE) is a strategy to improve the fitness of the microorganism in a challenging environment. H<sub>2</sub>O<sub>2</sub> stress improved lipid accumulation but increased the lag phase and lowered the growth rate. Therefore, ALE was performed to improve those parameters. The first cycle of ALE started with 10 mmol/L of H<sub>2</sub>O<sub>2</sub>. After 16 passages (ca. 30 generations), the lag phase decreased 11-fold (from 46 h to 4 h), and  $\mu_{\max}$  stabilised at  $0.045 \pm 0.003$  h<sup>-1</sup>, resulting in a 66% increase compared to the parental strain under the same conditions (Supplementary Figure S3A). Once the  $\mu_{\max}$  plateaued, the second cycle of ALE was started by increasing the selective pressure to 20 mmol/L of H<sub>2</sub>O<sub>2</sub> in the medium (Supplementary Figure S3B). As a result, after 15 passages (ca. 20 generations), the lag phase of yeast growth decreased from 30 to 5 h, and the  $\mu_{\max}$  improved to  $0.055 \pm 0.001$  h<sup>-1</sup>, representing a 22% increase compared to the first cycle. Although  $\mu_{\max}$  did not improve remarkably during the second cycle of ALE, the cells presented stronger pink colouration compared to the parental strain, indicating increased carotenoid accumulation. Therefore, the ALE experiment was halted, and the strains were characterised.

The adapted strain in presence of H<sub>2</sub>O<sub>2</sub> (20 mmol/L) showed a remarkable ca. 3-fold increase in carotenoid yield and titre, respectively, compared to the parental strain under the same condition during the initial screening experiments (Supplementary Table S6). Therefore, the adapted strain was further characterised under controlled environmental conditions in bioreactors and studied on the proteomics level.

### 3.5 Adapted strain under H<sub>2</sub>O<sub>2</sub> stress

Under H<sub>2</sub>O<sub>2</sub> stress, the adapted strain (ADA) exhibited a 70-hour shorter lag phase compared with the parental strain (PER) (Supplementary Figure S2C). Aeration and agitation in the bioreactor may increase the oxidative stress, which could explain the longer lag phase compared to initial screening experiments in shake flasks (mentioned above).

Some differences were noticed between the ADA and other conditions. Although the  $\mu$  in P1 did not show a significant difference, the  $r_{\text{CAR}}$  was increased by 4-fold in ADA compared with PER (Supplementary Table S1). However, in P2, the  $\mu$  in ADA was 2-fold reduced compared with the other conditions, but  $r_{\text{CAR}}$  remained 20% increased. The lower  $r_{\text{XYL}}$  in P2 could be related to the lower production of xylitol and arabitol in ADA compared to the other conditions probably due to a softer redox imbalance during xylose catabolism. The ADA showed ca. 2-fold increased yield of carotenoids and lipids under nitrogen-limiting phases (P2 and P3) compared to REF. In P3, the lipid yield was  $1.0 \pm 0.03$  g/g<sub>DCW</sub>, indicating that the gain of cell mass noted during this phase was mainly related to lipid accumulation.

The whole batch growth of ADA exhibited a 2.3-fold increase in carotenoid yield compared to PER; however, the lipid yield did not show a significant difference (Table 1).



### 3.6 Composition of carotenoids in biomass

*R. toruloides* mainly produces four carotenoids:  $\gamma$ -carotene,  $\beta$ -carotene, torulene and torularhodin (Mata-Gómez et al., 2014). The carotenoid profile was very similar in all the studied conditions. The  $\beta$ -carotene fraction decreased over time, whereas the opposite was observed for torulene (Supplementary Figure S3). Torularhodin was the most abundant fraction of carotenoids under all studied conditions. Growth of the parental strain under light irradiation and adapted strain cultivations showed torularhodin fraction higher than 50% during P2, which can be related to a stronger antioxidative property of this carotenoid, attributed to the presence of more double bonds in its chemical structure (Kot et al., 2018).

### 3.7 Proteomic results revealed the highest difference between nitrogen excess and -limiting conditions

Total protein measurements were combined with the absolute proteome analysis for the most relevant conditions. Therefore, samples from all three different growth phases from reference cultivation were analysed together with P2 of light-induced oxidative stress (LIG P2) and P1 and P3 of H<sub>2</sub>O<sub>2</sub>-induced oxidative stress for the adapted strain (ADA P1 and P3, respectively). At the latter three sampling points, the production of carotenoids and lipids was increased compared to the REF. One-third increased total protein content was measured for the nitrogen-excess condition during the P1 of a reference culture. All other conditions showed no significant differences in protein content with an average of 0.14 g/g<sub>DCW</sub> (Figure 4A). In differential expression analysis, proteome data were normalised to a constant protein mass, representing allocation differences for the individual proteins.

On average, more than 3,000 individual proteins were quantified under every condition studied (Supplementary Table S7). Principal component (PC) analysis clearly identified the biggest differences in the data set, which were determined by the switch into nitrogen limitation as indicated by the clear separation of the samples on the first PC, characterising 39% of the changes (Figure 4B). The second PC separated samples based on the use of the adapted strain under the oxidative stress environment (17% of the difference in the data). Altogether, 1,518 proteins showed significant (adj. p-value < 0.01) allocation changes under at least one of the environmental conditions (Figure 4C). Interestingly, no significant allocation (nor abundance) change was detected between the samples collected from reference or light-induced stress conditions under nitrogen limitation; however, significant differences in biomass composition were detected.

### 3.8 Translation and NADPH metabolism were most affected under nitrogen limitation

To understand the main differences in the dataset, gene set enrichment analysis (GSA) was used to identify classes of proteins that are significantly over-represented among the measured proteins and may have an association with a specific phenotype. A variation of GSA-based analysis was conducted. First, protein-Gene Ontology (GO) group relations were received from UniProt database. Second, protein-subsystem relationships were obtained from rhtoGEM, providing more specific information on various metabolic pathways present in *R. toruloides*. Given that PCA divided samples into four separate quadrants based on nitrogen availability and oxidative stress, we focussed on the comparison of these sample clusters throughout the study. Using UniProt-provided GO groups in GSA, 14 groups exhibited significant over-representation with an adj. p-value < 0.01 under the studied conditions (Supplementary Table S8). Most of these groups were related to protein translation, which were downregulated under nitrogen limitation and correlated with the lower specific growth rates under these conditions. A clear correlation between ribosome abundances and specific growth rate has been demonstrated previously for other organisms (Scott et al., 2014; Metzl-raz et al., 2017). When subsystems from rhtoGEM were

considered, significant upregulation was detected among carbon metabolism and its subgroups (glycolysis, gluconeogenesis, TCA cycle, glyoxylate and dicarboxylate metabolism; Supplementary Table S9). Interestingly, only the parental strain (under nitrogen-limiting conditions) showed overexpression in fatty acid degradation pathways and downregulation in amino acid biosynthesis pathways (with the exception of the tryptophan pathway, which was upregulated). Differences in the regulation of fatty acid degradation pathways could be responsible for the significantly increased lipid accumulation under the oxidative stress condition.

Based on the enzyme-metabolite relationships present in the rhtoGEM, reporter metabolites were analysed as the third variation of GSA, illustrating metabolites showing significant alterations among enzymes they interact with (Supplementary Table S10). When samples under nitrogen-limitation were compared to samples cultured under excess nitrogen, the most significant upregulation was detected among proteins in proximity to NAD<sup>+</sup>/NADH (Figure 4D). More than 65% of the proteins associated with NAD<sup>+</sup>/NADH showed increased allocation under nitrogen limitation (REF P2 and P3, LIG P2, ADA P3). In contrast, protein allocation decreased significantly for proteins related to NADP<sup>+</sup>/NADPH metabolism. Downregulation was predominantly noted in amino acid biosynthesis pathways, while NADPH consumption in lipid metabolism and the glutamate production pathway showed upregulation. Upregulation of enzymes in glutamate biosynthesis in response to nitrogen starvation have been demonstrated previously (Zhu et al., 2012; Tiukova et al., 2019b). Additionally, metabolites related to lipid synthesis (CoA, acetyl-CoA, acetate, and pyruvate) showed increased protein allocation during nitrogen limitation, while xylitol related proteins were downregulated (in all cases adj. p-val < 0.01).

### 3.9 Protein changes under the reference condition are well correlated with the simulated fluxes

In response to nitrogen limitation, proteins involved in central nitrogen metabolism, such as glutamate dehydrogenase (GDH, RHTO\_04650, RHTO\_07718) and glutamine synthetase (GLN, RHTO\_00673, RHTO\_00401) were upregulated (Figure 3A). This response has been previously reported under nitrogen limitation for *R. toruloides* grown in both glucose and xylose (Zhu et al., 2012; Tiukova et al., 2019b). Additionally, activation of autophagy process has been described as a direct response via TOR activation to recycle nitrogenous compounds (Zhu et al., 2012; Tiukova et al., 2019b). Although initially expressed at a low level, upregulation of autophagy-related proteins (RHTO\_05541, RHTO\_06526) was detected.

Proteins related to oxidative stress response showed upregulation for P2 and P3 (nitrogen limitation) compared to P1 (nitrogen excess). Catalase (CAT, RHTO\_01370), which breaks down hydrogen peroxide in the peroxisomal matrix, was the most upregulated protein with a 6-fold increase under nitrogen limitation. Recent reports in oleaginous microorganisms showed that ROS is an important signalling molecules in response to various stresses (Shi et al., 2017). Nitrogen depletion is an example of such stress, leading to the accumulation of ROS (Liu et al., 2012; Fan et al., 2014; Chokshi et al., 2017) and higher activities of catalase and other antioxidant enzymes, suggesting that lipid accumulation under nitrogen depletion is mediated by oxidative stress (Yilancioglu et al., 2014).

The highest carbon fluxes detected with GEM analysis were further assessed to understand the level of their regulation. Xylose uptake decreased under the nitrogen limitation in P2 and was lacking in P3, which was also reflected in downregulation of proteins involved in xylose assimilation. Proteins belonging to arabinol metabolism, such as L-xylulose reductase (LXR, RHTO\_00373) and D-arabinitol dehydrogenase (AD, RHTO\_07844), were also downregulated in P2 and P3.

Approximately 4-fold increased transketolase (TKT, RHTO\_03248) abundance compared to phosphoketolase (PK, RHTO\_04463) was consistent with the simulated increased flux through the transketolase reaction. However, PK levels increased more than 2-fold under nitrogen limitation, which was consistent with the increased flux levels under the mentioned conditions. Furthermore, the magnitude of the PK increase in this condition was 50-60% higher than TKT (Figure 3B). Carbon was

mainly channelled via TKT because it leans towards glycolysis and the oxidative branch of the PPP, which have been identified as the preferred pathway to regenerate NADPH (Lopes et al., 2020a, this study). GEM simulations revealed that xylose reductase (XR), fatty acid synthase and glutamate dehydrogenase consume most of the NADPH, which was regenerated by 6-phosphogluconate dehydrogenase (GND1), glucose-6-phosphate dehydrogenase (ZWF1), and L-xylulose reductase (LXR) (Figure 3A). With the exception of fatty acid synthase, these enzymes were also among the most abundant NADPH-dependent enzymes quantified.

In the oleaginous microorganism, cytosolic ATP-citrate lyase (ACL, RHTO\_03915) is considered an important enzyme as a source of acetyl-CoA (Ratledge and Wynn, 2002; Koutinas and Papanikolaou, 2011). Under nitrogen limitation, this enzyme was found in higher levels compared to P1 (4- and 2-fold increase for P2 and P3, respectively). Cytosolic acetyl-CoA can also be supplied by acetyl-CoA synthase (ACS1, RHTO\_08027), which was upregulated 2-fold in both P2 and P3, and by xylose metabolism via PK and phosphate acetyltransferase. However, FBA predicted that the majority of cytosolic acetyl-CoA originated from pyruvate via ACS1 and from PK and phosphate transacetylase (Figure 3B).

Acetyl-CoA can be transformed into malonyl-CoA, a substrate for FA synthesis, by the enzyme acetyl-CoA carboxylase (ACC1, RHTO\_02004), which was 3.4- and 1.6-fold upregulated in P2 and P3, respectively. Then, FAs are produced by fatty acid synthases FAS1 (RHTO\_02032) and FAS2 (RHTO\_02139), which were significantly upregulated in the nitrogen-limiting phases. Such findings were in accordance with the predicted higher fluxes from acetyl-CoA towards FA synthesis in P2 and P3. Additionally, acetyl-CoA can enter the mevalonate pathway (MEV) through acetyl-CoA C-acetyltransferase (ERG10, RHTO\_02048) to produce sterols and carotenoids. Despite the highest carotenoid content in P3, ERG10 and other proteins related to the MEV pathway, hydroxymethylglutaryl-CoA synthase (ERG13, RHTO\_02305) and reductase (HMG1, RHTO\_04045) were downregulated in P3. Proteins directly related to the carotenogenesis pathway were either not detected or did not show any significant difference. FBA predicted very low fluxes throughout carotenoid production.

### 3.10 Comparative proteomics at different growth phases under oxidative stresses

Although P2 of cultivation under light exposure (LIG\_P2) showed higher carotenoid production rates than REF\_P2, only two proteins (RHTO\_06480 and RHTO\_01160) were differentially expressed between those conditions (adj. p-value < 0.01). The functions of these proteins do not explain the higher production of carotenoids and lipids, suggesting that carotenoid production is regulated post-translationally.

Under oxidative stress, the adapted strain showed many similar changes to the reference condition while entering into nitrogen limitation but was also clearly differentiated as described by PC analysis (Figure 4B). Clustering of the significantly differentially expressed proteins was performed to understand the main changes compared to REF. Cluster showing significant differences among ADA and other conditions was enriched by the proteins from amino acid biosynthesis and the TCA cycle. Although TCA cycle was already upregulated under nitrogen-limiting conditions, more pronounced upregulation was detected in ADA, including the upregulation of citrate synthase (CIT1, RHTO\_06406), malate dehydrogenase (MDH, RHTO\_04363), and NADP<sup>+</sup>-dependent isocitrate dehydrogenase (IDH, RHTO\_04315). NADPH regeneration by GND1 (RHTO\_02788) in PPP was also upregulated. Similar to the REF under nitrogen limitation, FA biosynthesis was upregulated; however,  $\beta$ -oxidation, which is responsible for lipid degradation, remained lower and could at least partially explain the higher lipid accumulation under oxidative stress. Additionally, the carotenogenesis pathway was more activated since ERG10 and phytoene dehydrogenase (CRTI, RHTO\_04602) were 2- and 4-fold upregulated, respectively (Figure 5A).

As expected, enzymes involved in the oxidative stress response were upregulated in the ADA compared to REF; CAT exhibits a 16- and 5-fold increase for P1 and P3, respectively. For P3, other oxidative stress related proteins, such as S-(hydroxymethyl) glutathione dehydrogenase (RHTO\_03559) and glutathione S-transferase (RHTO\_00710), were also significantly upregulated (Supplementary Table S7).

### 3.11 Translation processes also play a crucial role in proteome allocation analysis

Absolute protein analysis allows comparisons of protein abundance levels between conditions and quantification of condition-dependent protein allocation patterns (Figure 5B). The top 100 of the most abundant proteins represented greater than 50% of the total proteome. Ribosomal proteins were the largest protein group, forming almost one-third of the total proteome in REF P1. However, these levels decreased to approximately 10% under the lower growth rate conditions in P3. For ADA, ribosomal protein allocation was already reduced to a lower level and decreased further, forming only 21% under the nutrient-limiting conditions in P3. As ribosomes are essential for achieving faster cell growth, the trade-off between allocation towards ribosomal proteins or energy generation pathways has been demonstrated previously (Nilsson and Nielsen, 2016; Sánchez et al., 2017, Kumar and Lahtvee, 2020). Interestingly, glycolysis was increased to the same extent under nitrogen-limiting conditions in the presence and absence of oxidative stress, but mitochondria and amino acid metabolism demonstrated significantly increased allocation for the ADA. The latter changes seemed to be responsible for the more efficient metabolism, providing higher biomass yields for adapted cells (Figure 5B and Table 1).

## 4 Discussion

Efficient microbial production of chemicals from sustainable resources is essential for the transition towards bioeconomy. *R. toruloides* has been considered as a potential microorganism to produce high-value products from biological resources, including hemicellulosic material, mainly composed of xylose. However, xylose metabolism in *R. toruloides* is still not completely understood. Only few studies have focussed on the metabolism of xylose assimilation in this oleaginous yeast (Jagtap and Rao, 2018; Tiukova et al., 2019b; Lopes et al., 2020a). Therefore, in our study, detailed physiology characterisation was combined with genome-scale modelling and absolute proteomics with the goal to investigate xylose metabolism in *R. toruloides* and use oxidative stress as a strategy to improve the production of lipids and carotenoids.

In this study, we demonstrated how *R. toruloides* growth on xylose exhibited three distinct phases, where most metabolic changes occurred after the transition into nitrogen limitation. Approximately 30% of consumed xylose by the parental strain accumulated into xylitol and arabitol during the first two growth phases probably to balance NADPH required for the growth or due to limitations in the abundance of xylulokinase. Although Jagtap and Rao (2018) reported that xylose conversion into D-arabitol by arabitol dehydrogenase regenerated one NAD<sup>+</sup>, the UniProt protein database for *R. toruloides* contains only sequences for LXR and for AD, lacking the levogyrous version of the latter. According to Fernandes and Murray (2010), fungi can produce both L- and D-arabitol. However, our proteomics analysis was able to detect only the enzymes for L-arabitol production, and the corresponding pathway was also used in our modelling approach (Supplementary Figure S1). Xylulokinase, an enzyme responsible for the last step in xylose metabolism, was not detected in our proteomic analysis, suggesting a low expression level that could limit xylose catabolism.

Under nitrogen-limiting conditions, increased flux via phosphoketolase was observed, corresponding well with enzyme upregulation. The phosphoketolase reaction saves carbon by producing an acetyl residue together with G3P directly from D-xylulose 5-P instead of requiring several reactions until pyruvate decarboxylation, potentially resulting in an increased product yield from substrate. However,

cells seem to prefer the alternative pathway of transketolase as it provides more carbon towards glycolysis and oxidative PPP, where the majority of NADPH is regenerated. The overexpression of phosphoketolase and phosphotransacetylase enzymes in *Y. lipolytica* resulted in 20-50% increased lipid production (Niehus et al., 2018). Therefore, the upregulation of phosphoketolase could be partially related to the increased accumulation of lipids and carotenoids at the aforementioned phases. NADPH regeneration is an important mechanism for xylose assimilation and lipid synthesis. ME is considered a key enzyme in the recycling of NADPH for lipid biosynthesis (Ratledge and Wynn, 2002). However, according to our proteomic data and simulations, the majority of the reducing power is generated in the oxidative branch of the PPP. These results are consistent with previous <sup>13</sup>C-labelling experiments with *Y. lipolytica* performed using glucose as the carbon source (Wasylenko et al., 2015). However, ME overexpression in *Rhodotorula glutinis* has led to a 2-fold increase in lipid accumulation (Li et al., 2013).

Induction of oxidative stress in *R. toruloides* cultivation was employed as a strategy to identify the metabolic changes that result in the higher levels of carotenoids and lipids. Surprisingly, under PER conditions, no changes in carotenoid production compared to REF were observed. However, the lipid content was two-fold increased, reaching a yield of 0.65 g/g<sub>DCW</sub>. The highest reported lipid production in *R. toruloides* is 0.675 g/g<sub>DCW</sub> (Li et al., 2007) using a rich medium with glucose as a carbon source. The review by Shi et al. (2017) reports that different authors found evidence for ROS stress enhancing the formation of lipid droplets as an important signalling molecule in response to nitrogen starvation in oleaginous microorganisms. Cells under oxidative stress might possess an even greater supply of NADPH by channelling sugar catabolism to PPP to stabilise the redox balance and ROS clearance (Kuehne et al., 2015). Enzymes related to oxidative stress were upregulated not only in ADA but also in REF under nitrogen limitation (P2 and P3). Tiukova et al. (2019b) reported not only the upregulation of oxidative stress related enzymes during the lipid accumulation phases but also the upregulation of the beta-oxidation representing an ATP sink. According to Xu et al. (2017), lipid oxidation leads to accumulation of oxidative and aldehyde species in *Y. lipolytica*, reducing production performance. The same authors showed that overexpressing enzymes from the oxidative stress defence pathway resulted in industrially relevant lipid production (lipid titre of 72.7 g/L and content of 81.4%) by synchronising lipogenesis to cell growth and mitigating lipotoxicity.

The successive application of H<sub>2</sub>O<sub>2</sub> in the cell through the ALE potentially overwhelmed the cellular antioxidant defence system, boosting carotenoid production. Increased accumulation of carotenoids and lipids was noted in all the phases of ADA compared to REF. In general, improved production can result from the upregulation of phosphoketolase, which led to a slightly more efficient production process. The mevalonate pathway was more activated by higher levels of ERG10, providing more precursors for carotenoid biosynthesis, which along with the upregulation of CRTI (enzyme of the carotenoid biosynthesis pathway) could explain the higher carotenoid levels. IDP1 and GND1 upregulation improved the capacity of NADPH regeneration. The upregulation of the oxidative stress defence pathway potentially diminished lipotoxicity. All the aforementioned factors combined with the downregulation of the  $\beta$ -oxidation could explain the higher lipid production by the adapted cells under oxidative stress (Figure 5A). No reasonable significant modifications in proteomics levels were observed under the light condition in contrast to the findings of Gong et al. (2019) for *R. glutinis*. However, irradiation increased final titres of carotenoids and lipids (75 and 40% improvement, respectively) likely due to post-translational mechanisms. Some fungi, such as *Mucor circinelloides*, harbour genes for carotenoid synthesis that are based on light-induced expression mechanisms (Quiles-Rosillo et al., 2005).

## 5 Conclusions

In this study, the detailed physiological characterisation of *R. toruloides* growth revealed that a considerable amount of xylose was converted into by-products, such as arabitol and xylitol. The accumulation of these by-products can be considered an overflow metabolism, contributing to the redox balancing during xylose catabolism. According to the simulations, the highest NADPH demand was related to substrate uptake and was about 5-fold higher compared with the NADPH levels required for lipid production. The main NADPH regeneration reactions were derived from the oxidative branch of PPP, and ME was underused.

Under nitrogen limitation, the parental strain showed some increased protein expression related to lipid synthesis (FAS complex) and oxidative stress response mechanisms. Interestingly, the adapted strain downregulated the FA degradation pathway, which combined with the upregulation of NADPH regeneration mechanisms and FA and carotenoid synthesis, led to better production performance. Additionally, this strain demonstrated increased allocation of proteins related to mitochondria and amino acid metabolism, potentially explaining the more efficient metabolism.

In general, a good correlation was noted between the predicted fluxes and determined protein abundances. Using data obtained in this study, we can design strategies of metabolic engineering to make the process economically viable by improving cell factory performance.

## 6 Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## 7 Author Contributions

MJP and NB share first co-authorship. MJP, NB, EAM and PJJ designed the experiments. MJP performed the experiments. IB performed the metabolic flux simulations. MJP, NB, IB and PJJ analysed the data. MJP, NB, IB, EAM and PJJ wrote and revised the manuscript.

## 8 Funding

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 668997 and the Estonian Research Council (grant PUT1488P). MJP would additionally like to acknowledge Coordination for the Improvement of Higher Education Personnel (Capes) and São Paulo Research Foundation (FAPESP, grant 2016/10636-8) and DORA Plus.

## 9 Acknowledgements

We thank the Proteomics Core Laboratory at University of Tartu for the proteome quantification.

## 10 References

- Adrio, J. L. (2017). Oleaginous yeasts: Promising platforms for the production of oleochemicals and biofuels. *Biotechnol. Bioeng.* 114, 1915–1920. doi:10.1002/bit.26337.
- BBC Research (2018). The Global Market for Carotenoids. <https://www.globenewswire.com/news-release/2018/05/21/1509412/0/en/Global-Carotenoids-Market-to-Reach-2-0-Billion-by-2022.html> [Accessed 20<sup>th</sup> February 2020].

- Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Statist. Soc. B.* 57, 298-300. doi:0035-9246/95/57289.
- Blomqvist, J., Pickova, J., Tilami, S. K., and Sabine, S. (2018). Oleaginous yeast as a component in fish feed. *Sci Rep*, 8, 1–8. doi:10.1038/s41598-018-34232-x.
- Bommareddy, R. R., Sabra, W., Maheshwari, G., and Zeng, A.-P. (2015). Metabolic network analysis and experimental study of lipid production in *Rhodospiridium toruloides* grown on single and mixed substrates. *Microb. Cell Fact.* 14, 36. doi:10.1186/s12934-015-0217-5.
- Bonturi, N., Crucello, A., Viana, A. J. C., and Miranda, E. A. (2017). Microbial oil production in sugarcane bagasse hemicellulosic hydrolysate without nutrient supplementation by a *Rhodospiridium toruloides* adapted strain. *Process Biochem.* 57, 16–25. doi:10.1016/j.procbio.2017.03.007.
- Bonturi, N., Matsakas, L., Nilsson, R., Christakopoulos, P., Miranda, E. A., and Berglund, K. A. (2015). Single cell oil producing yeasts *Lipomyces starkeyi* and *Rhodospiridium toruloides*: Selection of extraction strategies and biodiesel property prediction. *Energies* 8, 5040–5052. doi:10.3390/en8065040.
- Castañeda, M. T., Nuñez, S., Garelli, F., Voget, C., and De Battista, H. (2018). Comprehensive analysis of a metabolic model for lipid production in *Rhodospiridium toruloides*. *J. Biotechnol.* 280, 11–18. doi:10.1016/j.jbiotec.2018.05.010.
- Chokshi, K., Pancha, I., Ghosh, A., and Mishra, S. (2017). Nitrogen starvation - induced cellular crosstalk of ROS - scavenging antioxidants and phytohormone enhanced the biofuel potential of green microalga *Acutodesmus dimorphus*. *Biotechnol. Biofuels*, 1–12. doi:10.1186/s13068-017-0747-7.
- Coradetti, S. T., Pinel, D., Geiselman, G. M., Ito, M., Mondo, S. J., and Reilly, M. C. (2018). Functional genomics of lipid metabolism in the oleaginous yeast *Rhodospiridium toruloides*. *Elife* 7: e32110. doi:10.7554/eLife.32110.
- Dinh, H. V., Suthers, P. F., Chan, S. H. J., Shen, Y., Xiao, T., and Deewan, A. (2019). A comprehensive genome-scale model for *Rhodospiridium toruloides* IFO0880 accounting for functional genomics and phenotypic data. *Metab. Eng. Commun.* 9, e00101. doi:10.1016/j.mec.2019.e00101.
- Du, C., Li, Y., Guo, Y., Han, M., Zhang, W., and Qian, H. (2016). The suppression of torulene and torularhodin treatment on the growth of PC-3 xenograft prostate tumors. *Biochem. Biophys. Res. Commun.* 469, 1146–1152. doi:10.1016/j.bbrc.2015.12.112.
- Fan, J., Cui, Y., Wan, M., Wang, W., and Li, Y. (2014). Lipid accumulation and biosynthesis genes response of the oleaginous *Chlorella pyrenoidosa* under three nutrition stressors. *Biotechnol. Biofuels* 7, 1–14. doi:10.1186/1754-6834-7-17.
- Fernandes, S., and Murray, P. (2010). Metabolic engineering for improved microbial pentose fermentation. *Bioeng Bugs.* 1, 424–428. doi:10.4161/bbug.1.6.12724.
- Folch, J., Lees, M., and Sloane Stanley, G. (1957). A simple method for the isolation and purification of total lipides from animal tissues. *J Biol Chem.* 226(1):497-509.

- Gong, G., Liu, L., Zhang, X., and Tan, T. (2019). Multi-omics metabolism analysis on irradiation-induced oxidative stress to *Rhodotorula glutinis*. *Appl. Microbiol. Biotechnol.* 103, 361–374. doi:10.1007/s00253-018-9448-9.
- Gong, G., Zhang, X., and Tan, T. (2020). Simultaneously enhanced intracellular lipogenesis and  $\beta$ -carotene biosynthesis of *Rhodotorula glutinis* by light exposure with sodium acetate as the substrate. *Bioresour. Technol.* 295, 122274. doi:10.1016/j.biortech.2019.122274.
- Hu, J., and Ji, L. (2016). Draft Genome Sequences of *Rhodospiridium toruloides* Strains ATCC 10788 and ATCC 10657 with Compatible Mating Types. *Genome Announc.* 4(2): e00098-16. doi:10.1128/genomeA.00098-16.
- Irazusta, V., Nieto-Peñalver, C. G., Cabral, M. E., Amoroso, M. J., and De Figueroa, L. I. C. (2013). Relationship among carotenoid production, copper bioremediation and oxidative stress in *Rhodotorula mucilaginosa* RCL-11. *Process Biochem.* 48, 803–809. doi:10.1016/j.procbio.2013.04.006.
- Jagtap, S. S., and Rao, C. V (2018). Production of D -arabitol from D -xylose by the oleaginous yeast *Rhodospiridium toruloides* IFO0880. *Appl Microbiol Biotechnol*, 102:143–151. <https://doi.org/10.1007/s00253-017-8581-1>
- Kerkhoven, E. J., Lahtvee, P., and Nielsen, J. (2015). Applications of computational modeling in metabolic engineering of yeast. *FEMS Yeast Research*, 15, 1–13. doi:10.1111/1567-1364.12199.
- Kot, A. M., Błazejak, S., Gientka, I., Kieliszek, M., and Bryś, J. (2018). Torulene and torularhodin: “New” fungal carotenoids for industry. *Microb. Cell Fact.* 17, 1–14. doi:10.1186/s12934-018-0893-z.
- Kot, A. M., Błazejak, S., Kieliszek, M., Gientka, I., Bryś, J., and Reczek, L. (2019). Effect of exogenous stress factors on the biosynthesis of carotenoids and lipids by *Rhodotorula* yeast strains in media containing agro-industrial waste. *World J. Microbiol. Biotechnol.* 35, 1–10. doi:10.1007/s11274-019-2732-8.
- Koutinas, A. A., and Papanikolaou, S. (2011). *Biodiesel production from microbial oil*. Woodhead Publishing Limited, 177 - 198. doi:10.1533/9780857090492.2.177.
- Kuehne, A., Emmert, H., Lucius, R., Zamboni, N., Kuehne, A., and Emmert, H. (2015). Acute activation of oxidative pentose phosphate pathway as first-line response to oxidative stress in human skin cells. *Mol. Cell* 59, 359–371. doi:10.1016/j.molcel.2015.06.017.
- Kumar, R., and Lahtvee, P. (2020). Proteome overabundance enables respiration but limitation onsets carbon overflow. *bioRxiv [Preprint]* <https://doi.org/10.1101/2020.02.20.957662>
- Kumar, S., Kushwaha, H., Bachhawat, A. K., Raghava, G. P. S., and Ganesan, K. (2012). Genome sequence of the oleaginous red yeast *Rhodospiridium toruloides* MTCC 457. *Eukaryot. Cell* 11, 1083–1084. doi:10.1128/EC.00156-12.
- Lahtvee, P. J., Sánchez, B. J., Smialowska, A., Kasvandik, S., Elsemman, I. E., and Gatto, F. (2017). Absolute quantification of protein and mRNA abundances demonstrate variability in gene-specific translation efficiency in yeast. *Cell Syst.* 4, 495-504. doi:10.1016/j.cels.2017.03.003.



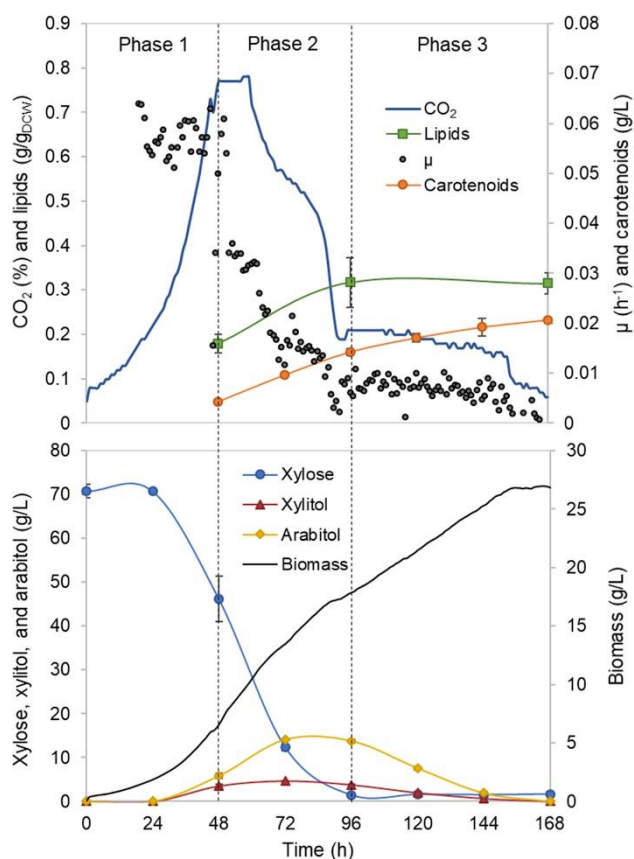
- Lee, J. J. L., Chen, L., Shi, J., Trzcinski, A., and Chen, W. N. (2014). Metabolomic profiling of *Rhodospiridium toruloides* grown on glycerol for carotenoid production during different growth phases. *J. Agric. Food Chem.* 62, 10203–10209. doi:10.1021/jf502987q.
- Li, Y., Zhao, Z. (Kent), and Bai, F. (2007). High-density cultivation of oleaginous yeast *Rhodospiridium toruloides* Y4 in fed-batch culture. *Enzyme Microb. Technol.* 41, 312–317. doi:10.1016/j.enzmictec.2007.02.008.
- Li, Z., Sun, H., Mo, X., Li, X., Xu, B., and Tian, P. (2013). Overexpression of malic enzyme (ME) of *Mucor circinelloides* improved lipid accumulation in engineered *Rhodotorula glutinis*. *Appl. Microbiol. Biotechnol.* 97, 4927–4936. doi:10.1007/s00253-012-4571-5.
- Liu, H., Zhao, X., Wang, F., Li, Y., Jiang, X., Ye, M., Zhao, Z., and Zou, H. (2009). Comparative proteomic analysis of *Rhodospiridium toruloides* during lipid accumulation. *Yeast*. 26, 553-566. doi: 10.1002/yea.1706.
- Liu, W., Huang, Z., Li, P., Xia, J., and Chen, B. (2012). Formation of triacylglycerol in *Nitzschia closterium* f. *minutissima* under nitrogen limitation and possible physiological and biochemical mechanisms. *J. Exp. Mar. Bio. Ecol.* 418–419, 24–29. doi:10.1016/j.jembe.2012.03.005.
- Lopes, H. J. S., Bonturi, N., and Miranda, E. A. (2020a). *Rhodotorula toruloides* single cell oil production using eucalyptus urograndis hemicellulose hydrolysate as a carbon source. *Energies* 13. doi:10.3390/en13040795.
- Lopes, H. J. S., Bonturi, N., Kerkhoven, E. J., Miranda, E. A., and Lahtvee, P. (2020b). C/N ratio and carbon source-dependent lipid production profiling in *Rhodotorula toruloides*. *Applied Microbial and Cell Physiology*, 104, 2639-2649. doi.org/10.1007/s00253-020-10386-5
- Cox, J. and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nature Biotechnology*, 26, 1367–1372. doi:10.1038/nbt.1511.
- Mata-Gómez, L. C., Montañez, J. C., Méndez-Zavala, A., and Aguilar, C. N. (2014). Biotechnological production of carotenoids by yeasts: an overview. *Microb. Cell Fact.* 13, 12. doi:10.1186/1475-2859-13-12.
- Metzl-raz, E., Kafri, M., Yaakov, G., Soifer, I., and Gurvich, Y. (2017). Principles of cellular resource allocation revealed by condition-dependent proteome profiling. *Elife*, 31;6. e28034. doi: 10.7554/eLife.28034.
- Morin, N., Calcas, X., Devillers, H., Durrens, P., Sherman, J., and Nicaud, J. (2014). Draft genome sequence of *Rhodospiridium toruloides* CECT1137, an oleaginous yeast of biotechnological interest. 2, 578–579. doi:10.1128/genomeA.00641-14.
- Niehus, X., Marie, A., Le, C., Sandoval, G., Nicaud, J. M., and Amaro, R. L. (2018). Engineering *Yarrowia lipolytica* to enhance lipid production from lignocellulosic materials. *Biotechnol. Biofuels*, 1–10. doi:10.1186/s13068-018-1010-6.
- Nilsson, A., and Nielsen, J. (2016). Metabolic Trade-offs in Yeast are Caused by F1F0-ATP synthase. *Nat. Publ. Gr.*, 0–11. doi:10.1038/srep22264.

- Park, Y., Nicaud, J., and Ledesma-amaro, R. (2017). The Engineering Potential of *Rhodospiridium toruloides* as a Workhorse for Biotechnological Applications. *Trends Biotechnol.* 1–14. doi:10.1016/j.tibtech.2017.10.013.
- Peterson, H., Kolberg, L., Kuzmin, I., Arak, T., and Adler, P. (2019). g:Profiler: a web server for functional enrichment analysis and conversions of gene lists. *Nucleic Acids Res.*, 47, 191–198. doi: 10.1093/nar/gkz369.
- Quiles-Rosillo, M. D., Ruiz-Vázquez, R. M., Torres-Martínez, S., and Garre, V. (2005). Light induction of the carotenoid biosynthesis pathway in *Blakeslea trispora*. *Fungal Genet. Biol.* 42, 141–153. doi:10.1016/j.fgb.2004.10.008.
- Ratledge, C., and Wynn, J. P. (2002). The biochemistry and molecular biology of lipid accumulation in oleaginous microorganisms. *Adv. Appl. Microbiol.* 51, 1–51. doi: 10.1016/s0065-2164(02)51000-5
- Raudvere, U., Kolberg, L., Kuzmin, I., Arak, T., Adler, P., Peterson, H., and Vilo, J. (2019). g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Research.* 47, 191–198. doi:10.1093/nar/gkz369.
- Sambles, C., Middelhaufe, S., Soanes, D., Kolak, D., Lux, T., Moore, K. (2017). Genomics Data Genome sequence of the oleaginous yeast *Rhodotorula toruloides* strain. *Genomics data.* 13, 1–2. doi:10.1016/j.gdata.2017.05.009.
- Sánchez, B. J., Lahtvee, P., Campbell, K., Kasvandik, S., and Yu, R. (2020). Benchmarking accuracy and precision of intensity-based absolute quantification of protein abundances in *Saccharomyces cerevisiae*. *bioRxiv [Preprint]*. doi: <https://doi.org/10.1101/2020.03.23.998237>
- Sánchez, B. J., Zhang, C., Nilsson, Lahtvee, P., A., Kerkhoven, E. J., and Nielsen, J. (2017). Improving the phenotype predictions of a yeast genome-scale metabolic model by incorporating enzymatic constraints. *Mol Syst Biol.* 1–16. doi:10.15252/msb.20167411.
- Scott, M., Klumpp, S., Mateescu, E. M., and Hwa, T. (2014). Emergence of robust growth laws from optimal regulation of ribosome synthesis. *Mol Syst Biol.* 10:747. doi: 10.15252/msb.20145379
- Shi, K., Gao, Z., Shi, T., Song, P., Ren, L., and Huang, H. (2017). Reactive oxygen species-mediated cellular stress response and lipid accumulation in oleaginous microorganisms: the state of the art and future perspectives. *Front Microbiol.* 8:793, 1-14. doi:10.3389/fmicb.2017.00793.
- Stahl, W., and Sies, Æ. H. (2007). Carotenoids and flavonoids contribute to nutritional protection against skin damage from sunlight. *Molecular Biotechnology.* 37, 26–30. doi:10.1007/s12033-007-0051-z.
- Tiukova, I. A., Prigent, S., Nielsen, J., Sandgren, M., and Kerkhoven, E. J. (2019a). Genome-scale model of *Rhodotorula toruloides* metabolism. *Biotechnol. Bioeng.* 116, 3396–3408. doi:10.1002/bit.27162.
- Tiukova, I., Brandenburg, J., Blomqvist, J., Sampels, S., Mikkelsen, N., Skaugen, M., Arntzen, M., Nielsen, J., Sandgren, M., and Kerkhoven, E. (2019b). Proteome analysis of xylose metabolism in *Rhodotorula toruloides* during lipid production. *Biotechnol Biofuels.* 12:137.

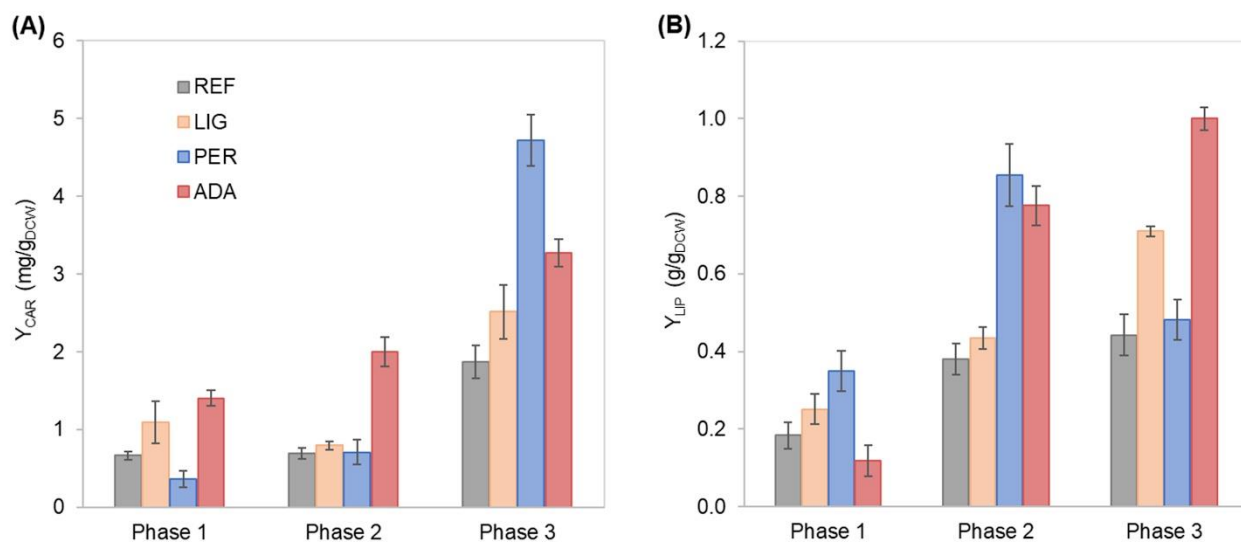
doi: 10.1186/s13068-019-1478-8

- Unrean, P., Khajeeram, S., and Champreda, V. (2017). Combining metabolic evolution and systematic fed-batch optimization for efficient single-cell oil production from sugarcane bagasse. *Renew. Energy* 111, 295–306. doi:10.1016/j.renene.2017.04.018.
- Väremo, L., Nielsen, J. and Nookaew, I. (2013). Enriching the gene set analysis of genome-wide data by incorporating directionality of gene expression and combining statistical hypotheses and methods. *Nucleic Acids Research*. 41 (8), 4378-4391. doi:10.1093/nar/gkt111.
- Wang, H., Marcis, S., Hermansson, D., Agren, R., Nielsen, J., and Kerkhoven, E. J. (2018a). RAVEN 2.0: A versatile toolbox for metabolic network reconstruction and a case study on *Streptomyces coelicolor*. *Plos Comput Biol*. 14, 1–17. doi:10.1371/journal.pcbi.1006541.
- Wang, Y., Zhang, S., Zhu, Z., Shen, H., Lin, X., Jin, X. (2018b). Systems analysis of phosphate-limitation-induced lipid accumulation by the oleaginous yeast *Rhodospiridium toruloides*. *Biotechnol. Biofuels* 11, 1–15. doi:10.1186/s13068-018-1134-8.
- Wasylenko, T. M., Ahn, W. S., and Stephanopoulos, G. (2015). The oxidative pentose phosphate pathway is the primary source of NADPH for lipid overproduction from glucose in *Yarrowia lipolytica*. *Metab. Eng.* 30, 27–39. doi:10.1016/j.ymben.2015.02.007.
- Weber, R. W. S., Anke, H., and Davoli, P. (2007). Simple method for the extraction and reversed-phase high-performance liquid chromatographic analysis of carotenoid pigments from red yeasts (Basidiomycota, Fungi). *J. Chromatogr. A* 1145, 118–122. doi:10.1016/j.chroma.2007.01.052.
- Xu, P., Qiao, K., and Stephanopoulos, G. (2017). Engineering Oxidative Stress Defense Pathways to Build a Robust Lipid Production Platform in *Yarrowia lipolytica*. *Biotechnol Bioeng.* 114, 1521–1530. doi:10.1002/bit.26285.
- Yang, X., Sun, W., Shen, H., Zhang, S., Jiao, X., and Zhao, Z. K. (2018). Expression of phosphotransacetylase in: *Rhodospiridium toruloides* leading to improved cell growth and lipid production. *RSC Adv.* 8, 24673–24678. doi:10.1039/c8ra03028f.
- Yilancioglu, K., Cokol, M., Pastirmaci, I., Erman, B., and Cetiner, S. (2014). Oxidative Stress is a Mediator for increased lipid accumulation in a newly isolated *Dunaliella salina* strain. *PLoS One*. 9 (3): e91957. doi:10.1371/journal.pone.0091957.
- Zhu, Z., Zhang, S., Liu, H., Shen, H., Lin, X., and Yang, F. (2012). A multi-omic map of the lipid-producing yeast *Rhodospiridium toruloides*. *Nat. Commun.* 3, 1112. doi:10.1038/ncomms2112.

## 11 Figures and Tables



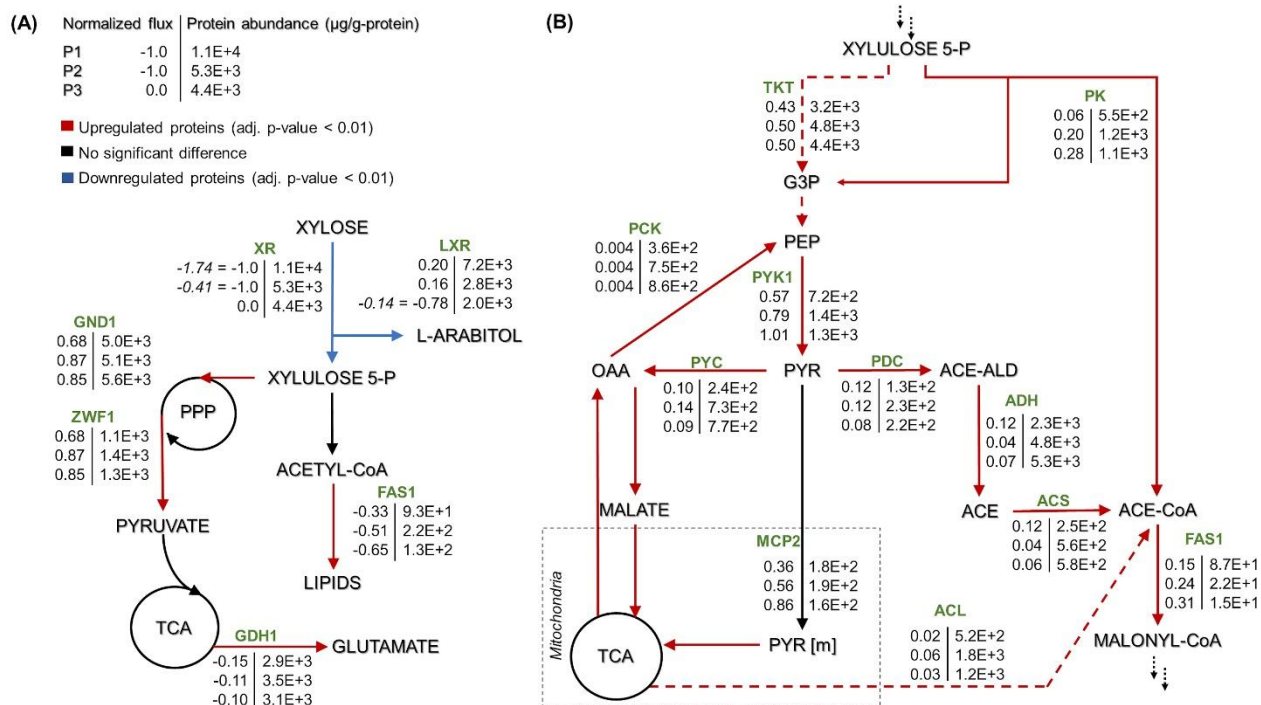
**Figure 1.** *R. toruloides* batch cultivation on xylose under the reference (optimal) environmental conditions for the parental strain. Dashed vertical lines define three observed growth phases. The specific growth rate ( $\mu$ ), biomass concentration, intracellularly accumulated lipid and carotenoid concentrations, extracellular metabolite profiles and CO<sub>2</sub> production profile in the outflow gas are presented. The values represent an average of three independent cultivation experiments; error bars represent standard deviation.



**Figure 2.** Carotenoid (A) and lipid (B) yields on biomass in each growth phase of *R. toruloides* parental strain under the reference condition (REF), light irradiation (LIG) and in the presence of hydrogen peroxide (PER), and the adapted strain under oxidative stress (ADA). Yields were calculated considering the production in each growth phase.

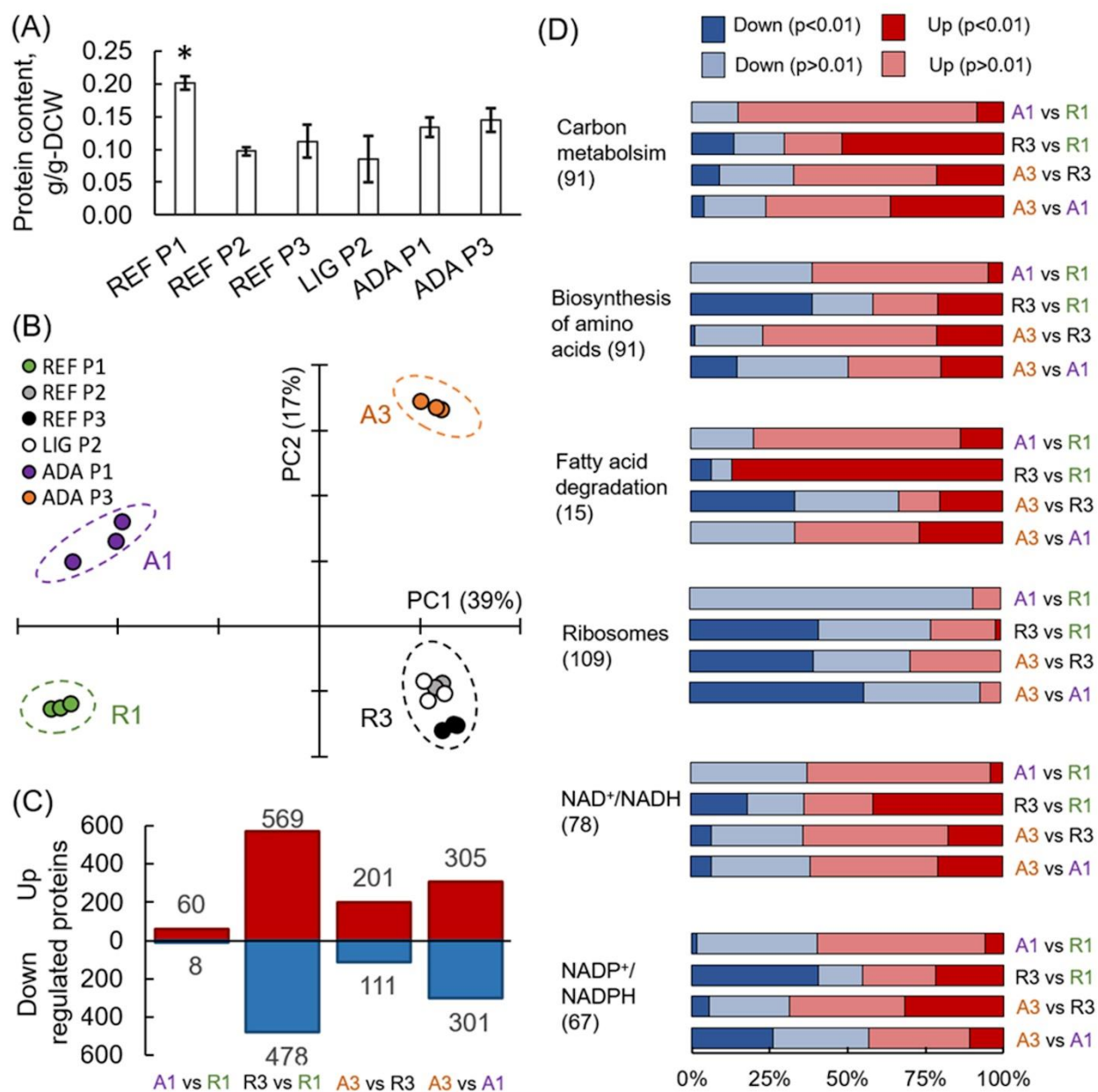
**Table 1.** Global titres, yields on biomass (Y), and volumetric production rate (q) of carotenoids and lipid production by *R. toruloides* parental strain under the reference condition (REF), light irradiation (LIG), hydrogen peroxide stress (PER), and the adapted strain under oxidative stress (ADA). These values were based on the final titres and the total fermentation time, representing the mean and standard deviation of triplicate experiments.

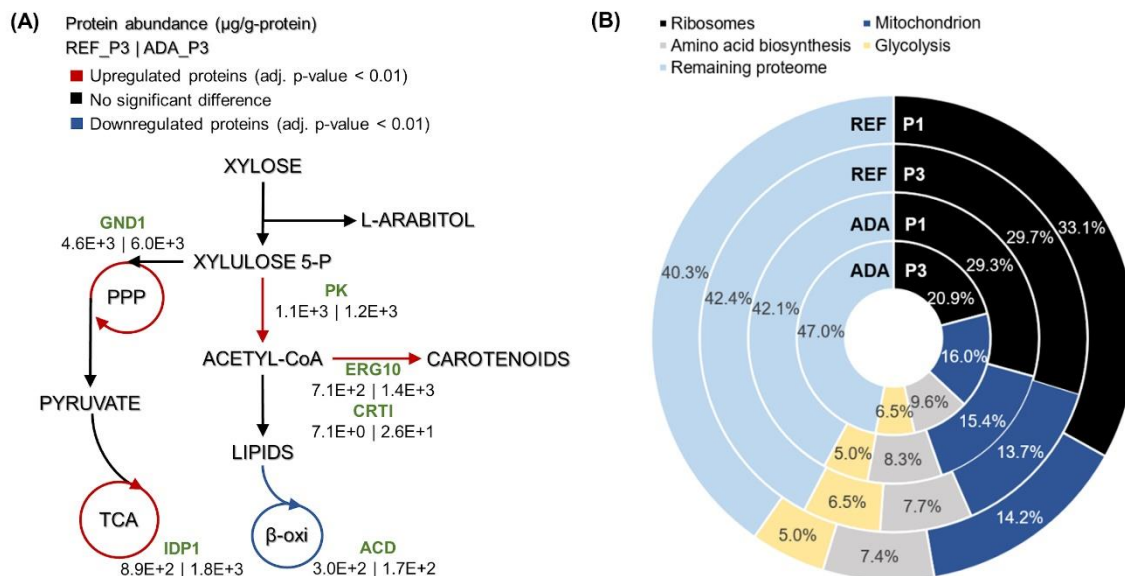
		REF	LIG	PER	ADA
<b>Carotenoids</b>	Titre (mg/L)	20.7 ± 0.5	36.2 ± 2.6	19.3 ± 0.2	44.0 ± 2.4
	$Y_{CAR}$ (mg/DCW)	0.85 ± 0.01	1.45 ± 0.14	0.89 ± 0.04	1.90 ± 0.13
	$q_{CAR}$ (mg/L.h)	0.12 ± 0.003	0.22 ± 0.02	0.06 ± 0.02	0.23 ± 0.01
<b>Lipids</b>	Titre (mg/L)	8.1 ± 0.3	11.1 ± 1.7	13.8 ± 0.9	13.3 ± 0.5
	$Y_{LIP}$ (g/g <sub>DCW</sub> )	0.33 ± 0.07	0.46 ± 0.12	0.65 ± 0.06	0.58 ± 0.07
	$q_{LIP}$ (g/L.h)	0.05 ± 0.003	0.07 ± 0.005	0.05 ± 0.002	0.07 ± 0.002



**Figure 3.** *R. toruloides* GEM-based fluxes (normalized to the substrate uptake) and measured protein abundances ( $\mu\text{g/g-protein}$ ) under the three growth phases of reference cultivation on xylose (P1-P3). The main NADPH regeneration and utilization fluxes are presented with positive and negative values, respectively (A). Representation of the central carbon metabolism illustrating xylulose 5-phosphate, pyruvate and acetyl-CoA branchpoints (B). The fluxes were normalized by the substrate uptake rate in each phase (the absolute substrate uptake values ( $\text{mmol/g}_{\text{DCW}}\cdot\text{h}$ ) are represented in italic in panel A (note that in P3, arabitol and xylitol were co-consumed)). The arrow colour indicates the significant protein allocation change under nitrogen limitation (P2 and P3) compared to nitrogen excess condition (P1); and the dashed arrow on panel B represent multiple reactions. All the fluxes and protein abundance are represented in the Supplementary Tables S3 and S7, respectively. G3P: glyceraldehyde 3-phosphate, PEP: phosphoenolpyruvate, PYR: pyruvate, OAA: oxaloacetate, ACE-ALD: acetaldehyde, TKT: transketolase, PK: phosphoketolase, PCK: phosphoenolpyruvate carboxykinase, PYK1: pyruvate kinase, PYC: pyruvate carboxylase, PDC: pyruvate decarboxylase, MCP2: mitochondria pyruvate carrier, ADH: alcohol dehydrogenase, ACS: acetyl-CoA synthase, ACL: ATP citrate lyase, and FAS1: fatty acid synthase. More abbreviations can be found in the Supplementary Table S4.

**Figure 4.** Proteomics results from *R. toruloides* studies on xylose under different environmental conditions (REF P1, P2, and P3; LIG P2; ADA P1 and P3). Total protein content in cells (A). Principal component analysis (B). Significantly (adj. p-value < 0.01) up- and downregulated proteins (C). Gene set enrichment analysis based on the proteomics data, where GO groups were received from Uniprot, enzyme-metabolite interactions and subgroups from the rhtoGEM (D). All previewed categories show significant difference at least in 2 comparisons (adj. p-value < 0.001). Number in brackets indicates proteins in each category. Panels (C) and (D) present comparisons between samples showing clear separation in PCA represented in (B).





**Figure 5.** Main differences in protein allocation between the adapted strain under oxidative stress (ADA) and the parental strain under reference condition (REF) studied under the nitrogen limitation (P3). The arrow colour indicates the change in protein allocation in ADA compared to REF (A). Proteome allocation into the most abundant metabolic groups for ADA and REF in P1 (nitrogen excess) and P3 (nitrogen limitation) (B). GND1, 6-phosphogluconate dehydrogenase; IDP1, isocitrate dehydrogenase; PK, phosphoketolase; ERG10, acetyl-CoA acetyltransferase; CRTI, phytoene dehydrogenase, ACD, acyl-CoA dehydrogenase.