



UNIVERSIDADE ESTADUAL DE CAMPINAS
Faculdade de Engenharia Química

MARINA JULIO PINHEIRO

**Estratégias para aumentar a produção de lipídeos e carotenoides na
Rhodotorula toruloides a partir da xilose**

**Strategies for increasing the production of lipids and carotenoids in
Rhodotorula toruloides from xylose**

CAMPINAS

2021

MARINA JULIO PINHEIRO

**Strategies for increasing the production of lipids and carotenoids in
Rhodotorula toruloides from xylose**

**Estratégias para aumentar a produção de lipídeos e carotenoides na
Rhodotorula toruloides a partir da xilose**

Tese apresentada à Faculdade de Engenharia Química da Universidade Estadual de Campinas como parte dos requisitos exigidos para a obtenção do título de Doutora em Engenharia Química

Orientador: Prof. Dr. Everson Alves Miranda

ESTE EXEMPLAR CORRESPONDE À VERSÃO FINAL DA TESE DEFENDIDA PELA ALUNA MARINA JULIO PINHEIRO E ORIENTADA PELO PROF. DR. EVERSON ALVES MIRANDA

CAMPINAS

2021

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca da Área de Engenharia e Arquitetura
Rose Meire da Silva - CRB 8/5974

P655s Pinheiro, Marina Julio, 1990-
Strategies for increasing the production of lipids and carotenoids in
Rhodotorula toruloides from xylose / Marina Julio Pinheiro. – Campinas, SP :
[s.n.], 2021.

Orientador: Everson Alves Miranda.
Tese (doutorado) – Universidade Estadual de Campinas, Faculdade de
Engenharia Química.

1. Rhodosporidium toruloides.. 2. Óleo microbiano. 3. Carotenóides. 4.
Biorrefinaria. 5. Proteômica. I. Miranda, Everson Alves, 1959-. II. Universidade
Estadual de Campinas. Faculdade de Engenharia Química. III. Título.

Informações para Biblioteca Digital

Título em outro idioma: Estratégias para aumentar a produção de lipídeos e carotenoides na Rhodotorula toruloides a partir da xilose

Palavras-chave em inglês:

Rhodosporidium toruloides

Microbial oil

Carotenoids

Biorefinery

Proteomic

Área de concentração: Engenharia Química

Titulação: Doutora em Engenharia Química

Banca examinadora:

Everson Alves Miranda [Orientador]

Andreas Karoly Gombert

Adriano Rodrigues Azzoni

Marisa Masumi Beppu

Roberto de Campos Giordano

Data de defesa: 24-11-2021

Programa de Pós-Graduação: Engenharia Química

Identificação e informações acadêmicas do(a) aluno(a)

- ORCID do autor: <https://orcid.org/0000-0002-5749-319>

- Currículo Lattes do autor: https://www.cnpq.br/cvlattesweb/PKG_MENU.men

Folha de Aprovação da Defesa de Tese de Doutorado defendida por **MARINA JULIO PINHEIRO** e aprovada em 24 de novembro de 2021 pela Comissão Examinadora constituída pelos doutores:

Prof. Dr. Everson Alves Miranda – Presidente e Orientador

FEQ/ UNICAMP

Videoconferência

Prof. Dr. Andreas Karoly Gombert

FEA / UNICAMP

Videoconferência

Profa. Dra. Marisa Masumi Beppu

FEQ / UNICAMP

Videoconferência

Dr. Adriano Rodrigues Azzoni

Escola Politécnica da USP

Videoconferência

Dr. Roberto de Campos Giordano

Centro de Ciências Exatas e de Tecnologia - DEQ

Videoconferência

A ATA da defesa com as respectivas assinaturas dos membros encontra-se no SIGA/Sistema de Fluxo de Dissertação / Tese e na Secretaria do Programa da Unidade.

*Aos meus pais Moacyr e Iolanda e
à minha querida irmã Mariana*

Agradecimentos

Aos meus pais, Moacyr e Iolanda, pelo amor, compreensão, incentivo, por sempre me fazer ver o lado leve da vida.

À minha irmã Mariana, minha pessoa, pela cumplicidade, amizade, maluquices e por sempre estar aqui.

Ao prof. Dr. Everson Alves Miranda, por ter aceito a orientação deste trabalho, pela amizade, confiança e pelos ensinamentos não só de engenharia química, mas também de filosofia, música e poesia.

À Paola e Lorena, por todo apoio nos momentos difíceis, pelos cafés-terapia, pelas viagens, histórias e risadas inesquecíveis, pelos Natais memoráveis longe de casa.

Ao Guilherme, Jhonny, Raul e Sophia, pelos Mokus e Spartas.

Ao Rafael e Nemailla, pela amizade, conselhos e disponibilidade em ajudar durante todo meu doutorado.

Aos amigos do LEBp, LIMBio e LEBC pela ajuda e momentos de descontração.

À todos que de alguma forma contribuíram para a realização deste trabalho.

À Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) - Processo 2016/10636-8 pelo financiamento do projeto de pesquisa, tornando possível a produção desta tese.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001.

To prof. Dr Petri-Jaan Lathvee, for supervising this work, for the trust and all the opportunities offered during my stay in Estonia.

To Eka, Tobias and Nemailla, for helping with GGA.

To Synbio group, for welcoming me into the team.

To Dora Plus Programme, for the financial support.

RESUMO

O uso de fábricas celulares para converter açúcares de biomassa lignocelulósica em produtos químicos, tais como oleoquímicos e aditivos alimentares, é essencial para o desenvolvimento de processos sustentáveis. *Rhodotorula toruloides* é uma levedura que metaboliza naturalmente uma ampla gama de substratos, incluindo hidrolisados lignocelulósicos, e os converte em lipídeos e carotenoides. Neste estudo, o estresse oxidativo induzido por peróxido de hidrogênio (H₂O₂) ou irradiação de luz e a evolução adaptativa em laboratório (ALE) foram aplicados para aumentar os níveis de produção de lipídeos e carotenoides na *R. toruloides*, e também com o objetivo de entender a regulação metabólica da formação destes compostos. Para tanto, realizou-se caracterização fisiológica detalhada em cultivos em batelada, utilizando xilose como único substrato, combinando-se com análises de proteômica absoluta e modelos metabólicos em escala genômica. Nos cultivos da *R. toruloides* em biorreator, a irradiação de luz branca resultou em aumento de 70% e 40% na produção de carotenoides e lipídeos, respectivamente, em comparação com a condição de referência (sem adição de estresse oxidativo). A presença de H₂O₂ aumentou a produção de lipídeos em 95%, atingindo um conteúdo lipídico de 65%; entretanto, não afetou a produção de carotenoides. Já a cepa obtida através da ALE, elevou a produção de carotenoides em 2 vezes, para 44 mg/L. Ao longo do crescimento da *R. toruloides*, foi possível identificar três fases distintas durante o consumo da xilose baseando-se nas medições *online* de crescimento e excreção de CO₂. Os dados de proteômica demonstraram que na fase de limitação de nitrogênio teve-se maior atividade do mecanismo de regeneração do cofator NADPH, da fosfoctolase e uma redução na atividade da beta-oxidação, correlacionando-se com o aumento do rendimento de lipídeos. A cepa adaptada demonstrou um aumento na alocação de proteínas relacionadas à mitocôndria e ao metabolismo de aminoácidos, potencial explicação para um metabolismo mais eficiente. Este trabalho apresenta uma visão holística da conversão de xilose em óleo microbiano e carotenoides por *R. toruloides*, visando a aplicação desta levedura em biorefinarias. A partir dos dados de proteômica identificou a baixa expressão das enzimas relacionadas à via de síntese dos carotenoides como um possível gargalo da produção destes compostos. Dessa forma, objetivou-se a superexpressão dos genes da via dos carotenoides nesta levedura, o que resultou em cepa com produção de carotenoides duas vezes maior que a cepa parental. Para tanto, uma plataforma padronizada para a construção de cassete de expressão para *R. toruloides* foi desenvolvida baseada no método de *Golden Gate Assembly*, visando tornar mais eficiente o processo de engenharia metabólica para essa levedura.

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, the oxidative stress induced by hydrogen peroxide or light irradiation, and adaptive laboratory evolution (ALE) were applied to improve the production levels of lipids and carotenoids in *R. toruloides*, and also, to understand the metabolic regulation of formation of these compounds. For that, a detailed physiology characterisation was carried out in batch cultivation, using xylose as sole carbon source, in combination with absolute proteomics and genome-scale metabolic models. In *R. toruloides* cultivation, the use of white light irradiation resulted in 70% higher carotenoid and 40% higher lipid content compared to reference condition. The presence of hydrogen peroxide did not affect carotenoid production but culminated in the highest lipid content of 0.65 g/g_{DCW}. The adapted strain showed improved fitness and 2.3-fold higher carotenoid content than the parental strain. During the xylose consumption in batch cultivation, three distinct growth phases were identified based on the online measurements of growth and CO₂ excretion. In the nitrogen limitation phase, the proteomic data demonstrated an improved NADPH regeneration, phosphoketolase activity and reduced beta-oxidation, correlating with increasing lipid yields. The adapted strain showed an increased allocation of proteins related to mitochondria and amino acid metabolism, potentially explaining the more efficient metabolism. From the system biology data, it was possible to identify the low expression of enzymes related to carotenoid synthesis, which can mean a bottleneck in the production of these compounds in *R. toruloides*. Therefore, the overexpression of the carotenogenic genes was carried out in this work, and our best-engineered strain produced twice more carotenoids than the parental strain. For the construction of expression cassettes for *R. toruloides*, a standardized platform was developed based on the Golden Gate Assembly method. This work presented a holistic view of xylose conversion into microbial oil and carotenoids by *R. toruloides*, in a process towards renewable and cost-effective production of these molecules and aiming to make the process of metabolic engineering for this yeast more efficient.

List of Figures

Figure 1.1. Flowchart representing the steps of the work plan.....	20
Figure 2.1. Schematic of xylose catabolism by <i>R. toruloides</i> . The dashed arrow represents multiple reactions.	25
Figure 2.2. Metabolic pathways involving the conversion of xylose into lipids, ergosterol and carotenoids by oleaginous yeasts. Legend: G3P, glyceraldehyde 3-phosphate; DHAP, dihydroxyacetone phosphate; DAG, diacylglyceride; TAG, triacylglycerides; CRTE, geranylgeranyl diphosphate synthetase; CRTI, phytoene dehydrogenase; CRTYB, phytoene synthetase.	28
Figure 2.3. Golden Gate Assembly Method. The fragments are amplified via PCR with the respective primers, inserting the BsaI restriction sites and pre-designed 4-nucleotide overhangs. Amplicons and the target vector are incubated with the enzyme BsaI and DNA ligase in the same reaction tube. Then, as digestion occurs, the DNA parts are linked through the annealing of complementary overhangs and the action of DNA ligase, assembling the final DNA sequence.....	31
Figure 3.1. Bioreactor equipped with a photoperiod.	37
Figure 3.2. Effect of light irradiation on cell growth (filled symbols) and sugar consumption (open symbols) by <i>R. toruloides</i> under reference condition (without irradiation, black circles), and under different light irradiations: white (grey triangles), red (red squares), and blue (blue diamonds).	38
Figure 3.3. Carotenoid titre (A) and composition (B) under white, red, or blue light irradiation compared to the reference condition (without irradiation).	40
Figure 4.1. Scheme of ALE methodology applied to improve the carotenoids production in <i>R. toruloides</i>	44
Figure 4.2. Changes in μ_{MAX} and lag phase time during the (A) first, (B) second, and (C) third cycle of the adaptive laboratory evolution with H ₂ O ₂ . Error bars represent the standard deviation from three measurements.....	46
Figure 4.3. Relative carotenoid proportion of parental strain grown under optimal growth condition and adapted strains from the cycle 1, 2 and 3 (AD1, AD2, and AD3, respectively) grown under oxidative stress condition (10, 20, and 50 mmol/L of H ₂ O ₂ , respectively).....	47
Figure 5.1. <i>R. toruloides</i> 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 (μ), cell mass concentration, intracellularly	

accumulated lipid and carotenoid concentrations, extracellular metabolite profiles and CO₂ production profile in the outflow gas are presented. The values represent an average of three independent cultivation experiments; error bars represent standard deviation. The red arrows represent the proteomic data points. 53

Figure 5.2. Carotenoid (A) and lipid (B) yields on cell mass 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. 55

Figure 5.3. *R. toruloides* GEM-based fluxes (normalized to the substrate uptake) and measured protein abundances (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 (mmol/g_{DCW}.h) are represented in italic in panel A (note that in P3, arabitol and xylitol were co-consumed)). The colour of enzymes name indicates the significant protein allocation change under nitrogen limitation (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 Pinheiro et al., 2020, 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 ACC1: acetyl-CoA carboxylase. 59

Figure 5.4. *R. toruloides* batch cultivation on xylose under light irradiation (A) and in presence of H₂O₂ (B), and adapted strain under oxidative stress (C). Dashed vertical lines define three observed growth phases. The specific growth rate (μ), biomass concentration, intracellularly accumulated lipid and carotenoid concentrations, extracellular metabolite profiles and CO₂ production profile in the outflow gas are presented. The values represent an average of three independent cultivation experiments; error bars represent standard deviation. The red arrows represent the proteomic data points. 61

Figure 5.5. Carotenoids composition during the batch cultivation of *R. toruloides* under the studied conditions: (A) the parental strain in the reference condition (REF), (B) under light irradiation (LIG), or (C) in presence of hydrogen peroxide (PER), and (D) the adapted strain under oxidative stress (ADA). 62

Figure 5.6. Proteomics results from *R. toruloides* studies on xylose under different environmental conditions (REF P1, P3, and SP; LIG P3; 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). 65

Figure 5.7. 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; CRTI, phytoene dehydrogenase, ACD, acyl-CoA dehydrogenase. 70

Figure 6.1. Assembly of GG library, in which all three promoters, four genes, two insertional units, and one terminator were cloned into pCR-XL-2-TOPO™ plasmid. The inserts are represented by the black blocks, the red and green blocks represent the 4 nt-overhangs, and the blue blocks, the BsaI recognition sites (A). The Golden Gate Assembly platform in a three transcription unit format constructed on destination vector, pGGA. InsUp/Down - sequence targeting insertion in the 5'/3' region; M - selection marker; P1/2/3 - promoter in transcription unit 1, 2, and 3, respectively; G1/2/3 - gene of interest in transcription unit 1, 2, and 3, respectively; T1/2/3 - terminator located in transcription unit 1, 2, and 3, respectively. For each pre-designed 4-nt overhangs were assigned a letter from A to M. NotI and SadI recognition sites for the release of the expression cassette for yeast transformation (B). Legend: ori – origin of replication; KanR – kanamycin resistance gene; CmR – chloramphenicol resistance gene. 78

Figure 6.2. Golden Gate Assembly bearing the carotenoid synthesis pathway to integration into *R. toruloides* genome. The multigene cassettes (3.1, 3.2, and 3.3) comprised three transcription units and selection marker (G418), flanked with KU70 integration sequences, constructed on a destination vector backbone (pGGA). 80

Figure 6.3. Cell growth of Rt 3.1, Rt 3.2 and Rt 3.3 (engineered strain) compared to Rt (parental strain). 80

Figure 6.4. Production of total carotenoids and carotenoid yield on biomass in each engineered strain, Rt 3.1, Rt 3.2 and Rt 3.3, and parental strain, Rt (A). Production of different carotenoids in each strain (B). 81

List of abbreviations

ACC1: acetyl-CoA carboxylase
ACC1: acetyl-CoA carboxylase
ACD: acyl-CoA dehydrogenase
ACE-ALD: acetaldehyde
ACE-P: acetyl-phosphate
ACL: ATP citrate lyase
ACL: ATP-citrate lyase
ACS: acetyl-CoA synthase
ACS1: acetyl-CoA synthase
ADH: alcohol dehydrogenase
ALE: adaptive laboratory evolution
CRTE: geranylgeranyl diphosphate synthase
CRTI: phytoene dehydrogenase
CRTYB: phytoene synthase
DAD: D-arabitol dehydrogenase
ERG10: acetyl-CoA C-acetyltransferase
ERG13: hydroxymethylglutaryl-CoA synthase
FAS: fatty acid synthases
FAs: fatty acids
G3P: glyceraldehyde 3-phosphate
G418: geneticin resistance gene
GGA: Golden Gate Assembly
GND1: 6-phosphogluconate dehydrogenase
H₂O₂: hydrogen peroxide
HMG1: hydroxymethylglutaryl-CoA reductase
IDP1: isocitrate dehydrogenase
KU70: DNA-dependent ATP-dependent helicase subunit 70
LAD: L-arabitol dehydrogenase
LXR: L-xylulose reductase
MCP2: mitochondria pyruvate carrier
ME: malic enzyme
OAA: oxaloacetate

pADH2: alcohol dehydrogenase 2 promoter
PCK: phosphoenolpyruvate carboxykinase
PDC: pyruvate decarboxylase
PEP: phosphoenolpyruvate
pGPD: glyceraldehyde-3-phosphate dehydrogenase promoter
PK: phosphoketolase
PPP: pentose phosphate pathway
pXYL1: xylose reductase promoter
PYC: pyruvate carboxylase
PYK1: pyruvate kinase
PYR: pyruvate
ROS: reactive oxygen species
SCO: single-cell oil
TAG: triacylglycerides
TCA: tricarboxylic acid cycle
TKT: transketolase
tNOS: nopaline synthase terminator
XDH: xylitol dehydrogenase
XK: xylulose kinase
XR: xylose reductase
Xylulose 5-P: xylulose 5-phosphate
ZWF1: glucose-6-phosphate dehydrogenase

Summary

RESUMO.....	6
ABSTRACT	7
CHAPTER 1 Introduction.....	17
1.1 Objective.....	19
1.2 Work plan	19
1.3 Text outline	21
CHAPTER 2 Literature review	22
2.1 Biorefineries.....	22
2.2 <i>Rhodotorula toruloides</i> as a cell factory.....	23
2.2.1 Xylose metabolism in <i>R. toruloides</i>	24
2.2.2 Synthesis of lipids and carotenoids by <i>R. toruloides</i>	25
2.3 Adaptive laboratory evolution	29
2.4 Metabolic engineering	30
2.5 Systems biology.....	32
CHAPTER 3 Effect of light irradiation on production of lipids and carotenoids in <i>R. toruloides</i>	35
3.1 Introduction.....	35
3.2 Materials and methods	36
3.2.1. Strain and inoculum.....	36
3.2.2. Analytical methods	36
3.2.3. Yeast cultivation	37
3.3 Results and discussion	38
3.4 Conclusion	40
CHAPTER 4 Adaptive laboratory evolution to improve the carotenoid production	41
4.1 Introduction.....	41
4.2 Materials and methods	42
4.2.1 Strain and yeast cultivation.....	42
4.2.2 Analytical methods	42
4.2.3 Adaptive laboratory evolution.....	43
4.2.4 Characterisation of parental and adapted strains under reference and oxidative stress conditions	43

4.3	Results and discussion	44
4.4	Conclusion	47

CHAPTER 5 Xylose metabolism and the effect of oxidative stress on lipid and carotenoid production in *R. toruloides*: insights for its future application in biorefinery 48

5.1	Introduction.....	48
5.2	Materials and methods	49
5.2.1	Strain and inoculum.....	49
5.2.2	Yeast cultivation	49
5.2.3	Quantification of dry cell mass, extracellular metabolites, carotenoids, total lipids and proteins	49
5.2.4	Proteome analysis	51
5.2.5	Genome-scale modelling	52
5.3	Results.....	52
5.3.1	Three distinct phases of <i>R. toruloides</i> growth on xylose.....	52
5.3.2	Understanding intracellular flux patterns among three growth phases	55
5.3.3	Impact of oxidative stress via light irradiation or the presence of H ₂ O ₂	58
5.3.4	Adapted strain under H ₂ O ₂ stress	59
5.3.5	Composition of carotenoids in cell mass	62
5.3.6	Proteomic results revealed the highest difference between nitrogen excess and -limiting conditions.....	62
5.3.7	Translation and NADPH metabolism were most affected under nitrogen limitation	64
5.3.8	Protein changes demonstrate similar trends with the simulated fluxes	66
5.3.9	Comparative proteomics at different growth phases under oxidative stresses...	68
5.3.10	Translation processes also play a crucial role in proteome allocation analysis..	69
5.4	Discussion.....	69
5.5	Conclusions.....	73

CHAPTER 6 Development of a dedicated Golden Gate Assembly platform for the nonconventional yeast *Rhodotorula toruloides* 74

6.1	Introduction.....	74
6.2	Materials and methods	75
6.2.1.	DNA sequences	75
6.2.2.	DNA amplification by polymerase chain reaction (PCR).....	76
6.2.3.	Golden Gate Assembly	76

6.2.4.	Bacterial transformation	77
6.2.5.	Yeast transformation.....	77
6.2.6.	Characterisation of the strains in shake flasks.....	79
6.2.7.	Analytical methods	79
6.3	Results and discussion	79
6.4	Conclusion	81
CHAPTER 7 Conclusions and suggestions for future work		83
7.1	Conclusions.....	83
7.2	Suggestions for future work.....	85
References.....		86
Appendix		96

CHAPTER 1

Introduction

Bioeconomy has emerged as a sustainable alternative to the production of chemical products and fuels. Its concept arose from the need to replace petroleum-based raw materials with renewable sources to obtain several products and make industrial processes cleaner and more efficient. This transition aims to reduce the world dependence on oil, climate change, and to ensure food security. It is estimated that the bioeconomy could mitigate the climate change in the order of 2.5 billion tons of CO₂ per year by 2030 (BIOECONOMY POLICY, 2020).

The expansion of industrial biotechnology drives the economic transition towards the bioeconomy. The market size of the biotechnology sector was valued at USD 497.2 billion in 2020 and is expected to reach USD 952.0 billion by 2027 (GLOBAL MARKET INSIGHTS, 2020). Biorefineries are the main strategy to boost the bioeconomy, since it targets to efficiently produce bio-based compounds, mainly the building blocks and fuels, from different feedstocks, such as lignocelluloses, algae, and various types of wastes (UBANDO; FELIX; CHEN, 2020a). However, biorefinery development depends on the enhancement of microbiological processes efficiency and finding robust microorganisms for a feasible facility.

Rhodotorula toruloides has been considered a promising microorganism to industrial processes. This microorganism can assimilate a wide variety of substrates, including industrial wastes containing toxic compounds, and is a natural producer of industrially important compounds, such as single-cell oils and carotenoids (PARK; NICAUD; LEDESMA-AMARO, 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 (BLOMQVIST et al., 2018; UNREAN; KHAJEERAM; CHAMPREDA, 2017; WEI; SIEWERS; NIELSEN, 2017; YANG et al., 2018). Carotenoids are important molecules for different industries, such as the food, chemical, pharmaceutical and cosmetics industries. In addition to its colourant 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 (DU et al., 2016; KOT et al., 2019; STAHL; SIES, 2007). 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).

Socool *et al.* (2017) reported a pilot-scale process for biodiesel production from single-cell oil of *R. toruloides*. The lipid productivity was 0.44 g/L/h in a low-cost culture medium (sugarcane juice and urea). Through a preliminary economic analysis, the authors showed that the production cost of microbial biodiesel was competitive with vegetal biodiesel. However, according to Koutinas *et al.* (2014), the price equality between microbial and vegetal oil would be possible if the lipid productivity was higher than 1.6 g/L/h and the substrate had no cost to the process.

Although *R. toruloides* stand out concerning production levels of lipids and carotenoids compared to other microorganisms (PARK; NICAUD; LEDESMA-AMARO, 2017), the yield of these compounds is still low, which makes the industrial application of this yeast unfeasible. A key for boosting the production levels of relevant compounds in *R. toruloides* is the development of efficient genetic engineering techniques combined with system biology analysis.

In the last decade, system biology was studied for a better understanding of the genomic organisation and metabolic pathways in *R. toruloides*. Complete genome sequences are available for several *R. toruloides* strains (CORADETTI et al., 2018; HU; JI, 2016a; KUMAR et al., 2012; MORIN et al., 2014; SAMBLES et al., 2017; ZHU et al., 2012). The metabolism of lipid accumulation was investigated (LIU et al., 2009) as well as the metabolism response to changes in the medium composition, such as nitrogen (ZHU et al., 2012) and phosphate limitation (WANG et al., 2018b) through the proteomic analysis. Multi-omics analysis allowed the investigation of the conversion of different substrates into lipids and carotenoids (JIE et al., 2014; TIUKOVA et al., 2019a; ZHU et al., 2012). Genome-scale modelling (GEM) also was applied to *R. toruloides* to provide a holistic view of metabolic fluxes (DINH et al., 2019; TIUKOVA et al., 2019a).

The understanding of microorganism metabolism helps the development of efficient strategies for metabolic engineering. However, the toolset for the genetic manipulation in *R. toruloides* is still very poor, resulting in a major obstacle to making this yeast a cell factory. The functionality of some tools was studied for *R. toruloides*, such as constitutive (LIU et al., 2013, 2016; WANG et al., 2016) and inducible promoters (JOHNS; LOVE; AVES, 2016; LIU et al., 2015), auxotrophic markers (YANG et al., 2008) and antibiotics resistance (LIN et al.,

2014). Recently, three different research teams developed the CRISPR/Cas9 system for gene deletion in *R. toruloides* (JIAO et al., 2019; OTOUPAL et al., 2019; SCHULTZ, CAO and ZHAO, 2019).

Despite recent advances in the field of metabolic engineering, efficient genetic manipulation tools for *R. toruloides* are still lacking due to two main factors: low homologous recombination efficiency and the lack of autonomously replicating sequence, which makes mandatory the integration of heterologous genes into the yeast genome.

In summary, there is still little knowledge about the metabolic pathways and their interconnections in *R. toruloides*, which makes the design of efficient strategies to improve yeast performance hard. In addition, lack of efficient and standardised genetic manipulation tools make the yeast transformation process very time-consuming. Thus, the use of *R. toruloides* as a cell factory depends on the development of efficient metabolic engineering techniques together with systems biology analysis, not only to increase the production levels of compounds naturally produced by this yeast but also to expand the range of products of industrial interest.

1.1 Objective

This work aimed to identify metabolic regulation of *R. toruloides* which trigger the increase in lipid and carotenoid from xylose through the imposition of oxidative stress conditions and system biology analysis. Oxidative stress was established by light irradiation or the presence of hydrogen peroxide (H_2O_2). The H_2O_2 was also used as a selective pressure to develop a more robust strain through adaptive laboratory evolution.

Furthermore, it was proposed to design a strategy to improve the carotenoid production in *R. toruloides* via metabolic engineering together with the development of a robust and standardised synthetic biology platform using Golden Gate Assembly methodology.

1.2 Work plan

This work was carried out according to the following steps (Figure 1.1):

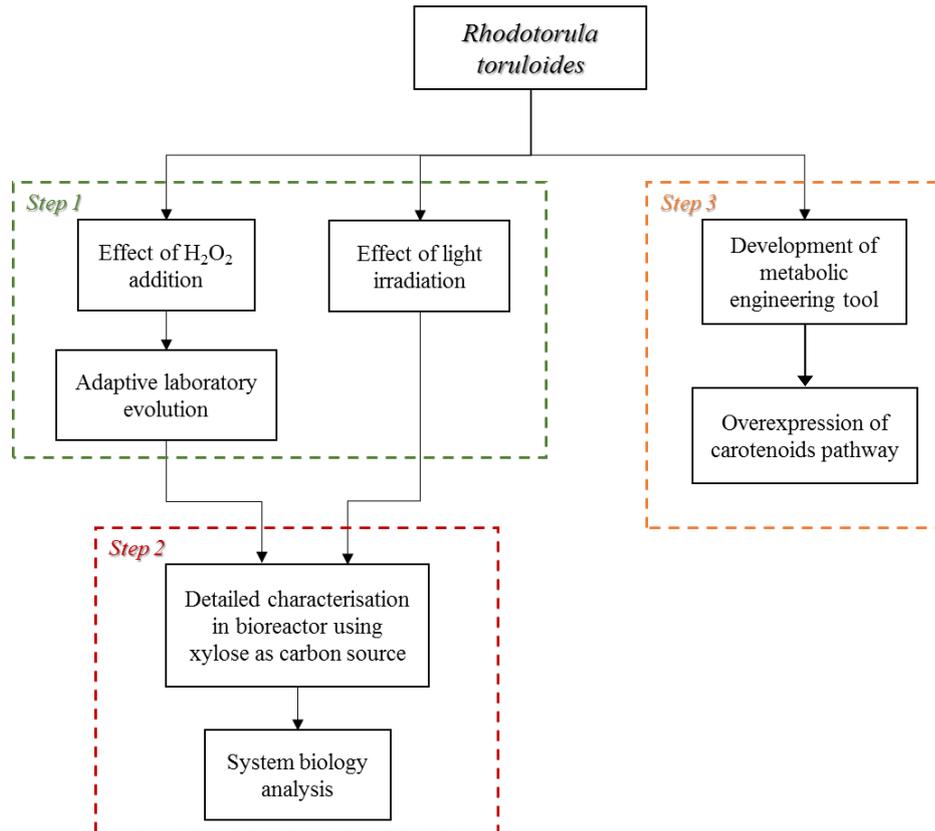
Step 1: Effect of oxidative stress on *R. toruloides* cultivation

In this step, the oxidative stress by addition of H_2O_2 or light irradiation in the *R. toruloides* cultivation was evaluated in terms of cell growth, substrate consumption and production of carotenoids and lipids. As a defence mechanism against the oxidative condition, the production of these compounds could be boosted, mainly the carotenoid synthesis due to

their strong antioxidant properties. H_2O_2 was added at the beginning of cultivation at a determined concentration or light irradiation was applied throughout the cultivation.

The adaptive laboratory evolution (ALE) was also applied using H_2O_2 as selective pressure to select cells with a favoured phenotype for the production of carotenoids.

Figure 1.1. Flowchart representing the steps of the work plan.



Step 2: Physiological characterization of *R. toruloides*

The understanding of microorganism metabolism is an efficient approach to boost its use at industrial levels. Therefore, this step consisted of a detailed physiological characterisation of *R. toruloides* using xylose as sole carbon source and under oxidative stress. These conditions were also analysed in terms of absolute proteomic data and genome-scale modelling. Xylose was the only carbon source used since it is the second most abundant sugar in lignocellulosic materials, potential substrates for biorefineries.

Step 3: Metabolic engineering to overexpress genes related to carotenoid production

The absence of an efficient genetic manipulation tool for *R. toruloides* has become one of the barriers to the development of economically viable processes for this yeast. Therefore, in this step, it was proposed to develop a standardized genetic manipulation system for *R. toruloides* using the Golden Gate Assembly method. This tool was used for the overexpression of genes related to the production of carotenoids (CRTE, CRTI and CRTYB).

1.3 Text outline

This work contains seven more chapters. Chapter 2, Literature review, comprises a survey of knowledge already reported in the literature related to the subject of the thesis. This review was used to contextualize and support the thesis hypothesis and to plan, execute and analyze the experiments. The experimental studies carried out in this work are presented in chapters 3 to 6. In each of them, a brief introduction is presented highlighting the subject and objective of the chapter, the main results and discussions, along with a conclusion.

Chapters 3 and 4 present the study of the effect of light irradiation and the presence of H₂O₂ on *R. toruloides* cultivation, respectively. Chapter 4 also reports the ALE application to obtain a strain with higher production of carotenoids. The study of system biology is presented in chapter 5 and was carried out in Synthetic Biology Laboratory at University of Tartu (Estonia) led by Dr Petri-Jaan Lahtvee. Chapter 6 presents the development of a metabolic engineering tool for *R. toruloides*, and its application to improve the production of carotenoids. Finally, Chapter 7 is presented the final conclusion along with suggestions for future work. In the end, the “References” used to base this work are listed.

CHAPTER 2

Literature review

2.1 Biorefineries

The bioeconomy has been growing as a fundamental role in the development of a sustainable society. This emerging concept defends an economy based on renewable sources, aiming to minimize environmental impacts and help overcome challenges such as food security and the energy crisis (UBANDO; FELIX; CHEN, 2020b).

Given this scenario, biorefineries are considered an important strategy for the expansion of the bioeconomy. The biorefinery, analogous to a refinery, is defined by the use of biomass as a raw material for the generation of a variety of fuels, energy, chemicals and high-value compounds. The objective is to maximize the use of raw material and minimize waste and emissions via the integration of physical-chemical, enzymatic or biological processes (CHERUBINI; WELLISCH; WILLKE, 2009).

Recently, there is an increasing effort to incorporate lignocellulosic biomass in biorefineries. This bioresource is the most abundant biological source, totalling a global production of 1.3 billion tons per year (BARUAH et al., 2018), and does not require extra arable land and does not compete with the food sector. Several lignocellulosic materials can be used in biorefinery, including agricultural residues (such as wheat and rice straw), forest residues, paper pulp and sugarcane bagasse (VALDIVIA et al., 2016).

Lignocellulosic biomass is mainly composed of three polymers: cellulose ($(C_6H_{10}O_5)_n$), hemicellulose ($[C_5H_8O_4]_m$), and lignin ($[C_9H_{10}O_3(OCH_3)_{0.9-1.7}]_x$) with a content of 30–60%, 20–40% and 15–25%, respectively (BARUAH et al., 2018). After thermal, acidic or enzymatic hydrolysis, fermentable sugars are obtained from this material. The cellulose fraction generates glucose (hexoses, called C6 sugars). The hemicellulosic fraction results mainly in pentoses (C5 sugars), mostly composed of xylose, and hexoses in smaller amounts, like mannose and galactose. However, these processes usually release toxic compounds that can inhibit microbial growth, such as 5-hydroxymethylfurfural (HMF), furfural, phenolic acids and acetic acid (BARUAH et al., 2018).

Some companies have commercial-scale facilities for the production of second-generation (2G) ethanol or value-added compounds from different lignocellulosic biomass. Some examples are Raizen, GranBio and POET-DSM (cellulosic ethanol), DSM (succinic acid, cellulosic ethanol), Dow-DuPont (1,3-propanediol, 1,4-butanediol), Clariant-Global bioenergies-INEOS (bio-isobutene), Braskem (ethylene, polypropylene), Amyris (farnesene), and several other potential players (CHANDEL et al., 2018).

Although traditional bioethanol production (first generation) is well known and established, the production of 2G ethanol from lignocellulosic biomass is still challenging concerning the efficient utilization of pentoses. *Saccharomyces cerevisiae* is unable to metabolize xylose naturally, and even after great progress in metabolic engineering research, the engineered strains still grow at slow rates on xylose (PATEL; SHAH, 2021). Among the microorganisms, yeasts are considered more favourable to large-scale production compared to microalgae, mainly due to their high growth rate that results in higher productivity (BHOSALE; GADRE, 2001; MATA-GÓMEZ et al., 2014).

To expand the product range and improve the efficiency of lignocellulose hydrolysate conversion, several non-conventional yeasts, which can naturally assimilate xylose and produce industrially important compounds, have been studied to avoid the intense metabolic engineering effort and establish an economically viable integrated biorefinery. *R. toruloides* has been shown a great potential for this application. Oliveira et al. (2021) studied the performance of nine non-conventional yeasts in hemicellulosic hydrolysates, and an *R. toruloides* strain highlighted due to the highest specific uptake rate of the total substrate, and the ability to grow in concentrated hydrolysates with high concentration of sugar (145 g/L) and inhibitor, showing no detectable lag phase.

2.2 *Rhodotorula toruloides* as a cell factory

Rhodotorula toruloides, also known as *Rhodospidium toruloides*, is a red yeast, belonging to the *Sporidiobolaceae* family, *Sporidiobolales* order, *Microbotryomycetes* class and Basidiomycota phylum. This yeast is heterothallic and classified as an oleaginous microorganism as it is able to accumulate more than 20% of its dry biomass in lipids. In nature, *R. toruloides* is found in a wide range of habitats including surfaces of leaves, soil, and seawater (SAMPAIO, 2011; RATLEDGE, 1991; WEN et al., 2020).

This yeast has been considered a promising microorganism for biotechnological applications due to its ability to grow at high cell density (LI; ZHAO; BAI, 2007) and show great growth in a wide range of temperature and pH (GADANHO; LIBKIND; SAMPAIO,

2006). It can also assimilate different carbon sources, including monosaccharides (hexoses and pentoses), disaccharides (sucrose, maltose, and cellobiose), polysaccharides (starch), alcohols (glycerol), and organic acids (acetic and lactic acid) (DINH et al., 2019; LOPES et al., 2020a; WEN et al., 2020). *R. toruloides* also exhibits tolerance to toxic compounds found in lignocellulosic hydrolysates (BONTURI et al., 2017; LOPES et al., 2020b; OLIVEIRA et al., 2021).

R. toruloides can accumulate a high lipid content, up to 68% of its biomass (LI; ZHAO; BAI, 2007), and also can synthesize carotenoids, a high-value compound which the selling price reached USD 350 - 7500/kg in 2017, depending on the type of carotenoid (GLOBENEWSWIRE, 2017). According to Koutinas et al. (2014), the valorization of co-products becomes essential for the economic viability of biodiesel production from microbial oils.

Furthermore, aiming at *R. toruloides* use in biorefineries, the efficient conversion of both hexoses and pentoses into products of interest is necessary. However, the knowledge about the catabolism of xylose by this yeast is little, which makes it difficult to design strategies to increase the efficiency of use of this substrate. Wiebe et al. (2012) reported that the lipid production in *R. toruloides* CBS 14 from xylose can be up to 40% less efficient compared to cultivation on glucose, which evidences the need to study the xylose assimilation.

2.2.1 Xylose metabolism in *R. toruloides*

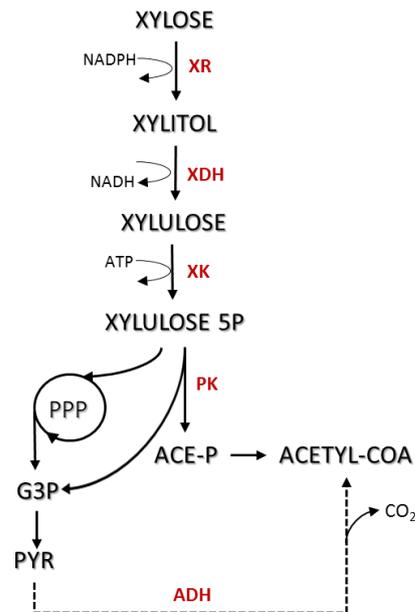
The conversion of xylose into lipids and carotenoids in *R. toruloides* is still poorly investigated. Some works report that xylose assimilation involves the NADPH- and NADH-dependent enzymes (Figure 2.1) (JAGTAP; RAO, 2018; LOPES et al., 2020a; TIUKOVA et al., 2019a).

Xylose reductase (XR) reduces xylose to xylitol, which is oxidized to xylulose by xylitol dehydrogenase (XDH). Then, xylulose is phosphorylated to xylulose 5-phosphate (xylulose 5-P) by the enzyme xylulose kinase (XK). This molecule can enter the pentose phosphate pathway (PPP) and it is converted to glyceraldehyde 3-phosphate (G3P). Through glycolysis, G3P originates pyruvate (PYR) which is decarboxylated to acetyl-CoA, a precursor for the synthesis of lipids and carotenoids.

R. toruloides has an alternative route to xylulose 5P, through the enzyme phosphoketolase (PK), originating acetyl-phosphate (ACE-P) and G3P. This route allows a higher efficiency in carbon metabolism since there is no loss of carbon in the form of CO₂ (a

loss that can represent a third of the carbon from the substrate) (PAPANIKOLAOU; AGGELIS, 2011; TIUKOVA et al., 2019a).

Figure 2.1. Schematic of xylose catabolism by *R. toruloides*. The dashed arrow represents multiple reactions.



Jagtap e Rao (2018) detected the production of D-arabitol by *R. toruloides* cultivated in xylose. The authors suggested that this molecule is produced from xylulose by NADH-dependent arabitol dehydrogenase. Lopes et al. (2020a) also identified the production of co-products during xylose catabolism in which one-third of the consumed substrate was converted to xylitol or arabitol. The accumulation of these co-products can be associated with a possible mechanism to keep the redox balance during xylose assimilation.

2.2.2 Synthesis of lipids and carotenoids by *R. toruloides*

Microbial lipids, also known as single-cell oil (SCO), composed mainly of TAGs, have great potential to be used as biodiesel and in the cosmetics, surfactants, cocoa butter and animal feed industries (BLOMQVIST et al., 2018; UNREAN; KHAJEERAM; CHAMPREDA, 2017; WEI; SIEWERS; NIELSEN, 2017; YANG et al., 2018). SCO has advantages over vegetable oils as raw material for biodiesel production: a faster production cycle, no need for large areas of arable land and is not affected by weather conditions (PARK; NICAUD; LEDESMA-AMARO, 2017).

In general, fatty acids produced by *R. toruloides* are palmitic (C16:0), stearic (C18:0), oleic (C18:1) and linoleic (C18:2) acids. This microorganism can also produce high-value lipids through genetic modifications. For example, fatty acids with very long chains such as erucic acid (C22:1) and nervonic acid (C24:1), used as important raw materials for the production of plastics, cosmetics, nylon and lubricants (FILLET et al., 2017).

Carotenoids represent one of the most attractive groups of pigments found naturally (ARMSTRONG; HEARST, 1996). In addition to their colouring property, these compounds are widely known for their antioxidant activity and for being a precursor of vitamin A (KRINSKY, 1989; BENDICH, OLSONT, 1988). They are used in food industries as dyes and nutritional supplements; in pharmaceutical industries as cosmetic additives and as diseases prevention by reducing the risk of degenerative diseases such as cancer, cardiovascular disease, multiple sclerosis and cataracts (BUZZINI et al., 2007; EDGE; MCGARVEY; TRUSCOTT, 1997; MATA-GÓMEZ et al., 2014).

The market demand for carotenoids has been supplied by extraction from fruits and vegetables, such as tomatoes and carrots. However, the largest production occurs through chemical synthesis (ARMSTRONG; HEARST, 1996; JL LEE; WN CHEN, 2016; MATA-GÓMEZ et al., 2014). In chemical synthesis, despite the low production cost and high purity, there is the generation of a large amount of polluting waste and the growing concern about the negative effects of synthetic products on human health (AUSICH, 1997; KOT et al., 2016; MATA-GÓMEZ et al., 2014). *R. toruloides* mainly produces four types of carotenoids: torularhodin, torulene, γ -carotene and β -carotene, the proportion of which is dependent on the strain and growing conditions (BUZZINI et al., 2007; CARDOSO et al., 2016; FRENGOVA; BESHKOVA, 2009).

The metabolic pathways of lipid and carotenoid synthesis have acetyl-CoA as a precursor (Figure 2.2). At high C/N ratios, the production of this precursor is induced resulting in higher yields of these products (BRAUNWALD et al., 2013; ZHANG; ZHANG; TAN, 2014; ZHU et al., 2012). The most important steps of these pathways are:

1. Under nitrogen limitation (at low C/N ratio), adenosine monophosphate (AMP) molecules undergo deamination, then NH_4^+ can be used as a nitrogen source. As a result, the activity of isocitrate dehydrogenase (IDH, responsible for the oxidation of isocitrate to alpha-oxoglutarate, α -KG, in the tricarboxylic acid cycle - TCA) is reduced, as this enzyme is dependent on AMP. Thus, by a reversible reaction, the excess of isocitrate is converted into citrate, which is exported via malate/citrate transport to the cytoplasm and is cleaved by enzyme ATP citrate lyase (ACL) in oxalacetate and acetyl-CoA.

Oxalacetate will be converted to pyruvate with the aid of the malic enzyme (EM). This enzyme is considered a key enzyme in oleaginous yeast as it is responsible for supplementing NADPH, a cofactor in great demand for lipid biosynthesis.

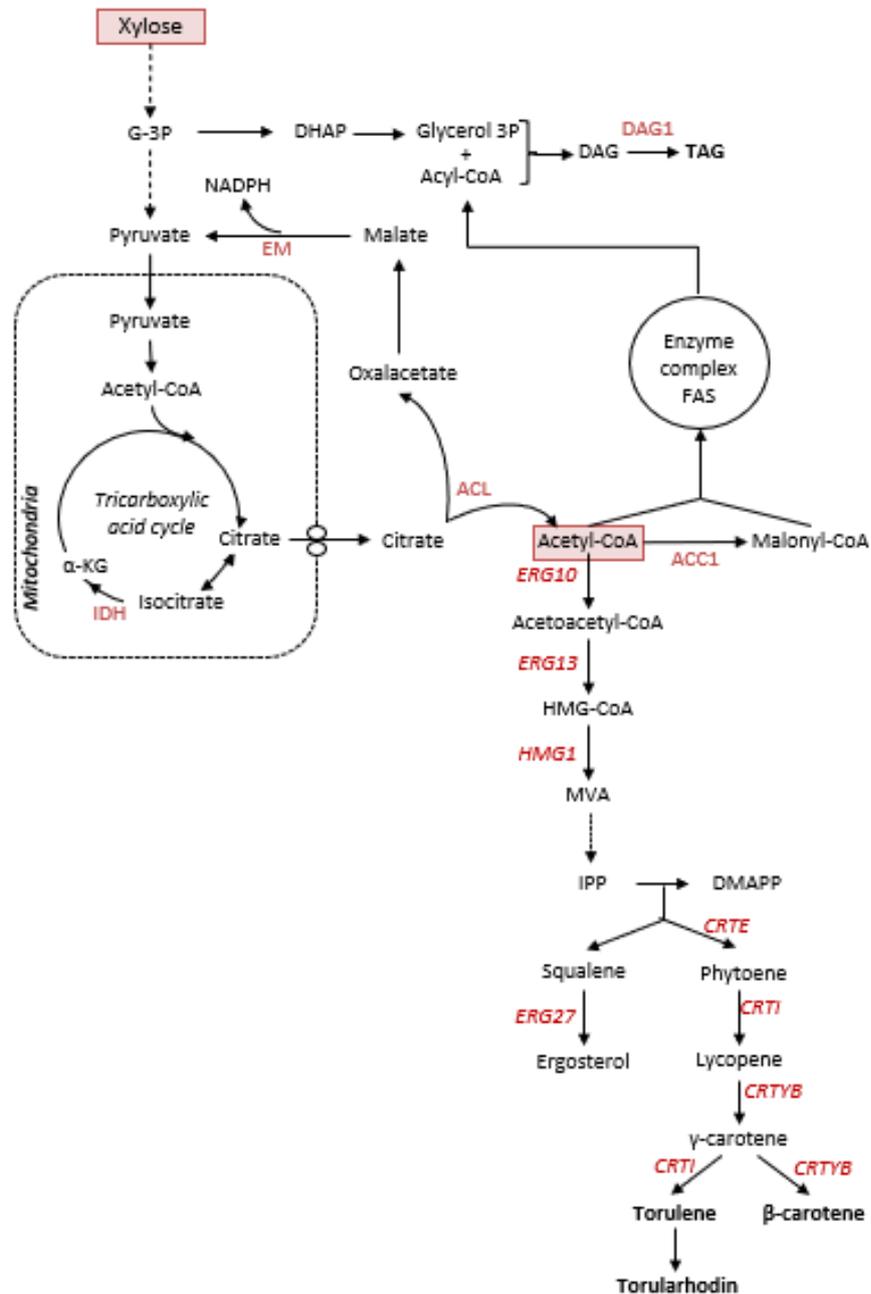
2. Acetyl-CoA can be carboxylated into malonyl-CoA by acetyl-CoA carboxylase (ACC1), which will be used in the synthesis of triacylglycerides that occurs in the fatty acid synthetase complex (FAS).
3. Another alternative pathway to acetyl-CoA is the synthesis of isoprenoids. Therefore, acetyl-CoA is converted into 3-hydroxymethylglutaryl-CoA (HMG-CoA) by acetyl-CoA C-acetyltransferase (ERG10) and HMG-CoA synthetase (ERG13). This molecule enters the mevalonate pathway and is reduced to mevalonate (MVA) by HMG reductase (HMG1). After some reactions, phytoene (the first carotenoid of the pathway) and squalene (the precursor of steroid production) are synthesised competitively.

Phytoene is converted into γ , β -carotene and torulene by phytoene dehydrogenase (CRTI) and phytoene synthetase (CRTYB), and torularhodin is formed by hydroxylation and oxidation reactions.

In addition to the C/N ratio, other cultivation factors can influence the production of lipids and carotenoids, such as oxidative stress. Some studies have described the increased production of carotenoids by inducing oxidative stress, such as the presence of hydrogen peroxide (H_2O_2) and light irradiation in cultivation. The presence of H_2O_2 at an initial concentration of 5 mmol/L resulted in a five-fold increase in carotenoid production by *Rhodotorula mucilaginosa* (IRAZUSTA et al., 2013). Under light irradiation (4000 lux), the production of carotenoids and lipids by *Rhodotorula glutinis* increased by 60% (GONG; ZHANG; TAN, 2020).

The oxidative stress response mechanism in yeast is still unclear. This stress condition is associated with the presence of reactive oxygen species (ROS), including H_2O_2 , superoxide (O_2^-) and hydroxyl radicals ($\text{HO}\cdot$). ROS are potent oxidants that can damage all cellular components such as DNA, lipids and proteins. Microorganisms have two types of defensive systems against oxidative damage: enzymatic and non-enzymatic.

Figure 2.2. Metabolic pathways involving the conversion of xylose into lipids, ergosterol and carotenoids by oleaginous yeasts. Legend: G3P, glyceraldehyde 3-phosphate; DHAP, dihydroxyacetone phosphate; DAG, diacylglyceride; TAG, triacylglycerides; CRTE, geranylgeranyl diphosphate synthetase; CRTI, phytoene dehydrogenase; CRTYB, phytoene synthetase.



The first is composed of enzymes such as superoxide dismutase which converts the superoxide radical into H_2O_2 and oxygen, and catalase and peroxidase which breaks down H_2O_2 into water and oxygen. The non-enzymatic system is directly involved in the elimination of ROS by the recycling of oxidized compounds, such as ascorbate, glutathione, alpha-tocopherol and carotenoids (IRAZUSTA et al., 2013).

In addition to being possible to increase the production of lipids and carotenoids by changing culture factors, another alternative to increase production is the development of super-producer strains based on adaptive evolution or metabolic engineering techniques.

2.3 Adaptive laboratory evolution

Adaptive laboratory evolution (ALE) has been successfully used to improve cell performance through an artificial selection of microbial cells grown under specific long-term growth conditions. Usually, the objective of this technique is to improve microbial growth on relevant substrates or tolerance to some stress (DRAGOSITS; MATTANOVICH, 2013).

During ALE, selective pressure is applied for a long period which allows cells with an advantageous phenotype (due to random genetic alterations) to become dominant in the microbial population. This strategy can run for weeks or even years. As an advantage over metabolic engineering, ALE is cheaper, simpler, and does not require detailed knowledge of cell metabolism (KOPPRAM; ALBERS; OLSSON, 2012; SAUER, 2001).

Mutagenesis is another widely used technique to select better phenotypes. In this strategy, the mutation can be induced by physical treatment (such as UV radiation) or chemical agents (such as nitrosoguanidine). Compared to ALE, the selective process in mutagenesis is faster due to high mutation rates; however, it generates a wide variety of mutants, making the screening step more expensive (SAUER, 2001).

Reyes et al. (2014) successfully used ALE to improve carotenoid production in a genetically modified strain of *Saccharomyces cerevisiae*. In this strategy, cells were periodically exposed to high concentrations of H_2O_2 (1 mol/L). After 30 cycles, the carotenoid content increased threefold (from 6 to 18 mg/gDCW). Transcriptomic analyzes showed that the regulation of genes directly involved in the synthesis of carotenoids had no significant changes. However, the ERG13 gene encoding hydroxymethylglutaryl-CoA synthase (HMG1) involved in the mevalonate pathway was upregulated. Additionally, the lipid synthesis pathway was also more active in the adapted strain.

2.4 Metabolic engineering

The lack of efficient genetic engineering tools for *R. toruloides* has been considered one of the main obstacles to establishing this yeast as a cell factory (PARK; NICAUD; LEDESMA-AMARO, 2017). Thus, the construction of an efficient and standardized platform for assembling the DNA sequence for *R. toruloides* can boost the genetic modifications of this yeast.

The last decade has seen a breakthrough in synthetic biology that has led to the creation of several robust gene assembly platforms that can be exploited for metabolic engineering. These platforms can be divided into three methods based on oligonucleotide bridging, sequence homology, and restriction enzymes (CHAO; YUAN; ZHAO, 2015).

The ligase cycling reaction (LCR) based on the oligonucleotide bridging method uses single-stranded oligonucleotides complementary to the ends of adjacent DNA parts. These DNA structures are joined by thermostable ligase through several temperature cycles of denaturation-annealing-ligation, assembling the complex DNA constructs (KOK et al., 2014).

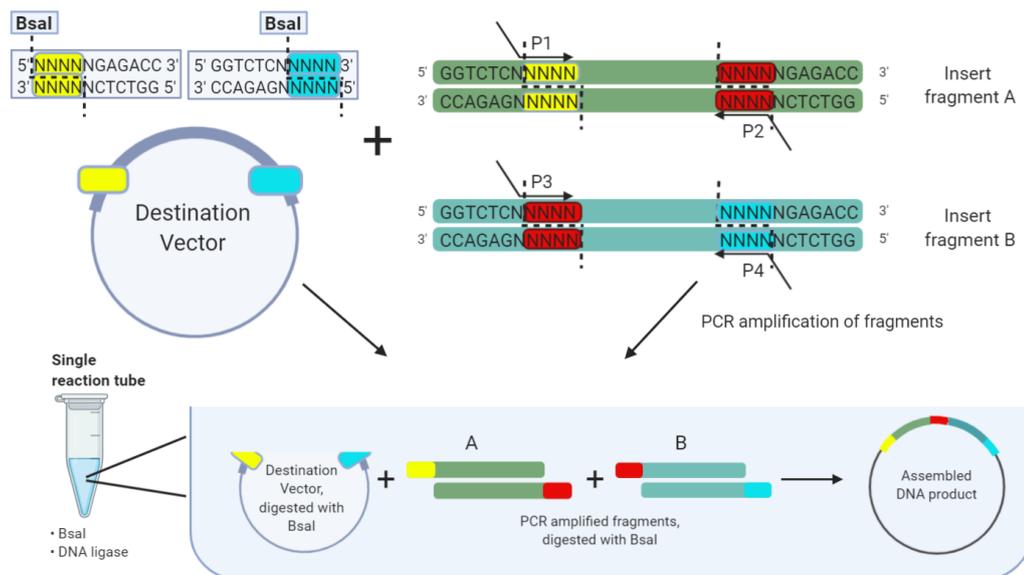
The SLIC method (sequence and ligation-independent cloning) and Gibson method are examples of methods based on in vitro homologous recombination and single-strand annealing. In the SLIC technique, T4 DNA polymerase is used to generate single-stranded overhangs, which anneal in a later step through the endogenous DNA repair mechanism, building the recombinant DNA sequence. On the other hand, Gibson method takes place in a single step and requires T5 exonuclease in combination with Phusion polymerase and DNA ligase, which can improve the efficiency of assembling several pieces of DNA. Furthermore, the higher temperature used in the Gibson method (50°C versus 37°C in the SLIC method) can limit the formation of secondary structures at the ends of the fragments. The main advantage of SLIC is the low cost of the T4 polymerase compared to the enzymes needed for the Gibson method. (GIBSON et al., 2009; LI; ELLEDGE, 2012).

Golden Gate Assembly (GGA) is one of the methods that use restriction enzymes for DNA sequence assembly and has been considered as one of the most robust techniques. The GGA strategy is to create a library of standardized DNA parts that can subsequently be assembled into an expression cassette in a single step (ENGLER; KANDZIA; MARILLONNET, 2008).

In this method, type II restriction enzymes bind to recognition sites (eg., GGTCTC for the BsaI enzyme) and cut the DNA generating fragments with overhangs (with four nucleotides, in the case of BsaI). After digestion, the fragments lose the restriction enzyme recognition site, making further cuts impossible. The assembly of the digested fragments occurs by annealing

the complementary overhang of the adjacent fragment through the enzyme T4 or T7 DNA ligase (Figure 2.3). GGA propose the possibility of standardizing DNA fragments, being able to exchange these fragments easily and thus making the diversification and optimization of engineering pathways a less laborious task. This technique has not yet been applied to *R. toruloides*; however, it was used to create a platform for *Yarrowia lipolytica* and its efficiency was proven by the overexpression of carotenogenic genes (CELINSKA et al., 2017).

Figure 2.3. Golden Gate Assembly Method. The fragments are amplified via PCR with the respective primers, inserting the BsaI restriction sites and pre-designed 4-nucleotide overhangs. Amplicons and the target vector are incubated with the enzyme BsaI and DNA ligase in the same reaction tube. Then, as digestion occurs, the DNA parts are linked through the annealing of complementary overhangs and the action of DNA ligase, assembling the final DNA sequence.



Source: Adapted from <https://international.neb.com/golden-gate/golden-gate>. Create on BioRender.com

In recent years, the emergence of studies aimed at improving the production of lipids in *R. toruloides* through metabolic engineering has been observed. Zhang et al. (2016a) used the LCR technique to construct the plasmid with the native genes acetyl-CoA carboxylase (ACC1) and diacylglycerol acyltransferase (DAG1). The engineered strain was able to produce two-fold more lipids (16.4 g/L lipid concentration) compared to the parental strain.

Overexpression of malic enzyme (ME) resulted in a 24% increase in lipid production (ZHANG et al., 2016b). In this case, the authors used the Gibson method to construct the expression cassette. In the work by Yang et al. (2018), a plasmid with the phosphotransacetylase gene (an enzyme involved in the phosphoketolase pathway) was constructed through assembly with restriction enzymes, EcoRI and SpeI. The strain transformed with this construction showed a 54% higher lipid production compared to the parental strain.

Metabolic engineering can facilitate the development of highly efficient cell factories to produce compounds of interest. However, for a successful application of this technique, it is necessary to understand the metabolic processes of a microorganism and how they are interconnected, making it possible to detect metabolic bottlenecks.

2.5 Systems biology

Systems biology aims to understand the metabolic regulation of organisms, through data generated from omic approaches and metabolic models, by which it is possible to have an overview of all biological processes in the cell and how they are interconnected with each other. The omic techniques include (PARK; NICAUD; LEDESMA-AMARO, 2017; DEBNATH et al., 2010):

- a) Genomics: the study of the complete genome of the organism from DNA sequencing to understand the structure, organization and functions of genes;
- b) Transcriptomics: the study of an RNA transcription set that identifies the genes that are being expressed by the microorganism. The comparison between transcriptomes allows the identification of differences in gene expression between different conditions;
- c) Proteomics: analyzes and quantifies the set of all proteins expressed by the cell. In this study it is possible to characterize the information flow inside the cell through protein expressions and to quantify the post-translational modifications;
- d) Phosphoproteomics: identifies proteins that contain a phosphate group as a post-translational modification. Protein phosphorylation is an important regulatory mechanism for protein function in cells, as it alters cell activity. Thus, it is considered a signal of response to external stimuli, as it can indicate the activation of a protein or metabolic pathway;
- e) Metabolomics: allows detecting and quantifying small molecules (such as amino acids, fatty acids, and carbohydrates) that are intermediates or end products of metabolic

processes. The detected metabolites allow the identification of active metabolic processes.

In recent years, systems biology including omics data has been investigated to understand the genomic organization and metabolic pathways in *R. toruloides*. Complete genome from several *R. toruloides* strains have been sequenced (CORADETTI et al., 2018; HU; JI, 2016a; KUMAR et al., 2012; MORIN et al., 2014; SAMBLES et al., 2017; ZHU et al., 2012) and many studies have reported the metabolic response of this yeast in different environmental factors.

The lipid formation process during different growth phases of cultivation was investigated through proteomic analysis (LIU et al, 2009). It was the first proteomic study that provided insights into how the nitrogen limitation results in lipid overproduction. Further, a study approaching the transcriptomic and proteomic analysis provided more details about the metabolism modification under nitrogen limitation (ZHU et al., 2012). Under this condition, the ribosome biogenesis and protein translation were suppressed and genes related to degradation/recycling of proteins or nitrogenous compounds were upregulated. In addition, proteins involved in lipogenesis were found at elevated levels and the molecules of pyruvate and acetyl-CoA were directed to fatty acid biosynthesis.

A metabolomic study has investigated the regulation for carotenoids production from glycerol in three different strains (LEE et al., 2014). The lower production of TCA intermediate metabolites results in the canalization of metabolic flux to carotenoids production.

A lignocellulosic hydrolysate-tolerate *R. toruloides* was investigated through transcriptomic and proteomic approaches (QI et al, 2017). Phosphoglucomutase (PGM) and phosphoenolpyruvate carboxykinase (Pck1) related to glycolysis and TCA cycle were upregulated, producing more energy to improve the stress tolerance. Aldehyde dehydrogenase (ALDH) and alcohol dehydrogenase (ADH) were also highly expressed, improving the tolerance to acetate present in the hydrolysates.

Multi-omics analyses were employed to investigate the phosphate limitation as an inducer of lipid production (WANG et al., 2018b). The phosphate limitation results in the upregulation of phosphate metabolism, RNA degradation and triacylglycerol synthesis and downregulation of ribosome biosynthesis and TCA cycle. This limitation led to dephosphorylation of AMP, resulting in the downregulation of isocitrate dehydrogenase (Idh1) and, consequently, in the citrate accumulation which is converting into acetyl-CoA, precursor of lipid synthesis. The NADPH is regenerated mainly by the malic enzyme since the flux through the PPP is reduced.

Proteome analysis was used to compare the differences between the conversion of xylose and glucose into lipids (TIUKOVA et al., 2019a). The xylose metabolism showed a higher level of enzymes involved in sugar transport, peroxisomal β -oxidation and oxidative stress response compared to glucose metabolism. NADPH is regenerated is mainly through PPP, since this pathway was upregulated during the lipid accumulation. Some enzymes involved in the carotenoids pathway were upregulated in xylose-grown cells (geranylgeranyl diphosphate synthase - crtE) and the stationary phase for both sugar metabolisms (phytoene dehydrogenase – crtI).

Genome-scale metabolic models (GEM) are another important tool to understand and provide a holistic view of metabolic networks in an organism. GEMs are derived from available genome sequences and can be used to identify potential targets for metabolic engineering and select more efficient pathways (KERKHOVEN; LAHTVEE; NIELSEN, 2015). Concerning *R. toruloides*, there are two GEMs available (DINH et al., 2019; TIUKOVA et al., 2019b). Lopes et al. (2020a) reported the first study in which physiological data from cultures of *R. toruloides* were combined with GEM. The authors showed that GEM analysis pointed out the malic enzyme and phosphoketolase pathway as possible targets to improve lipid production.

CHAPTER 3

Effect of light irradiation on production of lipids and carotenoids in *R. toruloides*

3.1 Introduction

Carotenoids are tetraterpenoids containing 40 carbon atoms in a long-conjugated chain with 9 to 11 double bonds. Due to this conjugated system, carotenoids have high reducing power and the ability to transfer electrons throughout the molecule, scavenging the free radicals by trapping peroxy radicals and quenching the singlet oxygen. This electron-rich system plays an important role in light absorption for photosynthetic organisms (TAN; NORHAIZAN, 2019). These molecules transfer this energy to chlorophylls, initiating the photosynthesis and, also regulating the flow of energy within the photosynthetic apparatus, protecting the cell against photooxidation caused by excess light absorption (POLIVKA AND FRANK, 2010).

Carotenoids also play an important protection mechanism against photo-oxidative damage in non-photosynthetic organisms, since the production of these compounds is increased under light irradiation conditions. Few works in the literature demonstrated the effect of light on carotenoid production in *Rhodotorula* genus. Zhang et al. (2014) reported that white light irradiation (800 $\mu\text{mol}/\text{m}^2\text{s}$) led to a two-fold increase in carotenoid production. Yen and Yang (2012) also reported the positive effect of light irradiation on biomass production in *Rhodotorula glutinis* cultivation (25 to 50% increase, depending on light colour). Gong et al. (2019) investigated the metabolism involved in irradiation-induced stress protection in *R. glutinis* through multi-omics analysis. One of the main findings of that work was the upregulation of the carotenoid biosynthesis pathway, confirming the carotenoids indeed are directly involved in yeast defence mechanism against oxidative stress. Therefore, the study presented in this chapter aimed to investigate the effect of different light colours (white, red and blue) on *R. toruloides* cultivation in terms of production of lipids and carotenoids.

3.2 Materials and methods

3.2.1. Strain and inoculum

Rhodotorula toruloides was obtained from Fundação André Tosello (Campinas, Brazil) as *Rhodospiridium toruloides* CCT 7815 and stored at -80°C in 10% (v/v) glycerol. Cultivation inoculum was prepared in YPD medium at 200 rpm and 28°C for 24 h (shaker TE-4200, Tecnal, Brazil). The cells were washed with 0.9% (v/v) NaCl solution before inoculation.

3.2.2. Analytical methods

Cell growth was monitored by measuring the optical density (OD) at 600 nm (spectrophotometer Quimis, Brazil). The dry cell mass (DCW) was determined gravimetrically after drying at 60°C for 24h. Reducing sugars were measured based on DNS method (MILLER, 1959).

Lipids were extracted according to an adaptation of the Folch method (FOLCH, 1956) described by Bonturi et al. (2015). Briefly, a mixture of chloroform and methanol (2:1 v/v) was added to dried cells. After 24 h, the solvent was evaporated in a rotary evaporator (TE-211, Tecnal, Brazil) and the total lipid content was determined gravimetrically.

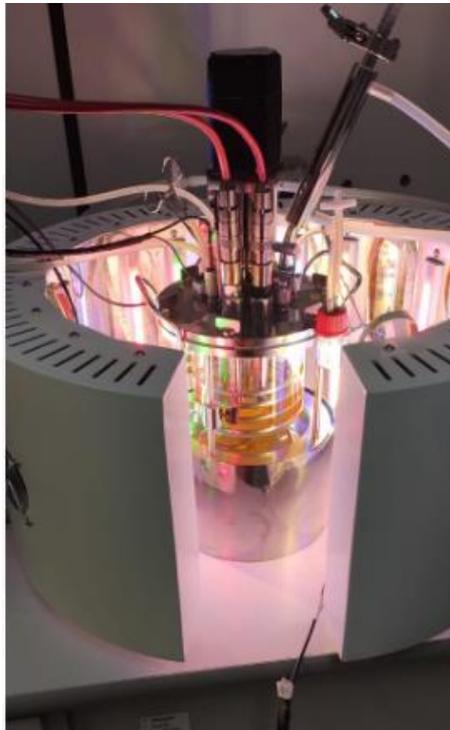
The carotenoid quantification was done based on the method described by Lee et al. (2014). Briefly, the cells from 2 mL of cultivation broth were harvested by centrifugation, washed twice in distilled water and resuspended in 1.0 mL of acetone. The cells were lysed with acid-washed glass beads (400 – 650 μm) by vortexing for three cycles (at full speed for 30 s). 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 under nitrogen gas and the remaining extract was resuspended in a known volume of acetone. Carotenoids were measured using HPLC (Thermo Fisher, USA) equipped with a reverse phase column RP-18 (Acclaim 120, C18 4.6×250 mm). 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 and the absorption spectra in the 190-800 nm photodiode array detector (modified from WEBER et al., 2007). All identified peaks were quantified using the β -carotene standard (Sigma Aldrich, USA). Detected peaks were identified according to the known carotenoid retention time profile (LEE et al., 2014; WEBER; ANKE; DAVOLI, 2007). The chromatogram obtained by HPLC analysis of the carotenoids produced by *R. toruloides* CCT7815 and the peaks spectrum are shown in Appendix Figure A1.

3.2.3. Yeast cultivation

Batch cultivations were performed in 1-L bioreactor (BioTec Pro, Tecnal, Brazil) equipped with a LED light system (BioTec LED D1, Tecnal, Brazil; Figure 3.1). The working volume was 500 mL and pH was controlled at 6.0 by the addition of 2 mol/L KOH. Dissolved oxygen was maintained at higher 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 light irradiation throughout the experiment. Different LED lights were used: warm white, red, or blue (the respective light absorption spectra are shown in Appendix Figure A2). Samples for dry cellular mass and extracellular metabolite analysis were collected every 24 h and for carotenoids and lipids were collected at the end of cultivation.

The medium composition was (per litre): 70.0 g xylose, 2.0 g yeast extract, 0.4 g $(\text{NH}_4)_2\text{SO}_4$, 2.5 g KH_2PO_4 , 2.0 g $\text{MgSO}_4 \cdot 7 \text{H}_2\text{O}$, and 1.0 mL trace metal solution (MEESTERS; HUIJBERTS; EGGINK, 1996), supplemented with 100 μl antifoam 204 (Sigma-Aldrich, USA).

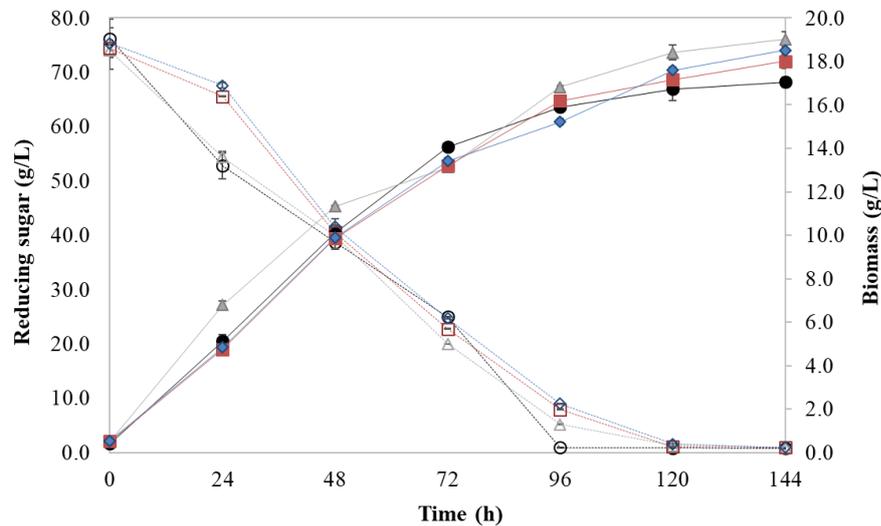
Figure 3.1. Bioreactor equipped with a photoperiod.



3.3 Results and discussion

The effect of three different irradiation conditions (white, red, and blue lights) on cell growth and production of lipids and carotenoids were investigated in *R. toruloides* cultivation and compared to the reference condition (without irradiation) (Figure 3.2 and Table 3.1). In these experiments, the light intensity was kept at $130 \mu\text{mol}/\text{m}^2\text{s}$ (ca. 9700 lux).

Figure 3.2. Effect of light irradiation on cell growth (filled symbols) and sugar consumption (open symbols) by *R. toruloides* under reference condition (without irradiation, black circles), and under different light irradiations: white (grey triangles), red (red squares), and blue (blue diamonds).



The stress caused by the light irradiation did not inhibit the cell growth and the sugar consumption profile under light irradiation was very similar to the reference condition (Figure 3.2). The maximum growth rate (μ_{MAX}) was not affected under white light irradiation compared to the reference condition (Table 3.1). However, under red and blue light irradiation, μ_{MAX} showed a small decrease (ca. 10%, significant difference according to Tukey test, p -value < 0.05). Interesting, Yen and Yang (2012) reported a considerable increase in μ_{MAX} of *R. glutinis* under light irradiation, even using a higher light intensity ($730 - 900 \mu\text{mol}/\text{m}^2\text{s}$) than the one used in this present study ($130 \mu\text{mol}/\text{m}^2\text{s}$).

The light irradiation increased slightly the biomass yield, Y_{XS} (Table 3.1). The white light condition showed the maximum final biomass, 19 g/L, representing a significant increase of 12% compared to the reference condition (Tukey test, p -value < 0.05). According to Gong

et al. (2019), metabolomics analysis showed that the pathway related to the TCA cycle were found more active under irradiation conditions in *R. glutinis*, which might be an explanation for the higher biomass yield under irradiation since the TCA cycle is tightly connected to amino acid synthesis and further biomass accumulation.

As expected, the carotenoid production increased in all irradiation conditions, reaching an increase of 92% under white light compared to the reference condition. The lipid production increased only under white light condition (32% compared to the reference condition). In red and blue irradiation conditions, the lipid production was not significantly different from the reference condition (Tukey test, p -value < 0.05).

Table 3-1. Effect of irradiation on *R. toruloides* batch cultivation in reference condition (without irradiation) and under different light irradiation.

Condition	μ_{MAX} (h ⁻¹)	Y _{XS} (g _{DCW} /g)	Lipid titre (g/L)	Carotenoid titre (mg/L)
Reference	0.104 ± 0.001	0.220 ± 0.002	8.5 ± 0.3	15.3 ± 0.6
White light	0.108 ± 0.001	0.254 ± 0.003	11.2 ± 0.1	30.3 ± 1.3
Red light	0.091 ± 0.002	0.238 ± 0.003	7.9 ± 0.3	20.7 ± 0.9
Blue light	0.092 ± 0.002	0.241 ± 0.002	7.4 ± 0.1	22.7 ± 1.2

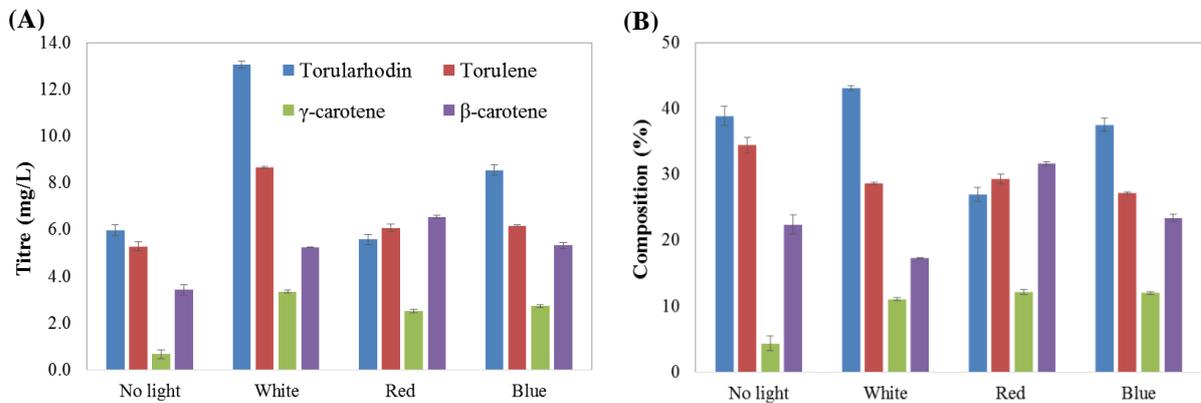
Gong et al. (2019) also reported an increase in the production of carotenoid and lipid by *R. glutinis* cultivated under white light irradiation. Through RNA-seq analysis, the carotenoid pathway was found significantly upregulated under this stress condition. Also, aldehyde dehydrogenase and acyl-CoA synthetase were found upregulated which means that the level of acetyl-CoA, a precursor for both lipids and carotenoids, was increased.

Interesting, each type of light affected the carotenoid composition differently (Figure 3.3). Although the white light irradiation enhanced the production of all types of carotenoids, this condition favoured the torularhodin production increasing its titre by 2.2-fold compared to the reference condition, representing 43% of total carotenoids. Torulene production also had a considerable increase of 63%. These types of carotenoids have attracted the attention of researchers due to their strong antioxidant properties and potential application against cancer and degenerative diseases (KOT et al., 2018).

However, red light irradiation condition favoured the β -carotene production, resulting in a titre increase of 90% compared to the reference condition. These findings are consistent with previous experiments with *R. glutinis*, in which red light irradiation resulted in a higher increase in β -carotene production compared to other colours of light (YEN; YANG, 2012).

Although the γ -carotene represented the less abundant type of carotenoids in *R. toruloides*, this carotenoid had a remarkable increase of 4- to 5-fold in irradiation conditions (Figure 3.3). Besides the antioxidant property, β - and γ -carotene can be converted into vitamin A.

Figure 3.3. Carotenoid titre (A) and composition (B) under white, red, or blue light irradiation compared to the reference condition (without irradiation).



3.4 Conclusion

Light irradiation successfully enhanced the production of carotenoids and lipids in *R. toruloides*, without affecting the cell growth. The white light was more effective in increasing the total carotenoid production, with a two-fold increase compared to the reference condition. Also, under this condition, lipid production increased by 32%. The carotenoid composition was dependent on the type of light. The torularhodin production was higher under white light irradiation, while the red light irradiation favoured the β -carotene synthesis. Therefore, light irradiation can be a tool to improve the production of carotenoids and lipids and also to control the carotenoid composition. In the industrial view, this strategy can be a challenge due to the complex design of stirred tank photobioreactor for yeast cultivation at large-scale. However, in the research view, light irradiation can be a tool to understand better the carotenoid synthesis.

CHAPTER 4

Adaptive laboratory evolution to improve the carotenoid production

4.1 Introduction

Artificial laboratory selection is a traditional approach to obtain microorganisms with industrially desirable phenotypes. This strategy is based on natural selection theory, which claims that an individual organism that better adapts to its environment due to random mutation is more likely to survive and generates more offspring with the same fitness features. Then these individuals become dominant in the population.

Adaptive laboratory evolution (ALE) is a strategy to select cells with improved phenotype by the application of a long-term selective pressure (DRAGOSITS; MATTANOVICH, 2013). This approach has been explored not only to develop strains with superior physiological features but also to delineate a more efficient metabolism via system biology analysis. The generated information can be used for a reverse metabolic engineering study.

The key of the ALE is to design and develop cultivation strategies that effectively select cells with desirable phenotypes. Usually, ALE experiments are performed in serial batch cultivation or chemostat. The advantages of chemostat cultivation are constant growth rates, tight control of nutrient supply and environmental conditions such as pH and oxygenation.

In batch cultivation, the variation of environment condition (such as the variation of the nutrients available throughout the cultivation) leads to a complex selection process and the fitness can be evaluated through several variables (such as stationary and lag phase length, or instantaneous growth rate). The complexity of defining the fitness in serial batch cultivation can be reduced by keeping the cells in a single growth phase, as the exponential phase. In these cases, the fitness can be evaluated through the maximum growth rate and the cells need to be transferred to a fresh medium during the exponential phase (LACROIX et al., 2016).

Another important parameter of serial passages batch culture is the passage size, which determines how much of the population is transfer to each subsequent batch culture. If the passage size is small, a beneficial mutation can be lost leading to a slowdown in the rate of

evolution or even halt it. A large population has more diversity, however, the small population favours the robustness due to a high mutation rate (ELENA et al., 2007; LACROIX et al., 2016). Studies reported that the passage size of 10% reached the fitness criteria faster than other population sizes (LACROIX et al., 2016).

Based on this information, in this study, the ALE strategy was designed to improve the carotenoid production in *R. toruloides*. ALE was performed on the sequential passages in shake flasks and the selective pressure was an oxidative stress inducer, hydrogen peroxide (H₂O₂), and was chosen based on the known antioxidant properties of carotenoids. The cells that had a higher carotenoid content must be more tolerant to the oxidative environment than others.

4.2 Materials and methods

4.2.1 Strain and yeast cultivation

Rhodotorula toruloides was obtained from Fundação André Tosello (Campinas, Brazil) as *Rhodospiridium toruloides* CCT 7815 and stored at -80°C in 10% (v/v) glycerol. The yeast was cultivated in shake flasks at 28°C , 200 rpm (incubator shaker TE-4200, Tecnal, Brazil). Cultivation inoculum was prepared in YPD medium at 200 rpm and 28°C for 24 h. The cells were washed with 0.9% (v/v) NaCl solution before inoculation. The medium composition was, per litre, 30.0 g xylose, 2.5 g glucose, 0.9 g yeast extract, 0.2 g (NH₄)₂SO₄, 1.5 g KH₂PO₄ and 0.9 g MgSO₄·7 H₂O.

4.2.2 Analytical methods

Cell growth was monitored by measuring the optical density (OD) at 600 nm (spectrophotometer Quimis, Brazil). The dry biomass (DCW) was determined gravimetrically.

A method modified from Lee et al. (2014) was used to extract and quantify the carotenoids, in which the cells from 2 mL of broth cultivation 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 (MP Biomedicals, USA) for three cycles (4 m/s for 20 s). 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, Germany), and the remaining extracts were resuspended in a known volume of acetone. Carotenoids were measured using Acquity UPLC (Waters, USA) equipped with a TUV detector (Waters, USA) and C18 column (BEH130, 1.7 μm , 2.1 x 100 mm, Waters, 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, USA). Detected peaks were identified according to the known carotenoid retention time profile (LEE et al., 2014; WEBER; ANKE; DAVOLI, 2007).

4.2.3 Adaptive laboratory evolution

Adaptive laboratory evolution (ALE) was performed by successive shake flask cultivations of *R. toruloides* in the presence of H₂O₂. The initial H₂O₂ concentration was 10, 20 and 50 mmol/L in the 1st, 2nd and 3rd cycle, respectively.

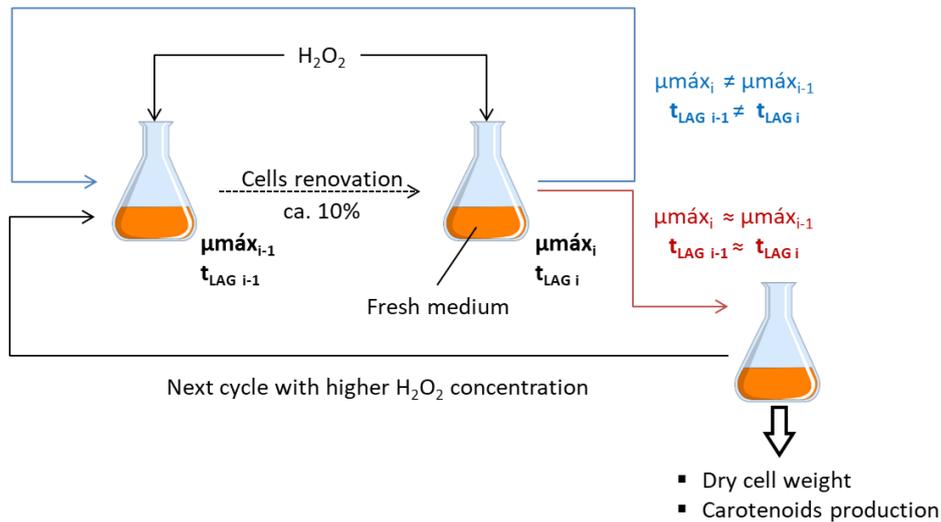
During the serial cultivations, the cells at the end of the exponential growth phase were harvested and transferred to a new flask with a fresh medium. In each cultivation, the maximum specific growth rate (μ_{MAX}) and the length of the lag phase were determined and used as fitness criteria. If these parameters were significantly different (Tukey test, p-value < 0.05) between two successive cultivations, the selective pressure was kept at the same concentration. Otherwise, the H₂O₂ concentration was increased, starting a new cycle (Figure 4.1). Therefore, the end of the ALE cycles was determined by the stabilisation of the μ_{MAX} and the length of the lag phase. Cells from each passage were stored in 10% (v/v) glycerol at -80°C.

The optical density at 600 nm (OD) was measured at least in four different time points during the exponential phase and the μ_{MAX} was calculated by determining the slope of the linear regression obtained by plotting the OD against cultivation time.

4.2.4 Characterisation of parental and adapted strains under reference and oxidative stress conditions

The parental and the adapted strains (AD1, AD2, and AD3) obtained from the first, second and third cycle of ALE, respectively, were characterised in shake flasks cultivation in terms of growth and carotenoid production. The cells used in this characterisation has been stored in 10% (v/v) glycerol at -80°C. The medium composition was the same used in ALE. The effect of H₂O₂ in the medium were evaluated in all strains. For parental strain, the initial H₂O₂ concentration was 10 and 20 mmol/L (higher concentration tolerated by the parental strain). For the adapted strains AD1, AD2, and AD3, the initial H₂O₂ concentration was 10, 20 and 50 mmol/L, respectively.

Figure 4.1. Scheme of ALE methodology applied to improve the carotenoids production in *R. toruloides*.



4.3 Results and discussion

The H_2O_2 stress in parental strain affected the cell growth profile increasing the lag phase length (data not shown) but did not affect the carotenoid production (Table 4.1).

The first cycle of ALE started with 10 mmol/L of H_2O_2 (Figure 4.2A). In the first passage was observed a sharp decrease of the lag-phase length accompanied by an increase in μ_{MAX} . After 16 passages (ca. 30 generations), the lag phase decreased 11-fold (from 46 h to 4 h), and μ_{MAX} stabilised at $0.045 \pm 0.003 \text{ h}^{-1}$, resulting in a 66% increase compared to the parental strain under the same conditions. Once the μ_{MAX} plateaued, the second cycle of ALE was started by increasing the selective pressure to 20 mmol/L of H_2O_2 in the medium (Figure 4.2B).

In the first passage, the lag-phase was once again long (30 h), but no significant change in μ_{max} was observed. However, after 15 passages (ca. 20 generations), the lag phase of yeast growth decreased from 30 to 5 h and the μ_{max} improved to $0.055 \pm 0.001 \text{ h}^{-1}$, representing a 22% increase compared to the first cycle.

Table 4-1. Effect of H₂O₂ stress on carotenoids production by parental strain and strains obtained in each cycle of ALE (AD1, AD2, and AD3, for cycles 1, 2 and 3, respectively).

Strain	Condition	Biomass (g/L)	Carotenoid titre (mg/L)	Y _{CAR} (mg/gDCW)	q _{CAR} (mg/L.h)
Parental strain	Control ⁽¹⁾	13.2 ± 0.2	12.2 ± 0.1	0.92 ± 0.03	0.10 ± 0.04
	10 mmol/L H ₂ O ₂ ⁽²⁾	13.6 ± 0.1	11.8 ± 0.1	0.87 ± 0.01	0.07 ± 0.02
	20 mmol/L H ₂ O ₂ ⁽³⁾	13.7 ± 0.2	11.0 ± 0.2	0.81 ± 0.05	0.06 ± 0.05
AD1	Control ⁽¹⁾	9.6 ± 0.1	13.4 ± 0.1	1.40 ± 0.01	0.11 ± 0.02
	10 mmol/L H ₂ O ₂ ⁽¹⁾	9.1 ± 0.3	14.4 ± 0.1	1.59 ± 0.02	0.12 ± 0.03
AD2	Control ⁽¹⁾	12.7 ± 0.1	29.8 ± 0.1	2.35 ± 0.01	0.25 ± 0.01
	20 mmol/L H ₂ O ₂ ⁽¹⁾	12.0 ± 0.2	36.9 ± 0.1	3.08 ± 0.03	0.31 ± 0.04
AD3	Control ⁽¹⁾	11.4 ± 0.1	49.7 ± 0.3	4.36 ± 0.02	0.41 ± 0.04
	50 mmol/L H ₂ O ₂ ⁽¹⁾	11.1 ± 0.1	49.9 ± 0.5	4.49 ± 0.03	0.42 ± 0.01

The superscript numbers (1, 2, and 3) represent the cultivation time: 120, 168, and 192h, respectively.

The third cycle has an initial H₂O₂ concentration of 50 mmol/L and was carried out through 10 passages (ca. 17 generations). Although μ_{\max} did not improve during this cycle (remained at $0.055 \pm 0.003 \text{ h}^{-1}$, Figure 4.2C), the lag phase length was reduced from 35 to 8 h. At the end of this cycle, the cells presented stronger pink colour compared to the parental strain, indicating increased carotenoid accumulation. Therefore, the ALE experiment was halted, and the strains were characterised.

In cycle 1, the carotenoid titre increased by 18% compared to parental strain under optimal growth condition (Table 4.1). Although this increase was significant (Tukey test, p-value < 0.05), it was not considerable. However, it might be related to the carotenoid composition, since the torularhodin fraction in AD1 was higher than in the parental strain, while the β -carotene was lower (Figure 4.3). This can be explained by the stronger antioxidant propriety of torularhodin and torulene compared to β -carotene, attributed to the presence of more double bonds in its chemical structure (KOT et al., 2018). So, the yeast defence mechanism against oxidative stress seems not only to involve the increase of the total carotenoid production, but also redirect the production to compounds of stronger antioxidant effect. The

same trend can be observed in AD2 and AD3, in which the torulene fraction increase gradually, reaching more than 40% in cycle 3.

In the last cycle, the carotenoid yield increased remarkably compared to parental strain under control condition: ca. 5-fold (Table 4.1). AD1 and AD2 strain grown under optimal growth condition presented a slight decrease in the carotenoids production when compared to the oxidative stress condition. This can indicate that the adapted cells could lose the fitness advance in absence of the stress condition. However, this loss was not observed in cycle 3 which might suggest that ALE caused constitutive modifications in this strain.

Figure 4.2. Changes in μ_{MAX} and lag phase time during the (A) first, (B) second, and (C) third cycle of the adaptive laboratory evolution with H_2O_2 . Error bars represent the standard deviation from three measurements.

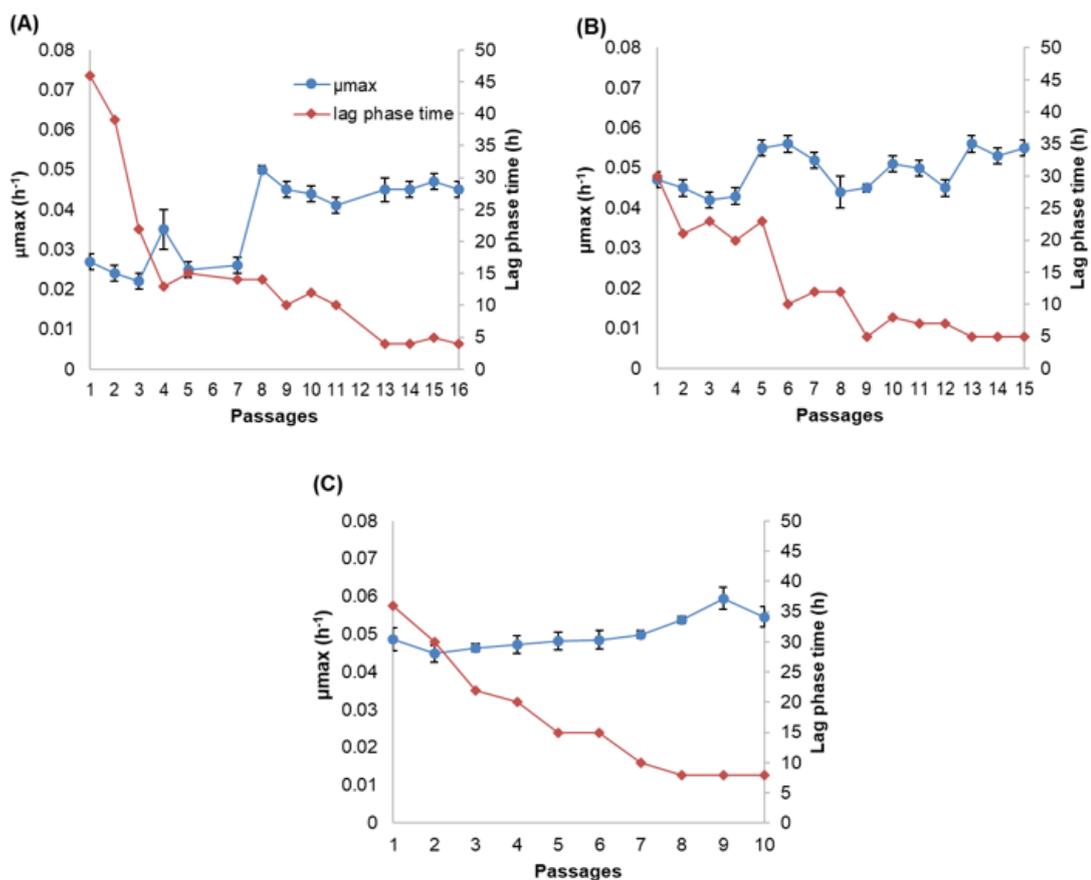
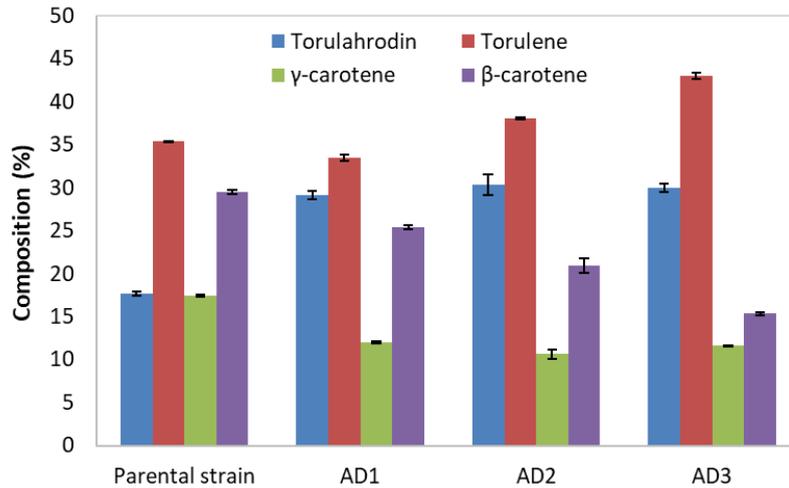


Figure 4.3. Relative carotenoid proportion of parental strain grown under optimal growth condition and adapted strains from the cycle 1, 2 and 3 (AD1, AD2, and AD3, respectively) grown under oxidative stress condition (10, 20, and 50 mmol/L of H₂O₂, respectively).



4.4 Conclusion

The successive exposure of cells to H₂O₂ through the ALE potentially overloaded the cellular antioxidant defence system, boosting carotenoid production. The adapted strain produced ca. 50 mg/L of total carotenoids (a 5-fold increase compared to parental strain). Also, this strain produced a higher proportion of torularhodin and torulene, two important carotenoids with potential anticancer activity. The evolutionary mechanism during the ALE can be analysed by system biology to elucidate the yeast metabolic networks.

CHAPTER 5

Xylose metabolism and the effect of oxidative stress on lipid and carotenoid production in *R. toruloides*: insights for its future application in biorefinery

5.1 Introduction

The use of cell factories to convert sugars from lignocellulosic biomass into fuels and chemicals, such as oleochemicals and food additives 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. 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 the metabolism 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). As demonstrated in chapters 3 and 4, lipid and carotenoid production is enhanced in the presence of oxidative stress induced by hydrogen peroxide (H₂O₂) and light irradiation.

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 quantitative proteomics and GEM analysis. Oxidative stress (H₂O₂ and light irradiation) and adaptive laboratory evolution were employed to improve lipid and carotenoid 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 using this yeast.

5.2 Materials and methods

5.2.1 Strain and inoculum

Rhodotorula toruloides was obtained from Fundação André Tosello (Campinas, Brazil) as *Rhodospiridium toruloides* CCT 7815 and stored at -80°C in 10% (v/v) glycerol. Also, the adapted strain AD2 obtained through adaptive laboratory evolution with H_2O_2 (described in chapter 4) was studied. Cultivation inoculum was prepared in YPX medium at 200 rpm and 30°C for 24 h. The cells were washed twice with 0.9% (v/v) NaCl before inoculation.

5.2.2 Yeast cultivation

Batch cultivations were performed in 1-L bioreactors (Applikon Biotechnology, 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, Germany), and optical density was monitored online using a Bug Lab BE3000 Biomass Monitor (Bug Lab, USA) at 1,300 nm. Data were collected and processed with BioXpert V2 software v. 2.95 (Applikon Biotechnology, 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 $(\text{NH}_4)_2\text{SO}_4$, 3.0 g KH_2PO_4 , 0.5 g $\text{MgSO}_4 \cdot 7 \text{H}_2\text{O}$, 1.0 mL vitamins solution and 1.0 mL trace metal solution (LAHTVEE et al., 2017), supplemented with 100 μl antifoam 204 (Sigma-Aldrich, USA). Samples for dry cellular mass, 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. Cell pellets were snap-frozen in liquid nitrogen, stored at -80°C , and further used for proteomics analysis.

5.2.3 Quantification of dry cell mass, extracellular metabolites, carotenoids, total lipids and proteins

Online turbidity measurements were calibrated by gravimetrically measuring the dry cellular mass (DCW) every 24 h. Extracellular metabolites in the broth were measured using HPLC (LC-2030C Plus, Shimadzu, Japan) equipped with a refractive index detector

(RID-20A, Shimadzu, Japan). Concentrations of xylose, organic acids and glycerol were measured using an HPX-87H column (Bio-Rad, USA) at 45°C, and 5 mmol/L H₂SO₄ served as the mobile phase with isocratic elution at 0.6 mL/min. Xylitol and arabitol were quantified using Rezex RPM-Monosaccharide column (Phenomenex, USA) at 85°C, and purified water (MilliQ Ultrapure Water System, Merck, Germany) was 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, 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, Germany), and the remaining extracts were resuspended in a known volume of acetone. Carotenoids were measured using Acquity UPLC (Waters, USA) equipped with a TUV detector (Waters, USA) and C18 column (BEH130, 1.7 µm, 2.1 x 100 mm, Waters, 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, USA). Detected peaks were identified according to the known carotenoid retention time profile (LEE et al., 2014; WEBER; ANKE; DAVOLI, 2007).

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 cells. 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 cells resuspended in Y-PER solution (Thermo Fisher, USA) in a 2-mL Eppendorf tube. This suspension was incubated at 30°C for 45 min. Then, glass beads were added to 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 cells 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).

5.2.4 Proteome analysis

Fully labelled cellular biomass was used as the internal standard in absolute proteome analysis and produced by cultivating *R. toruloides* in minimal mineral medium containing labelled heavy ¹⁵N, ¹³C-lysine (Silantes, 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 a 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, 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, USA) tips and were reconstituted in 0.5% trifluoroacetic acid (TFA). For separation, 2 µg of peptides were injected on an Ultimate 3000 RSLCnano system (Dionex, Sunnyvale, USA) coupled to a C18 cartridge trap-column in a backflush configuration and an analytical 50 cm x 75 µm emitter-column (New Objective, USA) in-house packed with 3 µm C18 particles (Dr Maisch, Germany). Eluted peptides were sprayed to a quadrupole-orbitrap Q Exactive Plus (Thermo Fisher Scientific, 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). 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 proteomics data were used for differential expression analysis. P-values were adjusted for quantitative data are presented in Pinheiro et al. (2020), Supplementary Table S7. Triplicated quantitative multiple testing using the Benjamini-Hochberg procedure (BENJAMINI, HOCKBERG, 1995). Additional data analysis included gene set enrichment analysis (carried out using PIANO platform; VÄREMO et al., 2013) and gene enrichment analysis (g:Profiler; RAUDVERE et al., 2019).

5.2.5 Genome-scale modelling

The intracellular flux patterns were predicted using the *R. toruloides* genome-scale metabolic network rhto-GEM version 1.2.1 and flux balance analysis (TIUKOVA et al., 2019a; LOPES et al., 2020a). Models and experimental rates used for the reference condition (REF_P1, REF_P2, and REF_P3) were uploaded on GitHub repository (https://github.com/SynBioUniTartu/Rhto_OxidativeStress). Random sampling simulations, quantitative proteomics data and the reference file (gmt format) used for the gene enrichment analysis (g: Profiler) are hosted on the same repository. The model was improved by adding carotenoids (β -carotene, γ -carotene, torulene and torularhodin) into the biomass pseudo-reaction, and adding D-arabitol production pathway (Jagtap et al. 2018). To allow the model to predict the production of either D- or L-arabitol, *in silico* pseudo-reactions were added, converting, both, D- and L-arabitol into an “artificial” arabitol without a specific isoform. Flux balance analysis was performed to calculate intracellular flux patterns using RAVEN Toolbox (WANG et al., 2018a) on MATLAB (The MathWorks Inc., USA), Gurobi solver (Gurobi Optimization Inc., 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. Cellular biomass composition was adjusted to the measured total protein, lipid and carotenoid content.

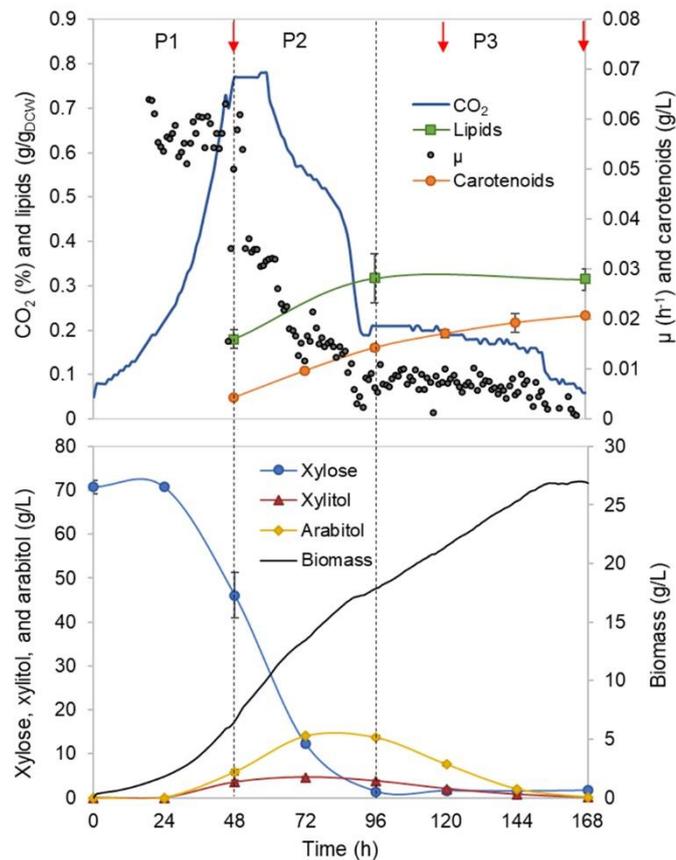
5.3 Results

5.3.1 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, CO₂ production and O₂ consumption was complemented by off-line analysis of sugars, alcohols and cell mass composition (total lipids, proteins and carotenoids; Figure 5.1).

Figure 5.1. *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 (μ), cell mass concentration, intracellularly accumulated lipid and carotenoid concentrations, extracellular metabolite profiles and CO₂ production profile in the outflow gas are presented. The values represent an average of three independent cultivation experiments; error bars represent standard deviation. The red arrows represent the proteomic data points.



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 was used as the sole carbon source. Arabitol, xylitol and CO₂ were the main fermentation by-products detected. Phase two (P2) started with a sudden decrease in the specific growth rate probably due to nitrogen limitation. At that point, approximately 23 g/L of xylose was consumed, and assuming

that the nitrogen was depleted, indicating a critical C/N ratio of 26 (mol/mol) 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 environmental changes during the batch growth.

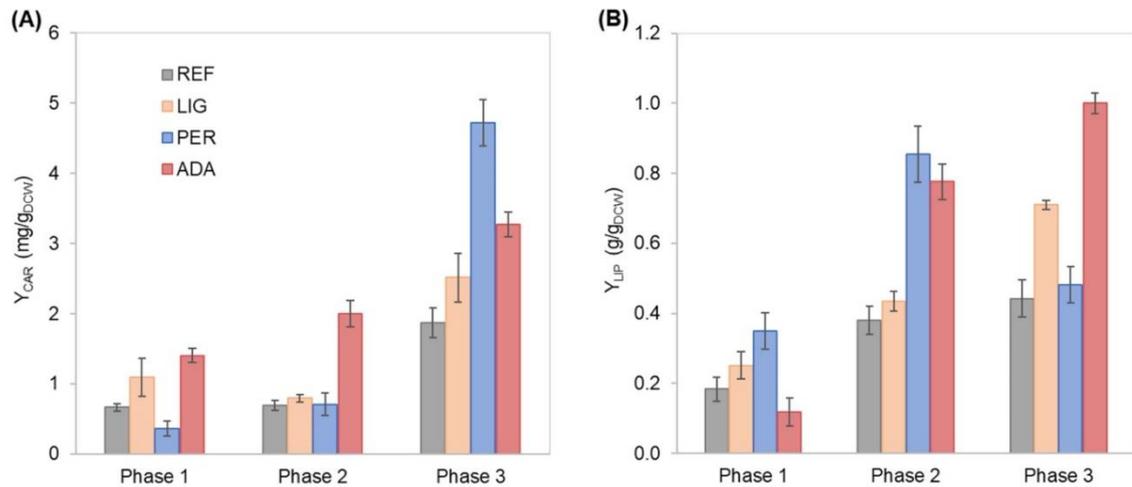
P1 comprised the highest specific growth rate (μ , $0.060 \pm 0.001 \text{ h}^{-1}$) and the specific xylose uptake rate (r_{XYL} , $1.74 \pm 0.12 \text{ mmol/g}_{\text{DCW}}\cdot\text{h}$) (Table 5.1), while no nutrient-level limitations were detected. During the exponential growth, approximately one-third of the consumed carbon was secreted as arabitol and xylitol. An additional one-third of the consumed carbon was secreted as CO_2 . Although the carotenoid content in P1 was low ($0.66 \pm 0.06 \text{ mg/g}_{\text{DCW}}$), the specific production rate (r_{CAR}) was the highest ($0.042 \pm 0.003 \text{ mg/g}_{\text{DCW}}\cdot\text{h}$) due to the higher μ (Figure 5.2A, Table 5.1). The biomass yield and lipid content 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 μ (average value $0.020 \pm 0.002 \text{ h}^{-1}$) and r_{XYL} (average value $0.41 \pm 0.06 \text{ mmol/g}_{\text{DCW}}\cdot\text{h}$). Compared to P1, the carotenoid content did not show a significant difference; however, r_{CAR} decreased 4-fold compared with P1 (Table 5.1). 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 5.2B). 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 cell mass was detected in this phase, increasing substantially to $1.87 \pm 0.21 \text{ mg/g}_{\text{DCW}}$; however, the r_{CAR} was the lowest of all phases (Figure 5.2 and Table 5.1). The lipid content remained at the same level as that noted in P2.

Also, 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 on cell mass 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 5.2, column REF).

Figure 5.2. Carotenoid (A) and lipid (B) yields on cell mass 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.



5.3.2 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 (LOPES et al., 2020a; TIUKOVA et al., 2019b). 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 by using linear programming based optimization under the provided constraints and selects the most ATP-efficient pathways for satisfying the pre-set constraints. However, the model does not take into account metabolic regulation nor limitations due to the capacity or

Table 5-1. Physiological parameters of *R. toruloides* growth in batch cultivation for three growth phases (P1-P3) under the studied environmental conditions: parental strain under the optimal condition (REF), light irradiation (LIG) and in the presence of hydrogen peroxide (PER), and the adapted strain under oxidative e stress (ADA).

Condition	Phase	μ (h ⁻¹)	r _{XYL} (mmol/gdcw.h)	r _{XYLITOL} (mmol/gdcw.h)	r _{ARABITOL} (mmol/gdcw.h)	r _{CAR} (mg/gdcw.h)	r _{LIP} (g/gdcw.h)
REF	P1	0.060 ± 0.001	-1.74 ± 0.12	0.228 ± 0.03	0.372 ± 0.04	0.042 ± 0.003	0.012 ± 0.001
	P2	0.020 ± 0.002	-0.41 ± 0.06	0.003 ± 0.004	0.068 ± 0.02	0.014 ± 0.003	0.007 ± 0.001
	P3	0.005 ± 0.000	-	-0.039 ± 0.01	-0.142 ± 0.03	0.010 ± 0.001	0.002 ± 0.002
LIG	P1	0.055 ± 0.008	-1.12 ± 0.08	0.174 ± 0.03	0.286 ± 0.05	0.058 ± 0.008	0.014 ± 0.003
	P2	0.022 ± 0.005	-0.50 ± 0.08	0.021 ± 0.01	0.120 ± 0.03	0.017 ± 0.003	0.010 ± 0.002
	P3	0.007 ± 0.001	-	-0.027 ± 0.001	-0.101 ± 0.00	0.016 ± 0.002	0.005 ± 0.003
PER	P1	0.047 ± 0.004	-0.50 ± 0.16	0.084 ± 0.03	0.173 ± 0.06	0.012 ± 0.004	0.012 ± 0.003
	P2	0.024 ± 0.003	-0.64 ± 0.15	0.022 ± 0.01	0.158 ± 0.04	0.017 ± 0.006	0.020 ± 0.004
	P3	0.007 ± 0.001	-	-0.069 ± 0.07	-0.354 ± 0.20	0.026 ± 0.012	0.003 ± 0.001
ADA	P1	0.046 ± 0.001	-0.65 ± 0.01	0.061 ± 0.04	0.110 ± 0.03	0.045 ± 0.004	0.004 ± 0.001
	P2	0.012 ± 0.002	-0.27 ± 0.03	0.001 ± 0.00	0.020 ± 0.05	0.024 ± 0.004	0.009 ± 0.000
	P3	0.004 ± 0.000	-	-0.002 ± 0.00	-0.062 ± 0.06	0.014 ± 0.001	0.004 ± 0.001

μ , specific growth rate; r_{XYL}, Specific xylose consumption rate; r_{XYLITOL}, Specific xylitol production/consumption rate; r_{ARABITOL}, Specific arabitol production/consumption rate; r_{CAR}, Specific carotenoid production rate; r_{LIP}, Specific lipid production rate

activity of enzymes, which should be taken into account when interpreting the modelling results. For the flux distribution comparison between the observed growth phases, fluxes were normalised by the total carbon uptake rate.

As to the best of our knowledge, there has not been a direct experimental validation of the arabitol isoform which is produced by *R. toruloides*, we used the GEM to predict it based on the provided stoichiometry of the reactions. Our simulations suggested L-arabitol

Table 5-2. Global titres, yields on cell mass (Y) and on substrate (Y_S), 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	
Carotenoids	Titre (mg/L)	20.7 ± 0.5	36.2 ± 2.6	19.3 ± 0.2	44.0 ± 2.4
	Y (mg/g _{DCW})	0.85 ± 0.01	1.45 ± 0.14	0.89 ± 0.04	1.90 ± 0.13
	q (mg/L.h)	0.12 ± 0.003	0.22 ± 0.02	0.06 ± 0.02	0.23 ± 0.01
	Y _S (mg/g _{sub})	0.29 ± 0.02	0.54 ± 0.04	0.29 ± 0.01	0.62 ± 0.04
Lipids	Titre (mg/L)	8.1 ± 0.3	11.1 ± 1.7	13.8 ± 0.9	13.3 ± 0.5
	Y (g/g _{DCW})	0.33 ± 0.07	0.46 ± 0.12	0.65 ± 0.06	0.58 ± 0.07
	q (g/L.h)	0.05 ± 0.003	0.07 ± 0.005	0.05 ± 0.002	0.07 ± 0.002
	Y _S (g/g _{sub})	0.12 ± 0.01	0.17 ± 0.03	0.21 ± 0.02	0.19 ± 0.01

production, as the L-arabitol production pathway regenerates one NADPH and one NADH, while D-arabitol production pathway regenerates only one NADH. NADPH supply is crucial for lipid production given that every elongation step of fatty acid (FA) synthesis requires the oxidation of two NADPH (WASYLENKO; AHN; STEPHANOPOULOS, 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. 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 (Appendix Figure A3). 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 transketolase (r_1049, r_1050) in P1 (Figure 5.3 B). 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 (Appendix Figure A3).

Our physiological data showed that the lipid content was higher under nitrogen-limitation phases (P2 and P3), and the carotenoids content 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.

5.3.3 Impact of oxidative stress via light irradiation or the presence of H₂O₂

Further, we were interested in how oxidative stress created by either 20 mmol/L H₂O₂ 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 (Figure 5.4 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 H₂O₂ stress (PER) in P1 (Table 5.1). The most significant difference was detected in the longer lag phase (approximately 90 h) shown in the PER condition (Figure 5.4 B). Moreover, the accumulation profiles of carotenoids and lipids showed altered behaviour compared to the reference (Figure 5.2). The stress caused by H₂O₂ negatively affected carotenoid production in P1. However, in P3, the total carotenoid content was the highest of all conditions (4.72 ± 0.47 mg/g_{DCW}); thus, the highest r_{CAR} (0.026 ± 0.012 mg/g_{DCW}.h) was reached in the third phase (Figure 5.2A, Table 5.1). Lipid production exhibited a different behaviour, presenting a higher content 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 ± 0.004 g/g_{DCW}.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 content (1.45 ± 0.14 mg/g_{DCW}) and 40% increased lipid content (0.46 ± 0.12 g/g_{DCW}) compared to the reference condition. H₂O₂ stress in the parental strain (PER) did not affect carotenoid production (compared to REF), and the production yield on cell mass was maintained at 0.89 ± 0.04 mg/g_{DCW}. Surprisingly, this condition showed the highest lipid content (0.65 ± 0.06 g/g_{DCW}), which was increased by two-fold compared with the reference condition (Table 5.1). The achieved lipid content was only slightly lower than the highest lipid content reported in the literature for *R. toruloides*; specifically, 0.68 g/g_{DCW} has been reported in fed-batch cultivation on a rich, glucose-based medium (LI; ZHAO; BAI, 2007).

5.3.4 Adapted strain under H₂O₂ stress

The performance of the adapted strain under oxidative stress (by the presence of 20 mmol/L H₂O₂, ADA) was evaluated and compared to the parental strain under the same stress condition (PER). ADA exhibited a 70-hour shorter lag phase compared with the PER (Figure 5.4 C). 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 in chapter 4).

Figure 5.3. *R. toruloides* GEM-based fluxes (normalized to the substrate uptake) and measured protein abundances (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 (mmol/g_{DCW}.h) are represented in italic in panel A (note that in P3, arabitol and xylitol were co-consumed)). The colour of enzymes name indicates the significant protein allocation change under nitrogen limitation (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 Pinheiro et al., 2020, 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 ACC1: acetyl-CoA carboxylase.

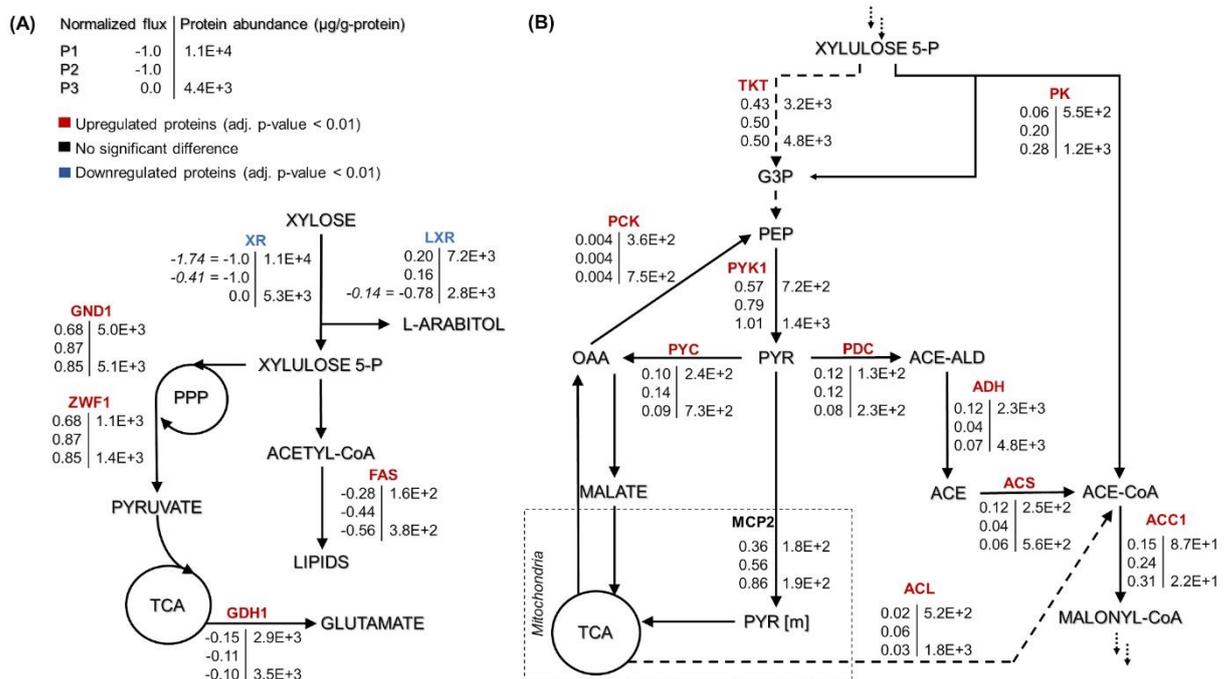
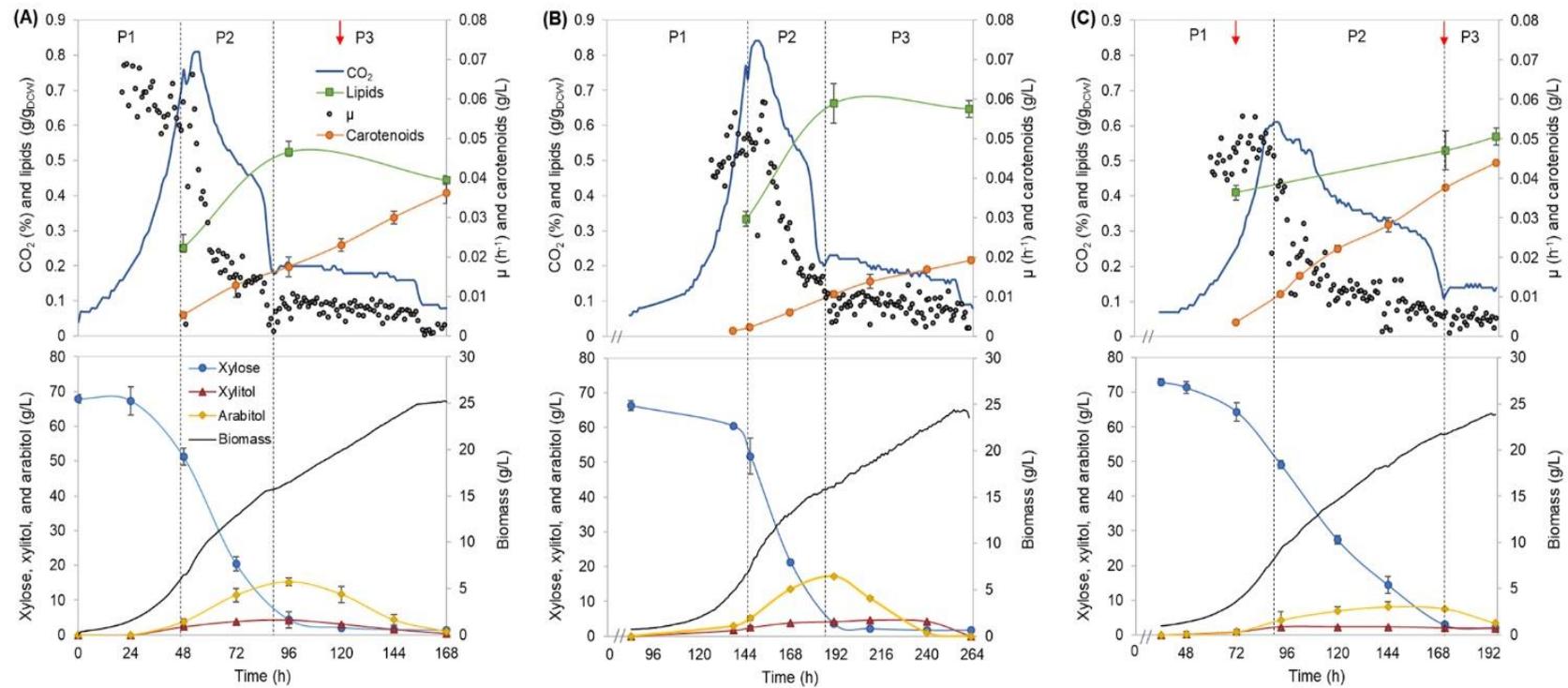


Figure 5.4. *R. toruloides* batch cultivation on xylose under light irradiation (A) and in presence of H₂O₂ (B), and adapted strain under oxidative stress (C). Dashed vertical lines define three observed growth phases. The specific growth rate (μ), biomass concentration, intracellularly accumulated lipid and carotenoid concentrations, extracellular metabolite profiles and CO₂ production profile in the outflow gas are presented. The values represent an average of three independent cultivation experiments; error bars represent standard deviation. The red arrows represent the proteomic data points.



Some differences were noticed between the ADA and other conditions. Although the μ in P1 did not show a significant difference, the r_{CAR} was increased by 4-fold in ADA compared with PER (Table 5.1). However, in P2, the μ in ADA was 2-fold reduced compared with the other conditions, but r_{CAR} remained 20% increased. The lower r_{XYL} in P2 could have resulted in the lower production of xylitol and arabitol in ADA compared to the other conditions is probably due to a softer redox imbalance during xylose catabolism. The ADA showed ca. 2-fold increased content of carotenoids and lipids under nitrogen-limiting phases (P2 and P3) compared to REF. In P3, the lipid content was 1.0 ± 0.03 g/g_{DCW}, 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 on cell mass compared to PER; however, the lipid yield did not show a significant difference (Table 5.2).

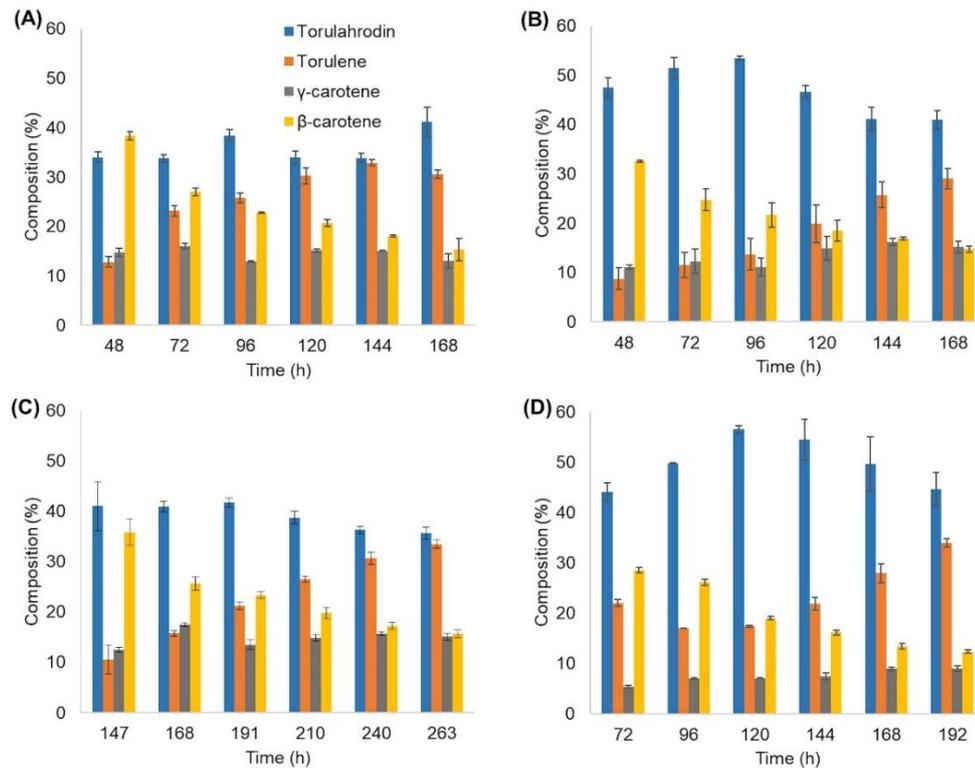
5.3.5 Composition of carotenoids in cell mass

R. toruloides mainly produces four carotenoids: γ -carotene, β -carotene, torulene and torularhodin (MATA-GÓMEZ et al., 2014). The carotenoid profile was very similar in all the studied conditions. The β -carotene fraction decreased over time, whereas the opposite was observed for torulene (Figure 5.5). 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).

5.3.6 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 three different growth phases (P1, P3, and SP- an early stationary phase) from the reference cultivation were analysed, illustrating conditions under the nitrogen excess and two samples from nitrogen limitation conditions. Additionally, P3 of light-induced oxidative stress (LIG P3) and P1 and P3 of H₂O₂-induced oxidative stress for the adapted strain (ADA P1 and P3, respectively). For LIG condition, P3 represents the nitrogen-limitation phase with a higher carotenoid titre and specific production rate than REF P3.

Figure 5.5. Carotenoids composition during the batch cultivation of *R. toruloides* under the studied conditions: (A) the parental strain in the reference condition (REF), (B) under light irradiation (LIG), or (C) in presence of hydrogen peroxide (PER), and (D) the adapted strain under oxidative stress (ADA).



In ADA P1, the adapted strain presented a higher carotenoid yield on cell mass (Y_{CAR}) compared to parental strain under reference condition (REF P1). While P3 (ADA P3) showed the highest lipid content (g/g_{DCW}) compared to all conditions, indicating that during the phase, almost all carbon taken up by the cells ended in the lipid composition (Figure 5.2B).

Almost doubled 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 $0.24 g/g_{DCW}$ (Figure 5.6A). In differential expression analysis, proteome data were normalised to a constant protein mass, representing allocation differences for the individual proteins.

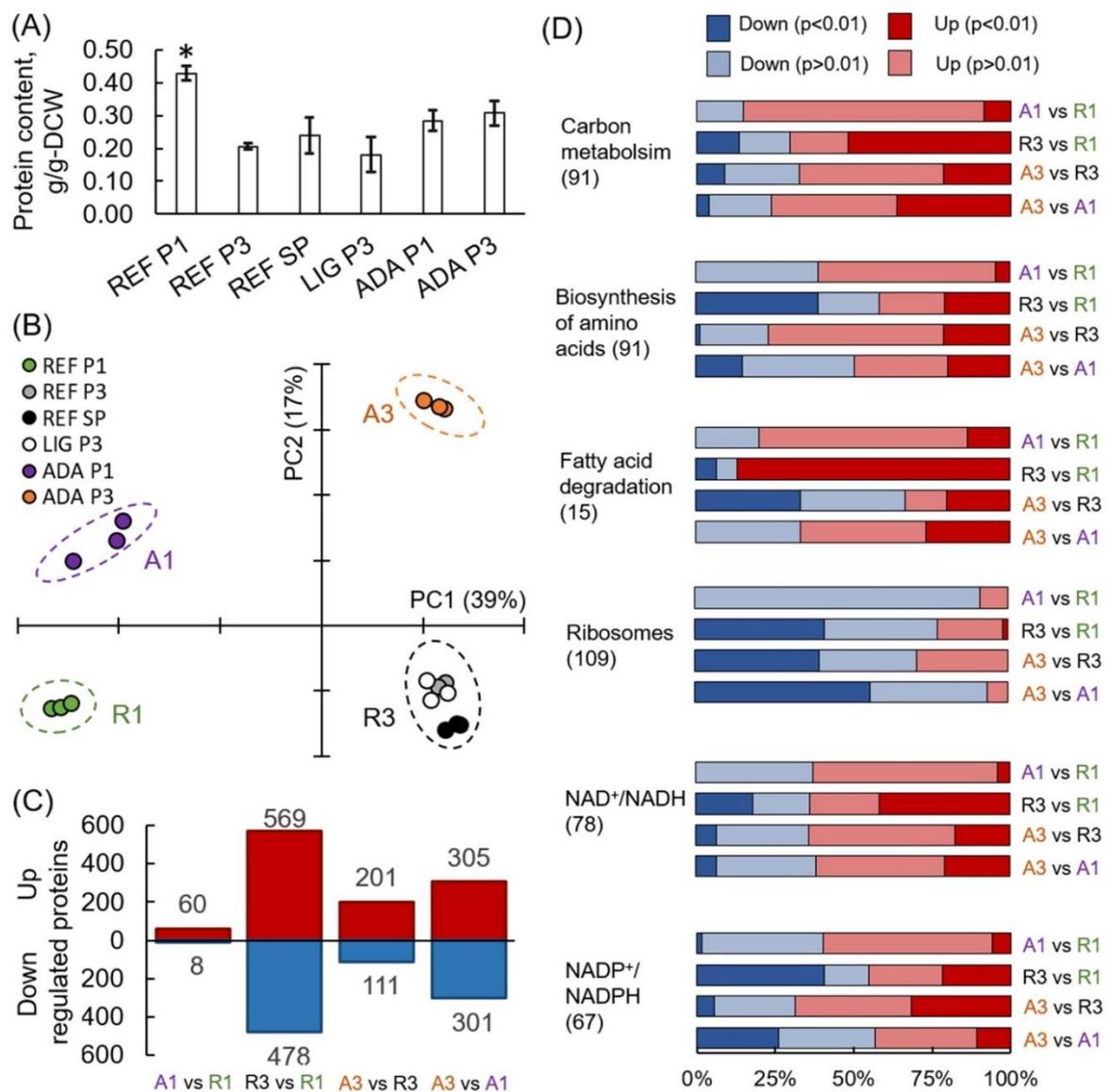
On average, more than 3,000 individual proteins were quantified under every condition studied (PINHEIRO et al., 2020, Supplementary Table S7). Principal component (PC) analysis clearly identified the biggest differences in the data set, which were determined by the switch

into nitrogen limitation as indicated by the clear separation of the samples on the first PC, characterising 39% of the changes (Figure 5.6B). 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 5.6C).

5.3.7 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 (PINHEIRO et al., 2020, 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 (Figure 5.6D). A clear correlation between ribosome abundances and specific growth rate has been demonstrated previously for other organisms (METZL-RAZ et al., 2017; SCOTT et al., 2014). When subsystems from rhtoGEM were considered, significant upregulation was detected among carbon metabolism and its subgroups (glycolysis, gluconeogenesis, TCA cycle, glyoxylate and dicarboxylate metabolism) (PINHEIRO et al., 2020, 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.

Figure 5.6. Proteomics results from *R. toruloides* studies on xylose under different environmental conditions (REF P1, P3, and SP; LIG P3; 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).



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 (PINHEIRO et al., 2020, 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^+/NADH (Figure 5.6D). More than 65% of the proteins associated with NAD^+/NADH showed increased allocation under nitrogen limitation (REF P3 and SP, LIG P3, ADA P3). In contrast, protein allocation decreased significantly for proteins related to $\text{NADP}^+/\text{NADPH}$ 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 (TIUKOVA et al., 2019a; ZHU et al., 2012). 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).

5.3.8 Protein changes demonstrate similar trends with the simulated fluxes

The global comparison of differentially expressed proteins with their corresponding fluxes between the nitrogen excess and limitation conditions demonstrated similar trends for the 62% of cases (PINHEIRO et al., 2020, Supplementary Table S11). 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 5.3A). This response has been previously reported under nitrogen limitation for *R. toruloides* grown in both glucose and xylose (TIUKOVA et al., 2019a; ZHU et al., 2012). Additionally, activation of autophagy process has been described as a direct response via TOR activation to recycle nitrogenous compounds (TIUKOVA et al., 2019a; ZHU et al., 2012). 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 P3 and SP (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 (PINHEIRO et al., 2020, Supplementary Table S7). Recent reports in oleaginous microorganisms showed that ROS is an important signalling

molecule in response to various stresses (SHI et al., 2017). Nitrogen depletion is an example of such stress, leading to the accumulation of ROS (CHOKSHI et al., 2017; FAN et al., 2014; LIU et al., 2012) and higher activities of catalase and other antioxidant enzymes, suggesting that lipid accumulation under nitrogen depletion is mediated by oxidative stress (YILANCIUGLU et al., 2014).

The highest carbon fluxes detected with GEM analysis were further assessed to understand the level of their regulation. In P3 and SP, there is no longer xylose in the system, which was reflected in the downregulation of proteins involved in xylose assimilation (Table 5.1). Proteins belonging to arabitol metabolism, such as L-xylulose reductase (LXR, RHTO_00373), and D-arabitol dehydrogenase (DAD, RHTO_07844), and L-identol 2-dehydrogenase, identified by Bommaredy et al. (2015) as L-arabitol dehydrogenase (LAD, RHTO_01629) were all downregulated under nitrogen- limiting conditions.

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 5.3B). 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 5.3A). Except for 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 (KOUTINAS; PAPANIKOLAOU, 2011; RATLEDGE; WYNN, 2002). Under nitrogen limitation, this enzyme was found in higher levels compared to P1 (4- and 2-fold increase for P3 and SP, respectively). Cytosolic acetyl-CoA can also be supplied by acetyl-CoA synthase (ACS1, RHTO_08027), which was upregulated 2-fold in nitrogen-limiting phases, and by xylose metabolism via PK and phosphate acetyltransferase. FBA predicted that the majority of cytosolic acetyl-CoA originating from

pyruvate via ACS1 (in nitrogen-excess phase) and from PK and phosphate transacetylase (in nitrogen-limiting phase) (Figure 5.3B).

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 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 under nitrogen limitation, ERG10 and other proteins related to the MEV pathway, hydroxymethylglutaryl-CoA synthase (ERG13, RHTO_02305) and reductase (HMG1, RHTO_04045) were downregulated. 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.

5.3.9 Comparative proteomics at different growth phases under oxidative stresses

Although P3 of cultivation under light exposure (LIG_P3) showed higher carotenoid production rates than REF_P3 (Figure 5.2, Table 5.1), only two proteins (RHTO_06480 and RHTO_01160) were differentially expressed between those conditions (adj. p-value < 0.01), suggesting a post-translational regulation of the carotenoid production pathways under the studied environmental conditions.

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 5.6B). 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⁺-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, β -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 phytoene dehydrogenase (CRTI, RHTO_04602) were 3-fold upregulated (Figure 5.7A). As expected, enzymes involved in the oxidative stress response were upregulated in the ADA compared to REF; CAT exhibits a 16-fold increase for P1.

5.3.10 Translation processes also play a crucial role in proteome allocation analysis

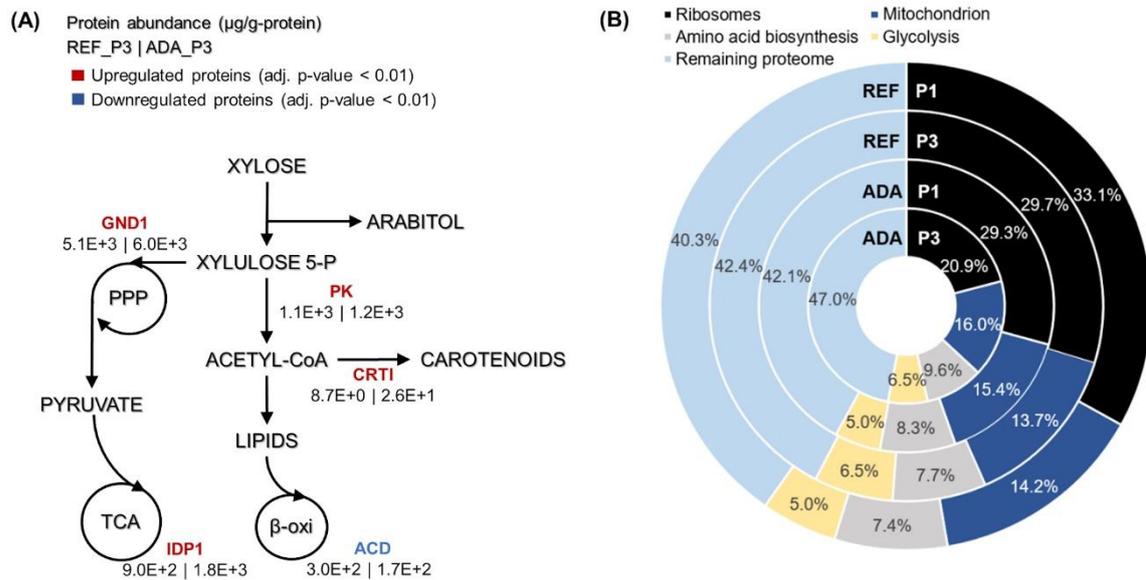
Absolute protein analysis allows comparisons of protein abundance levels between conditions and quantification of condition-dependent protein allocation patterns (Figure 5.7B). The top 100 of the most abundant proteins represented greater than 50% of the total proteome. Ribosomal proteins were the largest 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, NIELSEN, 2016; SÁNCHEZ et al., 2017; KUMAR, 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 yields on cell mass for adapted cells (Figure 5.7B and Table 5.1).

5.4 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 a 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 quantitative 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.

Figure 5.7. 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; CRTI, phytoene dehydrogenase, ACD, acyl-CoA dehydrogenase.



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. According to Fernandes and Murray (2010), fungi can produce both L- and D-arabitol. Jagtap and Rao (2018) reported that D-xylulose is converted to D-arabitol by DAD, although the methodology employed by the authors is not able to discriminate between the L and D isoforms. Same for the methodology employed in this work. The UniProt protein database for *R. toruloides* contains only sequences

for LXR and DAD, lacking the levogyrous version of the latter. Bommaredy et al. (2015) assembled the metabolic network for *R. toruloides* based on previous omics studies and identified that the enzyme LAD, identified as L-idoitol dehydrogenase by Uniprot, is responsible for arabitol production in this species. The GEM constructed by Tiukova et al. (2019) contains only the L-arabitol production pathway. In this work, we updated the model including also the pathway described in Jagtap and Rao (2018) and our simulations showed that the L-isoform was preferred. Forcing the model for the production of only D-arabitol did not reflect our experimental data, unless the cofactor for DAD was NADPH instead of NADH. An *in silico* analysis of the amino acids sequences of the LXR, LAD, and DAD by the NCBI conserved domain search tool (<https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>) showed that those enzymes could use both or either NADH and NADPH. Our proteomics analysis was able to detect enzymes for both pathways, LXR, LAD, and DAD (Appendix Figure A3). Seiboth et al. (2003) reported that the LAD can partially compensate for the xylose metabolism in mutants knocked out for xylitol dehydrogenase in *Hypocrea jecorina*. Curiously, 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. The available knowledge on the xylose metabolism in *R. toruloides* is poor and further studies aiming at revealing the involved enzymes along with their co-factors and reactions are needed.

Under nitrogen-limiting conditions, increased flux via phosphoketolase was observed, corresponding well with the detected 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 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; 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 ¹³C-labelling experiments with *Y. lipolytica* performed using

glucose as the carbon source (WASYLENKO; AHN; STEPHANOPOULOS, 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 content of 0.65 g/g_{DCW}. The highest reported lipid production in *R. toruloides* is 0.675 g/g_{DCW} (LI; ZHAO; BAI, 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. 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 H₂O₂ 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 carotenoid biosynthesis pathway was more activated by higher levels of CRTI what 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 β -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 (QUILES-ROSILLO et al., 2005).

5.5 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 model simulations, which does not take into account metabolic regulation or enzyme capacity constraints, 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.

CHAPTER 6

Development of a dedicated Golden Gate Assembly platform for the nonconventional yeast *Rhodotorula toruloides*

6.1 Introduction

In the past decade, systems biology studies have shed light on the genomic organization and metabolic pathways of *R. toruloides*. Complete genome sequences of different strains have been determined, helping to identify pathways of interest for strain engineering, discover the response of the organism towards environmental changes, and investigate the metabolism of lipid accumulation (HU; JI, 2016b; KUMAR et al., 2012; MORIN et al., 2014; ZHU et al., 2012). Basic tools for genetic manipulation have been developed or studied, such as constitutive (LIU et al., 2013, 2016; WANG et al., 2016) and inducible promoters (JOHNS; LOVE; AVES, 2016; LIU et al., 2015), and auxotrophic (YANG et al., 2008) and antibiotic-based (LIN et al., 2014) selectable markers.

Despite these advances, there is still a lack of advanced genome engineering tools, which plays an essential role in developing *R. toruloides* as a workhorse for biotechnological applications. This is attributed to several factors: firstly, *R. toruloides* has a very high GC content (~62%) (SAMBLES et al., 2017), which hinders PCR reactions due to higher chances of forming secondary structures, and hence reduces polymerase efficiency and accuracy, hampers the gene synthesis and the design of specific primers without reaching an extremely high annealing temperature (LIN et al., 2014). Another factor is the lack of known ARS elements, meaning plasmids cannot be used to express heterologous genes and, instead, have to be integrated into the genome (TSAI et al., 2017).

Furthermore, integrating heterologous elements into the genome of *R. toruloides* is low efficient, as the favoured mechanism of repairing DNA double strand breaks (DSBs) is non-homologous end joining (NHEJ) over homologous recombination (HR) (KRAPPMANN, 2007; LIU et al., 2017). In the case of NHEJ, the DSBs are repaired without a DNA template, which usually results in random insertions and/or deletions into

the DNA. However, the presence of DNA template homologous to the sequence flanking the DSB location could guide the repairing by HR. (PARDO; GÓMEZ-GONZÁLEZ; AGUILERA, 2009). NHEJ and HR operate competitively in the cell, therefore, the gene targeting efficiency in *R. toruloides* could be enhanced by inhibition or elimination of the NHEJ pathway, forcing the heterologous DNA to be integrated via HR. The KU70/80 heterodimer is the main component of the NHEJ pathway in eukaryotes and exists in organisms ranging from fungi to humans (DALEY et al., 2005; LIEBER, 2011). Deleting the gene *KU70* (part of the KU70/80 complex) improves the frequency of gene deletion and integration without negatively affecting the growth parameters of *R. toruloides* (MEI et al., 2014).

Besides the availability of sequenced genomes, promoters, markers, and yeast transformation methods, an efficient system of assembling pathways is also needed. Therefore, this study aimed at the development of a platform to assemble DNA for *R. toruloides* based on Golden Gate Assembly (GGA) methodology in order to fill the gap of advanced metabolic engineering tools for this yeast. This platform was used for simultaneous deletion of *KU70* and overexpression of three native genes of the carotenoids pathway: geranylgeranyl diphosphate synthase (*CRTE*), phytoene dehydrogenase (*CRTI*), and phytoene synthase (*CRTYB*).

According to the proteomic data (presented in chapter 5), the protein CRTI was found upregulated in adapted strain, which can partially explain the higher carotenoid yield compared to the parental strain. The others protein (*CRTE* and *CRTYB*) were not detected in proteomic analysis, which might mean that the expression levels were very low. Therefore, the overexpression of these genes is a potential strategy to enhance the carotenoid production in *R. toruloides*.

6.2 Materials and methods

6.2.1. DNA sequences

The primer sequences used for removing internal BsaI cutting sites as well as to amplify all parts with their respective overhangs and BsaI recognition sites for the GGA can be found in Appendix Table A1. The genomic DNA from *R. toruloides* CCT 0783 (Coleção de Culturas Tropicais, Fundação André Tosello, Campinas, Brazil; synonym IFO10076) was used as a template for the amplification of insertional region, promoters, and genes from the carotenoids pathway. The genomic extraction was performed according to Lõoke et al. (2011). Five hundred

base pairs fragments upstream and downstream the KU70 gene (DNA-dependent ATP-dependent helicase subunit 70, RHTO_06014) were used for the insertional region. The promoters from the genes glyceraldehyde-3-phosphate dehydrogenase (GPD, RHTO_03746), alcohol dehydrogenase 2 (ADH2, RHTO_03062), and xylose reductase (XYL1, RHTO_03963) were obtained by amplifying 786, 725, and 746 bp upstream the starting codon of the gene and called pGPD, pADH2, and pXYL, respectively (DÍAZ et al., 2018; SUN et al., 2017). The native genes used were geranylgeranyl diphosphate synthase (*CRTE*, RHTO_02504), phytoene dehydrogenase (*CRTI*, RHTO_04602), and phytoene synthase (*CRTYB*, RHTO_04605). Nopaline synthase terminator (tNOS) from *Agrobacterium tumefaciens* (GenBank No. MF116010) was used. The geneticin resistance gene (G418) was codon-optimized for *R. toruloides* using an in house tool (<https://github.com/SynBioUniTartu/R.toruloides>) and synthesized by IDT (Leuven, Belgian). G418 was further assembled with pXYL and tNOS, herein called marker (M). All the sequences were analysed in silico to find internal BsaI sites, to be eventually eliminated using assembly PCR technique.

6.2.2. DNA amplification by polymerase chain reaction (PCR)

All sequences were amplified via PCR using high-fidelity Phusion DNA Polymerase or Platinum™ SuperFi™ Green PCR Master Mix (Thermo Fisher Scientific, USA). All reactions were set up and performed according to the manufacturer's instructions. Following PCR, the reactions with Phusion polymerase were stained with 6X TriTrack DNA Loading Dye (Thermo Fisher Scientific, USA). The samples were loaded on a 1% agarose gel and the electrophoresis was run at a constant voltage of 120 V. The fragments sizes were estimated using GeneRuler 1 kb Plus DNA Ladder (Thermo Fisher Scientific, USA). For GGA assembly, the amplicons with the correct size were excised from the gel and purified via FavorPrep™ GEL/PCR Purification Kit (Favorgen, Taiwan). The fragments were quantified using Thermo Fisher's NanoDrop 1000 Spectrophotometer and verified by sequencing.

6.2.3. Golden Gate Assembly

The Golden Gate library comprised the storage of every single part (promoters, genes, marker, insertional region, and terminator) into a pCR-XL-2-TOPO™ plasmid with kanamycin-resistance gene (Thermo Fisher Scientific, USA) (Figure 6.1A). The reaction for TOPO™ Cloning was assembled as follows: 1 µL of pCR-XL-2-TOPO™ Vector, X µL of

insert required for 5:1 molar ratio of insert: vector, 1 μL of Salt Solution (Thermo Fisher Scientific, USA), up to 6 μL of deionized H_2O . The mixture was incubated at 25°C for 1 hour.

The final assembly of the GGA platform is the multigene cassette with three transcriptional units and a selection marker flanked with insertional regions (Figure 6.1B). The reaction for GGA was prepared as follows: 75 ng of pGGA plasmid, 150 ng of each insert (into TOPO™ plasmid), 2 μL of T4 DNA Ligase Buffer, 1 μL of T7 DNA Ligase, 1 μL of BsaI-HFv2 (New England Biolabs, USA), up to 20 μL of nuclease-free H_2O . The reaction was performed on a thermocycler: 30 cycles of 37°C for 5 min and 16°C for 5 min; 50°C for 5 min and 80°C for 5 min.

6.2.4. Bacterial transformation

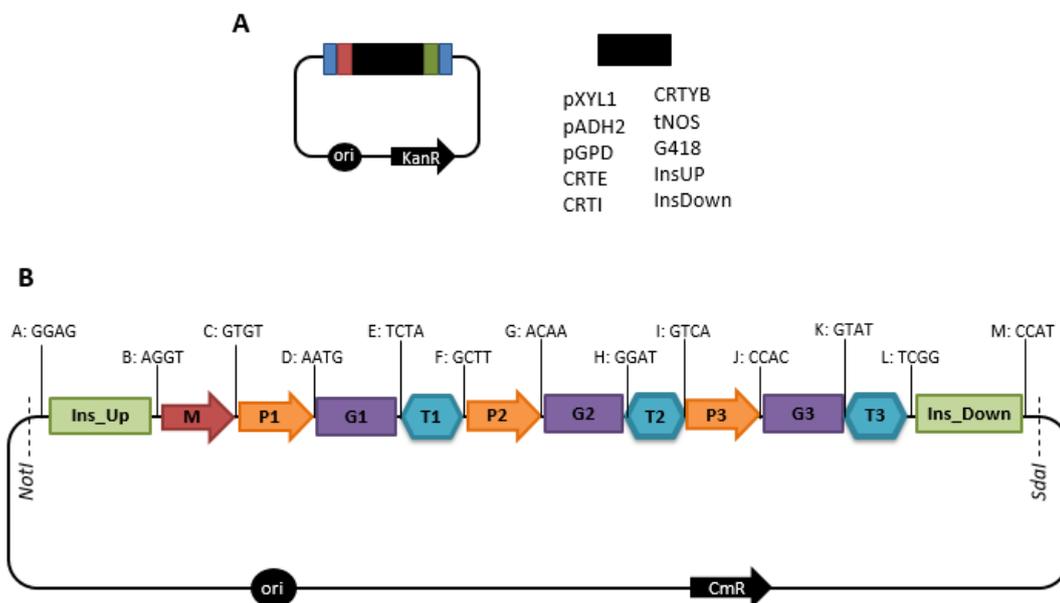
The plasmids were propagated in *Escherichia coli* DH5 α . Cell solution (50 μL) was mixed to 2 μL of GGA or TOPO reaction and resuspended gently. After keeping the mixture on ice for 30 min, the heat shock was done at 42°C for 2 min and then it was replaced on ice for 1 min. 400 μL of LB media (10 g/L tryptone, 5 g/L yeast extract, 10 g/L NaCl) was added to the transformation reaction and the tube was incubated at 37°C , 220 rpm for 1 h. The cells were harvested at 6000 rpm for 1 min and plated on LB plates containing either 100 $\mu\text{g}/\text{mL}$ kanamycin or 25 $\mu\text{g}/\text{mL}$ chloramphenicol and incubated overnight at 37°C .

6.2.5. Yeast transformation

The deletion or expression cassettes were digested with NotI and SdaI (New England Biolabs, USA) before using in the yeast transformation. The transformation was performed using a modified lithium acetate (LiAc) protocol. *R. toruloides* was inoculated in 10 mL of YPD medium and incubated overnight at 30°C and 200 rpm. The culture was diluted to an OD600 of ~ 0.2 in 10 mL of fresh YPD. This culture was incubated at 30°C and 200 rpm until reaching an OD600 of ~ 0.8 . Cells were harvested in a 50 mL tube by centrifugation at 4,000 g for 10 min, washed with 25 mL of sterile water, and resuspended in 1 mL 100 mmol/L LiAc at pH 7.5 and 500 μL of the suspension was transferred to a 1.5 mL microfuge tube and then centrifuged at 8,000 g for 1 min. The cell pellet was resuspended, in the following order, 240 μL of PEG 4000 (50% w/v), 36 μL of 1 mol/L LiAc (pH 7.5), 24 μL of sterile water, 10 μL of salmon sperm DNA (10 mg/mL, preboiled at 100°C for 10 min), and 50 μL of transforming DNA (0.1-10 μg). The solution was mixed by pipetting and incubated at 30°C for 30 min. Then, 34 μL of dimethyl sulfoxide (DMSO) was added and the mixture was heat-shocked at 42°C for 15

min. The transformed cells were harvested at 3000 g for 30 s and resuspended in 2 mL of YPD. The culture was transferred to a 15 mL tube, which was incubated overnight with shaking at 30 °C. The following day the suspension was plated onto YPD plates containing G418 and incubated at 30 °C for two days.

Figure 6.1. Assembly of GG library, in which all three promoters, four genes, two insertional units, and one terminator were cloned into pCR-XL-2-TOPO™ plasmid. The inserts are represented by the black blocks, the red and green blocks represent the 4 nt-overhangs, and the blue blocks, the BsaI recognition sites (A). The Golden Gate Assembly platform in a three transcription unit format constructed on destination vector, pGGA. InsUp/Down - sequence targeting insertion in the 5'/3' region; M - selection marker; P1/2/3 - promoter in transcription unit 1, 2, and 3, respectively; G1/2/3 - gene of interest in transcription unit 1, 2, and 3, respectively; T1/2/3 - terminator located in transcription unit 1, 2, and 3, respectively. For each pre-designed 4-nt overhangs were assigned a letter from A to M. NotI and SadI recognition sites for the release of the expression cassette for yeast transformation (B). Legend: ori – origin of replication; KanR – kanamycin resistance gene; CmR – chloramphenicol resistance gene.



6.2.6. Characterisation of the strains in shake flasks

The strains were characterised in terms of growth and carotenoid production. Mineral medium containing, per litre, 30.0 g carbon source, 0.8 g (NH₄)₂SO₄, 3.0 g KH₂PO₄, 0.5 g MgSO₄·7 H₂O, 1.0 mL vitamins solution and 1.0 mL trace metal solutions were used (LAHTVEE et al., 2017). The C/N ratio of the medium was 80 (mol/mol). The cells were incubated at 30 °C and 200 rpm for 120 h. Samples for OD600 were taken every 24 h, and for carotenoids and cell mass quantification were taken at the end of cultivation.

6.2.7. Analytical methods

Cell growth was monitored by measuring the OD600 (spectrophotometer Shimadzu, Japan). The dry cell mass (DCW) was determined gravimetrically after drying at 60°C for 24h.

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, 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, Germany), and the remaining extracts were resuspended in a known volume of acetone. Carotenoids were measured using Acquity UPLC (Waters, USA) equipped with a TUV detector (Waters, USA) and C18 column (BEH130, 1.7 µm, 2.1 x 100 mm, Waters, 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, USA). Detected peaks were identified according to the known carotenoid retention time profile (LEE et al., 2014; WEBER; ANKE; DAVOLI, 2007). Carotenoids were extracted and quantified using a modified method from Lee (2014).

6.3 Results and discussion

Three different multigene cassettes were assembled (3.1, 3.5, and 3.6) based on the GGA platform, changing the promoters position (Figure 6.2). These assemblies were integrated into *R. toruloides* genome leading to simultaneous deletion of KU70 and overexpression of the carotenoid pathway. After confirming the successful integration of the insertion cassette, by colony PCR and sequencing, the corrected transformants were characterised. The engineered

strains showed a similar profile of growth compared to the parental strain (Figure 6.3). However, these strains indeed showed an increased carotenoid production (Figure 6.4).

Figure 6.2. Golden Gate Assembly bearing the carotenoid synthesis pathway to integration into *R. toruloides* genome. The multigene cassettes (3.1, 3.2, and 3.3) comprised three transcription units and selection marker (G418), flanked with KU70 integration sequences, constructed on a destination vector backbone (pGGA).

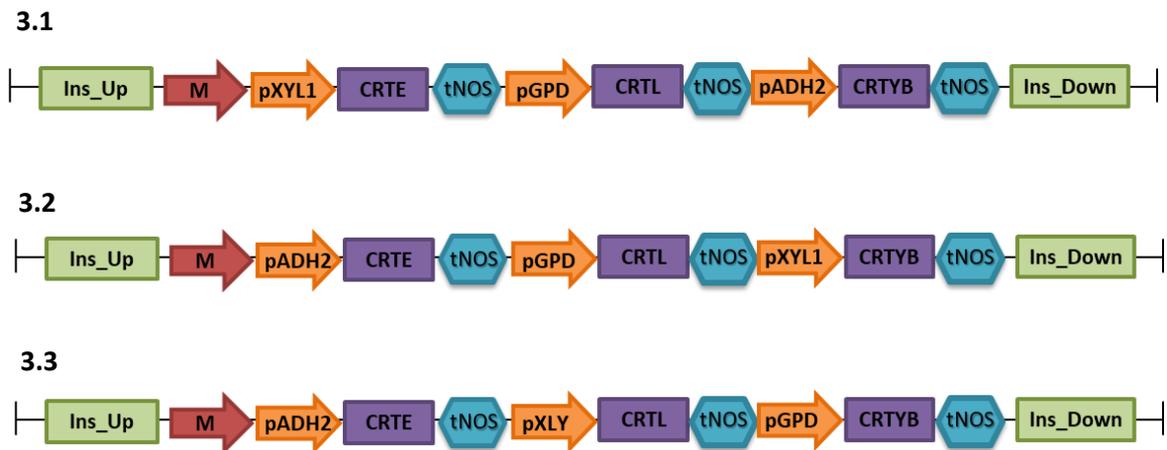


Figure 6.3. Cell growth of Rt 3.1, Rt 3.2 and Rt 3.3 (engineered strain) compared to Rt (parental strain).

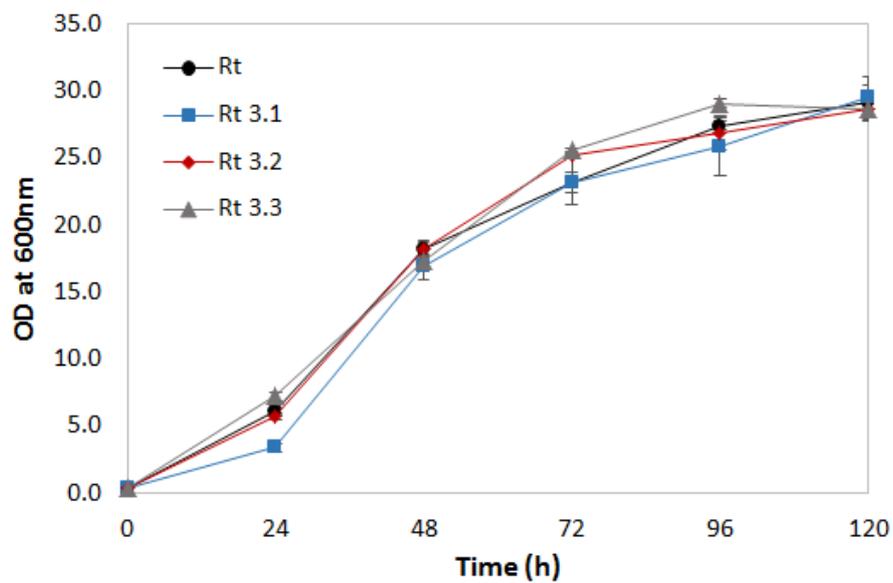
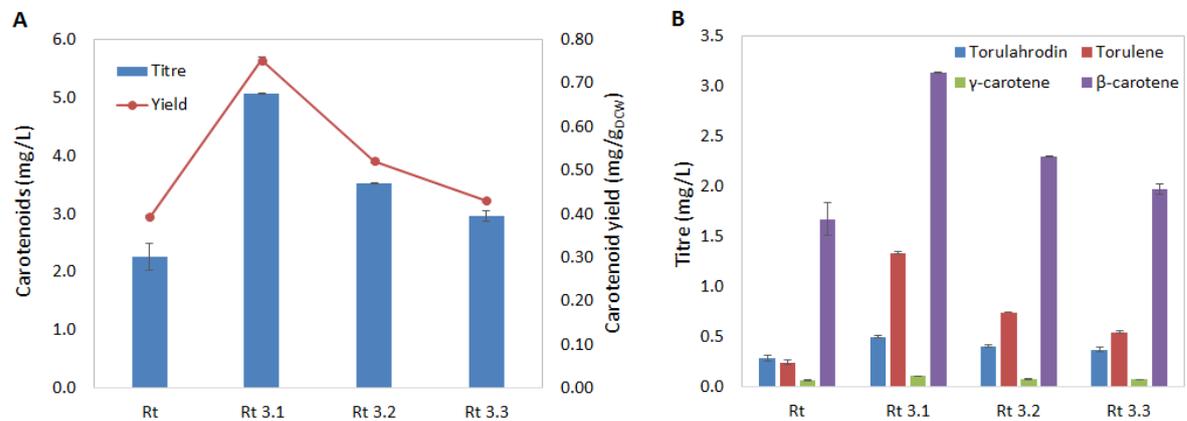


Figure 6.4. Production of total carotenoids and carotenoid yield on biomass in each engineered strain, Rt 3.1, Rt 3.2 and Rt 3.3, and parental strain, Rt (A). Production of different carotenoids in each strain (B).



Rt 3.1 showed the highest increase of carotenoid production (2.2-fold) compared to the parental strain, followed by Rt 3.2 and Rt 3.3 (1.5- and 1.3-fold, respectively, Figure 6.4A). These differences in the increased production can be associated with the promoter position in the final assembly. As expected and confirmed by proteomic data (showed in chapter 5), XLY1 was the third protein most abundant detected during the xylose consumption, indicating that the promoter pXYL can be more active than pGPD and pADH2 (41st and 89th protein most abundant, respectively). In assembly 3.1, pXYL was associated with *CRTE*, the first gene in the carotenoid pathway (Figure 2.2), which directed more carbon flux to this pathway, increasing the total carotenoid production.

For assemblies 3.2 and 3.3, *CRTE* was under control of the pADH2, which showed a lower efficiency in increasing flux through the carotenoid pathway. The promoter pXYL was associated with *CRTYB* and *CRTI*, for assemblies 3.2 and 3.3, respectively, suggesting that Rt 3.2 would produce more β -carotene and Rt 3.3 more torulene and torularhodin. However, this difference was not clear since these genes are involved in more than one reaction during the carotenoid biosynthesis (Figure 6.4 B).

6.4 Conclusion

This study has presented the development of a standardized assembly platform for *R. toruloides*, based on Golden Gate Assembly method. The modularity of this platform allowed a different combination in the expression cassette. The methodology was used here for

overexpressing the genes from the carotenoid pathway. The best-engineered strain of *R. toruloides* obtained in this study produced 5.1 mg/L of carotenoids (2.2-fold higher than the parental strain). To our knowledge, this is the first report of a successful improvement in carotenoid production in *R. toruloides* through metabolic engineering.

This technology can accelerate the construction of engineered strains with the expansion of the building blocks library and enriches the molecular biology toolbox for *R. toruloides*, allowing a faster advance in the improvement of yeast performance aiming its industrial use.

CHAPTER 7

Conclusions and suggestions for future work

7.1 Conclusions

R. toruloides has been considered a potential cell factory, mainly for the industrial production of lipids and carotenoids from a wide variety of substrates. However, the efficiency of biosynthesis of these compounds is still considered low, being a major obstacle for a viable industrial process. Therefore, this work has presented efficient strategies to increase the yield of lipids and carotenoids in *R. toruloides*, combined with the study of the synthesis metabolism of these compounds from xylose, aiming at the application of this yeast in biorefineries.

The induction of an oxidative environment was successfully applied to improve the production of carotenoids and lipids in *R. toruloides*. In this work, the light irradiation resulted in a 30 to 100% increase in carotenoid production and up to 36% increase in lipid production, depending on the type of light and the cultivation conditions. Also, the carotenoid composition could be directed with the use of different light colours.

The addition of H₂O₂ at the beginning of parental *R. toruloides* batch cultivation did not affect the carotenoid production, however, boosted the lipid content to 0.65 g/g_{DCW} (two-fold higher compared to reference condition). This yield has the potential to increase even more with optimization of cultivation conditions. The highest lipid content of *R. toruloides* reported in the literature was 0.67 g/g_{DCW}, in fed-batch cultivation with the discontinuous feeding of pure carbon source, an optimized condition for lipid accumulation. Our results indicate that H₂O₂ use can be explored as a low-cost chemical to improve lipid yield.

The successive application of H₂O₂ through the ALE might overload the defence mechanism against oxidative stress, forcing the increase of carotenoid production. The strain obtained at the end of the ALE was able to accumulate five-fold more carotenoids than the parental strain, demonstrating a successful case of ALE application to improve secondary metabolite production. In bioreactor batch cultivation, the adapted strain also showed a 73% increase in lipid content compared to the parental strain.

The higher lipid production in the oxidative stress conditions might be associated with the overexpression of enzymes involved in the defence mechanism against oxidative stress and the increase in carotenoid content, resulting in a more efficient defence mechanism that can mitigate lipotoxicity and thus favour lipid accumulation.

Besides the upregulation of oxidative stress response proteins, our proteomic data of the adapted strain also showed lower expression levels of proteins related to fatty acid degradation and greater activity of the NADPH regeneration mechanism, which contributed to higher lipid content in this strain. In addition, the carotenoid biosynthesis pathway was more activated by higher levels of CRTI which could explain the higher carotenoid levels.

The physiological characterisation and proteomic data indicated the accumulation of xylitol and arabitol as a bottleneck for efficient xylose catabolism, which could be related to the maintenance of redox balance and the low level expression of xylulokinase. A deeper understanding of these overflow metabolism pathways is essential to improve the efficiency of the conversion of xylose into lipids and carotenoids.

The information from system biology presented in this work contributes to clarifying the molecular mechanism of *R. toruloides* behind the physiological features, metabolism and response to stress conditions, and is essential for developing efficient strategies of metabolic engineering.

Here we demonstrated the first application of Golden Gate Assembly methodology in *R. toruloides*. The developed platform was successfully applied to construct a multigene cassette for overexpression of the carotenoid pathway and the modularity of the system allowed a simple test of a different combination of promoters. The engineered strains of *R. toruloides* was able to produce up to two times more carotenoids compared to the parental strain.

The GGA system is a powerful synthetic biology tool and can accelerate the metabolic engineering process in *R. toruloides* to improve its natural capacities and expand the product range, bringing its use as a cell factory closer to reality.

This work presented different strategies to improve the production of lipids and carotenoids in *R. toruloides*, besides enriching the knowledge about this potential yeast in terms of physiology and metabolism and developing a tool to boost the metabolic engineering process.

7.2 Suggestions for future work

Based on the results presented in this work, it is suggested as possible future works:

1. Studies involving fed-batch cultivation with the adapted and engineered strains obtained in this study, in order to increase cell density and, consequently, the production of carotenoids and lipids;
2. Cultivation of the adapted and engineered strain in lignocellulosic hydrolysate to assess the tolerance of these strains to this complex medium and production of carotenoids and lipids for the application of this yeast in biorefineries;
3. Evaluate xylose catabolism by removing possible arabitol production pathways by deleting LXR and DAD genes and inserting a more efficient heterologous xylulokinase gene. It is expected that the yeast metabolism will rearrange itself to maintain the redox balance without the production of by-products;
4. Overexpress the phosphoketolase pathway and the malic enzyme. Thus, the carbon flux can be directed from PPP to the phosphoketolase pathway and NADPH regeneration can be supplied by the malic enzyme;
5. To reduce the costs of industrial production of carotenoids and fatty acid, it would be important to develop a strain that secretes these compounds, thus reducing the cost of the recovery and purification process.

References

- ADRIO, J. L. Oleaginous yeasts: Promising platforms for the production of oleochemicals and biofuels. **Biotechnology and Bioengineering**, v. 114, n. 9, p. 1915–1920, 2017.
- ARMSTRONG, G. A; HEARST, J. E. Carotenoids 2: Genetics and molecular biology of carotenoid pigment biosynthesis. **FASEB Journal**, v. 10, n. 2, p. 228–237, 1996.
- AUSICH, R. L. Commercial opportunities for carotenoid production by biotechnology. **Pure and Applied Chemistry**, v. 69, n. 10, p. 2169–2174, 1997.
- SAMPAIO, P. *Rhodospiridium Banno* (1967). **The Yeasts, a Taxonomic Study**, n. 1967, 2011.
- BARUAH, J. et al. Recent trends in the pretreatment of lignocellulosic biomass for value-added products. **Frontiers in Energy Research**, v. 6, n. DEC, p. 1–19, 2018.
- BBC Research (2018). *The Global Market for Carotenoids*. Available online at: <<https://www.globenewswire.com/news-release/2018/05/21/1509412/0/en/Global-Carotenoids-Market-to-Reach-2-0-Billion-by-2022.html>>. Accessed on 20th February 2020.
- BENDICH, A.; OLSONT, J. A. Biological actions of Carotenoids. **The FASEB Journal**, v. 3, n. 8, p. 1927–1932, 1988.
- BENJAMINI, Y., AND HOCHBERG, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. **Journal of Royal Statistical Society**, v. 57, p.298–300, 1995.
- BHOSALE, P. B.; GADRE, R. V. Production of beta-carotene by a mutant of *Rhodotorula glutinis*. **Applied Microbiology and Biotechnology**, v. 55, n. 4, p. 423–427, 2001.
- BIOECONOMY POLICY, 2020. Available at: <<https://ec.europa.eu/research/bioeconomy/index.cfm?pg=policy&lib=bbpp>>. Accessed on 20th June 2020.
- BLOMQVIST, J. et al. Oleaginous yeast as a component in fish feed. **Scientific Reports**, 8, p. 1–8, 2018.
- BONTURI, N. et al. Single cell oil producing yeasts *Lipomyces starkeyi* and *Rhodospiridium toruloides*: Selection of extraction strategies and biodiesel property prediction. **Energies**, v. 8, n. 6, p. 5040–5052, 2015.
- BONTURI, N. et al. Microbial oil production in sugarcane bagasse hemicellulosic hydrolysate without nutrient supplementation by a *Rhodospiridium toruloides* adapted strain. **Process Biochemistry**, v. 57, p. 16–25, 2017.
- BRAUNWALD, T. et al. Effect of different C/N ratios on carotenoid and lipid production by *Rhodotorula glutinis*. **Applied Microbiology and Biotechnology**, v. 97, n. 14, p. 6581–6588, 2013.
- BUZZINI, P. et al. Carotenoid profiles of yeasts belonging to the genera *Rhodotorula*, *Rhodospiridium*, *Sporobolomyces*, and *Sporidiobolus*. **Canadian Journal of Microbiology**, v. 53, n. 8, p. 1024–31, 2007.

CARDOSO, L. A. C. et al. Improvement of *Sporobolomyces ruberrimus* carotenoids production by the use of raw glycerol. **Bioresource Technology**, v. 200, p. 374–379, 2016.

CELINSKA, E. et al. A synthetic biology approach to transform *Yarrowia lipolytica* into a competitive biotechnological producer of β -carotene. **Biotechnology and Bioengineering**, v. 115, n. 2, p. 464–472, 2018.

CELINSKA, E. et al. Golden Gate Assembly system dedicated to complex pathway manipulation in *Yarrowia lipolytica*. **Microbial Biotechnology**, v.10, p. 450 - 455, 2017.

CHANDEL, A. K.; GARLAPATI, V. K.; SINGH, A. K. Detoxification of lignocellulosic hydrolysates for improved bioethanol production. **Biofuel Production-Recent Developments and Prospects**, v. 10, p. 225–246, 2011.

CHAO, R.; YUAN, Y.; ZHAO, H. Recent advances in DNA assembly technologies. **FEMS Yeast Research**, v.15, p. 1–9, 2015.

CHERUBINI, F.; WELLISCH, M.; WILLKE, T. Toward a common classification approach for biorefinery systems. **Biofuels, Bioproducts and Biorefinery**, 2009.

CHOKSHI, K. et al. Biotechnology for Biofuels Nitrogen starvation-induced cellular crosstalk of ROS-scavenging antioxidants and phytohormone enhanced the biofuel potential of green microalga *Acutodesmus dimorphus*. **Biotechnology for Biofuels**, p. 1–12, 2017.

CHRISTOPHE, G. et al. Production of oils from acetic acid by the oleaginous yeast *Cryptococcus curvatus*. **Applied Biochemistry and Biotechnology**, v. 167, n. 5, p. 1270–1279, 2012.

CORADETTI, S. T. et al. Functional genomics of lipid metabolism in the oleaginous yeast *Rhodospiridium toruloides*. **eLife**, v. 7, p. e32110, 2018.

COX, J., AND MANN, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. **Nature Biotechnology**, v. 26, p. 1367–1372, 2008.

DALEY, J. M. et al. Nonhomologous End Joining in Yeast. **Annual Review of Genetics**, v. 39, n. 1, p. 431–451, 2005.

DÍAZ, T. et al. Combining evolutionary and metabolic engineering in *Rhodospiridium toruloides* for lipid production with non-detoxified wheat straw hydrolysates. **Applied Microbiology and Biotechnology**, v. 102, p. 3287–3300, 2018.

DINH, H. V et al. A comprehensive genome-scale model for *Rhodospiridium toruloides* IFO0880 accounting for functional genomics and phenotypic data. **Metabolic Engineering Communications**, v. 9, p. e00101, 2019.

DRAGOSITS, M.; MATTANOVICH, D. Adaptive laboratory evolution – principles and applications for biotechnology TL - 12. **Microbial Cell Factories**, v. 12, p. 64, 2013.

DU, C. et al. Biochemical and Biophysical Research Communications The suppression of torulene and torularhodin treatment on the growth of PC-3 xenograft prostate tumors *. **Biochemical and Biophysical Research Communications**, v. 469, p. 1146–1152, 2016.

EDGE, R.; MCGARVEY, D. J.; TRUSCOTT, T. G. The carotenoids as anti-oxidants--a review.

Journal of Photochemistry and Photobiology, v. 41, n. 3, p. 189–200, 1997.

ELENA, S. F. et al. Effects of population size and mutation rate on the evolution of mutational robustness. **The Society for the Study of Evolution**, p. 666–674, 2007.

ENGLER, C.; KANDZIA, R.; MARILLONNET, S. A One Pot , One Step , Precision Cloning Method with High Throughput Capability. **PlusOne**, v. 3, e3647, 2008.

FALLIS, A. **Molecular Diagnostics: Promises and Possibilities**. London: Springer Science, 2010.

FAN, J. et al. Lipid accumulation and biosynthesis genes response of the oleaginous *Chlorella pyrenoidosa* under three nutrition stressors. **Biotechnology for Biofuels**, v. 7, n. 1, p. 1–14, 2014.

FERNANDES, S.; MURRAY, P. Metabolic engineering for improved microbial pentose fermentation. **Bioengineered Bugs**, v. 1, n. 6, p. 424–428, 2010.

FILLET, S. et al. Engineering *Rhodospiridium toruloides* for the production of very long-chain monounsaturated fatty acid-rich oils. **Applied Microbiology and Biotechnology**, v.101, p. 7271–7280, 2017.

FOLCH, J.; LEES, M.; SLOANE STANLEY, G. Simple method for the isolation and purification of total lipids from animal tissues. **Journal of Biological Chemistry**, v. 226, p. 497 – 509, 1957.

FRENGOVA, G. I.; BESHKOVA, D. M. Carotenoids from *Rhodotorula* and *Phaffia*: Yeasts of biotechnological importance. **Journal of Industrial Microbiology and Biotechnology**, v. 36, n. 2, p. 163–180, 2009.

GADANHO, M.; LIBKIND, D.; SAMPAIO, J. P. Yeast diversity in the extreme acidic environments of the Iberian pyrite belt. **Microbial Ecology**, v. 52, n. 3, p. 552–563, 2006.

GIBSON, D. G. et al. Enzymatic assembly of DNA molecules up to several hundred kilobases. **Nature Methods**, v. 6, n. 5, p. 12–16, 2009.

GLOBAL MARKET INSIGHT, 2020. Available at: <<https://www.gminsights.com/industry-analysis/biotechnology-market>>. Accessed on 20th June 2020.

GLOBENEWSWIRE, 2017. Available at < <https://www.globenewswire.com/news-release/2019/10/15/1929461/0/en/Global-Carotenoids-Market-is-expected-to-reach-USD-3-59-billion-by-2025-Fior-Markets.html>>. Accessed 20th June 2020.

GONG, G. et al. Multi-omics metabolism analysis on irradiation-induced oxidative stress to *Rhodotorula glutinis*. **Applied Microbiology and Biotechnology**, p. 361–374, 2019.

GONG, G.; ZHANG, X.; TAN, T. Bioresource Technology Simultaneously enhanced intracellular lipogenesis and β -carotene biosynthesis of *Rhodotorula glutinis* by light exposure with sodium acetate as the substrate. **Bioresource Technology**, v. 295, n. September 2019, p. 122274, 2020.

HU, J.; JI, L. Draft Genome Sequences of *Rhodospiridium toruloides* Strains ATCC 10788 and ATCC 10657 with Compatible Mating Types. **Genome Announcements**, v. 4, n. 2, p. e00098-16, 2016a.

- HU, J.; JI, L. Draft Genome Sequences of *Rhodospiridium toruloides* Strains ATCC 10788 and ATCC 10657 with Compatible Mating Types. **American Society for Microbiology**, v. 4, n. 2, p. 1083–1084, 2016b.
- IRAZUSTA, V. et al. Relationship among carotenoid production, copper bioremediation and oxidative stress in *Rhodotorula mucilaginosa* RCL-11. **Process Biochemistry**, v. 48, n. 5–6, p. 803–809, 2013.
- JAGTAP, S. S.; RAO, C. V. Production of D-arabitol from D-xylose by the oleaginous yeast *Rhodospiridium toruloides* IFO0880. **Applied Microbiology and Biotechnology**, p. 143–151, 2018.
- JIAO, X. et al. Developing a CRISPR/Cas9 System for Genome Editing in the Basidiomycetous Yeast *Rhodospiridium toruloides*. **Biotechnology Journal**, v. 14, n. 7, p. 1–7, 2019.
- JIE, J. et al. Metabolomic profiling of *Rhodospiridium toruloides* grown on glycerol for carotenoid production during different growth phases. **Journal of Agricultural and Food Chemistry**, v.62, p. 10203-10209, 2014.
- JL LEE, J.; WN CHEN, W. The Production, Regulation and Extraction of Carotenoids from *Rhodospiridium toruloides*. **Journal of Molecular and Genetic Medicine**, v. 10, n. 2, p. 14–16, 2016.
- JOHNS, A. M. B.; LOVE, J.; AVES, S. J. Four inducible promoters for controlled gene expression in the oleaginous yeast *Rhodotorula toruloides*. **Frontiers in Microbiology**, v. 7, n. OCT, p. 1–12, 2016.
- KERKHOVEN, E. J.; LAHTVEE, P.; NIELSEN, J. Applications of computational modeling in metabolic engineering of yeast. **FEMS Yeast Research**, v. 15, p. 1–13, 2015.
- KOK, S. DE et al. Rapid and Reliable DNA Assembly via Ligase Cycling Reaction. **ACS Synthetic Biology**, v. 3, p. 97 - 106, 2014.
- KOPPRAM, R.; ALBERS, E.; OLSSON, L. Evolutionary engineering strategies to enhance tolerance of xylose utilizing recombinant yeast to inhibitors derived from spruce biomass. **Biotechnology for Biofuels**, v. 5, n. 1, p. 32, 2012.
- KOT, A. M. et al. *Rhodotorula glutinis*—potential source of lipids, carotenoids, and enzymes for use in industries. **Applied Microbiology and Biotechnology**, v. 100, n. 14, p. 6103–6117, 2016.
- KOT, A. M. et al. Torulene and torularhodin: “New” fungal carotenoids for industry? **Microbial Cell Factories**, v. 17, n. 1, p. 1–14, 2018.
- KOT, A. M. et al. Effect of exogenous stress factors on the biosynthesis of carotenoids and lipids by *Rhodotorula* yeast strains in media containing agroindustrial waste. **World Journal of Microbiology and Biotechnology**, v. 35, n. 10, p. 1–10, 2019.
- KOUTINAS, A. A. et al. Design and techno-economic evaluation of microbial oil production as a renewable resource for biodiesel and oleochemical production. **Fuel**, v. 116, p. 566–577, 2014.
- KOUTINAS, A. A., PAPANIKOLAOU, S. Biodiesel production from microbial oil. sawston: Woodhead Publishing Limited, 177–198, 2011.

- KRAPPMANN, S. Gene targeting in filamentous fungi: the benefits of impaired repair. **Fungal Biology Reviews**, v. 21, n. 1, p. 25–29, 2007.
- KRINSKY, N. I. Antioxidant functions of carotenoids. **Free Radical Biology and Medicine**, v. 7, n. 6, p. 617–635, 1989.
- KUEHNE, A. et al. Activation of Oxidative Pentose Phosphate Pathway as First-Line Response to Oxidative Stress in Human Skin Cells. **Molecular Cell**, v. 59, n. 3, p. 359–371, 2015.
- KUMAR, S. et al. Genome sequence of the oleaginous red yeast *Rhodospiridium toruloides* MTCC 457. **Eukaryotic Cell**, v. 11, n. 8, p. 1083–1084, 2012.
- LAHTVEE, P. J. et al. Absolute Quantification of Protein and mRNA Abundances Demonstrate Variability in Gene-Specific Translation Efficiency in Yeast. **Cell Systems**, v. 4, n. 5, p. 495–504.e5, 2017.
- LEE, J. J. L. et al. Metabolomic profiling of *Rhodospiridium toruloides* grown on glycerol for carotenoid production during different growth phases. **Journal of Agricultural and Food Chemistry**, v. 62, n. 41, p. 10203–10209, 2014.
- LI, M. Z.; ELLEDGE, S. SLIC: a method for sequence- and ligation-independent cloning. **Methods in Molecular Biology**, v. 852, 2012.
- LI, Y.; ZHAO, Z. (KENT); BAI, F. High-density cultivation of oleaginous yeast *Rhodospiridium toruloides* Y4 in fed-batch culture. **Enzyme and Microbial Technology**, v. 41, n. 3, p. 312–317, 2007.
- LI, Z. et al. Overexpression of malic enzyme (ME) of *Mucor circinelloides* improved lipid accumulation in engineered *Rhodotorula glutinis*. **Applied Microbiology and Biotechnology**, v. 97, p. 4927–4936, 2013.
- LIEBER, M. R. The mechanism of DSB repair by the NHEJ. **Annual Review of Biochemistry**, v. 79, n. 3, p. 181–211, 2011.
- LIN, X. et al. Functional integration of multiple genes into the genome of the oleaginous yeast *Rhodospiridium toruloides*. **FEMS Yeast Research**, v. 14, n. 4, p. 547–555, 2014.
- LIU, H. et al. Fast and efficient genetic transformation of oleaginous yeast *Rhodospiridium toruloides* by using electroporation. **FEMS Yeast Research**, v. 17, n. 2, p. 1–11, 2017.
- LIU H. et al. Comparative proteomic analysis of *Rhodospiridium toruloides* during lipid accumulation. **Yeast**, v. 26, p. 553–566, 2009.
- LIU, J.; HUANG, J.; CHEN, F. Microalgae as feedstocks for biodiesel production, biodiesel feedstocks and processing technologies, InTech, 2011. Available online at: <<http://www.intechopen.com/books/biodiesel-feedstocks-and-processingtechnologies/microalgae-as-feedstocks-for-biodiesel-production>>. Accessed on Set/2020.
- LIU, W. et al. Journal of Experimental Marine Biology and Ecology Formation of triacylglycerol in *Nitzschia closterium* f. *minutissima* under nitrogen limitation and possible physiological and biochemical mechanisms. **Journal of Experimental Marine Biology and Ecology**, v. 418–419, p. 24–29, 2012.

- LIU, Y. et al. Characterization of glyceraldehyde-3-phosphate dehydrogenase gene RtGPD1 and development of genetic transformation method by dominant selection in oleaginous yeast *Rhodospiridium toruloides*. **Applied Microbiology and Biotechnology**, v. 97, n. 2, p. 719–729, 2013.
- LIU, Y. et al. Engineering an efficient and tight d-amino acid-inducible gene expression system in *Rhodospiridium/Rhodotorula* species. **Microbial Cell Factories**, v. 14, n. 1, p. 1–16, 2015.
- LIU, Y. et al. Developing a set of strong intronic promoters for robust metabolic engineering in oleaginous *Rhodotorula* (*Rhodospiridium*) yeast species. **Microbial Cell Factories**, v. 15, n. 1, p. 1–9, 2016.
- LOPES, H. et al. C/N ratio and carbon source-dependent lipid production profiling in *Rhodotorula toruloides*. **Applied Microbiology and Biotechnology**, v.104, p. 2639-2649 2020a.
- LOPES, H. J. S.; BONTURI, N.; MIRANDA, E. A. *Rhodotorula toruloides* single cell oil production using eucalyptus urograndis hemicellulose hydrolysate as a carbon source. **Energies**, v. 13, n. 4, 2020b.
- LÕOKE, M.; KRISTJUHAN, K.; KRISTJUHAN, A. Extraction of genomic DNA from yeasts for PCR-based applications. **Biotechniques**, v. 50, n. 5, p. 325–328, 2011.
- COX, J.; MANN, M. MaxQuant individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. **Nature Biotechnology**, v. 26, n. 12, p. 1367–1372, 2008.
- MATA-GÓMEZ, L. C. et al. Biotechnological production of carotenoids by yeasts: an overview. **Microbial Cell Factories**, v. 13, n. January, p. 12, 2014.
- MEESTERS, P.; HUIJBERTS, G.; EGGINK, G. High-cell-density cultivation of the lipid accumulating yeast *Cryptococcus curvatus* using glycerol as a carbon source. **Applied Microbiology and Biotechnology**, v. 45, p. 575–579, 1996.
- MEI, C. et al. Molecular characterization of KU70 and KU80 homologues and exploitation of a KU70-deficient mutant for improving gene deletion frequency in *Rhodospiridium toruloides*. **BMC Microbiology**, v. 14, n. 1, p. 1–10, 2014.
- METZL-RAZ, E. et al. Principles of cellular resource allocation revealed by condition-dependent proteome profiling. **eLife**, 6:e28034, p. 1–21, 2017.
- MILLER, G. L. Use of Dinitrosalicylic Acid Reagent for Determination of Reducing Sugar. **Analytical Chemistry**, v. 31, n. 3, p. 426–428, 1959.
- MORIN, N. et al. Draft Genome Sequence of *Rhodospiridium toruloides* CECT1137, an Oleaginous Yeast of Biotechnological Interest. **Genome Announcements**, v. 2, n. 4, p. 578–579, 2014.
- NIEHUS, X. et al. Biotechnology for Biofuels Engineering *Yarrowia lipolytica* to enhance lipid production from lignocellulosic materials. **Biotechnology for Biofuels**, p. 1–10, 2018.
- NILSSON, A.; NIELSEN, J. Metabolic Trade-offs in Yeast are Caused by F1F0-ATP synthase. **Nature Publishing Group**, p. 0–11, 2016.
- OLIVEIRA, P. M. et al. Screening and Growth Characterization of Non-conventional Yeasts

- in a Hemicellulosic Hydrolysate. **Frontiers in Bioengineering and Biotechnology**, v. 9, p. 1–13, 2021.
- OTOUPAL, P. B. et al. Multiplexed CRISPR-Cas9-based genome editing of *Rhodospiridium toruloides*. **MSphere**, v. 4, 2019.
- PAPANIKOLAOU, S.; AGGELIS, G. Lipids of oleaginous yeasts. Part I: Biochemistry of single cell oil production. **European Journal of Lipid Science and Technology**, v. 113, p. 1031–1051, 2011.
- PARDO, B.; GÓMEZ-GONZÁLEZ, B.; AGUILERA, A. DNA double-strand break repair: How to fix a broken relationship. **Cellular and Molecular Life Sciences**, v. 66, n. 6, p. 1039–1056, 2009.
- PARK, Y. K.; NICAUD, J. M.; LEDESMA-AMARO, R. The Engineering Potential of *Rhodospiridium toruloides* as a Workhorse for Biotechnological Applications. **Trends in Biotechnology**, p. 1–14, 2017.
- PATEL, A.; SHAH, A. R. Integrated lignocellulosic biorefinery : Gateway for production of second generation ethanol and value added products. **Journal of Bioresources and Bioproducts**, v. 6, n. 2, p. 108–128, 2021.
- PETERSON, H. et al. g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). **Nucleic Acids Research** , v. 47, p. 191–198, 2019.
- PINHEIRO, M. J. et al. Xylose Metabolism and the Effect of Oxidative Stress on Lipid and Carotenoid Production in *Rhodotorula toruloides*: Insights for Future Biorefinery. **Frontiers in Bioengineering and Biotechnology**, v. 8, p. 1–15, 2020.
- POLÍVKA, T.; FRANK, H. A. Molecular Factors Controlling Photosynthetic Light-Harvesting by Carotenoids. **Accounts of Chemical Research**, v. 43, p. 1125 – 1134, 2010.
- QI, F. et al. Integrative transcriptomic and proteomic analysis of the mutant lignocellulosic hydrolyzate-tolerant *Rhodospiridium toruloides*. **Engineering in Life Science**, v. 17, p. 249–261, 2017.
- QUILES-ROSILLO, M. D. et al. Light induction of the carotenoid biosynthesis pathway in *Blakeslea trispora*. **Fungal Genetics and Biology**, v. 42, p. 141–153, 2005.
- RATLEDGE, C. Microorganisms for lipids. **Acta Biotechnologica**, v. 11, n. 5, p. 429–438, 1991.
- RATLEDGE, C.; WYNN, J. P. The Biochemistry and Molecular Biology of Lipid Accumulation in Oleaginous Microorganisms. **Advances in Applied Microbiology**, v. 51, p. 1–51, 2002.
- RAUDVERE, U. et al. g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). **Nucleic Acids Research**, v. 47, p. 191–198, 2019.
- REYES, L. H.; GOMEZ, J. M.; KAO, K. C. Improving carotenoids production in yeast via adaptive laboratory evolution. **Metabolic Engineering**, v. 21, p. 26 – 33, 2014.

- SALARI, R. Investigation of the Best *Saccharomyces cerevisiae* Growth Condition. **Electronic Physician**, v. 9, p. 3592-3597, 2017.
- SAMBLES, C. et al. Genome sequence of the oleaginous yeast *Rhodotorula toruloides* strain. **Genomics Data**, v. 13, p. 1–2, 2017.
- SÁNCHEZ, B. J. et al. Improving the phenotype predictions of a yeast genome-scale metabolic model by incorporating enzymatic constraints. **Molecular System Biology**, v.13, p. 1–16, 2017.
- SÁNCHEZ, B. J. et al. Benchmarking accuracy and precision of intensity-based absolute quantification of protein abundances in *Saccharomyces cerevisiae*. **Proteomics**, v. 21, 2000093, 2021.
- SAUER, U. Evolutionary engineering of industrially important microbial phenotypes. **Advances in biochemical engineering/biotechnology**, v. 73, p. 129–169, 2001.
- SCHULTZ, J. C; CAO, M.; ZHAO, H. Development of a CRISPR / Cas9 system for high efficiency multiplexed gene deletion in *Rhodospiridium toruloides*. **Biotechnology and Bioengineering**, v. 116, p. 2103 - 2109, 2019.
- SCOTT, M. et al. Emergence of robust growth laws from optimal regulation of ribosome synthesis. **Molecular Sytem Biology**, p. 1–14, 2014.
- SHI, K. et al. Reactive Oxygen Species-Mediated Cellular Stress Response and Lipid Accumulation in Oleaginous Microorganisms: The State of the Art and Future Perspectives. **Frontiers in Microbiology**, v. 8, p. 1–9, 2017.
- SOCOOL, C. R. et al. Pilot scale biodiesel production from microbial oil of *Rhodospiridium toruloides* DEBB 5533 using sugarcane juice: Performance in diesel engine and preliminary economic study. **Bioresource Techonology**, v. 223, p. 259 -268, 2017.
- STAHL, W.; SIES, Æ. H. Carotenoids and Flavonoids Contribute to Nutritional Protection against Skin Damage from Sunlight. **Molecular Biotechnology**, v.37, p. 26–30, 2007.
- SUN, W. et al. Homologous gene targeting of a carotenoids biosynthetic gene in *Rhodospiridium toruloides* by *Agrobacterium*-mediated transformation. **Biotechnology Letters**, v. 39, n. 7, p. 1001–1007, 2017.
- TAN, B. L.; NORHAIZAN, M. E. Carotenoids: How effective are they to prevent age-related diseases? **Molecules**, v. 24, p. 1801, 2019.
- TIUKOVA, I. et al. Proteome analysis of xylose metabolism in *Rhodotorula toruloides* during lipid production. **bioRxiv**, p. 601930, 2019a.
- TIUKOVA, I. A. et al. Genome-scale model of *Rhodotorula toruloides* metabolism. **Biotechnology and Bioengineering**, v. 116, n. 12, p. 3396–3408, 2019b.
- TSAI, Y. Y. et al. Development of a sufficient and effective procedure for transformation of an oleaginous yeast, *Rhodospiridium toruloides* DMKU3-TK16. **Current Genetics**, v. 63, n. 2, p. 359–371, 2017.

- UBANDO, A. T.; FELIX, C. B.; CHEN, W. H. Biorefineries in circular bioeconomy: A comprehensive review. **Bioresource Technology**, v. 299, n. November 2019, 2020b.
- UNREAN, P.; KHAJEERAM, S.; CHAMPREDA, V. Combining metabolic evolution and systematic fed-batch optimization for efficient single-cell oil production from sugarcane bagasse. **Renewable Energy**, v. 111, p. 295–306, 2017.
- VALDIVIA, M. et al. Biofuels 2020: Biorefineries based on lignocellulosic materials. **Microbial Biotechnology**, v. 9, n. 5, p. 585–594, 2016.
- VÄREMO, L., NIELSEN, J., AND NOOKAEW, I. Enriching the gene set analysis of genome-wide data by incorporating directionality of gene expression and combining statistical hypotheses and methods. **Nucleic Acids Research**, v. 41, p. 4378–4391, 2013.
- VIZCAÍNO, J. A. et al. 2016-update of the PRIDE database and its related tools. **Nucleic Acids Research**, v. 44, p.447–456, 2016.
- WANG, H. et al. RAVEN 2.0: A versatile toolbox for metabolic network reconstruction and a case study on *Streptomyces coelicolor*. **Plos Computational Biology**, v. 14, p. 1–17, 2018a.
- WANG, Y. et al. Cloning and evaluation of different constitutive promoters in the oleaginous yeast *Rhodospiridium toruloides*. **Yeast**, v. 33, p. 99–106, 2016.
- WANG, Y. et al. Systems analysis of phosphate-limitation-induced lipid accumulation by the oleaginous yeast *Rhodospiridium toruloides*. **Biotechnology for Biofuels**, v. 11, p. 1–15, 2018b.
- WASYLENKO, T. M.; AHN, W. S.; STEPHANOPOULOS, G. The oxidative pentose phosphate pathway is the primary source of NADPH for lipid overproduction from glucose in *Yarrowia lipolytica*. **Metabolic Engineering**, v. 30, p. 27–39, 2015.
- WEBER, R. W. S.; ANKE, H.; DAVOLI, P. Simple method for the extraction and reversed-phase high-performance liquid chromatographic analysis of carotenoid pigments from red yeasts (Basidiomycota, Fungi). **Journal of Chromatography A**, v. 1145, n. 1–2, p. 118–122, 2007.
- WEI, Y.; SIEWERS, V.; NIELSEN, J. Cocoa butter-like lipid production ability of non-oleaginous and oleaginous yeasts under nitrogen-limited culture conditions. **Applied Microbiology and Biotechnology**, v. 101, n. 9, p. 3577–3585, 2017.
- WEN, Z. et al. *Rhodospiridium toruloides* - A potential red yeast chassis for lipids and beyond. **FEMS Yeast Research**, v.20, p. 1–12, 2020.
- WIEBE, M. G. et al. Lipid production in batch and fed-batch cultures of *Rhodospiridium toruloides* from 5 and 6 carbon carbohydrates. **BMC Biotechnology**, v. 12, n. 1, p. 26, 2012.
- XU, P. Engineering oxidative stress defense pathways to build a robust lipid production platform in *Yarrowia lipolytica*. **Biotechnology and Bioengineering**, v. 114, n. 7, p. 1521–1530, 2017.
- YANG, F. et al. Identification of the orotidine-5'-monophosphate decarboxylase gene of the oleaginous yeast *Rhodospiridium toruloides*. **Yeast**, v. 25, n. 9, p. 623–630, set. 2008.
- YANG, X. et al. Expression of phosphotransacetylase in: *Rhodospiridium toruloides* leading

to improved cell growth and lipid production. **RSC Advances**, v. 8, n. 43, p. 24673–24678, 2018.

YEN, H. W.; YANG, Y. C. The effects of irradiation and microfiltration on the cells growing and total lipids production in the cultivation of *Rhodotorula glutinis*. **Bioresource Technology**, v. 107, p. 539–541, 2012.

YILANCIOGLU, K. et al. Oxidative stress is a mediator for increased lipid accumulation in a newly isolated *Dunaliella salina* strain. **PlusOne**, v. 9, e91957., 2014.

ZHANG, S. et al. Engineering *Rhodospiridium toruloides* for increased lipid production. **Biotechnology and Bioengineering**, v. 113, 2016a.

ZHANG, S. et al. Metabolic engineering of the oleaginous yeast *Rhodospiridium toruloides* IFO0880 for lipid overproduction during high-density fermentation. **Applied Microbiology and Biotechnology**, v. 100, n. 21, p. 9393–9405, 2016b.

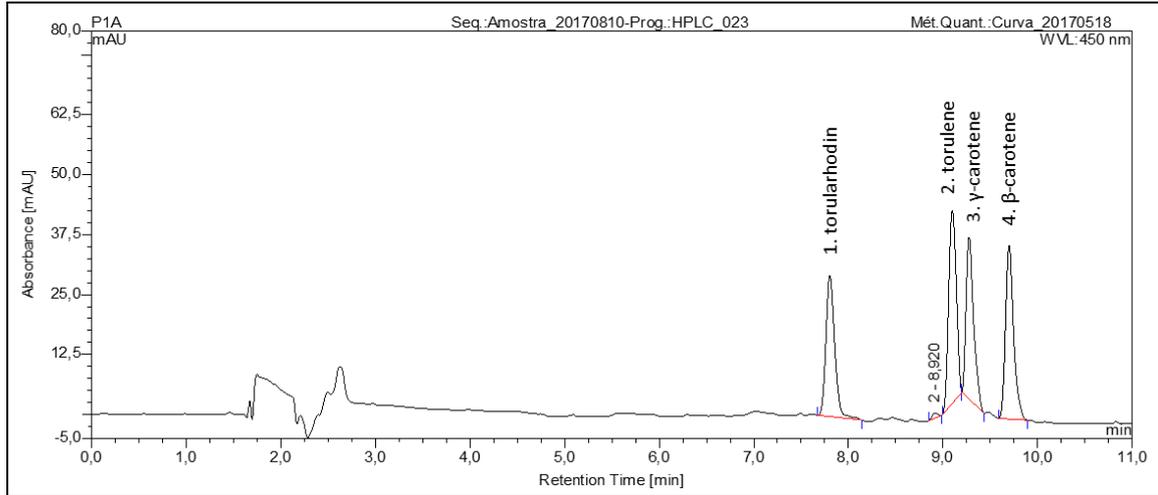
ZHANG, Z.; ZHANG, X.; TAN, T. Bioresource Technology Lipid and carotenoid production by *Rhodotorula glutinis* under irradiation / high-temperature and dark / low-temperature cultivation. **Bioresource Technology**, v. 157, p. 149–153, 2014.

ZHU, Z. et al. A multi-omic map of the lipid-producing yeast *Rhodospiridium toruloides*. **Nature Communications**, v. 3, p. 1112, 2012.

Appendix

Figure A1. (A) HPLC profile of the carotenoids produced by *R. toruloides* CCT7915 identifying the following carotenoids: torularhodin, 1; torulene, 2; γ -carotene, 3; β -carotene, 4. (B) Absorbance spectra of each identified carotenoids.

(A)



(B)

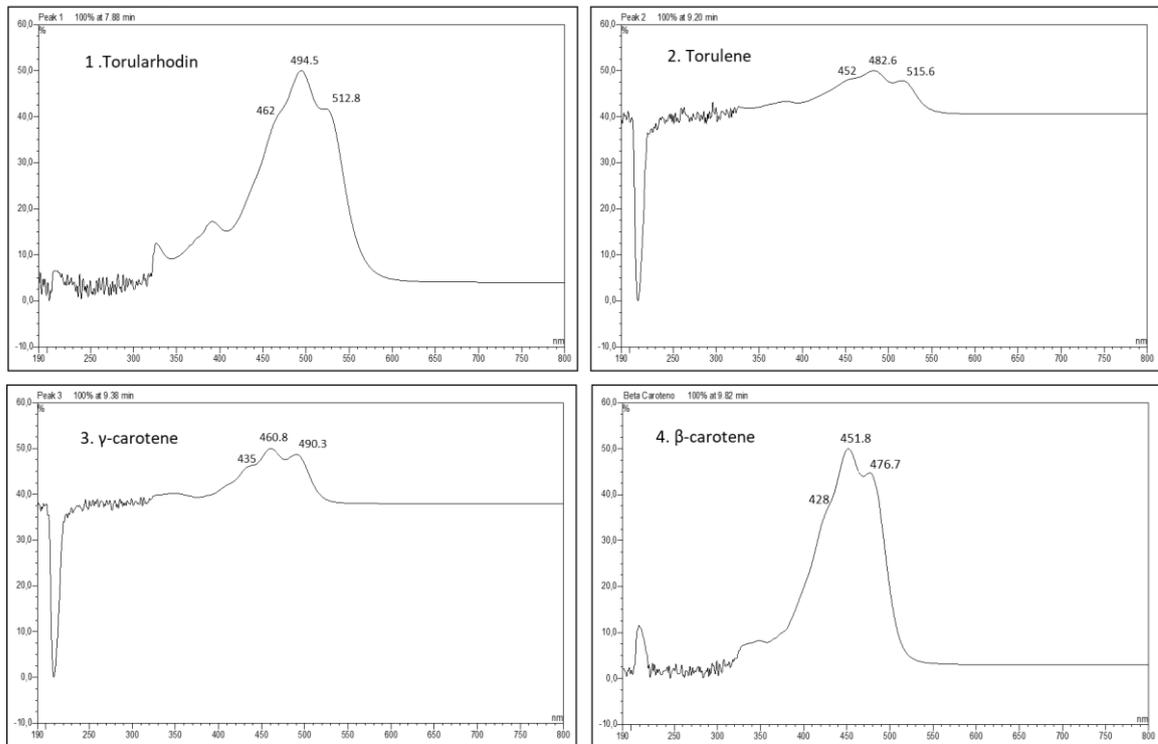
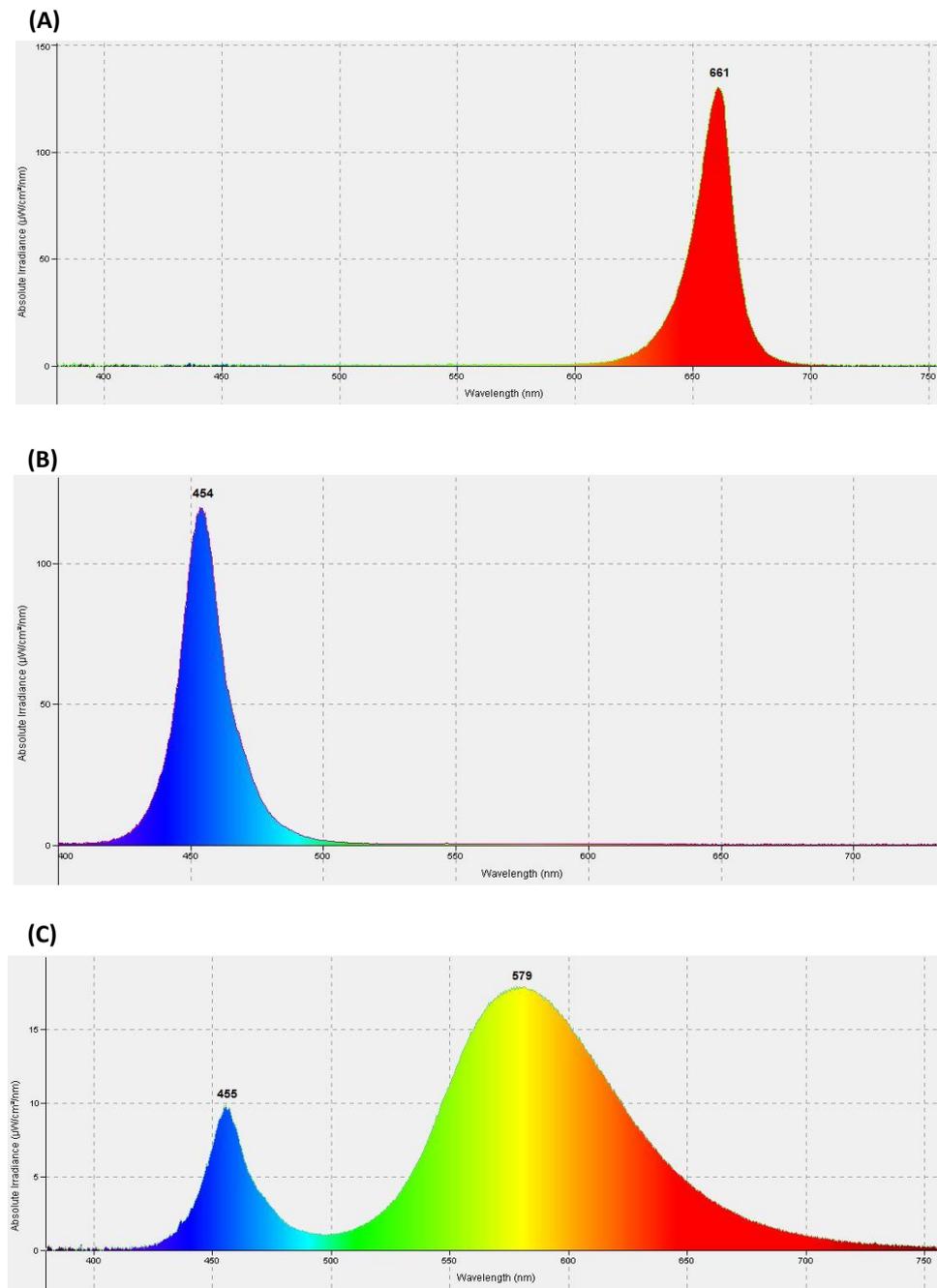
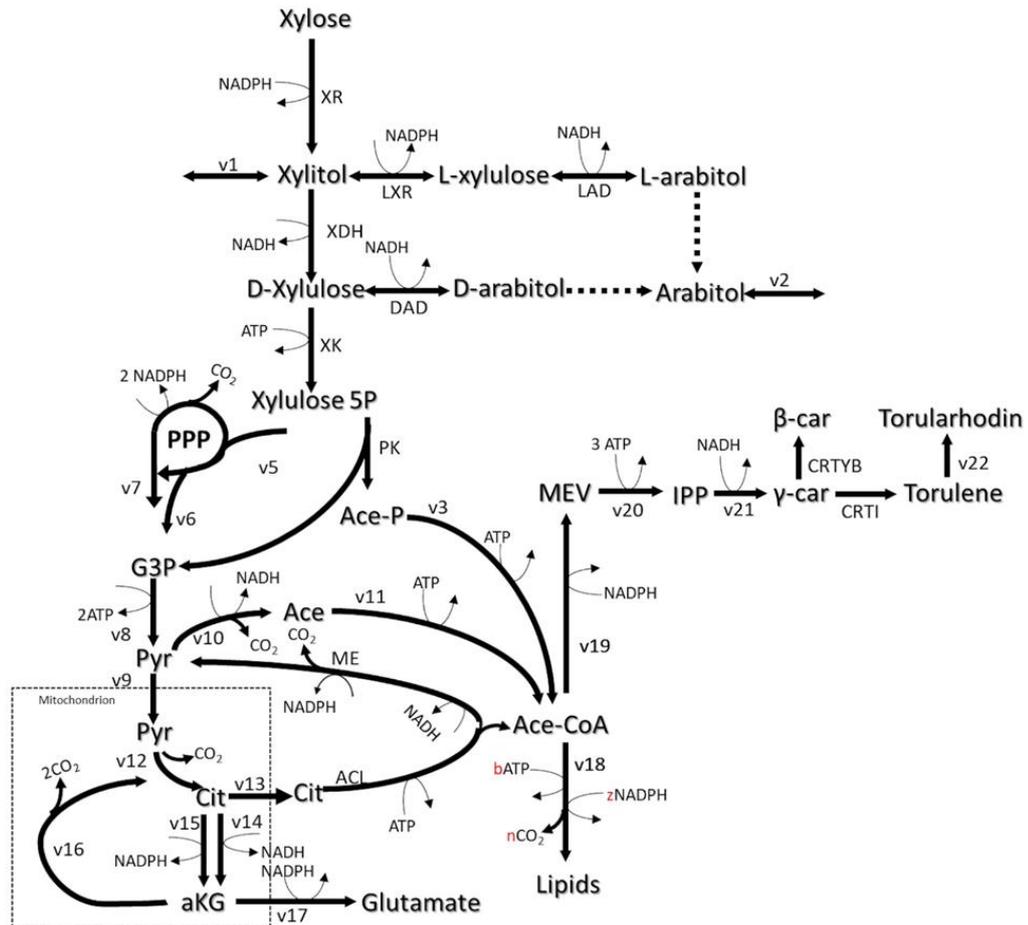


Figure A2. Absorbance spectrum for (A) red, (B) blue, and (C) white light.



Source: Information obtained from Tecnal, Brazil.

Figure A3. Schematic representation of the metabolic pathways for lipid and carotenoid synthesis from xylose in *R. toruloides*. The full caption is in the table below, which also presents the reaction code (rxnID) used in GEM.



Legend	rxnID	Enzyme name
ACL	y200003	ATP-citrate lyase
AD	t_0881	arabinitol 4-dehydrogenase
CRTI	t_0007	phytoene synthase
CRTYB	t_0006	phytoene synthase
LXR	t_0882	L-xylulose reductase
ME	t_0027	malic enzyme (NADP) cytoplasmic
PK	t_0081	phosphoketolase
PPP	r_0091	6-phosphogluconolactonase
	r_0466	glucose 6-phosphate dehydrogenase
	r_0889	phosphogluconate dehydrogenase
v1	r_2105	xylitol transport
v10	r_2116	acetaldehyde dehydrogenase
	r_0959	pyruvate decarboxylase
v11	r_0112	acetyl-CoA synthetase

v12	r_0300	citrate synthase
	r_0961	pyruvate dehydrogenase
v13	r_1112	AKG transporter, mitochondrial
v14	r_0280	cis-aconitate(3-) to isocitrate
	r_0302	citrate to cis-aconitate(3-)
	r_0658	isocitrate dehydrogenase (NAD+)
v15	r_2131	isocitrate dehydrogenase
v16	r_0451	fumarase
	r_0713	malate dehydrogenase
	r_0831	oxoglutarate dehydrogenase
	r_0832	oxoglutarate dehydrogenase (lipoamide)
	r_1021	succinate dehydrogenase (ubiquinone-6)
	r_1022	succinate-CoA ligase (ADP-forming)
v17	r_0471	glutamate dehydrogenase (NADP)
v18	r_0109	acetyl-CoA carboxylase, reaction
	r_2140	fatty-acyl-CoA synthase (n-C16:0CoA)
	r_2141	fatty-acyl-CoA synthase (n-C18:0CoA)
v19	r_0103	acetyl-CoA C-acetyltransferase
	r_0558	hydroxymethylglutaryl CoA reductase
	r_0559	hydroxymethylglutaryl CoA synthase
v2	r_1876	L-arabinitol transport
v20	r_0735	mevalonate kinase (atp)
	r_0739	mevalonate pyrophosphate decarboxylase
	r_0904	phosphomevalonate kinase
v21	r_0373	farnesyltransferase
	t_0002	phytoene synthase
	t_0003	phytoene dehydrogenase
	t_0004	phytoene dehydrogenase
v21	t_0005	phytoene synthase
	t_0008	torularhodin synthase
v22	t_0008	torularhodin synthase
v3	t_0082	phosphate transacetylase
v5	r_1048	transaldolase
v5	r_1050	transketolase 2
v6	r_1049	transketolase 1
v7	r_0450	fructose-bisphosphate aldolase
	r_0886	phosphofructokinase
v8	r_0366	enolase
	r_0486	glyceraldehyde-3-phosphate dehydrogenase
	r_0492	glycerol-3-phosphate dehydrogenase (NAD)
	r_0892	phosphoglycerate kinase
	r_0893	phosphoglycerate mutase
	r_0962	pyruvate kinase

v9	r_1138 r_2034	D-lactate/pyruvate antiport pyruvate transport
XDH	r_1092	xylitol dehydrogenase (D-xylulose-forming)
XK	r_1094	xylulokinase
XR	r_1093	xylose reductase

Table A1. List of oligonucleotides used to assemble the Golden Gate platform. The list includes the sequence of the pair of oligonucleotides used to amplify a specific part of DNA. In the "Sequence" column, the lowercase base pairs are for enzyme anchoring, the BsaI recognition site is underlined, and the 4-nucleotide overhangs are in italics.

Sequence	Purpose
gcatGGTCTCAACGGTGTCCGTATTCTA CATCGACG	Amplification of promoter from XYL1 gene for P1 position
atgcGGTCTCACATTTCGACATGGCGTGT ATTCTG	
gcatGGTCTCAGCTTTGTCCGTATTCTAC ATCGACG	Amplification of promoter from XYL1 gene for P2 position
atgcGGTCTCATTGTCGACATGGCGTGT ATTCTG	
gcatGGTCTCAGTCATGTCCGTATTCTAC ATCGACG	Amplification of promoter from XYL1 gene for P3 position
atgcGGTCTCAGTGGCGACATGGCGTGT ATTCTG	
gcatGGTCTCAACGGTGTGACTGATCTG GTGTTGTTCTGA	Amplification of promoter from GPD gene for P1 position
atgcGGTCTCACATTTGGAGTTCGACGTT CTCCTCGC	
gcatGGTCTCAGCTTTGTGACTGATCTG GTGTTGTTCTGA	Amplification of promoter from GPD gene for P1 position
atgcGGTCTCATTGTTGGAGTTCGACGTT CTCCTCGC	Amplification of promoter from GPD gene for P2 position
gcatGGTCTCAGTCATGTGACTGATCTG GTGTTGTTCTGA	Amplification of promoter from GPD gene for P3 position
atgcGGTCTCAGTGGTGGAGTTCGACGT TCTCCTCGC	

gcatGGTCTCAACGGCGGCTGAGGCTTC CCCGACG	Amplification of promoter from ADH2 gene for P1 position
atgcGGTCTCACATTTGTGACTGTCGGA GACGTGGCAGC	
gcatGGTCTCAGCTTCGGCTGAGGCTTC CCCGACG	Amplification of promoter from ADH2 gene for P2 position
atgcGGTCTCATGTTGTGACTGTCGGA GACGTGGCAGC	
gcatGGTCTCAGTCACGGCTGAGGCTTC CCCGACG	Amplification of promoter from ADH2 gene for P3 position
atgcGGTCTCAGTGGTGTGACTGTCGGA GACGTGGCAGC	
gcatGGTCTCATCTACGTTCAAACATTTG GCAATAAAGTTTC	Amplification of terminator from Nos gene for T1 position
atgcGGTCTCAAAGCCCCGATCTAGTAA CATAGATGACA	
gcatGGTCTCAGGATCGTTCAAACATTT GGCAATAAAGTTT	Amplification of terminator from Nos gene for T2 position
atgcGGTCTCATGACCCCGATCTAGTAA CATAGATGACA	
gcatGGTCTCAGTATCGTTCAAACATTTG GCAATAAAGTTT	Amplification of terminator from Nos gene for T3 position
atgcGGTCTCAACTCCCCGATCTAGTAA CATAGATGACA	
gcatGGTCTCAACAAATGCGCCCGCTTG CAC	Amplification of CRTI gene for G2 position
atgcGGTCTCAATCCTCAACCGCGCAGG TACATC	Amplification of CRTI gene for G2 position
CGTTAAAGATCTCGTCAAACAGTCGC GGG	Site-directed mutagenesis for internal BsaI recognition site removal
GTTTGACGAGATCTTTAACGATCTTGG G	
gcatGGTCTCACCATGGGCGGACTGG ACTACTGG	Amplification of CRTYB gene for G3 position
atgcGGTCTCAATACTCACAGCGCCTGC CACG	

ATGGGAAGACCGACGGCCCA	Site-directed mutagenesis for internal BsaI recognition site removal
CCGTCGGTCTTCCCATCCTCCTC	
gcatGGTCTCTGTATTACTAGTAGCGGCC GCTGC	Amplification of plasmid vector for G3 fragment insertion for Level I constructs
gcatGGTCTCTGTGGCTCTAGAAGCGGC CGCGA	
gcatGGTCTCAAGGTTGTCCGTATTCTAC ATCGACG	Amplification of Marker for Level II constructs
atgcGGTCTCAGTGGCGACATGGCGTGT ATTCTG	
gcatGGTCTCACCATGGGCAAGGAGA AGACC	Amplification of G418 gene
atgcGGTCTCAATACCTAGAAGAAGCTCG TCGAGCATGAGGT	
gcatGGTCTCAGTATCGTTCAAACATTTG GCAATAAAGTTT	Amplification of terminator from Nos gene for Marker
atgcGGTCTCACCGTCCCGATCTAGTAA CATAGATGACA	
acggGGTCTCTACGGTACTAGTAGCGGC CGCTG	Amplification of plasmid vector for Marker insertion for Level I construct
acggGGTCTCTACCTCTCTAGAAGCGGC CGCGA	
atgcGGTCTCATAGATCAGACTTTGGGA AGCTCGTGC	Amplification of CRTE gene for G1 position in Level II constructs
gcatGGTCTCAAATGTCGCTGGACTGGT ACGACAAC	
gcatGGTCTCAGGAGCCGCCTCCTCCAC CTCAGCAAC	Amplification of upstream insertional sequence for Level II/III constructs
atgcGGTCTCAACCTTCATCGTCGGCGA TGAGGAGGAC	Amplification of upstream insertional sequence for Level III constructs
agagcaGGTCTCCTCGGGATGTGGCGAC CGGGCGCG	Amplification of downstream insertional sequence for Level III constructs
atgcGGTCTCAATGGGTAGACCGTTTCG GGCGCGAC	Amplification of downstream insertional sequence for Level II/III constructs

agagcaGGTCTCCAGGTTGTCCGTATTCT ACATCG	Amplification of Marker for Level III constructs
agagcaGGTCTCCACACCCGATCTAGTAA CATAGA	