Phosphoketolase gene knockout by CRISPR/Cas9 in nonconventional yeast *Rhodotorula toruloides*

Bachelor thesis

Student: Kristjan Pals
Student code: 212996 LAAB
Supervisor: Alīna Reķēna
Early-Stage Researcher, Department of Chemistry and Biotechnology
Study program: Applied Chemistry, Food and Gene Technology

Tallinn 2024

Fosfoketolaasi geeni väljalülitamine CRISPR/Cas9 meetodil mittekonventsionaalses pärmis *Rhodotorula toruloides*

Bakalaureusetöö

Üliõpilane: Kristjan Pals
Üliõpilaskood: 212996 LAAB
Juhendaja: Alīna Reķēna
Keemia ja biotehnoloogia instituut, doktorant-nooremteadur
Õppekava: Rakenduskeemia, toidu- ja geenitehnoloogia

Tallinn 2024

Declaration

Hereby I declare that I have compiled the paper independently and all works, important standpoints and data by other authors have been properly referenced and the same paper has not been previously been presented for grading.

Author: Kristjan Pals
[Signature, 29. 05. 2024]

The paper conforms to requirements in force.
Supervisor: Alīna Reķēna
[Signature, 29. 05. 2024]

Contents

Abbreviations

ACC1 - acetyl-CoA carboxylase 1 coding gene
agRNA - anti guide RNA
AMP - adenosine monophosphate
ATP - adenosine triphosphate
CRISPR - clustered regularly interspaced short palindromic repeats
crRNA - CRISPR RNA
DGA1 - diacylglycerol acyltransferase coding gene
DMSO - dimethyl sulfoxide
GFP - green fluorescent protein
HDR - homology-directed repair
FAS - fatty acid synthase complex
gDNA - genomic DNA
gRNA - guide RNA
LB - lysogeny broth medium
NHEJ - non-homologous end joining
OD 600 - optical density measured at 600 nm
PAM - protospacer adjacent motif
PCR - polymerase chain reaction
PEG - polyethylene glycol
PGI1 - glucose-6-phosphate isomerase coding gene
SDS - sodium dodecyl sulfate
TAG - triacylglyceride
TALEN - transcription activator like effector nuclease
ZFN - Zinc-finger nuclease
XPK - phosphoketolase
YOGE - yeast oligo-mediated genome engineering
YPD - yeast extract peptone dextrose medium

Introduction

The sustainability of fossil fuel production and usage as well as concerns about food security have been the central socio-economic and environmental challenges for several years. One way of reducing the environmental impact is to use microbial cell factories for production of fuel and other chemicals currently produced from oil or feed crops such as sugarcane or oil palm. Microbial cell factories are not competing with resources used for food production (Koutinas et al., 2014) since it is possible to use biological waste or biomass hydrolysate as a substrate.

Non-conventional yeasts like *Rhodotorula toruloides* are naturally capable of consuming several different carbon sources derived from lignocellulosic hydrolysates like glucose, xylose and acetate, even without the need to purify these substrates (Bonturi et al., 2017). To improve the yields of cell factories even further and therefore make them more economically viable it is essential to develop strains to produce higher titers of product and to be able to utilize wider variety of substrates.

Rhodotorula toruloides, an oleaginous yeast, is capable of accumulating lipids to at least 67.5% dry cell mass (Li et al., 2007). That fact alone makes it a good option to use in a cell factory. Additionally, *R. toruloides* is capable of growing under a range of environmental stressors (pH, salts, H_2O_2) present in lignocellulosic hydrolysates (Wen et al., 2020). This makes it a potential yeast to further study its metabolic pathways to improve its characteristics.

In this thesis the role fo phosphoketolase in *Rhodotorula toruloides* was investigated by knocking out the phosphoketolase gene using the CRISPR/Cas9 method. The chosen target was the phosphoketolase gene which impacts the lipid accumulation in several ways, although the exact mechanism is not yet fully understood. After the phosphoketolase knockout the genotype of mutants was verified with Sanger sequencing and compared with the wild type strain to evaluate the effect of the knockout. Growth characteristics of confirmed mutants with genotype change in phosphoketolase gene were examined by cultivating them in mini-bioreactors and comparing results with the wild type strain. The goal was to understand the impact of knockout of phosphoketolase on *R. toruloides* growth rate by comparing the specific growth rates of the mutants to specific growth rate of the wild type strain.

1. Literature review

1.1 Yeast-based cell factories

Yeast-based cell factories can offer a sustainable alternative to many commodities currently made from fossil fuels or feedstock. This would be beneficial for environmental sustainability as well as global food security. Possible products range from biofuels to pharmaceuticals (Donot et al., 2014). First published results of producing lipids from yeast are from 1878 (Von Nägeli, 1878). Subsequently over the decades different yeasts and substrates were examined for their potential of lipid production. First biochemical explanation describing lipid accumulation in oleaginous yeasts was published in 1979. It was determined, that lipid synthesis is caused by the accumulation of ATP and the depletion of AMP in nitrogen-limited cells (Botham and Ratledge, 1979). During the last decades the amount of oleaginous yeast research has increased significantly. The goal is to replace palm and soybean oil with a more sustainable oil source and develop a way to produce second generation biofuels.

Small scale commercial production of yeast based cell factories started in 1939, when several industrial processes were developed for protein production using yeasts. These processes remained on a smaller scale, between 10 and 60 tons per year (Lundin, 1950). In 1971 British and Italian joint venture called Italproteine built first large-scale (100000 tonnes per year) factory to produce protein using *Yarrowia lipolytica*. Unfortunately the production did not start due lack of regulatory approval of the product (Groenewald et al., 2014). First attempt to commercialise lipid production using *Cutaneotrichosporon oleaginosus* to produce cocoa butter equivalent was made in 1980s, but it did not succeed financially due to drop in product price and competition from palm oil based products (Abeln and Chuck, 2021).

1.2 *Rhodotorula toruloides*

Rhodotorula toruloides is an oleaginous yeast which is able to synthesize lipids and carotenoids. Lipogenesis is a secondary metabolic process and takes place when there is excess amount of carbon in substrate and at the same time there is a scarcity of nitrogen for protein synthesis (Papanikolaou and Aggelis, 2011). Under such conditions the cellular growth is limited and carbon is used for lipid synthesis (Ageitos et al., 2011). Regardless of the carbon source, higher carbon to nitrogen ratios increase lipid yields but decrease the specific growth rate (Lopes et al., 2020).

The bioproducts produced by *Rhodotorula toruloides* are also of great interest, especially lipids, enzymes, and carotenoids. Wild-type and engineered strains of *Rhodotorula toruloides* are truly oleaginous. They produce lipids at high titer making them promising organisms for the production of lipid-based chemicals such as biofuels, lubricants, surfactants, solvents, waxes, creams, and adhesives (Ageitos et al., 2011; Zhang et al., 2016a, 2016b). *Rhodotorula toruloides* can metabolize various substrates, including non-detoxified lignocellulosic hydrolysates (Bonturi et al., 2017; Papanikolaou and Aggelis, 2011). The ability to produce lipids and high-value co-products from a low-cost substrate is an important economic consideration when producing biodiesel from microbial oils (Koutinas et al., 2014).

1.2.1 Key genes involved in lipogenesis

Transcriptomic and proteomic analysis of *Rhodotorula toruloides* in nitrogen limited conditions has identified over 2000 genes with altered transcript abundance and over 500 genes with altered protein abundance during lipid accumulation phase. The central part of lipogenesis is fatty acid

synthase complex (FAS) where palmityl-CoA is formed from acetyl-CoA and malonyl-CoA, with NADPH as a co-factor (Zhu et al., 2012).

In *R. toruloides*, the FAS consists of two subunits: the β-subunit (Fas1), containing acyltransferase (AT) and enoyl reductase (ER) domains, and the α-subunit (Fas2), which includes two acyl carrier proteins (ACP), dehydratase (DH), ketoacyl reductase (KR), ketoacyl synthase (KS), malonyl/palmitoyl transferase (MPT), and phosphopantetheine transferase (PPT) (Zhu et al., 2012). Since FAS is a complex molecular machine, most of the research on improving lipid yield is dedicated to different enzymes which are upstream from FAS and therefore are able to provide increased flux of substrates to FAS. The other limiting factor to fatty acid biosynthesis is the availability of NADPH. It is shown that when NADPH is provided at higher rates, the lipid accumulation lasts longer (Ratledge and Wynn, 2002). *In silico* studies have shown that one successful strategy is using growth-linked triacylglyceride (TAG) production which means that mutants are forced to carry flux through TAG reactions to reach optimal growth rates. Another way of increasing lipid production is knocking out the glucose-6-phosphate isomerase coding gene (PGI1) (Castañeda et al., 2018). Increased lipid production can also be achieved by overexpression of acetyl-CoA carboxylase (ACC1) and diacylglycerol acyltransferase (DGA1) expressing genes (Zhang et al., 2016b).

1.2.2 The role of phosphoketolase

One of several genes likely impacting lipid yield is phosphoketolase (XPK) which catalyses the cleavage of sugar phosphates. This includes the conversion of D-xylulose 5-phosphate to glyceraldehyde 3-phosphate and acetylphosphate offering a more efficient way to supply acetyl-CoA for fatty acid biosynthesis (Lopes et al., 2020). Studies of oleaginous yeast *Yarrowia lipolytica* have shown, that overexpression of XPK leads to improved lipid accumulation. The mechanism behind this improvement is that synthesis of acetyl-CoA via XPK is more efficient than the route via glycolysis pathway - 2.3 fewer moles of glucose is needed to syntesize one mole of triolein (Kamineni et al., 2021).

XPK also plays a role under nitrogen-limited growth phase. It is observed, that phosphoketolase reaction saves carbon by converting D-xylulose 5-P directly to acetyl residue and glycerol-3-phosphate when compared to alternative pathway via transketolase until pyruvate decarboxylation which requires several reactions. However, the alternative pathway via transketolase is preferred by the cells since it distributes more carbon to oxidative pentose phosphate pathway (Pinheiro et al., 2020). This conclusion is contested by a proteome study (Tiukova et al., 2019) where proteomic analysis was not able to determine which pathway was actually used for xylulose-5-phosphate metabolisation.

These examples show, that the role of phosphoketolase in cell growth as well as lipid accumulation is complex and full understanding requires more studies on the topic.

1.2.3 Advances in *Rhodotorula toruloides* genome engineering

One of the ways to change the metabolic behaviour of the cell is to switch off the gene which encodes the enzyme catalysing certain reaction in a metabolic pathway. This can be achieved by introducing a frame shift in the coding region of the gene causing the premature ending of the translation by premature stop codon or the change in resulting amino acid sequence which changes the structure of the enzyme causing it to no longer catalyse the original reaction. There are various methods for creating a yeast mutant with a knocked out gene: Zinc-finger nucleases (ZFN) (Wang et al., 2013), yeast oligo-mediated genome engineering (YOGE) (DiCarlo et al., 2013), transcription

activator like effector nucleases (TALEN) (Li et al., 2011), traditional selection marker-based method (Koh et al., 2014) and CRISPR/Cas9 (Schultz et al., 2022).

Zinc-finger nucleases were the first programmable nucleases. This method uses proteins designed to bind to specific sequence of DNA and place a nuclease protein to the cleavage site to make a cut. This method is suitable for producing gene knockouts. The advantage of ZFNs is that they can target any DNA sequence, but the downside is the complexity and cost required to design and produce the needed zinc finger proteins (Chandrasegaran, 2017). The method is reported to work in *Saccharomyces cerevisiae* (Wang et al., 2013), but so far there are no reported cases of *Rhodotorula toruloides* modified by ZFNs.

TALENs are also one possible way to introduce a gene knockout, since the gene editing mechanism is simiar to ZFNs. The method also uses nucleases to cleave DNA as well as proteins to bind to specific cleavage site. Using TALEN method for gene knockout has been reported in oleaginous yeast *Yarrowia lipolytica* (Rigouin et al., 2017), but not yet in *Rhodotorula toruloides*.

Yeast oligo-mediated genome engineering is a method which allows to introduce knockouts and mutations, although the method is more suitable for site-directed mutagenesis (DiCarlo et al., 2013). Since none of the methods described above have been reported to be used in *R. toruloides*, the only available option was to use the CRISPR/Cas9 method, which is described in more detail in chapter 1.3.

1.3 CRISPR/Cas9

CRISPR, short for Clustered Regularly Interspaced Short Palindromic Repeats, is adaptive immune system in prokaryotes that uses RNA-guided nuclease to cut double-stranded DNA. The system contains the cluster of Cas genes, noncoding RNAs and a set of repetitive elements. These elements are interspaced by sequences of DNA targets known as protospacers. Each protospacer is next to protospacer adjacent motif (PAM) sequence. There are three types (I, II, and III) of CRISPR/Cas systems, which have all these parts, but differ in the ways the target sequence is detected and cleaved (Ran et al., 2013).

One of the most studied CRISPR systems is the type II system with nuclease Cas9 (Jinek et al., 2012), CRISPR RNA (crRNA) array which encodes the guide RNAs containing 20 nt guide sequence and a partial direct repeat which guides Cas9 to 20 bp DNA target based on complementarity (Ran et al., 2013). The system requires target DNA to be preceded by 5’-NGG-3’ PAM sequence (Jinek et al., 2012).

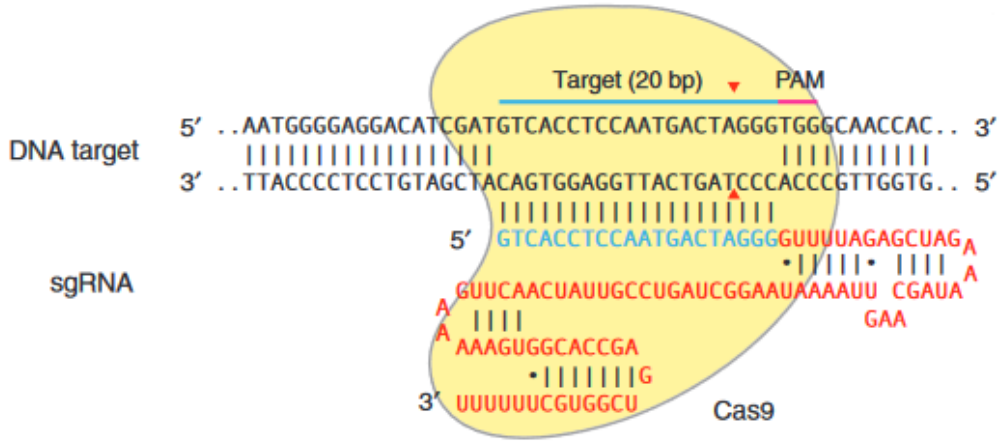


Figure 1. RNA-guided Cas9 nuclease (yellow) is targeted to genomic DNA (blue) by single stranded gRNA consisting of 20 nt guide sequence (blue) and scaffold (red) paired with target sequence, directly upstream from PAM. Adapted from (Ran et al., 2013).

The CRISPR/Cas system can be used to disable a gene by introducing a cleavage site in the coding region of the gene. After nuclease has cut both strands of the DNA, the cellular repair mechanisms activate and start repairing the DNA. There are two main ways for the DNA damage repair to occur: the error-prone non-homologous end joining (NHEJ) process or homology-directed repair (HDR) process. HDR requires a template of DNA to be present and in case there is none the NHEJ mechanism is used (Ran et al., 2013). NHEJ is known to cause errors in DNA repairing (inserting or deleting one or more nucleotides) it is suitable to use if the goal is to disrupt a gene (Otoupal et al., 2019) since it is likely to cause a frame shift in the gene which makes the enzyme encoded by the gene dysfunctional. NHEJ is the prevailing mechanism in *R. toruloides* [Koh et al. 2014].

1.4 Overview of microbial growth characterisation techniques

One of the most common ways to characterise microbial growth is to measure the specific growth rate of the culture. The growth rate can be measured as a change of optical density (OD) in the culture over time and then is used to calculate specific growth rate based on the readout (Mytilinaios et al., 2012).

The culture is grown in a liquid medium. The simplest way to evaluate OD from liquid medium is to grow the culture in shake flasks in the incubator and to take samples manually and measure the OD after every specific time period. For more automated approach bioreactors can be used. There are many different types available, but one of the common features is automated measuring of key parameters (OD, pH, CO_2 emission) during the fermentation.

The fermentation for strain characterisation is done in a batch process (Koch, 1881) which means inoculating the fixed volume of medium and not adding any more substrate during the fermentation. To minimise the variables affecting the growth rate the culture is usually grown in a chemically defined medium, where the amount of all the substrates is strictly controlled (Zhang, 1999).

2. Aims of the thesis

Rhodotorula toruloides has unique metabolic properties that enable a naturally high biosynthesis of intracellular lipids. Phosphoketolase is part of this metabolic network and its is shown to behave differently depending on growth conditions and growth phase. Exact role and impact of phosphoketolase to lipogenesis is not yet fully understood. The goal of this work is to obtain single gene knockout in *Rhodotorula toruloides* to elucidate the role of phosphoketolase during lipogenesis. This thesis has two goals:

- To perform a CRISPR/Cas9-mediated phosphoketolase knockout in yeast *Rhodotorula toruloides.*
- To evaluate the role of phosphoketolase on *R. toruloides* growth characteristics.

3. Materials and methods

3.1 Strains

Rhodotorula toruloides strain IFO0880 (Coradetti et al., 2018) was used as a parental strain for a knockout experiment. Yeast cultures from -80 °C stock were inoculated in 50 mL of YPD in shake flasks and grown overnight at 30 °C with stirring at 200 rpm. Yeast grown on YPD agar plates was incubated at least two days at 30 °C after plating and stored at 4 °C for not more than 4 weeks. YPD agar plates with G418 antibiotic were used for yeast transformation.

The plasmid pPBO.202 in *E. coli* was obtained from Joint BioEnergy Institute (jbei.org), part ID JBEI223791 and stored in -80 °C. "TOP10 competent cells" (Thermo Scientific) strain of *E. coli* was used for vector cloning. *E. coli* was grown in 3 mL liquid lysogeny broth (LB) at 37 °C with stirring. For storage and strain screening, LB agar plates were used. For all applications, the *E. coli* cultures were incubated overnight at 37 °C.

3.2 20-nucleotide target sequence choice

For the design of guide RNAs targeting XPK gene in the *R. toruloides* IFO0880 genome (Coradetti et al., 2018), the online tool CCTop (Stemmer et al., 2015) was used and target at convenient position with high efficacy score was chosen. Since this tool does not yet cover *Rhodotorula* species, *Ganoderma lucidum*, being the only representative of the Basidiomycota division, as the reference genome for sgRNA target site evaluation was chosen. Candidate CRISPR/Cas9 target (Table 1.) was identified using the following criteria: 23-mers ending with NGG (*Streptococcus pyogenes*), efficacy score by CRISPRater (Labuhn et al., 2018) above 0.74 (high efficacy), set custom overhangs CGCA and AAAC. To identify potential off-target sites in the *R. toruloides* IFO0880 genome, the 23 nucleotide sequence (with 2 upstream nucleotides because 25 is the minimum sequence length to get a positive query) of gRNA was used as a query in a JGI Mycocosm Blast search of the *R. toruloides* IFO0880 reference genome sequence v4.0 (Coradetti et al., 2018) with parameters: "E=1e5; Filter=True BLOSUM62". Alignments were manually examined.

Table 1. Guide RNAs for CRISPR/Cas9 targeting a cleavage site in exon 1 of XPK gene (transcript ID 13510) in R. toruloides strain IFO0880.

Name	Type	Sequence	GC content, %
1573	gRNA	5'-**CGCA**GATGCAGAGGAAGTTGACCA-3'	54.2
1574	agRNA	5'-**AAAC**TGGTCAACTTCCTCTGCATC-3'	45.8

3.3 CRISPR/Cas9 vector cloning

Oligonucleotides 1573 and 1574 were prepared for cloning by diluting each in distilled water to 0.25 mM concentration in reaction volume 20 µL and hybridised in thermocycler (Eppendorf Nexus Mastercycler) using the program described in Table 2.

Table 2. Thermocycler program for oligonucleotide hybridisation.

Temperature, °C	Time, min
95	1
90	1
85	1
80	1
75	1
70	5
65	5
60	5
55	5
50	5
45	5
40	5

After hybridisation the guide RNAs (Figure 2, nucleotides in bold are the sticky end overhangs) were ligated into pPBO.202 plasmid using BsaI restriction enzyme to make a cut in the plasmid and T4 ligase enzyme to ligate the hybridised guide RNA into the plasmid. BsaI restriction enzyme (Thermo Scientific) recognises 5'-GGTCTCN-3' and 5'-NNNNNGAGACC-3' cutting sites and in pPBO.202 plasmid produces two cuts with sticky end overhangs (Figure 2).

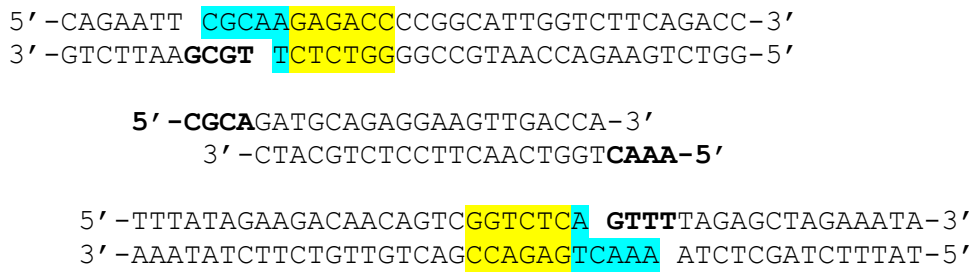


Figure 2. BsaI cutting sites (yellow marks the recognition sequence, cyan marks the N nucleotides) of pPBO.202 plasmid next to hybridised guide RNA. Matching overhangs are marked with bold.

Oligonucleotides were cloned into a purified pPBO.202 plasmid in a 200 µL PCR tube. Each reaction of 5 µL contained 0.25 µL BsaI (Eco31) restriction enzyme (Thermo Scientific), 0.25 µL T4 ligase (Thermo Scientific), 0.5 µL T4 ligase buffer (Thermo Scientific), 0.5 µL hybridised oligomers, 0.5 µL pPBO.202 plasmid stock (10 ng / µL) and 3 µL distilled water. Ligation mixture was digested in thermocycler (Eppendorf Nexus Mastercycler) using the program described in Table 3.

Table 3. Thermocycler program for pPBO.202 plasmid digestion-ligation.

	Temperature, °C	Time, min	Number of cycles
Initial restriction	37	10	1
Ligation	16	5	5
Restriction	37	5	
Final restriction	37	5	1

After digestion pPBO.202 plasmids with ligated gRNAs were transformed into *E. coli* cells. The bacterial culture was pipetted over ice into 1.5 mL tubes, 50 µL to each tube. 5 µL of digested ligation mixture was added to each tube and the tubes were left on ice for 20 minutes. After that the heat shock was done to enter plasmids into the bacterial cells: the tubes containing bacterial culture and ligation mixture were kept in thermoblock for 1 minute at 42 °C and then on ice for 2 minutes. After heat shock 100 µL of chilled liquid LB culture was added to each tube and incubated at 37 °C for 45 minutes. After incubation the culture was plated to Petri dishes containing LB agar and kanamycin and incubated overnight. Single colonies from plates were inoculated into 4.5 mL liquid LB with 4.5 µL kanamycin and incubated overnight at 37 °C to increase the copy number of plasmids. After incubation, bacterial colonies were screened for fluorescence under UV light. The plasmid contains a gene for green fluorescent protein, which gets removed when the guide RNA ligation is successful. The plasmids were extracted using the FavorPrep Plasmid Extraction kit (Ping Tung, Taiwan) according to manufacturer specifications.

3.4 Yeast transformation

Resulting plasmids were transformed into *R. toruloides* according to the lithium acetate method described in (Bonturi et al., 2022). Yeast culture in liquid YPD medium was grown overnight at 30 °C in shake incubator. Next morning, the culture was diluted to OD 600 of 0.2 and divided into 15 mL tubes with 5 mL of culture in each. The tubes with the culture were incubated at 30 °C until OD 600 reached 0.8. After that the culture was transferred to 50 mL tubes, centrifuged for 10 minutes at 3000 g, growth medium was removed and cells were resuspended in 25 mL of sterile water and centrifuged again for 10 minutes at 3000 g. After that the water was removed and the cells were resuspended in 1 mL 100 mM LiAc per 8 OD units of culture. 500 µL of suspension was transferred to 1.5 mL tube for each transformation reaction. The tubes were centrifuged for 30 s at 3000 g and the supernatant was removed.

To each tube was added 240 µL polyethylene glycol (PEG) 4000 (50% weight to volume ratio), 36 µL 1.0 M LiAc, 24 µL sterile water, 10 µL Salmon Sperm DNA (10 mg/mL), 50 µL extracted plasmid DNA. 10 µg of the vector was used for each transformation reaction of 800 µL resulting in vector concentration of 12,5 ng/µL. The mixture was vortexed until the pellet dissolved and incubated for 30 minutes at 30 °C. After incubation 34 µL dimethyl sulfoxide (DMSO) was added to each tube and heat shock was done at 42 °C for 15 minutes. After heat shock the tubes were centrifuged for 30 s at 3000 g, supernatant was removed. This step was done twice. The cells were then transferred to 15 mL tubes with 2 mL liquid YPD and incubated overnight. After incubation the tubes were centrifuged, 1 mL of supernatant was removed and the rest was plated to YPD agar plates with G418 antibiotic 200 ug/mL and incubated at 30 °C for 2 days. The transformants were restreaked to new plates to get single cell colonies and replated on a fresh YPD plate.

3.5 Screening

Screening was performed in two stages. Primary screening was performed using diagnostic PCR check for the presence of Cas9 gene in the genome. Secondary screening was done with preparative PCR in combination with Sanger sequencing. Preparatory PCR stage was necessary to provide sufficient nucleotide concentration of the selected region from XPK gene to be sequenced. Samples with unsuccessful nucleotide amplification were excluded from sequencing. Same *R. toruloides* transformant strains were used for all the screenings as well as the wild type for negative control. Primary screening was performed on 2nd generation transformants, while secondary was done on 3rd generation transformants.

Firstly, transformants were used for genomic DNA extraction as described in (Lõoke et al., 2011). Biomass was picked from selected colonies on a plate and transferred to PCR tubes. 100 µL 0.2 M LiAc with 1% sodium dodecyl sulfate (SDS) was added and the mixture was heated for 10 minutes at 75 °C in a thermocycler (Eppendorf Nexus Mastercycler). After that the mixture was transferred to 1.5 mL tubes, 300 µL 96% ethanol was added and vortexed for 30 s. Tubes were centrifuged for 3 minutes at 21000 g, supernatant was discarded, the pellet was resuspended in 300 µL 70% ethanol and centrifuged for 3 minutes at 21000 g. Supernatant was discarded and the pellets were dried at 37 °C for an hour to evaporate the remaining ethanol. 50 µL distilled water was added to the pellet, vortexed and centrifuged for 1 minute at 21000 g. Nucleic acid concentration in supernatant was measured with NanoDrop One (Thermo Scientific).

Diagnostic PCRs were done in 200 µL PCR tubes using DreamTaq 2x Green PCR Master Mix (Thermo Scientific). Reaction mixture contained 1 µL forward primer (10 mM), 1 µL reverse primer (10 mM) (Table 4.), 0.5 µL template DNA, 2.5 µL distilled water and 5 µL polymerase master mix. PCR reactions were performed using the thermocycler programs (Table 5). Annealing temperatures were calculated by manufacturer's online calculator (Thermo Fisher online calculator).

Table 4. Primers for Cas9 gene screening.

Primer ID	Type	Target	Sequence
1446	Forward primer	Cas9 gene	5'-GGAGTCGCGGGACGCCAAC-3'
1449	Reverse primer	Cas9 gene	5'-ACACGTTGGCGTCCCGCGA-3'

Table 5. Thermocycler program for a diagnostic PCR of Cas9 gene screening with DreamTaq 2x Green PCR Master Mix (Thermo Scientific).

	Temperature, °C	Time, min	Number of cycles
Initial denaturation	95	5	1
Denaturation	95	0.5	35
Annealing	68	0.5	
Extension	72	1.5	
Final extension	72	2	1

PCR products were visualised and separated on agarose gel. The gel contained 1.25% agarose (Invitrogen) in 1x TAE buffer. 8 µL Atlas Clearsight (Bioatlas) DNA stain was added to visualise the DNA. Quick-Load Purple 2-Log DNA Ladder (New England Biolabs) was added to evaluate the length of amplicons. Gel electrophoresis was performed at 160 V for 20 minutes.

For secondary screening the DNA was extracted from transformants as described in (Lõoke et al., 2011). Preparative PCR reaction for was performed in 200 µL PCR tubes using Platinum SuperFi 2x DNA polymerase master mix (Thermo Scientific). Reaction mixture contained 2 µL forward primer for XPK gene (10 mM), 2 µL reverse primer for XPK gene (10 mM) (Table 6.), 2 µL template DNA, 4 µL distilled water and 10 µL Platinum SuperFi polymerase master mix (Thermo Scientific). Thermocycler program for Platinum SuperFi PCR reaction is described in Table 7. Annealing temperatures were calculated by manufacturer's online calculator (Thermo Fisher online calculator).

Table 6. Primers for XPK gene screening.

Primer ID	Type	Target	Sequence
1343	Forward primer	Exon 1 of XPK gene	5'-CCTCCCTCTCACTCTTGCAC-3'
1344	Reverse primer	Exon 1 of XPK gene	5'-CACGCTGTCCAGTCAAAGAA-3'

Table 7. Thermocycler program for a preparative PCR of XPK gene screening with Platinum SuperFi polymerase master mix (Thermo Scientific).

Step	Temperature, °C	Time, min	Number of cycles
Initial denaturation	95	3	1
Denaturation	95	0.5	35
Annealing	63	0.5	
Extension	72	1	
Final extension	72	5	1

PCR products were visualised and separated on agarose gel. The gel contained 1.25% agarose (Invitrogen) in 1x TAE buffer. 8 µL Atlas Clearsight (Bioatlas) DNA stain was added to visualise the DNA. Quick-Load Purple 2-Log DNA Ladder (New England Biolabs) was added to evaluate the length of amplicons. Gel electrophoresis was performed at 160 V for 20 minutes.
Samples with signal at 550 bp length had DNA extracted from the gel using FavorPrep FavorPrep GEL/PCR Purification Kit (Ping Tung, Taiwan) according to manufacturer specifications and extracted DNA was used for Sanger sequencing. Nucleic acid concentration was measured with NanoDrop One (Thermo Scientific). For future use the positive knockout strains were grown in liquid YPD overnight, diluted to a final concentration of 60% glycerol and stored at -80°C.

3.6 Growth characterisation

R. toruloides mutants as well as the wild type pre-culture from -80°C stock was inoculated in 50 mL YPD at 30°C and 200 rpm overnight. Overnight cultures were centrifuged at 3000 g for 5 minutes, supernatant was removed and the cells were washed with distilled water. This cycle was repeated three times in total. After washing stage the cell culture was diluted to OD 600 value of 0.8 and 1 mL of that was inoculated to 10 mL liquid chemically defined Delft medium (Lahtvee et al., 2017) in 50 mL spin tube. The medium contained 20 g/L xylose as a sole carbon substrate, minerals, trace elements and vitamins (Table 8). High carbon to nitrogen ratio of (80 mol/mol) was chosen so the switch from exponential growth phase to lipid accumulation phase would take place at relatively low values of OD 600 to make the measurements more accurate.

Table 8. Composition of medium for mutant strain characterisation according to Lahvtee et al. 2017. Trace elements and vitamins solution prepared according to (Verduyn et al., 1992).

Component	Amount, per liter of medium
D-xylose	20 g
$(NH_4)_2SO_4$	5 g
0.5 g of $MgSO_4$ x $7H_2O$	0.5 g
KH_2PO_4	3 g
K_2HPO_4	5.2 g
Trace elements solution 1000X	1 mL
Vitamin solution 1000X	1 ml

Yeast strains were cultivated in a batch regime at 30 °C in aerated reverse spin regime using Biosan RTS-8 Plus (Riga, Latvia). Starting parameters for growth characterisation: OD 600 0.8, pH 7, 3 replicates of each strain. The mini-bioreactors were equipped with real-time optical density measuring at 600 nm (OD 600), calibrated to optical density of up to 55 using the same yeast cultures (in-built 8-point calibration). For manual optical density measurements at the beginning and at the end of growth experiment cell density meter (Fisher Scientific) was used. Experiment was continued until no increase in OD 600 was observed. Wild type strain in the same condition was used as a negative control. The growth phases were determined graphically by plotting the change of OD 600 values on a logarithmic scale over time. The increase of OD 600 was measured automatically every half an hour and resulting growth curves were used to determine the growth phases and specific growth rate. To find the best fit for OD 600 measurements to be used in specific growth rate calculations, the specific growth rate was calculated over extended period of time (Equation 1.) in both exponential an nitrogen-limited growth phases.

$$t_{growthrate} = \frac{\ln OD_{end} - \ln OD_{start}}{t_{end} - t_{start}}$$

Equation 1. The formula used to compute average growth rate over a time window preceding the elapsed time value at $t_{growthrate}$ for each 30-minute interval. OD_{end} – OD 600 value at the end of time window, OD_{start} – OD 600 value at the end of time window, t_{end} – elapsed time at the end of time window, t_{start} – elapsed time at the start of time window.

Specific growth rate was computed from OD 600 data using MS Excel (Redmond, Washington, USA) by fitting exponential trend line to selected growth period. The specific growth rate was expressed as the exponent of the trend line equation (Equation 2.).

$$OD = t * e^{\mu}$$

Equation 2. Exponentian trendline equation where OD – optical density, t – elapsed time and μ – specific growth rate, represented on a logarithmic scale as a slope of the trendline.

Student's t-test was used to determine whether the differences in specific growth rates of mutant strains were statistically significantly different from the specific growth rate of the wild type strain. MS Excel's built-in function was used to calculate the p-values of 2-tail Student's t-test.

4. Results

4.1 CRISPR/Cas9 strategy

CRISPR/Cas9 system in *R. toruloides* used in this study works by integrating the genes on the plasmid pPBO.202 into cell's genomic DNA. The integration is achieved by using selectable drug marker cassette that encodes nourseothricin resistance (NAT^r). Functional elements of the plasmid are placed under several promoters. Elements under pSNR52 promoter are ribozyme, guide RNA and guide RNA handle. Cas9 is expressed by pGADPH promoter and NAT^r under pTUB2 promoter. All of these three regions are followed by terminators. (Supplementary Figure 1.) (Otoupal et al., 2019). This design results in the expression of Cas9 nuclease combined with guide RNAs which then proceed to locate and cut the target site.

Additionally several selectable marker sequences are present on the plasmid which are used for screening purposes. G418 and bacterial kanamycin antibiotic resistancy genes are used to screen for transformants since only the cells containing the gene are able to grow in the presence of antibiotic. Green fluorescent protein (GFP) gene upstream GDP1 is used to screen for correct gRNA ligation into plasmid after the cloned vector is transformed to *E. coli*. The restriction sites (BsaI) for gRNA are placed at both ends of a GFP coding gene (Supplementary Figure 1.), so successful ligation means the GFP is no longer present afterwards and transformants containing the plasmid with gRNA can be screened under UV light. That makes it possible to pick *E. coli* colonies with negative response of fluorescence under UV light for further plasmid copy number increasing and plasmid extraction.

4.2 Confirmation of phosphoketolase knockouts in *R. toruloides*

16 yeast colonies were present on the YPD plate with added G418 antibiotic after the first yeast transformation step. All the colonies on the plate were assumed to be successful transformants, since the integration vector contained an antibiotic resistancy gene (G418). Each transformant was passed through a single colony stage and screened for a Cas9 gene integration (primary screening) in the genome because that is a necessary precondition for CRISPR/Cas9 construct to be expressed in a cell. For first screening diagnostic PCR check confirmed that all colonies, except #16, showed at least faint signal for amplicon size of 1.1 kilobase pair confirming Cas9 gene integration (Figure 3).

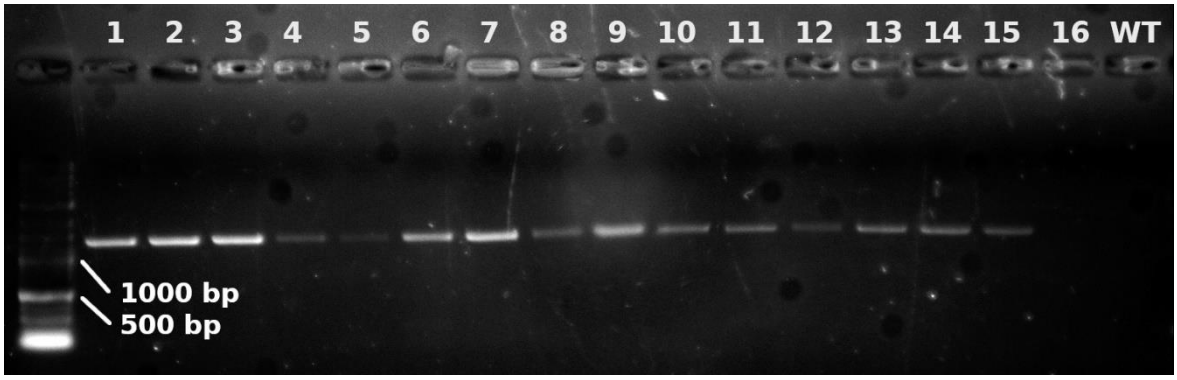


Figure 3. Agarose gel image showing the genotyping results of Cas9 gene fragment (1.1 kbp) amplification in *R. toruloides* IFO0880 transformant colonies (#1 - #16) and the wild type (WT).

Secondary screening was done to confirm the knockout genotype. For that the genomic DNA was extracted from the *R. toruloides* mutants, fragments of the XPK gene for up to 3 parallel samples were amplified using the preparative PCR method. Samples with strong band on the gel at 550 bp

region had the DNA extracted from the gel (Supplementary figure 2.), nucleotide concentration measured (Supplementary table 2.). Three samples #3B, #7A, #7B were dropped from the sequencing due to unsuccessful PCR amplification step causing no genomic material to be present on gel (Supplementary Table 2. and Supplementary Figure 2.). Sequencing confirmed insertion or deletion in 13 out of 16 transformants (Table 9.). Detailed results and alignment with *R. Toruloides* IFO0880 wild type are presented in Supplementary materials section S4, demonstrating that the sequencing had a good quality readout.

Table 9. Summary of sequencing results for the *R. toruloides* IFO0880 3rd generation XPK knockouts.

Transformant	Change, bp	Name
1	+1	XPK-Δ+1
2	-2	XPK-Δ-2
3	-8	XPK-Δ-8
4	+1	XPK-Δ+1
5	+1	XPK-Δ+1
6	+1	XPK-Δ+1
7	not sequenced	
8	-11	XPK-Δ-11
9	+1	XPK-Δ+1
10	-7	XPK-Δ-7
11	-1	XPK-Δ-1
12	not sequenced	
13	-10	XPK-Δ-10
14	+1	XPK-Δ+1
15	+1	XPK-Δ+1
16	not sequenced	

The Sanger sequencing confirmed that transformants display a variety of different changes in the genotype, ranging from insertion of 1 bp to deletion of 11 bp. 1 bp insertion was observed for transformants #1, #4, #5, #6, #9, #14 and #15. 2 bp deletion was observed in transformant #2. 8 bp deletion was observed in transformant #3. 11 bp deletion was observed in transformant #8. 7 bp deletion was observed in transformant #10. 1 bp deletion was observed in transformant #11. 10 bp deletion was observed in transformant #13.

All observed indels caused a frame shift (Figure 5.), which means that the protein translated from XPK gene should have different primary structure downstream from cleavage site and would be dysfunctional and effectively behave as if the XPK gene would not exist. Six of the mutants showed 1 bp insertion (strain IFO0880 XPK-Δ+1) to the genome. This mutation is caused by Cas9 protein

creating a 1 nucleotide 5’ overhang which is filled in by DNA polymerase and then ligated (Lemos et al., 2018).

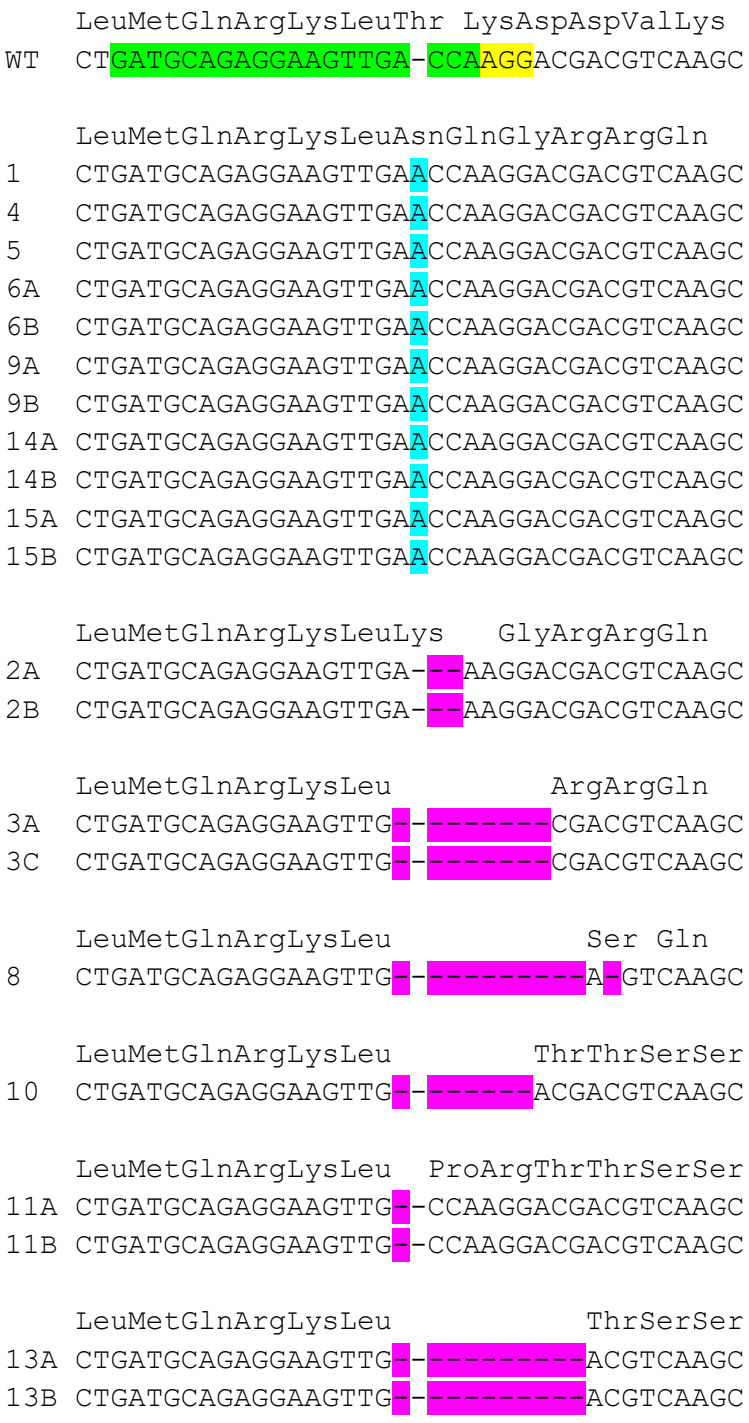


Figure 4. Genotype comparison of the wild type *R. toruloides* IFO0880 XPK gene region in exon #1 with the sequenced knockout transformants of 3rd generation (1 - 6, 8 - 11, 13 - 15). Matching indels are grouped together. Green marks the guide RNA sequence, yellow marks the PAM sequence, cyan marks the inserted nucleotide, magenta marks the deleted nucleotides. Translated amino acid sequence is presented above nucleotide sequence. A-C refer to replicates of the same transformant.

4.3 Growth characteristics of XPK knockout strains

Since XPK-Δ+1 is a different mechaism causing the mutation, these mutants were left out of the characterisation. Strains IFO0880 XPK-Δ-2, XPK-Δ-8, XPK-Δ-11, XPK-Δ-7, XPK-Δ-1, XPK-Δ-10 and the wild type were grown in a chemically defined medium at nitrogen limitation (C/N ratio of 80 (mol/mol) with xylose 20 g/L as a sole carbon source to compare their specific growth rate during the exponential growth and nitrogen-limitation phases with the wild type strain of *R. toruloides.* Xylose was chosen as a substrate since phosphoketolase is part of the 5-carbon sugar assimilation pathway and any significant change in metabolism should be more pronounced in that case. The growth phases were determined graphically by plotting the change of OD 600 values on a logarithmic scale over time (Figure 6.). The phase change from exponential growth phase to nitrogen-limited growth phase for all strains can be observed between 45 to 50 hours after inoculation. The final OD 600 values ranged from 35 to 45 (Supplementary Table 2.). For exponential growth phase, the 10-hour time window with highest average growth rate between 20 and 45 hours from inoculation for calculating the specific growth rate was selected. For nitrogen-limited phase a 12-hour window with the highest R^2 value between 45 h and 70 h after inoculation was selected, since this phase lasted longer and the curve fitting was more accurate.

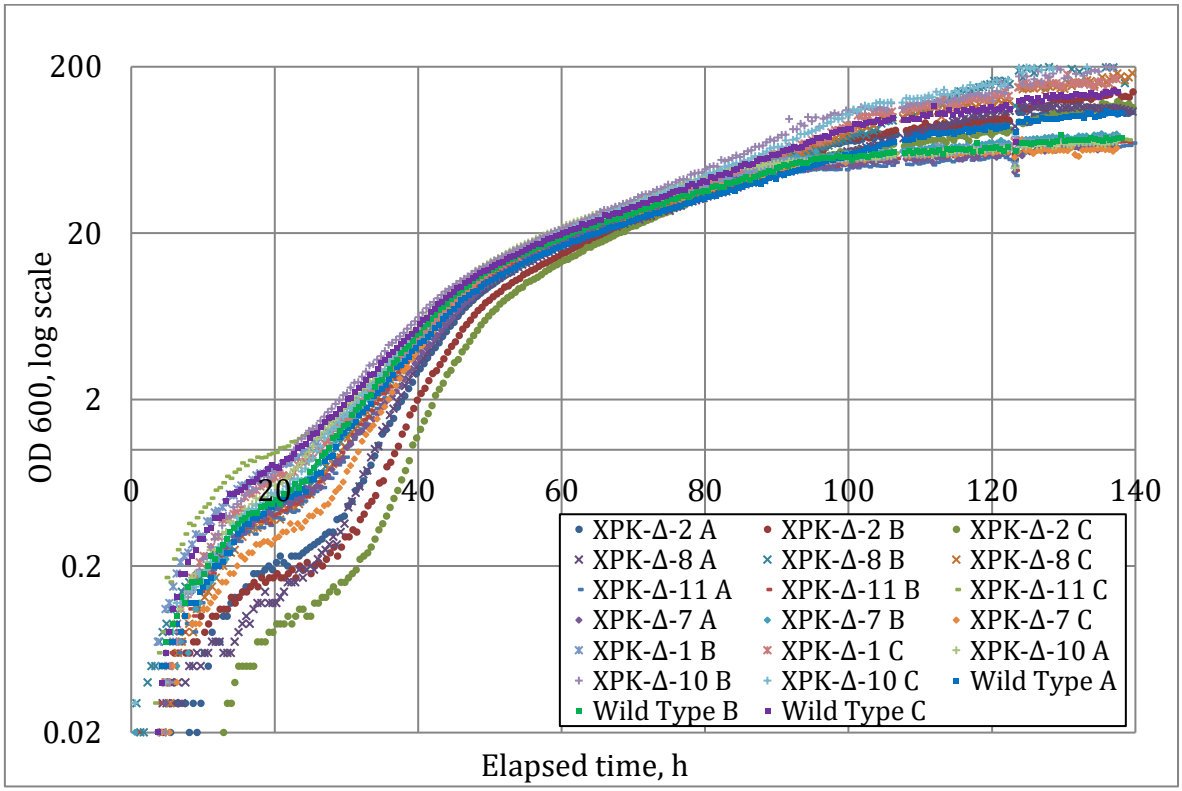


Figure 5. Growth curves of *R. toruloides* IFO0880 phosphoketolase knockout strains (XPK-Δ-2, XPK-Δ-8, XPK-Δ-11, XPK-Δ-7, XPK-Δ-1, XPK-Δ-10) and the wild type cultivated in chemically defined medium with xylose (20g/L) as a single carbon source.

During the exponential growth phase the values of specific growth rate ranged from 0.105 h^{-1} to 0.235 h^{-1} (Table 10., Supplementary figure 6.), which is close to reported specific growth rate of 0.14 h^{-1} measured in in similar experimental conditions (Monteiro De Oliveira et al., 2021). Student's 2-tailed t-test confirmed statistically significant ($p < 0.05$) difference of specific growth rate for XPK-Δ-2 mutant in exponential growth phase ($p = 0.003$). These results are inconsistent,

since all mutants should behave identically under similar conditions. The reason for inconsistency may stem from the inaccuracies in preparing the growth medium.

Table 70. Specific growth rates (h^{-1}) and 2-tail Student t-test p-values of R. toruloides IFO0880 mutant strains (XPK-Δ #2 - #10) compared against the wild type during exponential growth phase cultivated in a chemically defined medium with xylose as a sole carbon substrate (20 g/L) at starting C/N ratio of 80 mol/mol.

	XPK-Δ-2	XPK-Δ-8	XPK-Δ-11	XPK-Δ-7	XPK-Δ-1	XPK-Δ-10	wild type
Sample A	0.197	0.224	0.145	0.132		0.113	0.134
Sample B	0.211	0.134	0.140	0.125	0.114	0.112	0.131
Sample C	0.235	0.143	0.109	0.175	0.118	0.120	0.105
Average	0.215	0.167	0.131	0.144	0.116	0.115	0.123
2-tail t-test p-value	0.003	0.220	0.624	0.326	0.590	0.435	

During the nitrogen-limited growth phase the values of specific growth rate ranged from 0.047 h^{-1} to 0.053 h^{-1} (Table 11., Supplementary figure 8.) which is close to reported specific growth rate of 0.07 h^{-1} measured in in similar experimental conditions (Monteiro De Oliveira et al., 2021).

Table 11. Specific growth rates (h^{-1}) and 2-tail Student t-test p-values of R. toruloides IFO0880 mutant strains (XPK-Δ #2 - #10) compared against the wild type during nitrogen-limited growth phase cultivated in a chemically defined medium with xylose as a sole carbon substrate (20 g/L) at starting C/N ratio of 80 mol/mol.

	XPK-Δ-2	XPK-Δ-8	XPK-Δ-11	XPK-Δ-7	XPK-Δ-1	XPK-Δ-10	wild type
Sample A	0.051	0.047	0.053	0.052		0.054	0.047
Sample B	0.051	0.047	0.052	0.052	0.053	0.052	0.047
Sample C	0.054	0.046	0.053	0.051	0.052	0.053	0.050
Average	0.052	0.047	0.053	0.052	0.053	0.053	0.048
2-tail t-test p-value	0.029	0.261	0.007	0.015	0.030	0.007	

Student's 2-tailed t-test confirmed statistically significant ($p < 0.05$) difference of specific growth rate for all mutants (p range from 0.007 to 0.030) except XPK-Δ-8 ($p = 0.261$) in nitrogen-limited phase confirming statistically significant change in specific growth rate compared to wild type. Average increase of the growth rate of mutants with statistically significant difference is 10% when compared to the wild type strain.

5. Conclusions

CRISPR/Cas9 method was used to create a phosphoketolase knockout strains of *Rhodotorula toruloides*. The mutant strains were then investigated for the change in genotype and growth rate compared to the wild type strain.

Sequencing confirmed 13 mutant strains out of 16 strains produced with random integrative transformation. 3 of the mutants were dropped from sequencing due to failed PCR. This means that all sequenced strains were confirmed as mutants. CRISPR/Cas9 method utilizing random integration is suitable for producing knockout mutants of XPK in *R. toruloides.*

Characterisation of knockout strains in a chemically defined medium with xylose as a sole carbon source showed a statistically significant change in one of the mutants in exponential growth phase and in all mutants except one in nitrogen-limited growth phase.

Repeated experiment is needed to evaluate the effect of XPK knockout in exponential growth phase, since the results of this work show statistically significant differences in growth rates where there should be no such differences. Most likely reason for that is limited accuracy in growth medium preparation.

Nitrogen-limited growth phase showed statistically significant increase in specific growth rates in all the mutants except one. The increase of the growth rate among the mutants with statistically significant difference is 10%. This can be useful in biotechnological applications where lipids are the product of *R. toruloides* based cell factories. Increased specific growth rate in lipid accumulation phase can mean less time needed for fermentation to complete.

Annotatsioon

Rhodotorula toruloides on mittekonventsionaalne õline pärm, mis on suuteline kasvama erinevatel substraatidel ja tootma lipiide, karotenoide ja muid aineid mida on võimalik biotehnoloogiliselt toota. Käesolev töö uuris *Rhodotorula toruloides*'e füsioloogiat kasutades selleks CRISPR/Cas9 meetodit et välja lülitada fosfoketolaasi geen. Meetod andis tulemuseks mitu erineva genotüübiga mutanti mille kasvukiiruse iseloomustamiseks kasvatati neid mini-bioreaktorites ksüloosi sisaldavas keemiliselt määratletud söötmes. Bioreaktori eksperimendist saadud kasvukõverate põhjal määrati mutantide erikasvukiirused eksponentsiaalses ja lämmastikupiiranguga kasvufaasis.

Eksponentsiaalses kasvufaasis varieerusid erikasvukiirused märgatavalt, aga ei olnud statistiliselt olulised välja arvatud ühe mutandi puhul. Lämmastikupiiranguga kasvufaasis oli varieeruvus tunduvalt madalam ja kõigil mutantidel peale ühe oli erikasvukiiru ligi 10% kõrgem kui looduslikul organismil.

Töö eesmärgid täitusid – mutantide loomine õnnestus ja muutust nende füsioloogias lämmastikupiiranguga kasvufaasis oli võimalik demonstreerida. Erikasvukiiruse kõrge varieeruvus eksponentsiaalses kasvufaasis viitab vajadusele tulevikus bioreaktori katset täpsemini ette valmistada, et seda vähendada. Kõrgem erikasvukiirus lämmastikupiiranguga kasvufaasis aitab tulevikus täpsemini selle kasvufaasi kestvuse põhjal eksperimendi kogukestvust ennustada.

Abstract

Rhodotorula toruloides is nonconventional oleaginous yeast capable of growing on various substrates and producing lipids, carotenoids and other biotechnologically interesting products. This thesis evaluated the the effect of a phosphoketolase knockout on cell growth of *Rhodotorula toruloides* by knocking out phosphoketolase gene using CRISPR/Cas9 method. The knockout method produced several mutant strains with different genotypes which were grown in a chemically defined medium using xylose as a single carbon source in mini-bioreactors. The growth curves from bioreactor experiment were used to determine specific growth rates in exponential and nitrogen-limited growth phases.

The specific growth rates showed significant variance, but no statistically significant change for mutants in exponential growth phase except in case of one mutant strain. Nitrogen-limited growth phase showed statistically significant increase of 10% with low variance in specific growth rate over wild type strain.

The aims of the thesis were achieved – mutants were successfully created and the change in the mutant physiology in nitrogen-limited growth phase was demonstrated. The variance in specific growth rate in exponential growth phase points to improvements needed in future experiment preparation. The confirmed increase of specific growth rate in nitrogen-limited phase can be useful for planning the expected length of future experiments.

Acknowledgements

I would like to offer my sincere gratitude to my advisor Alīna Reķēna for her patience and insight in guiding me through the work in the laboratory and writing this thesis. I would also like to thank everybody in Bioengineering lab for offering their support and knowledge every time I asked for it.

References

Abeln, F., Chuck, C.J., 2021. The history, state of the art and future prospects for oleaginous yeast research. Microb Cell Fact 20, 221. https://doi.org/10.1186/s12934-021-01712-1

Ageitos, J.M., Vallejo, J.A., Veiga-Crespo, P., Villa, T.G., 2011. Oily yeasts as oleaginous cell factories. Appl Microbiol Biotechnol 90, 1219–1227. https://doi.org/10.1007/s00253-011-3200-z

Bonturi, N., Crucello, A., Viana, A.J.C., Miranda, E.A., 2017. Microbial oil production in sugarcane bagasse hemicellulosic hydrolysate without nutrient supplementation by a Rhodosporidium toruloides adapted strain. Process Biochemistry 57, 16–25. https://doi.org/10.1016/j.procbio.2017.03.007

Bonturi, N., Pinheiro, M.J., De Oliveira, P.M., Rusadze, E., Eichinger, T., Liudžiūtė, G., De Biaggi, J.S., Brauer, A., Remm, M., Miranda, E.A., Ledesma-Amaro, R., Lahtvee, P.-J., 2022. Development of a dedicated Golden Gate Assembly Platform (RtGGA) for *Rhodotorula toruloides*. Metabolic Engineering Communications 15, e00200. https://doi.org/10.1016/j.mec.2022.e00200

Botham, P.A., Ratledge, C., 1979. A Biochemical Explanation for Lipid Accumulation in Candida 107 and Other Oleaginous Micro-organisms. Journal of General Microbiology 114, 361–375. https://doi.org/10.1099/00221287-114-2-361

Castañeda, M.T., Nuñez, S., Garelli, F., Voget, C., De Battista, H., 2018. Comprehensive analysis of a metabolic model for lipid production in *Rhodosporidium toruloides*. Journal of Biotechnology 280, 11–18. https://doi.org/10.1016/j.jbiotec.2018.05.010

Chandrasegaran, S., 2017. Recent advances in the use of ZFN-mediated gene editing for human gene therapy. Cell Gene Therapy Insights 3, 33–41. https://doi.org/10.18609/cgti.2017.005

Coradetti, S.T., Pinel, D., Geiselman, G.M., Ito, M., Mondo, S.J., Reilly, M.C., Cheng, Y.-F., Bauer, S., Grigoriev, I.V., Gladden, J.M., Simmons, B.A., Brem, R.B., Arkin, A.P., Skerker, J.M., 2018. Functional genomics of lipid metabolism in the oleaginous yeast *Rhodosporidium toruloides*. eLife 7, e32110. https://doi.org/10.7554/eLife.32110

DiCarlo, J.E., Conley, A.J., Penttilä, M., Jäntti, J., Wang, H.H., Church, G.M., 2013. Yeast Oligo-Mediated Genome Engineering (YOGE). ACS Synth. Biol. 2, 741–749. https://doi.org/10.1021/sb400117c

Donot, F., Fontana, A., Baccou, J.C., Strub, C., Schorr-Galindo, S., 2014. Single cell oils (SCOs) from oleaginous yeasts and moulds: Production and genetics. Biomass and Bioenergy 68, 135–150. https://doi.org/10.1016/j.biombioe.2014.06.016

Groenewald, M., Boekhout, T., Neuvéglise, C., Gaillardin, C., Van Dijck, P.W.M., Wyss, M., 2014. *Yarrowia lipolytica* : Safety assessment of an oleaginous yeast with a great industrial potential. Critical Reviews in Microbiology 40, 187–206. https://doi.org/10.3109/1040841X.2013.770386

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., Charpentier, E., 2012. A Programmable Dual-RNA–Guided DNA Endonuclease in Adaptive Bacterial Immunity. Science 337, 816–821. https://doi.org/10.1126/science.1225829

Kamineni, A., Consiglio, A.L., MacEwen, K., Chen, S., Chifamba, G., Shaw, A.J., Tsakraklides, V., 2021. Increasing lipid yield in *Yarrowia lipolytica* through phosphoketolase and phosphotransacetylase expression in a phosphofructokinase deletion strain. Biotechnol Biofuels 14, 113. https://doi.org/10.1186/s13068-021-01962-6

Koch, R., 1881. Methods for the study of pathogenic organisms.

Koh, C.M.J., Liu, Y., Moehninsi, Du, M., Ji, L., 2014. Molecular characterization of KU70 and KU80 homologues and exploitation of a KU70-deficient mutant for improving gene deletion frequency in Rhodosporidium toruloides. BMC Microbiol 14, 50. https://doi.org/10.1186/1471-2180-14-50

Koutinas, A.A., Chatzifragkou, A., Kopsahelis, N., Papanikolaou, S., Kookos, I.K., 2014. Design and techno-economic evaluation of microbial oil production as a renewable resource for biodiesel and oleochemical production. Fuel 116, 566–577. https://doi.org/10.1016/j.fuel.2013.08.045

Labuhn, M., Adams, F.F., Ng, M., Knoess, S., Schambach, A., Charpentier, E.M., Schwarzer, A., Mateo, J.L., Klusmann, J.-H., Heckl, D., 2018. Refined sgRNA efficacy prediction improves large- and small-scale CRISPR–Cas9 applications. Nucleic Acids Research 46, 1375–1385. https://doi.org/10.1093/nar/gkx1268

Lahtvee, P.-J., Sánchez, B.J., Smialowska, A., Kasvandik, S., Elsemman, I.E., Gatto, F., Nielsen, J., 2017. Absolute Quantification of Protein and mRNA Abundances Demonstrate Variability in Gene-Specific Translation Efficiency in Yeast. Cell Systems 4, 495-504.e5. https://doi.org/10.1016/j.cels.2017.03.003

Lemos, B.R., Kaplan, A.C., Bae, J.E., Ferrazzoli, A.E., Kuo, J., Anand, R.P., Waterman, D.P., Haber, J.E., 2018. CRISPR/Cas9 cleavages in budding yeast reveal templated insertions and strand-specific insertion/deletion profiles. Proc. Natl. Acad. Sci. U.S.A. 115. https://doi.org/10.1073/pnas.1716855115

Li, T., Huang, S., Zhao, X., Wright, D.A., Carpenter, S., Spalding, M.H., Weeks, D.P., Yang, B., 2011. Modularly assembled designer TAL effector nucleases for targeted gene knockout and gene replacement in eukaryotes. Nucleic Acids Research 39, 6315–6325. https://doi.org/10.1093/nar/gkr188

Li, Y., Zhao, Z. (Kent), Bai, F., 2007. High-density cultivation of oleaginous yeast Rhodosporidium toruloides Y4 in fed-batch culture. Enzyme and Microbial Technology 41, 312–317. https://doi.org/10.1016/j.enzmictec.2007.02.008

Lõoke, M., Kristjuhan, K., Kristjuhan, A., 2011. Extraction of genomic DNA from yeasts for PCR-based applications. BioTechniques 50, 325–328. https://doi.org/10.2144/000113672

Lopes, H.J.S., Bonturi, N., Kerkhoven, E.J., Miranda, E.A., Lahtvee, P.-J., 2020. C/N ratio and carbon source-dependent lipid production profiling in *Rhodotorula toruloides*. Appl Microbiol Biotechnol 104, 2639–2649. https://doi.org/10.1007/s00253-020-10386-5

Lundin, H., 1950. Fat synthesis by micro-organisms and its possible applications in industry. Journal of the Institute of Brewing 56, 17–28. https://doi.org/10.1002/j.2050-0416.1950.tb01516.x

Monteiro De Oliveira, P., Aborneva, D., Bonturi, N., Lahtvee, P.-J., 2021. Screening and Growth Characterization of Non-conventional Yeasts in a Hemicellulosic Hydrolysate. Front. Bioeng. Biotechnol. 9, 659472. https://doi.org/10.3389/fbioe.2021.659472

Mytilinaios, I., Salih, M., Schofield, H.K., Lambert, R.J.W., 2012. Growth curve prediction from optical density data. International Journal of Food Microbiology 154, 169–176. https://doi.org/10.1016/j.ijfoodmicro.2011.12.035

Otoupal, P.B., Ito, M., Arkin, A.P., Magnuson, J.K., Gladden, J.M., Skerker, J.M., 2019. Multiplexed CRISPR-Cas9-Based Genome Editing of *Rhodosporidium toruloides*. mSphere 4, e00099-19. https://doi.org/10.1128/mSphere.00099-19

Papanikolaou, S., Aggelis, G., 2011. Lipids of oleaginous yeasts. Part II: Technology and potential applications. Euro J Lipid Sci & Tech 113, 1052–1073. https://doi.org/10.1002/ejlt.201100015

Pinheiro, M.J., Bonturi, N., Belouah, I., Miranda, E.A., Lahtvee, P.-J., 2020. Xylose Metabolism and the Effect of Oxidative Stress on Lipid and Carotenoid Production in Rhodotorula toruloides: Insights for Future Biorefinery. Front. Bioeng. Biotechnol. 8, 1008. https://doi.org/10.3389/fbioe.2020.01008

Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A., Zhang, F., 2013. Genome engineering using the CRISPR-Cas9 system. Nat Protoc 8, 2281–2308. https://doi.org/10.1038/nprot.2013.143

Ratledge, C., Wynn, J.P., 2002. The Biochemistry and Molecular Biology of Lipid Accumulation in Oleaginous Microorganisms, in: Advances in Applied Microbiology. Elsevier, pp. 1–52. https://doi.org/10.1016/S0065-2164(02)51000-5

Rigouin, C., Gueroult, M., Croux, C., Dubois, G., Borsenberger, V., Barbe, S., Marty, A., Daboussi, F., André, I., Bordes, F., 2017. Production of Medium Chain Fatty Acids by *Yarrowia lipolytica:* Combining Molecular Design and TALEN to Engineer the Fatty Acid Synthase. ACS Synth. Biol. 6, 1870–1879. https://doi.org/10.1021/acssynbio.7b00034

Schultz, J.C., Mishra, S., Gaither, E., Mejia, A., Dinh, H., Maranas, C., Zhao, H., 2022. Metabolic engineering of *Rhodotorula toruloides* IFO0880 improves C16 and C18 fatty alcohol production from synthetic media. Microb Cell Fact 21, 26. https://doi.org/10.1186/s12934-022-01750-3

Stemmer, M., Thumberger, T., Del Sol Keyer, M., Wittbrodt, J., Mateo, J.L., 2015. CCTop: An Intuitive, Flexible and Reliable CRISPR/Cas9 Target Prediction Tool. PLoS ONE 10, e0124633. https://doi.org/10.1371/journal.pone.0124633

Thermo Fisher online calculator: https://www.thermofisher.com/ee/en/home/brands/thermo-scientific/molecular-biology/molecular-biology-learning-center/molecular-biology-resource-library/thermo-scientific-web-tools/tm-calculator.html

Tiukova, I.A., Brandenburg, J., Blomqvist, J., Sampels, S., Mikkelsen, N., Skaugen, M., Arntzen, M.Ø., Nielsen, J., Sandgren, M., Kerkhoven, E.J., 2019. Proteome analysis of xylose metabolism in Rhodotorula toruloides during lipid production. Biotechnol Biofuels 12, 137. https://doi.org/10.1186/s13068-019-1478-8

Von Nägeli, 1878. Ueber die Chemische Zusammensetzung der Hefe. J. Prakt. Chem. 17, 403–428. https://doi.org/10.1002/prac.18780170139

Wang, L., Lin, J., Zhang, T., Xu, K., Ren, C., Zhang, Z., 2013. Simultaneous Screening and Validation of Effective Zinc Finger Nucleases in Yeast. PLoS ONE 8, e64687. https://doi.org/10.1371/journal.pone.0064687

Wen, Z., Zhang, S., Odoh, C.K., Jin, M., Zhao, Z.K., 2020. *Rhodosporidium toruloides* - A potential red yeast chassis for lipids and beyond. FEMS Yeast Research 20, foaa038. https://doi.org/10.1093/femsyr/foaa038

Zhang, J., 1999. Chemically defined media for commercial fermentations.

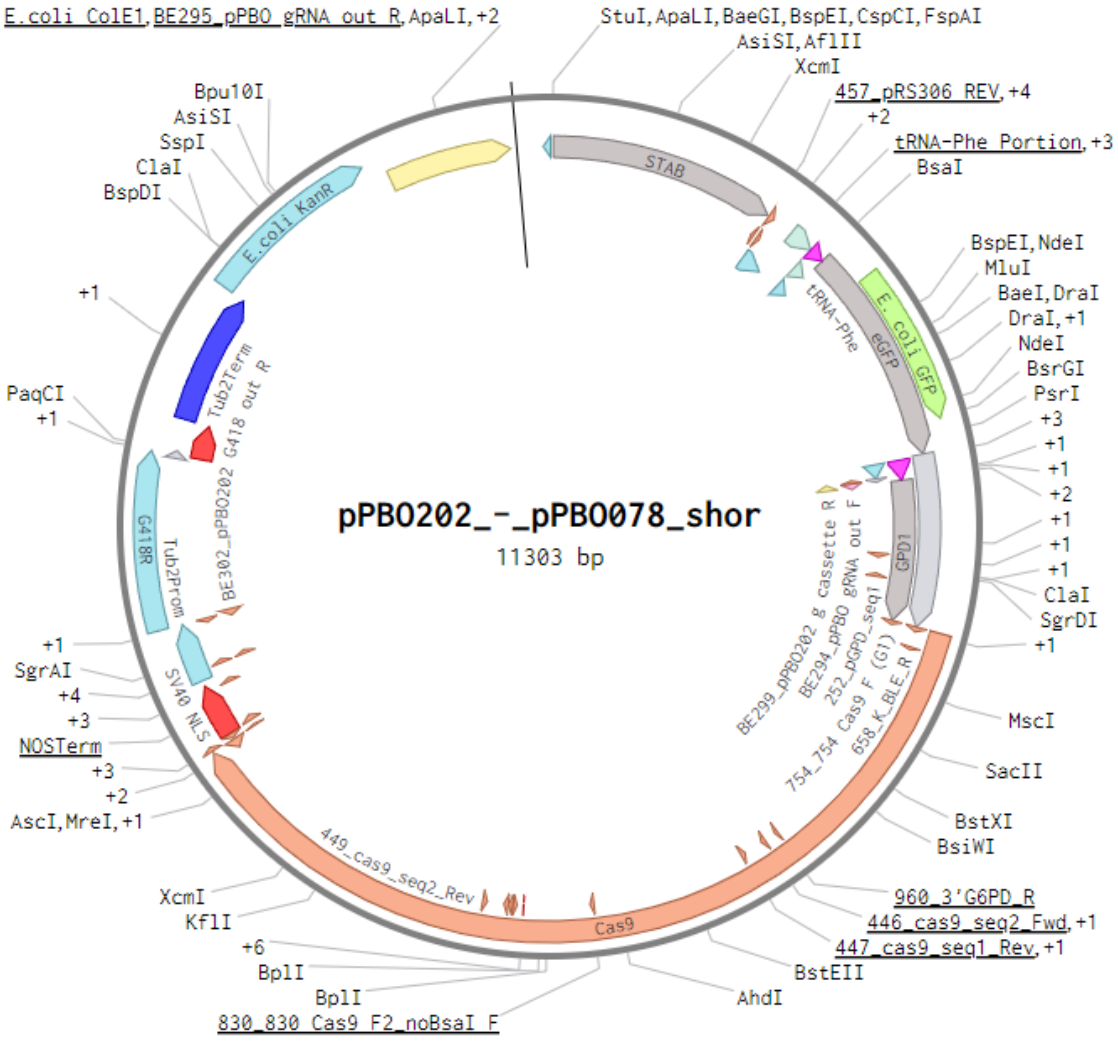
Zhang, S., Ito, M., Skerker, J.M., Arkin, A.P., Rao, C.V., 2016a. Metabolic engineering of the oleaginous yeast Rhodosporidium toruloides IFO0880 for lipid overproduction during high-density fermentation. Appl Microbiol Biotechnol 100, 9393–9405. https://doi.org/10.1007/s00253-016-7815-y

Zhang, S., Skerker, J.M., Rutter, C.D., Maurer, M.J., Arkin, A.P., Rao, C.V., 2016b. Engineering *Rhodosporidium toruloides* for increased lipid production. Biotech & Bioengineering 113, 1056–1066. https://doi.org/10.1002/bit.25864

Zhu, Z., Zhang, S., Liu, H., Shen, H., Lin, X., Yang, F., Zhou, Y.J., Jin, G., Ye, M., Zou, H., Zhao, Z.K., 2012. A multi-omic map of the lipid-producing yeast Rhodosporidium toruloides. Nat Commun 3, 1112. https://doi.org/10.1038/ncomms2112

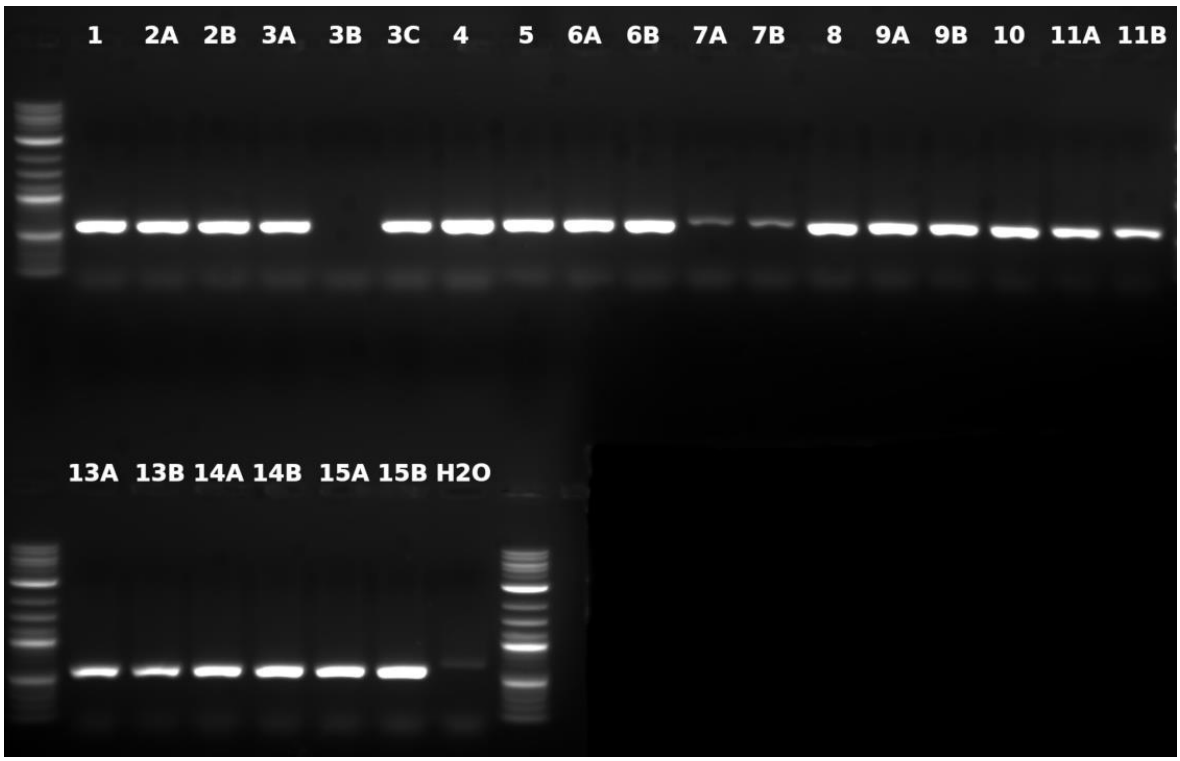
Supplementary materials

S1 pPBO.202 plasmid map



Supplementary figure 1. Map of pPBO.202 plasmid.

S2 Preparative PCR for XPK gene of 3rd generation transformants



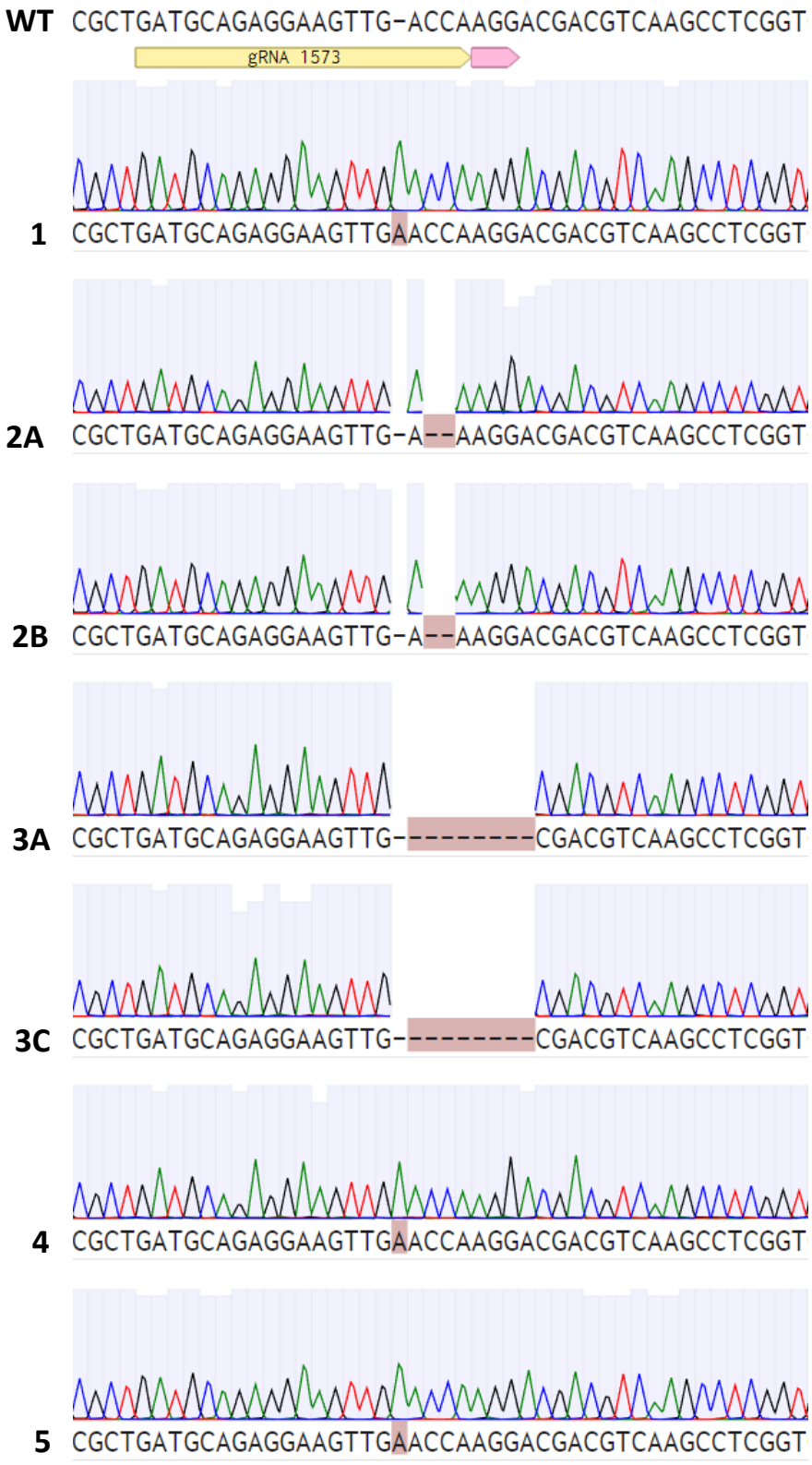
Supplementary figure 2. Preparative PCR results on agarose gel for *R. toruloides* IFO0880 3rd generation transformants amplifying genomic DNA fragment from XPK gene.

S3 Preparative PCR nucleotide concentration

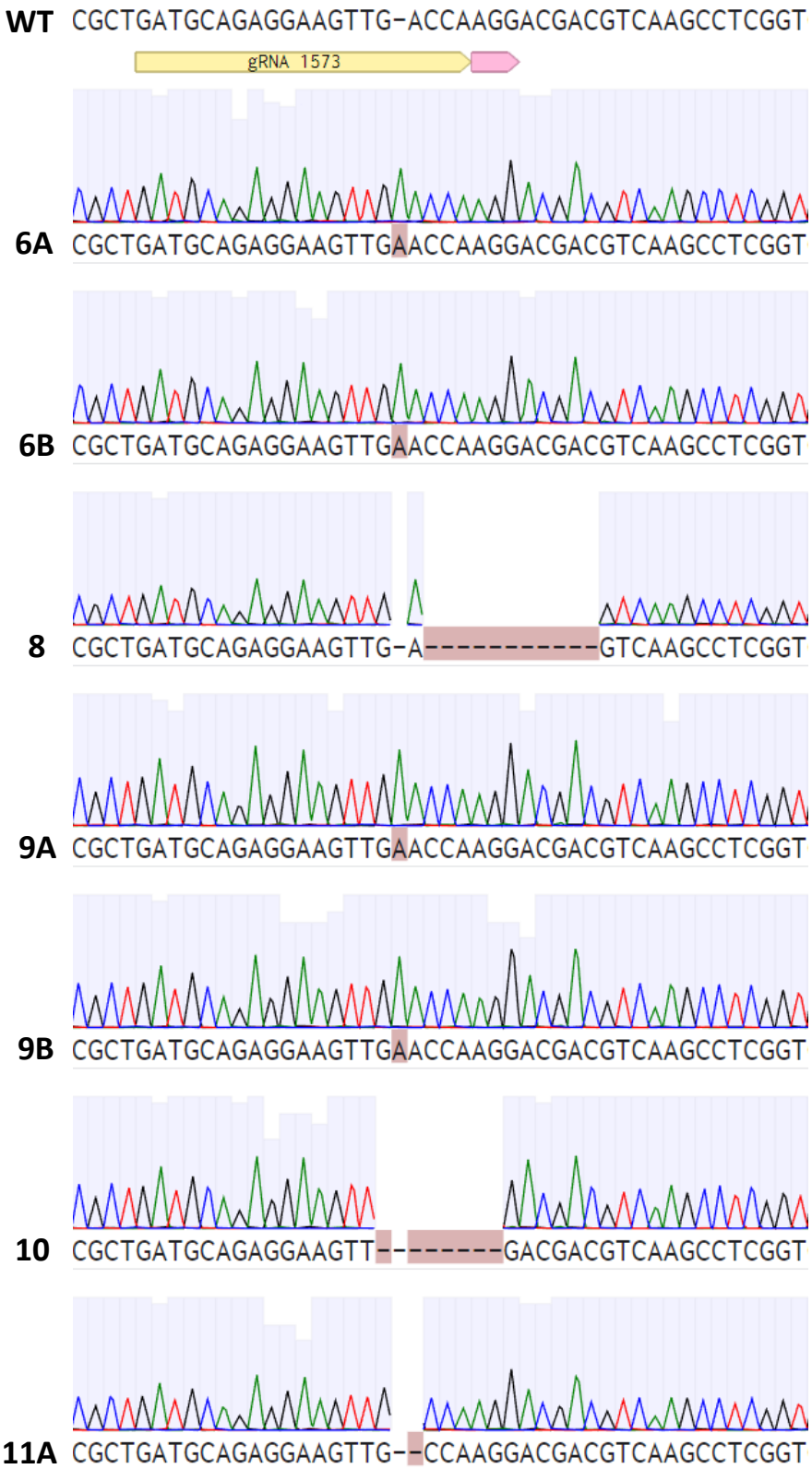
Supplementary table 1. Nucleotide concentration of XPK gene segment from preparative PCR extracted from gel.

Sample	Nucleotide concentration, ng/µL
1	14.9
2A	13.9
2B	14.0
3A	11.5
3C	9.1
4	19.4
5	11.7
6A	11.3
6B	12.7
7A	-0.1
7B	0.5
8	10.1
9A	9.7
9B	9.7
10	9.8
11A	8.0
11B	6.9
13A	7.4
13B	6.2
14A	10.1
14B	10.4
15A	8.8
15B	12.2

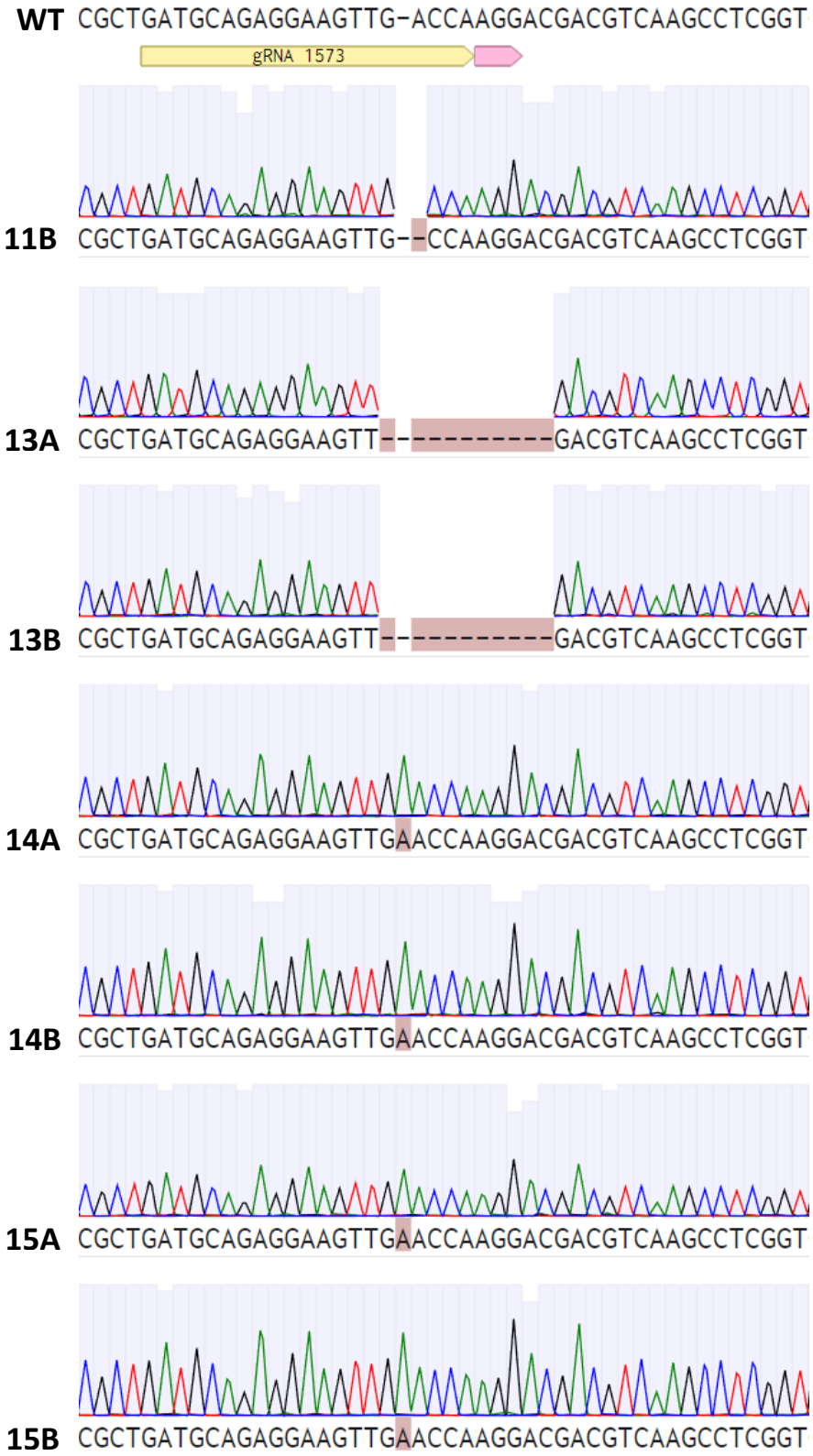
S4 Sanger sequencing results of 3rd generation transformants



Supplementary figure 3. Genotype comparison of wild type *R. toruloides* IFO0880 strain exon #1 region of XPK gene with sequenced transformants (1 – 5). Alignment includes read quality data. gRNA is annotated with yellow, PAM sequence with pink. Inserted or deleted nucleotides in mutant strains are marked with red.

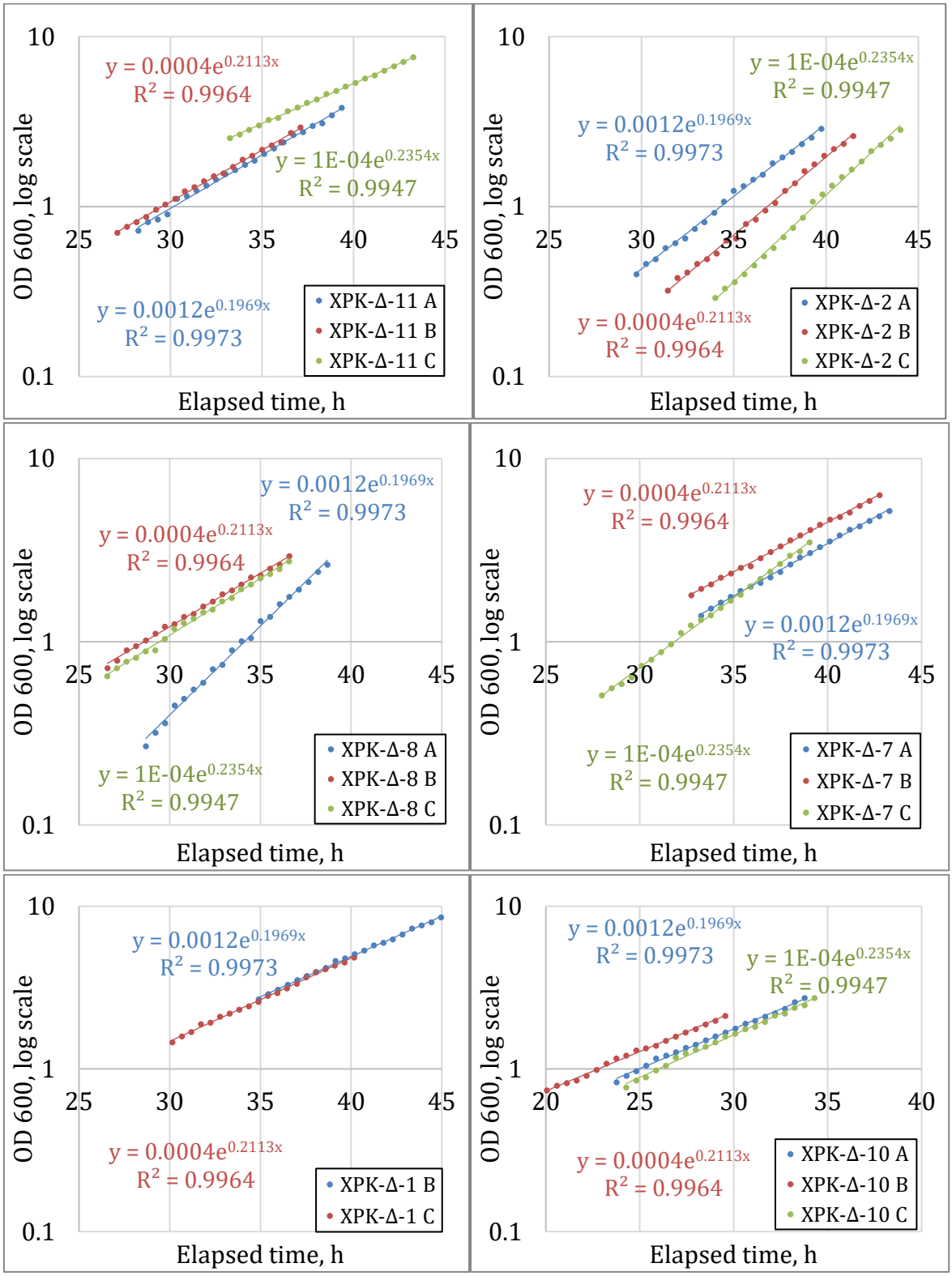


Supplementary figure 4. Genotype comparison of wild type *R. toruloides* IFO0880 strain exon #1 region of XPK gene with sequenced transformants (6A – 11A). Alignment includes read quality data. gRNA is annotated with yellow, PAM sequence with pink. Inserted or deleted nucleotides in mutant strains are marked with red.

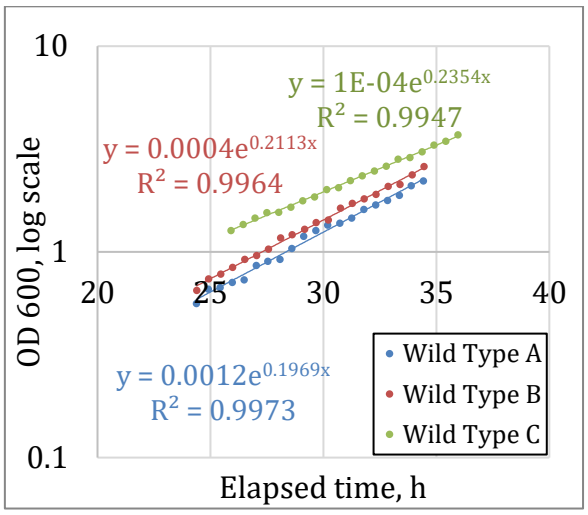


Supplementary figure 5. Genotype comparison of wild type *R. toruloides* IFO0880 strain exon #1 region of XPK gene with sequenced transformants (11B – 15B). Alignment includes read quality data. gRNA is annotated with yellow, PAM sequence with pink. Inserted or deleted nucleotides in mutant strains are marked with red.

S5 Growth rate calculation of knockouts during the exponential phase

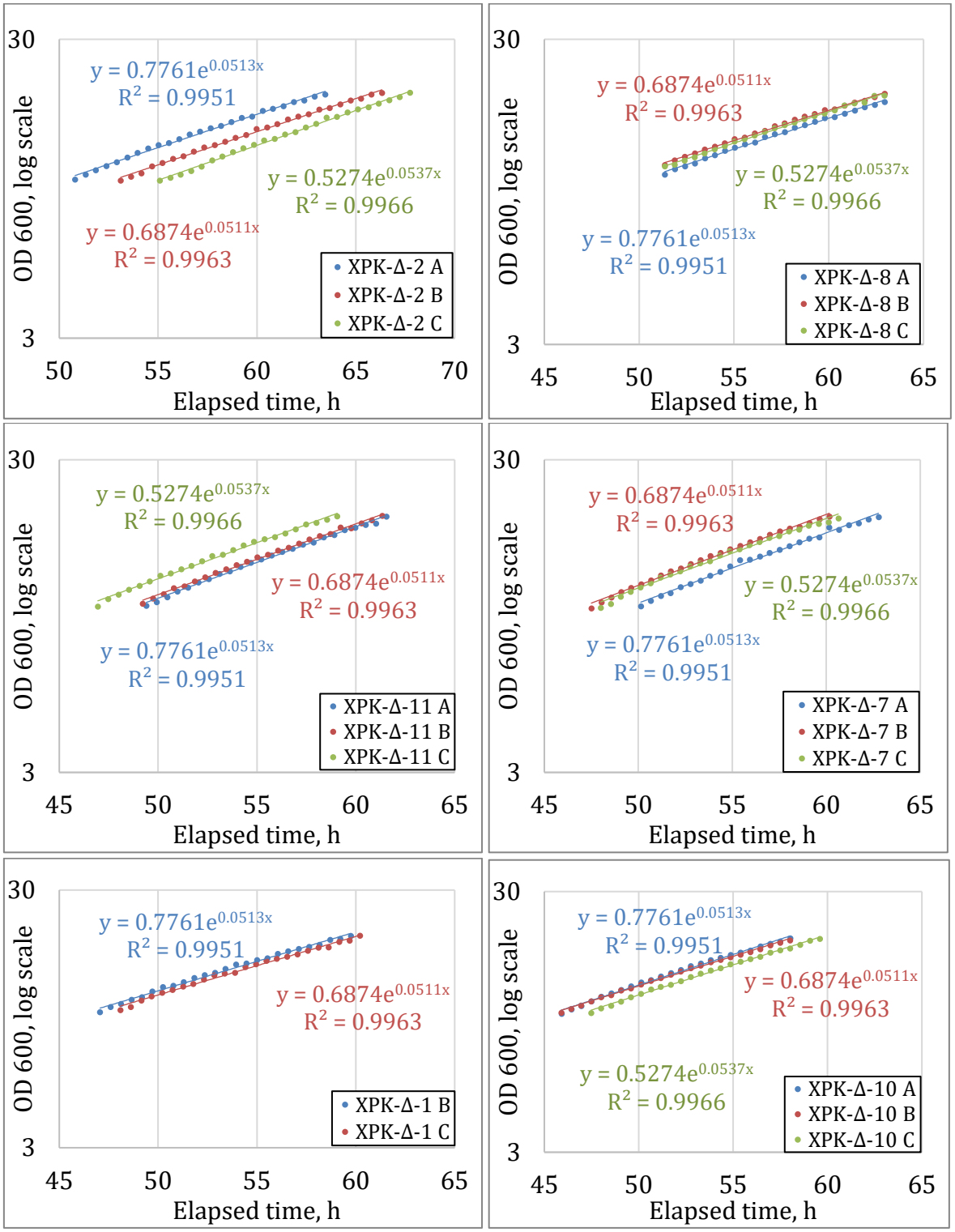


Supplementary figure 6. Specific growth rates of *R. toruloides* IFO0880 XPK knockout mutants in exponential growth phase, expressed as OD 600 measurements, presented at logarithmic scale. Exponential trendline added to 10-hour time window with highest average specific growth rate between 20 and 45 hours from inoculation and specific growth rate calculated from that trendline.

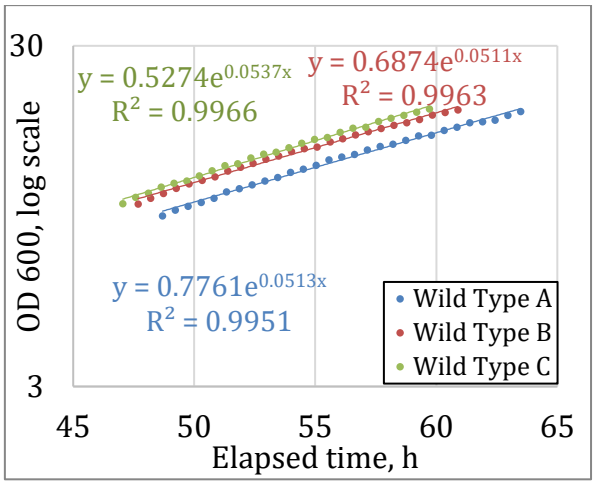


Supplementary figure 7. Specific growth rates of *R. toruloides* IFO0880 wild type strains in exponential growth phase, expressed as OD 600 measurements, presented at logarithmic scale. Exponential trendline added to 12-hour time window with highest average specific growth rate between 20 and 45 hours from inoculation and specific growth rate calculated from that trendline.

S6 Growth rates in nitrogen-limited phase



Supplementary figure 8. Specific growth rates of *R. toruloides* IFO0880 XPK knockout mutants in nitrogen-limited growth phase, expressed as OD 600 measurements, presented at logarithmic scale. Exponential trendline added to 12-hour time window with highest average specific growth rate between 45 and 70 hours from inoculation and specific growth rate calculated from that trendline.



Supplementary figure 9. Specific growth rates of *R. toruloides* IFO0880 wild type strains in nitrogen-limited growth phase, expressed as OD 600 measurements, presented at logarithmic scale. Exponential trendline added to 12-hour time window with highest average specific growth rate between 45 and 70 hours from inoculation and specific growth rate calculated from that trendline.

S7 Final OD 600 values

Supplementary table 2. Manually measured OD 600 values at the end of the experiment of *R. toruloides* IFO0880 XPK knockout mutants and the wild type IFO0880 strain cultivated in chemically defined medium. Standard deviation refers to three biological replicates, except in case of XPK-Δ-1 where only two replicates are present

	Replicate				
Strain	A	B	C	Average	SD
XPK-Δ-2	43	37	34	38.0	4.6
XPK-Δ-8	44	45	42	43.7	1.5
XPK-Δ-11	32	34	40	35.3	4.2
XPK-Δ-7	34	38	37	36.3	2.1
XPK-Δ-1		36	44	40.0	5.7
XPK-Δ-10	35	41	41	39.0	3.5
wild type	38	36	39	37.7	1.5

Lisa
rektori 07.04.2020 käskkirjale nr 1-8/17

Lihtlitsents lõputöö reprodutseerimiseks ja lõputöö üldsusele kättesaadavaks tegemiseks[1]

Mina, Kristjan Pals

1. Annan Tallinna Tehnikaülikoolile tasuta loa (lihtlitsentsi) enda loodud teose „Phosphoketolase gene knockout by CRISPR/Cas9 in nonconventional yeast *Rhodotorula toruloides*", mille juhendaja on Alīna Reķēna,

1.1 reprodutseerimiseks lõputöö säilitamise ja elektroonse avaldamise eesmärgil, sh Tallinna Tehnikaülikooli raamatukogu digikogusse lisamise eesmärgil kuni autoriõiguse kehtivuse tähtaja lõppemiseni;

1.2 üldsusele kättesaadavaks tegemiseks Tallinna Tehnikaülikooli veebikeskkonna kaudu, sealhulgas Tallinna Tehnikaülikooli raamatukogu digikogu kaudu kuni autoriõiguse kehtivuse tähtaja lõppemiseni.

2. Olen teadlik, et käesoleva lihtlitsentsi punktis 1 nimetatud õigused jäävad alles ka autorile.

3. Kinnitan, et lihtlitsentsi andmisega ei rikuta teiste isikute intellektuaalomandi ega isikuandmete kaitse seadusest ning muudest õigusaktidest tulenevaid õigusi.

---

29. 05. 2024

[1] *Lihtlitsents ei kehti juurdepääsupiirangu kehtivuse ajal vastavalt üliõpilase taotlusele lõputööle juurdepääsupiirangu kehtestamiseks, mis on allkirjastatud teaduskonna dekaani poolt, välja arvatud ülikooli õigus lõputööd reprodutseerida üksnes säilitamise eesmärgil. Kui lõputöö on loonud kaks või enam isikut oma ühise loomingulise tegevusega ning lõputöö kaas- või ühisautor(id) ei ole andnud lõputööd kaitsvale üliõpilasele kindlaksmääratud tähtajaks nõusolekut lõputöö reprodutseerimiseks ja avalikustamiseks vastavalt lihtlitsentsi punktidele 1.1. jq 1.2, siis lihtlitsents nimetatud tähtaja jooksul ei kehti.*