CHALMERS
UNIVERSITY OF TECHNOLOGY

Adaptive evolution of *Yarrowia lipolytica* for osmotic and saline tolerance

Improving tolerance towards osmotic and saline stress for sustainable production of biodiesel

Master's thesis within the Biotechnology Master Program

JOHN HELLGREN

Department of Biology and Biological Engineering
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2017

Master's thesis 2017

Adaptive evolution of *Yarrowia lipolytica* for osmotic and saline tolerance

Improving tolerance towards osmotic and saline stress for sustainable production of biodiesel

John Hellgren

CHALMERS
UNIVERSITY OF TECHNOLOGY

Department of Biology and Biological Engineering
Division of Systems and Synthetic Biology
Chalmers University of Technology
Gothenburg, Sweden 2017

Adaptive evolution of *Yarrowia lipolytica* for osmotic and saline tolerance
Improving tolerance towards osmotic and saline stress for sustainable production of biodiesel
JOHN HELLGREN

Supervisors: Joakim Norbeck and Eduard Kerkhoven
Department of Biology and Biological Engineering
Division of Systems and Synthetic Biology
Chalmers University of Technology

Examiner: Joakim Norbeck
Department of Biology and Biological Engineering
Division of Systems and Synthetic Biology
Chalmers University of Technology

Master's Thesis 2017
Department of Biology and Biological Engineering
Division of Systems and Synthetic Biology
Chalmers University of Technology
SE-412 96 Gothenburg
Telephone +46 31 772 1000

Cover: Fluorescent microscope picture of lipid accumulating *Yarrowia lipolytica* stained with Nile Red.

Typeset in LaTeX
Gothenburg, Sweden 2017

JOHN HELLGREN
Department of Biology and Biological Engineering
Chalmers University of Technology

Abstract

We need renewable resources to allow sustainable production of fuels. By using lipid accumulating yeast, biodiesel can be produced in a sustainable way from resources that were previously not used, for example lignocellulose-based sources such as agricultural waste. However, for this process to be profitable, the tolerance of the yeast needs to be improved. This project aims to improve the tolerance of the oleaginous yeast *Yarrowia lipolytica* towards osmotic and saline stress by using the method adaptive laboratory evolution. This method has previously been shown to be efficient in constructing strains to tolerate new conditions without the need of prior knowledge. After evolving *Y. lipolytica* for 220 generation in minimal medium containing 1.4 M NaCl, an improved performance in the same medium was observed, along with evolved cross-tolerance towards low pH. This indicates that this adaptive evolution of *Y. lipolytica* resulted in improved ionic tolerance rather than pure osmotic tolerance. The evolved strains were sent for whole genomic sequencing to find out which mutations that caused this phenotype. During this project, two previous CRISPR/Cas9 strategies were combined and adapted for efficient markerless reverse engineering. When genome data arrives, this strategy will be used for reconstruction of candidate mutations to find out which mutations are important for the observed phenotype. The gained knowledge from this evolution experiment can later be used for constructing a robust industrial strain that efficiently converts lignocellulose-based material to biodiesel, allowing sustainable production of fuels.

Keywords: *Yarrowia lipolytica*, adaptive laboratory evolution, osmotic, saline, tolerance, biodiesel.

Acknowledgements

After spending two summers in a row at Sysbio participating in the iGEM competition, I knew that this was the place where I'd want to continue working. Therefore, I chose to conduct my 10-month master's thesis work here at Sysbio. I want to thank my supervisor and examiner Joakim Norbeck and my secondary supervisor Eduard Kerkhoven for allowing me to do this master's thesis project and for always being there to provide me with useful advice and guidance when I needed it.

When I was not in the lab, I spend a lot of time in the master's office, called the dungeon. I'd like to thank all the other members of the dungeon, along with all the master students of Sysbio and Indbio for great discussions, tips and tricks and fika, especially Dóra Vitay and Gustaf Edman.

I would also like to thank all the research engineers at Sysbio for their incredible work to maintain a functional lab and their problem-solving skills.

Finally, I would like to thank everyone working at Sysbio for providing this great and friendly atmosphere, where you could always ask questions to anyone without hesitation.

John Hellgren, Gothenburg, June 2017

Contents

1
Introduction

One option for sustainable production of fuels is to produce biodiesel from vegetable oils. However, the profit of this process is heavily dependent on several factors, such as seasons, climate and the current situation of the market [1]. As a better alternative, microbes can be used to produce biodiesel from plants. This microbial solution has the advantages of being seasonal independent, have a fast growth rate and does not take up as much arable land [1]. However, one important factor is not to use a sugar-based feedstock, as that competes with the food production industry, which could lead to increasing prices and decreased availability of food in poor regions [2].

The competition with the food production could be avoided if the feedstock were based on lignocellulose instead, for example agricultural waste [2]. The problem with using a feedstock based on lignocellulose is that it requires extensive pre-treatments before the sugars can be utilized by microbes through fermentation. The treatments often cause the release of various inhibitors from the wood material, which negatively effects the fermentation process [3].

A promising candidate for biodiesel production is the oleaginous yeast *Yarrowia lipolytica*. The wild type can accumulate up to 36 % of dry cell weigh as lipids when growing on glucose. This production can be increased to 77 % (30.8 g/l lipids) in batch cultivation and 73 % (85 g/l lipids) in fed-batch cultivation through metabolic engineering and optimization of cultivation conditions [4]. However, the tolerance of *Y. lipolytica* needs to be improved to achieve these levels in lignocellulose-derived media [5].

1.1 Aims

The main goal of this study is to evolve *Y. lipolytica* to increase its tolerance towards high osmotic and saline stress using adaptive laboratory evolution. The final evolved strain will be characterized and the mechanism of adaption discovered by connecting the improved phenotype to genome data obtained through whole genome sequencing. The findings of this study can then be transferred to an industrial strain, contributing to the construction of a robust strain capable of efficient conversion of lignocellulose waste to biodiesel.

2

Theory

2.1 *Yarrowia lipolytica*

Y. lipolytica is a dimorphic yeast that can either grow as yeast single cells or as filamentous hyphae depending on environmental conditions [6]. An example of the different morphologies can be seen in Figure 2.1.

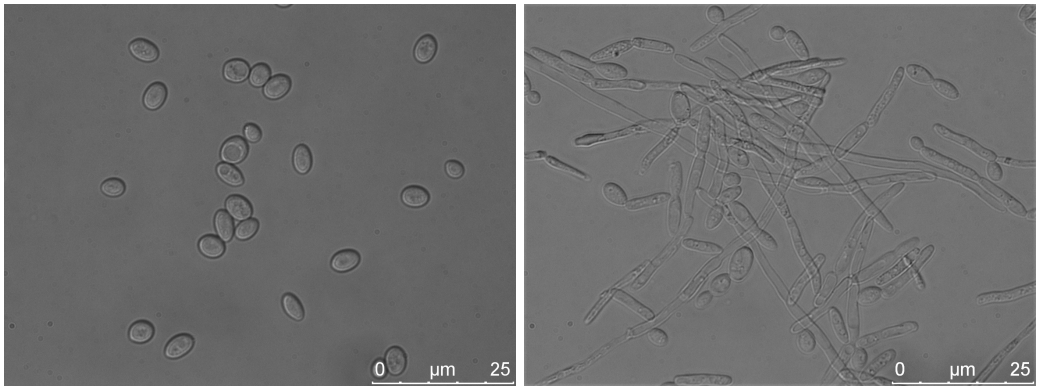

Figure 2.1: Morphology of *Y. lipolytica* in complex medium (left) and minimal medium (right)

Wild type *Y. lipolytica* can utilize a wide range of substrates, including glucose, fructose, glycerol and hydrophobic substrates such as alkenes, fatty acids, fats and oils [7, 8], in contrast with *Saccharomyces cerevisiae*, which prefers glucose as the carbon source [9]. *Y. lipolytica* even grows faster on glycerol (0.3 h^{-1}) than on glucose (0.24 h^{-1}) [9]. The ability to utilize hydrocarbons is of special interest for use of *Y. lipolytica* for bioremediation of soil polluted with oil or industrial waste containing oils [10]. In the industry, *Y. lipolytica* has been used in large-scale production of single cell protein and citric acid [11]. Furthermore, with its native ability to secrete proteins, mainly proteases and lipases, *Y. lipolytica*'s efficient secretion system has been exploited for heterologous protein production [11].

In contrast to *S. cerevisiae*, which both respire and ferment, *Y. lipolytica* has a obligatory respiratory metabolism [12] and requires oxygen to survive [13]. The respiratory chain of *Y. lipolytica* is more similar to mammalian cells than *S. cerevisiae*, and presence of the mitochondrial complex I makes *Y. lipolytica* an attractive model of mitochondrial complex I studies [14].

2.1.1 Lipid accumulation

Important for biodiesel production is the ability of the yeast to produce a large amount of lipids [15]. *Y. lipolytica* is an oleaginous yeast, which means that it accumulates lipids

to a level of more than 20 % of its biomass [4]. Wild type *Y. lipolytica* can accumulate typically up to 36 % lipids, where for example the non-oleaginous yeast *S. cerevisiae* can accumulate less than 15 % [16]. The lipid accumulation of *Y. lipolytica* is extensively studied and is therefore used as a model oleaginous yeast [16]. Lipid accumulation is often induced by cultivating the cells in nitrogen limiting conditions, which diverts the carbon flux towards lipid synthesis [16]. The large amounts of lipids produced are stored mainly as triacylglycerols (TAG) in lipid droplets [17], as free fatty acids could be toxic for the cell [4].

The assimilation of TAGs can be seen in Figure 2.2. It begins with acylation of glycerol-3-phosphate (G-3-P), first by G-3-P acyltransferase (*SCT1*) into lysophosphatidic acid (LPA) and secondly by lysophosphatidic acid acyltransferase (*SLC1*) into phosphatidic acid (PA). PA is dephosphorylated by phosphatidic acid phosphohydrolase (*PAP*, *PAH1* in *S. cerevisiae*), forming diacylglycerol (DAG). The last acyl is added either by an acyl-CoA-dependent pathway, where addition of acyl-CoA is catalyzed by diacylglycerol acyltransferase (*DGA1,2*) and steryl ester synthase (*ARE1,2*), or through an acyl-CoA-independent pathway, where addition of glycerophospholipid (PL) is catalyzed by phospholipid diacylglycerol acyltransferase (*LRO1*) [16].

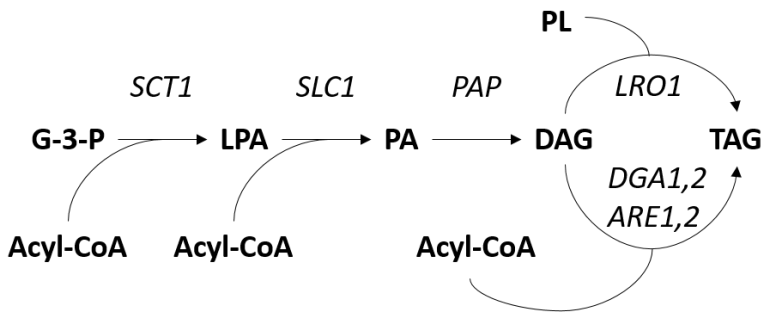

Figure 2.2: Formation of TAG in *Y. lipolytica* from G-3-P. G-3-P: glycerol-3-phosphate; LPA: lysophosphatidic acid; PA: phosphatidic acid; DAG: diacylglycerol; PL: glycerophospholipid; TAG: triacylglycerols; *SCT1*: G-3-P acyltransferase; *SLC1*: lysophosphatidic acid acyltransferase; *PAP*: phosphatidic acid phosphohydrolase; *LRO1*: phospholipid diacylglycerol acyltransferase; *DGA1,2*: diacylglycerol acyltransferase; *ARE1,2*: steryl ester synthase. Picture adapted from [4, 16]

Lipid accumulation can be verified *in vivo* by staining of lipid droplets with e.g. Nile red.

2.2 Osmotic stress

One kind of stress that the cells experience in lignocellulose-derived medium is osmotic stress [18]. Osmotic stress is caused by changes in external water activity and is a stress factor that cells need to cope with to be able to survive when conditions change. Decreased water activity is caused by an increase of osmolarity outside the cell, which could be the result of a high concentration of salts or sugars. This kind of stress is called hyper-osmotic stress and cellular water is lost due to passive diffusion, which leads to cell shrinking, arrest of cellular functions and denaturation of biomolecules [19]. The loss of cell volume can be as high as 50 % when *S. cerevisiae* are exposed to 1 M sodium chloride (NaCl) [20].

In contrast, hypo-osmotic stress is the result of high water activity, i.e. low concentration of salts, which leads to an influx of water that causes cell swelling and an increase of turgor pressure. This can result in bursting of cells; however, the cell wall of a yeast cell prevents

that from happening [19].

2.2.1 Compatible solute

A common way to handle hyper-osmotic stress is for the cell to produce organic osmolytes which keeps the osmotic pressure inside the cell equal to the outside environment and thus helps the cell to retain its water [21]. In fungi, polyols are commonly used as an osmolyte, where the extensively studied yeast *S. cerevisiae* uses glycerol as their osmolyte (or compatible solute as it does not interfere with cellular functions) when growing on glucose [22]. The information of how *Y. lipolytica* copes with osmotic stress is limited, but observations in *S. cerevisiae* are relevant, as many pathways are possible conserved in yeast.

Glycerol is produced from dihydroxyacetone phosphate (DHAP) in two steps in *S. cerevisiae*, see Figure 2.3. In the first step DHAP is converted into glycerol-3-phosphate (G-3-P) by glycerol-3-phosphate dehydrogenase, which has two isozymes (Gpd1 and Gpd2). Deletion of *GPD1* results in moderate sensitivity to osmotic stress, while *GPD2* deletion does not cause any increase in sensitivity, showing that Gpd1 plays the most important role during osmotic stress [22]. However, deletion of both isozymes is required to obtain strong sensitivity to osmotic stress [22]. The second step in glycerol production is conversion of G-3-P into glycerol by glycerol-3-phosphatase (Gpp1,2). The expression of *GPD1*, *GPP1* and *GPP2* is induced by osmotic stress [22]. There is also another route to glycerol through dihydroxyacetone (DHA), but it is not contributing to the glycerol production during osmotic stress [22]. Interestingly, *S. cerevisiae* does not produce glycerol when respiring on ethanol, but instead produces trehalose to combat osmotic stress, showing that the response is carbon source dependent [23].

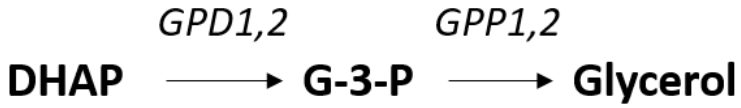

Figure 2.3: Glycerol production from DHAP in *S. cerevisiae*. DHAP: dihydroxyacetone phosphate; G-3-P: glycerol-3-phosphate; *GPD*: glycerol-3-phosphate dehydrogenase; *GPP*: glycerol-3-phosphatase

Y. lipolytica has a unique metabolism of glycerol, where G-3-P production competes with glycerol production, possible contributing to its high levels of TAG formed from G-3-P [15]. In contrast to *S. cerevisiae*, *Y. lipolytica* only possess one gene coding for Gpd and is lacking the *GPP* gene [15]. This could indicate that *Y. lipolytica* uses another route for glycerol production, or is using another compatible solute for protection against hyperosmotic stress. The use of a different compatible solute is supported by the observation that erythritol production can be increased by exposing *Y. lipolytica* to osmotic stress when growing on glycerol, indicating that erythritol is being used as a compatible solute in *Y. lipolytica* under these conditions [24].

2.2.2 High osmolarity glycerol pathway

The pathway which utilizes the most control over glycerol accumulation in *S. cerevisiae* is the high osmolarity glycerol (HOG) mitogen-activated protein kinase (MAPK) pathway, and is summarized by Hohmann (2015) [22]. Briefly, it consists of two branches: the Sln1 and the Sho1 branch, see Figure 2.4. Sln1 is a sensor histidine kinase, located in the plasma membrane, that acts as a negative regulator. Sln1 is active and autophosphorylated during normal conditions and keeps Ssk1 inactive by phosphorylation via the phosphotransfer protein Ypd1. During osmotic stress, Sln1 is inactivated and the inhibition of Ssk1 is relieved, causing Ssk1 to activate Hog1 through a phosphorylation cascade involving Ssk2, Ssk22 and Pbs2. *Sln1*$\triangle$ or *ypd1*$\triangle$ is lethal as the HOG pathway then never turns off [22].

The Sho1 branch is a positive regulator of HOG-pathway and is thought to sense osmotic changes through the membrane proteins Hkr1 and Msb2. During osmotic stress, Hkr1 and Msb2 interact with the membrane scaffold protein Sho1 and the G-protein Cdc42, causing activation of Hog1 through a phosphorylation cascade involving Ste20-Cla4, Ste11 and Pbs2 [22].

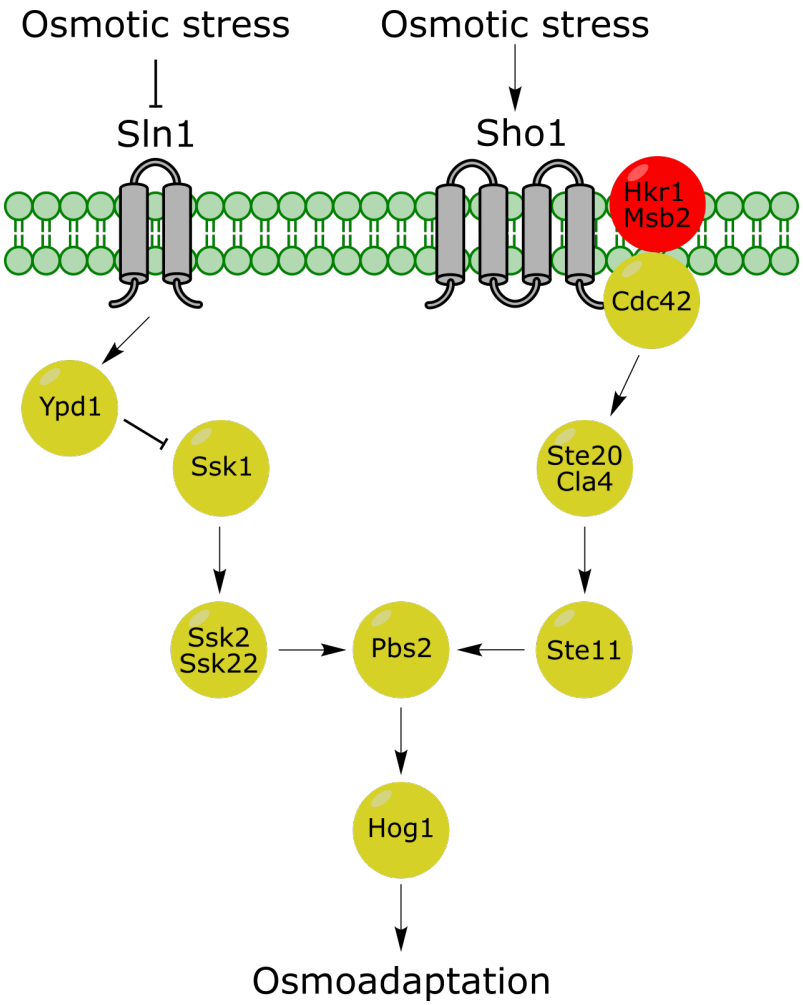

Figure 2.4: Sln1 and Sho1 branch of the HOG-pathway in *S. cerevisiae*. The activity of Sln1 is inhibited by osmotic stress, resulting in Hog1 activation through a phosphorylation cascade, which results in osmoadaptation through increased glycerol production and uptake, and reduced water loss. The Sho1 branch is activated by osmotic stress and leads to phosphorylation cascade, which activates Hog1 in the end. Both branches are connected through Pbs2. Picture adapted from [22]

Activated Hog1 in glucose conditions has targets both in the cytosol and the nucleus, resulting in up-regulation of *GPD1* and *GPP2* for increased glycerol production and up-regulation of *STL1* for increased glycerol uptake. Hog1 also increases the activity of Gpd1 and Pfk26 for increased glycolytic flux towards glycerol. Activated Hog1 also limits the glycerol efflux by inhibiting the glycerol transporter Fps1. Furthermore, water loss in limited by down-regulation of *AQY2*, coding for an aquaporin [22]. Hog1 is still activated during osmotic stress when grown on ethanol, even though *S. cerevisiae* does not accumulation glycerol in those conditions [23].

Y. lipolytica has predicted homologues for both the Sln1 and Sho1 branch, however, Hkr1 and Msb2 is not yet predicted to exist in *Y. lipolytica* [25]. *Ste11*$\triangle$ in *Y. lipolytica* can still handle osmotic stress, indicating that other pathways might active during osmotic stress, e.g. the Sln1 branch [26].

2.2.3 Ionic stress

Apart from causing osmotic stress, NaCl also induces an ionic stress as Na^+ ions replace K^+ ions in biomolecules, inducing toxicity [19]. To cope with this, cells living in environments with high levels of NaCl must spend a lot of energy to maintain low levels of cytosolic Na^+, while keeping sufficient amounts of K^+ [27]. This is controlled by restricting uptake of Na^+, high efflux of surplus Na^+ and compartmentalization of ions into organelles [27].

The efflux of Na^+ in *S. cerevisiae* is mostly achieved by two different systems in the plasma membrane: the Ena Na^+ - ATPases and the Nha1 Na^+/H^+ antiporter [27]. The *ENA* genes are transcribed from the *ENA/PMR2* locus, which contains five copies for *S. cerevisiae* S288C (*ENA1-5*) and only one single allele (*ENA6*) for *S. cerevisiae* CEN.PK stains [28]. This difference is responsible for the hypersensitivity of CEN.PK stains towards Na^+, potentially caused by low transcript levels and/or impaired kinetic properties of *ENA6*, resulting in insufficient efflux of Na^+ [28]. All sequenced yeast species contain *ENA* homologues [29].

While the Ena transporters are the most important for detoxification of Na^+ during higher pH, the Nha1 Na^+/H^+ antiporter is responsible detoxification at acidic pH [27]. *Y. lipolytica* possesses two different Nha transporters, where *Yl*Nha1 is responsible for efflux of K^+ and *Yl*Nha2 for efflux of Na^+ [30]. The HOG-pathway is coupled to ion stress resistance, as Hog1 controls the activity of Nha1 and induces the expression of *ENA1* through downstream targets [27].

2.2.4 Tolerance

Y. lipolytica has previously been shown to tolerate concentrations of NaCl up to 2 M [30] compared to *S. cerevisiae*, which tolerates concentrations up to 1.5 M [27]. The difference could partly be explained by the overrepresentation of ion transporter in genome of *Y. lipolytica* compared to other yeast species [31].

2.3 Adaptive laboratory evolution

Even though yeast cells are evolved to adapt and survive stressful environments [32], it would be interesting to find out if this capability can be improved for better performance at high osmotic pressure, using NaCl as the stressor. This could be achieved through adaptive laboratory evolution (ALE). In an ALE of microorganisms, the course of evolution is guided by the scientist by the choice of selective pressure. The cells are grown in a condition that is either stressful or contains nutrients that the organisms cannot utilize fully, resulting in a slower growth. By prolonged exposure to these conditions, there is a chance a cell acquires a beneficial mutation, which could result in a higher growth rate than the other cells in the population. The increased growth rate will allow the mutant to outcompete the cells lacking the beneficial mutation and take over the population [33].

At the end of the ALE, genomic DNA of the evolved strains are isolated and whole genomes sequenced using next-generation DNA sequencing (NGS) techniques. By evolving several populations in parallel, shared deviations in genomic sequence from the wild type could be potential candidate mutations that cause the improved phenotype. To verify the genotype-phenotype connection, the identified mutations are reconstructed through

genetic engineering in wild type strains and its phenotype compared with evolved strains. Potential findings could possible include improvements to the mechanisms of osmotic and ion stress tolerance.

In section 2.3.1 - 2.3.3 three previous studies using ALE for adaptation of *S. cerevisiae* to environmental stress are summarized to demonstrate what can be achieved using ALE.

2.3.1 Temperature

Three populations of *S. cerevisiae* CEN.PK113-7D were evolved by Caspeta et al. (2014) for >300 generations (90 days) at 39.5 ± 0.3 °C in minimal media with daily dilutions (after 4-7 generations) into fresh medium. The cultures were diluted 6-10 times to obtain an optical density (OD_{600}) of 0.2. The final evolved strains showed an average 1.57 ± 0.11 times increase in specific growth rate (μ) at 40 °C compared to wild type [34].

After whole genome sequencing, apart from a lot of different mutations, one nonsense mutation in *ERG3* was found in all lineages. This mutation created one premature stop codon for each strain at position Gln^{50}, Tyr^{66} or Tyr^{185} respectively. This caused the membrane sterol composition to change from ergosterol to fecosterol, which is a more bent sterol. This could have an important role for membrane fluidity at elevated temperatures, which also have been seen in thermophiles [34]. Reconstruction of the stop codon at Tyr^{185} in wild type showed up to 86 % of the specific growth rate compared to evolved strains [34].

2.3.2 pH

Fletcher et al. (2017) evolved *S. cerevisiae* CEN.PK113-7D in three different setups for three different kinds of low pH stress, using five replicate shake flasks for each. The ALE was conducted in minimal medium, with glucose as the carbon source, using either the inorganic acid HCl (Glu-HCl) or the organic acid lactic acid (Glu-LA) for the low pH stress. A third setup also contained lactic acid, but used raffinose as the carbon source (Raf-LA) to be able to isolate adaptations specific to lactic acid by comparing the results of the two evolved cell lines [35].

The three different ALE started at pH 4, with 88 mM lactic acid for Glu-LA and Raf-LA, and was conducted for 10 weeks. One colony from each population was isolated and characterized further. After 281 generations, the five Glu-HCl strains were down to pH 2.8 and showed 31.3 % increase in maximum average growth rate. Glu-LA evolved for 312 generations and the 5 strains showed 200 % increase in maximum average growth rate at pH 2.8 with 0.3 M lactic acid. The five strains of Raf-LA showed 11.5 % increase in maximum average growth rate at pH 3.2 with 0.2 M lactic acid after 277 generations of ALE [35].

Whole genomic sequencing showed that disruption of ergosterol biosynthesis (*ERG5*) and iron update (*FRE1*) was share between all five isolated strains from Glu-HCl. Construction of *erg5*△ and *fre1*△ resulted in an improved growth rate of 20 % and 34 % respectively compared to the parental strain during batch fermentation in bioreactors. However, the best performing evolved strain showed an increase of approximate 56 % in growth rate, indicating that other factors contribute to the phenotype of the evolved strain [35].

Four of the lactic acid evolution (two for each carbon source) showed adaptation by formation of clumps. This was a result from defect transcriptional activator (Ace2), which activates the expression of genes involved in separation of mother and daughter cell during cell division. Construction of *ace2*△ resulted in 5.3-fold increase in growth rate compared to wild type during batch fermentation in bioreactors. The evolved strain had growth rate of roughly three times higher than the *ace2*△ strain, also indicating that other factors are important for the evolved phenotype [35].

2.3.3 Saline

S. cerevisiae BY4741 was evolved as 3 replicate cell lines in complex medium containing 2 % peptone, 1 % yeast extract, 2 % galactose and 0.5 M NaCl (YPGN) by Dhar et al. (2011) for 300 generations. For each serial transfer the culture was diluted 1000 times into fresh medium. The transfers were performed every 24 hours (10 generations) and was repeated 30 times, resulting in an ALE of 30 days [36].

The evolved lines showed 8-12 % increased growth rate in YPGN, along with a smaller increase in growth rate in YPG (without NaCl). After whole genome sequencing of the three populations and the ancestral strain, the only common change was found to be a mutation in population two that caused the amino acid change G230D in *MOT2*. *MOT2* is involved in transcriptional regulation and post-translational modifications. However, reconstruction of this mutation only explained 25.3 % of the increased fitness for population two. Furthermore, reconstruction of G230D in *MOT2* also resulted in a fitness increase without NaCl [36].

2.4 Genetic tools

To be able to reconstruct and validate the genetic changes found in the evolved strains, genetic tools are required. The downside with using *Y. lipolytica* is the high frequency of non-homologous end-joining (NHEJ) mechanism for repair of double stranded breaks (DSB) of DNA, resulting in random integration of exogenous DNA rather than site specific integrations through homologous recombination (HR) repair [37]. Furthermore, the use of NHEJ for DSB DNA repair results in nonspecific insertions or deletions [38]. This can, however, be avoided by knocking out *KU70*, which is important for NHEJ mediated DNA repair, and thus forcing the cells to use HR instead of NHEJ [39]. Knocking out *KU70* did not result in any growth effects on *Y. lipolytica* [40], but have been reported to cause negative growth effects and sensitivity to UV-light in *Pichia pastoris* [41]. The strains used in this study have previously been made *KU70* deficient by integration of the hygromycin B (hph) resistance cassette into the *KU70* locus [42].

2.4.1 CRISPR/Cas9

An efficient genetic tool to construct genetic changes is the Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)/CRISPR-associated protein 9 (Cas9) system. CRISPR/Cas9 is a two-component system, based on a system used by bacteria and archaea for immunity against viral DNA, that has been exploited to be used in genetic engineering for precise modifications [43]. The endonuclease Cas9 is guided by a synthetic single guide RNA (sgRNA) to a 20 bp region of the chromosome complementary to a part of the sgRNA sequence and makes a double stranded cut [43]. Furthermore, the 20 bp region needs to be followed by a protospacer adjacent motif (PAM) for Cas9 to be able to

cut [43]. This cut needs to be repaired so the cells can survive and outcome can be guided by the design of the construct. If no homologous DNA fragments exists, the cell will choose NHEJ for the DNA repair, which could result in gene disruption by the mutation that arise in the process [44]. Another possibility is to repair the cut by HR, which is assisted by providing the cells with homologous DNA fragments, resulting in either deletions or insertions depending on the design [44].

For efficient gene editing, the intracellular levels of sgRNA needs to be high enough and properly folded [45]. If the sgRNA is incorrectly processed from the primary transcript, the Cas9 can be targeted to the wrong site as the recognition site for Cas9 is 5′-N20-NGG-3′, where N20 matched the 5′ end of the sgRNA and NGG is the PAM [44]. This can be avoided by, for example, using promoters with known transcription start sites [46], flanking sgRNA with ribozymes (catalytic RNA) [47] or exploiting the endogenous tRNA processing machinery [48].

Transformed *Y. lipolytica* should be allowed 2-4 days of outgrowth in selective medium to allow efficient expression and utilization of the CRISPR system [44]. The disruption efficiency is very low without the outgrowth step [38]. Due to limits of usable markers in *Y. lipolytica*, the one plasmid design from Gao et al. (2016) was chosen instead of the two-plasmid design by Schwartz et al. (2016).

3
Methods

3.1 Strains, media and culture conditions

The *Y. lipolytica* strains used in this study are shown in Table 3.1 and are all derived from *Y. lipolytica* W29 [42]. FKP355 is a leucine auxotroph and will be used for reverse engineering. The ALE will be conducted on the strain FKP391, which is a prototrophic strain, to avoid the need to add leucine to the medium during the ALE. T67C is a delft adapted control kindly provided by Xiaojun Ji. JHYl001 is constructed during this study to be sensitive towards hygromycin B, later to be used as a marker for further genetic engineering. All strains were cryopreserved in 25 % sterile glycerol at -80 °C, using a cooling unit for slow freezing.

Table 3.1: Strains

Strains	Characteristics	Reference
FKP355	*matA, xpr2-332, axp-2, ku70::hph+, leu2-270*	[42]
FKP391	*matA, xpr2-332, axp-2, ku70::hph+, leu2-270::leu2+*	[42]
T67C	Delft adapted FKP391	Xiaojun Ji (unpublished)
JHYl001	*matA, xpr2-332, axp-2, ku70△, leu2-270*	This work

The minimal (Delft) medium used in this study contained 7.5 g $(NH_4)_2SO_4$, 14.4 g KH_2PO_4, 0.5 g $MgSO_4 \bullet 7H_2O$, 20 g glucose, 1 mL vitamin stock (50 mg/L D-Biotin, 1 g/L D-Pantothenic acid hemicalcium salt, 1 g/L Thiamin-HCl, 1 g/L Pyridoxin-HCl, 1 g/L, Nicotinic acid, 200 mg/L 4-aminobenzoic acid, 25 g/L m-Inositol, pH = 6.5) and 1 mL trace metal solution (3 g/L $FeSO_4 \bullet 7H_2O$, 4.5 g/L $ZnSO_4 \bullet 7H_2O$, 4.5 g/L $CaCl_2 \bullet 2H_2O$, 1 g/L $MnCl_2 \bullet 4H_2O$, 300 mg/L $CoCl_2 \bullet 6H_2O$, 300 mg/L $CuSO_4 \bullet 5H_2O$, 400 mg/L $Na_2MoO_4 \bullet 2H_2O$, 1 g/L H_3BO_3, 100 mg/L KI, 19 g/L $Na_2EDTA \bullet 2H_20$, pH = 4) per liter.

The complex medium used for yeast cultivation was yeast-peptone-dextrose (YPD) medium, which contains 10 g/L yeast extract, 20 g/L peptone from meat and 20 g/L glucose. *E. coli* DH5α was grown in lysogeny broth (LB), containing 10 g/L peptone from casein (tryptone), 10 g/L NaCl and 5 g/L yeast extract. YPD and LB plates contained 20 and 16 g/L agar, respectively.

Transformed *E. coli* was selected on LB with 100 µg/mL ampicillin (amp) plates. Synthetic complete (SC) dropout plates without leucine was used for selection of *Y. lipolytica* and contained 6.9 g/L yeast nitrogen base (YNB) without amino acids, 0.69 g/L complete supplement mix (CSM) drop out without leucine, 20 g/L glucose and 20 g/L agar. Agar and glucose was autoclaved separately and the rest sterile filtered.

All cultivations were performed, unless specified, in 100 mL shake flasks containing 20 mL medium on rotary shakers at 200 revolutions per minute (rpm) at 30 °C. OD_{600} was measured using an GENESYS™ 20 Visible Spectrophotometer (Thermo Ficher Scientific) and samples diluted to match the linear range (0.05 - 0.3) when necessary. Technical replicates in this study are from the same preculture and biological replicates from different precultures.

3.2 Pilot cultivation

Several cultivation trials were performed to decide the starting stress level for the ALE. The preculture for all pilot trials were grown overnight in 50 mL falcon tubes containing 5 mL YPD medium before washed and resuspended in the final medium and diluted to OD_{600} of 0.1. The concentrations used in the pilot trial ranged from 0.6 - 2 M NaCl.

3.3 Adaptive laboratory evolution

One colony of FKP391 was grown overnight in a 50 mL falcon tube containing 5 mL YPD medium. The overnight culture was washed twice with Delft medium containing 1.2 M NaCl before diluted to an OD_{600} of 0.1 in 5 shake flasks with the final culture volume of 20 mL Delft medium containing 1.2 M NaCl. The cultures were diluted every 3-4 generation (1-1.5 days) to an OD_{600} of 0.05-0.1 in fresh medium. The old set of populations was kept on shaker as a backup until next dilution, in case the newly diluted culture did not grow. The number of generations was calculated as $\log_2(\text{final } OD_{600}/\text{initial } OD_{600})$, assuming exponential growth between every dilution. The stress level was increased to 1.4 M NaCl after 10 generations.

The evolution was paused after 100 generations, with the populations cryopreserved in -80 °C for 2 weeks. The populations were recovered in Delft medium containing 0.7 M NaCl for a few generations before the stress level was increased back to 1.4 M NaCl. The evolution continued in Delft medium containing 1.4 M NaCl until the total number of generations reached 220 and all 5 populations cryopreserved at -80 °C.

The purity of the populations during the ALE was ensured by weekly microscopically checks, colony PCR (cPCR) of Internal Transcribed Sequence (ITS) region [49] (after genomic extraction according to section B.1) and cryopreservation of half the population. The ITS PCR products of the final evolved populations were sent for sequencing for additional verification.

3.4 Evaluation of strains

From each evolved population, 10 colonies were isolated by plating on YPD with 1.2 M NaCl plates. These 50 colonies were characterized in the Bioscreen (section 3.4.1) to select the 3 best performing colonies from each population. 5 colonies from each population was cryopreserved at -80 °C after overnight growth in YPD. Some of the best performing strains were also characterized in shake flasks (section 3.4.2) and checked for lipid accumulation (section 3.4.3).

3.4.1 Bioscreen

The Microbiology Reader Bioscreen C (Oy Growth Curves Ab Ltd) was used for high throughput screening of growth characteristics in multiwell plates with the following shaking settings: continuous shaking, maximum amplitude and fast speed. The temperature was kept at 30 °C and automatic measurements taken every 15 minutes. All raw data was blank corrected (mean blank of that time point) and corrected according to equation 3.1 from Warringer et al. (2003) to adjust for the non-linearity at higher cell densities [50]

$$OD_{cor} = OD_{obs} + OD_{obs}^2 \times 0.449 + OD_{obs}^3 \times 0.191 \tag{3.1}$$

where OD_{cor} is the corrected optical density and OD_{obs} the observed optical density.

OD_{cor} was converted to OD_{600} with the equation 3.2 by adjusting for the path length inside the bioscreen wells

$$OD_{600} = \frac{OD_{cor}}{1.32 \times \frac{V}{r^2\pi}} \tag{3.2}$$

where V is the volume of the well (mL) and r is the radius of the well (cm).

In the first round of screening, 10 colonies from each population and wild type FKP391 was inoculated from YPD with 1.2 M NaCl plates into the multiwell plate containing 145 µL Delft medium containing 1.4 M NaCl, using 2 replicates, according to Table A.3 - A.4. The first screening was conducted for 135 hours (≈ 5.5 days).

The five best performing colonies of each population was selected for the second round bioscreen. This time precultures were made in Delft medium containing 1.4 M NaCl and grown for 2 days before spun down, resuspended in fresh medium, diluted to an OD_{600} of 0.1 and loaded 145 µL into the multiwell plates as 5 technical replicates, according to Table A.5 - A.6. Furthermore, two biological replicates were used for the FKP391 control.

3.4.2 Shake flask characterization

Some of the best performing strains was also characterized in shake flasks as the aeration and stirring is not optimal in multiwell cultivation settings. Colonies from YPD plates were grown for 1 day in Delft medium, spun down (3000 g, 5 min) and resuspended in medium appropriate for that trial. The conditions tested is shown in Table 3.2.

Table 3.2: Concentrations of the 5 different stressors tested during the shake flask characterization

NaCl (M)	Sorbitol (M)	LiOAc (mM)	Low pH	Temperature (°C)
1.4	2.5	50, 200	2.15, 2.27	35.6

The pH and temperature levels were chosen from the end point levels of Xiaojun Ji's adaptive evolution experiment of *Y. lipolytica* FKP391 for low pH (pH = 2.15) and high temperature (35.6 °C) tolerance. The pH = 2.27 medium was kindly provided by Xiaojun Ji. 50 mM glycine was added to the low pH medium for increased buffering capacities.

3.4.3 Lipid accumulation

To make sure that the increased performance of the evolved strains in high salt medium was not gained at the cost of lost lipid accumulation ability, the evolved strains were tested for lipid accumulation with and without NaCl. This was done in nitrogen limiting (3.75 g/L $(NH_4)_2SO_4$) Delft medium to promote lipid accumulation. Precultures were grown in Delft medium for 1 day, spun down, resuspended in N-limited Delft medium, diluted to an OD_{600} of 0.1 in shake flasks and let grow until stationary phase (roughly 3 days). FKP391 grown in normal Delft medium for 1 day was used as negative control, as these conditions should not induce lipid accumulation.

100 - 300 µL samples (depending on OD_{600}) were washed with phosphate-buffered saline (PBS) and stained with 0.2 µL Nile Red for 15 min in the dark to allow measurement of lipid accumulation. The cells were spun down and resuspended in PBS and lipid accumulation was measured using a Fluorescent microscope by overlaying bright field pictures (TL-DIC, exposure 30 ms, gain 3.1, intensity 56) with pictures taken with YFP filter (FLUO, excitation filter 510/20 nm (500–520 nm), barrier filter 560/40 nm (540–580 nm), 30 ms exposure, 3.1 gain, intensity 1).

3.5 Whole genome sequencing

Two different methods for genomic extraction were used in this study: phenol:chloroform extraction for the strains and the Amresco provided kit for the populations. The genomic DNA of all 15 strains and wild type FKP391 was extracted by growing overnight in 50 mL falcon tubes containing 5 mL YPD, made into spheroplasts by zymolyase treatment, cell lysis with SDS and extraction of the polar phase twice with phenol:chloroform:isoamyl alcohol and once with chloroform. Extracted DNA was precipitated with 2.5 volumes of pure Ethanol, DNA pellet washed twice with cold 70 % ethanol and resuspended in 1/10 TE-buffer. A more detailed protocol can be found in section B.2.

Genomic DNA from the populations was extracted by inoculation of the cryostocks into shake flasks containing 20 mL Delft medium and cultivated for 36 hours before genomic DNA extraction using the yeast genomic DNA purification kit (Amresco, Solon, OH).

DNA concentration was determined with Qubit 3.0 Fluorometer (Thermo fisher Scientific), DNA purity with NanoDrop 2000 (Thermo Fisher Scientific) and DNA quality by electrophoresis (0.7 % agarose gel, GelRed, 1X TAE buffer, GeneRuler 1 kb DNA Ladder) and imaging with Gel Doc XR+ (Bio-Rad).

Genomic DNA was sent to Joint Genome Institute (Walnut Creek, CA, USA) for whole genome sequencing using Illumina sequencing.

3.6 Plasmid construction

Plasmids used in this study are shown in Table 3.3. All plasmids used are replicative in *Y. lipolytica* and express Cas9 and the sgRNA. *LEU2* is used for selection in *Y. lipolytica* and AmpR for *E. coli*. The purpose of these plasmids is set up a markerless genome editing tool, which will be utilized for the reverse engineering when genome data of the evolved strain arrives.

Table 3.3: CRISPR/Cas9 plasmids used in this study

Plasmids	Characteristics	Reference
pCAS1yl-trp	Pol II expressed guide, ribozyme processing, targets *TRP1*, NHEJ repair	[38]
pCRISPRyl	Pol III expressed guide, tRNA processing, no target	[44]
pCAS2yl-erg3	Pol II expressed guide, ribozyme processing, targets *ERG3*, HR repair	This work
pCAS3yl	Pol III expressed guide, tRNA processing, no target	This work
pCAS4yl-hph	Pol III expressed guide, tRNA processing, targets *HPH*, HR repair	This work

All primers used are shown in Table A.1 and was ordered from Eurofins Genomics. The CRISPR guides used are shown in Table A.2 and was generated with the Benchling (Benchling, Inc.), were high specificity scores was considered most important to avoid off-targets. Gibson Assembly design generated with NEBuilder (web tool, New England Biolabs) with overlaps of 30 bp between all fragments. The annealing part of the all primers was optimized by use of Tm Calculator (web tool, Thermo Fisher Scientific). Plasmids and their sequences were obtained from Addgene. The genome of *Y. lipolytica* W29 (GenBank: LJBI00000000.1) was used for design of primers for amplification of genomic DNA.

Plasmids were extracted using GeneJET Plasmid Miniprep Kit (Thermo Fisher Scientific). DNA fragments for the constructs were generated through polymerase chain reaction (PCR) using Phusion High-Fidelity DNA Polymerase. Both PCR products and restriction digested plasmids was verified on 1 % agarose gels in 1X Tris-acetate-EDTA (TAE) buffer with GeneRuler 1 kb DNA Ladder as reference. All fragments were purified with GeneJET Gel Extraction Kit (Thermo Fisher Scientific) prior the Gibson Assembly reaction. Enzymes and buffers was purchased from Thermo Fisher Scientific.

3.6.1 pCAS2yl-erg3

To test the CRISPR/Cas9 system from Gao et al. (2016) [38], it was designed to knock out *ERG3*, which might lead to increased temperature tolerance as it did in *S. cerevisiae* [34]. The sgRNA of pCAS1yl-trp was modified to target *ERG3*, along with the addition of 500 bp repair fragments to the plasmid, which was homologous to upstream and downstream of *ERG3*, to promote knock-out by HR. pCAS1yl-trp was cut with MssI and SalI to remove the old guide targeting *TRP1*. The guide was exchanged through two PCR reactions of the sgRNA casette: gRNA-erg3-up (upstream of the old guide) and gRNA-erg3-dw (downstream of the old guide), where the new guide was added as an overlap between both fragments, using pCAS1yl-trp as template. The repair fragments, erg3-up and erg3-dw, was amplified from genomic DNA of *Y. lipolytica* extracted according to section B.1. The four fragments were generated with corresponding primers in Table A.1 and were assembled with the cut vector with Gibson Assembly for 2 hours to generate pCAS2yl-erg3 (12 951 bp).

3.6.2 pCAS3yl

The casette expressing the sgRNA from a polymerase II promoter was exchanged to the hybrid SCR1-tRNA^{Gly} polymerase III promoter system used in the pCRISPRyl plasmid created by Schwartz et al. (2016). pCAS1yl-trp was cut with MssI and SalI to remove the old guide targeting *TRP1* and the hybrid polymerase III promoter extracted from pCRISPRyl through PCR with pol3 FW/RV primers (Table A.1). After Gibson Assembly

of the two fragments, the intermediate plasmid pCAS3yl was generated. This plasmid does not direct the expressed Cas9, but will be the basis of all future constructs.

3.6.3 pCAS4yl-hph

pCAS3yl was modified to knock out hygromycin B resistance gene (*HPH*) to be able to reuse that marker for further genetic engineering. This was done by exchanging the sgRNA and adding 1000 bp repair fragments homologous to upstream and downstream of *HPH*, using the same strategy as for pCAS2yl-erg3. pCAS3yl was cut with MssI and SalI. Four PCR fragments, hph-up and hph-dw (amplified from genomic DNA of *Y. lipolytica*), gRNA-hph-up (amplified from pCRISPRyl) and gRNA-hph-dw (amplified from pCAS3yl), were generated with corresponding primers in Table A.1. The four fragments were assembled with the cut vector with Gibson Assembly for 2 hours to generate pCAS4yl-hph (13 153 bp).

3.6.4 *E. coli* transformation

Assembled plasmids were transformed to chemically competent *E. coli* DH5α by heat shock. Frozen competent cells were thawed 20-30 min on ice, mixed with 1-5 µL plasmid DNA (>10 ng), placed on ice 20-30 more minutes, heat shocked 42 °C for 45 seconds, placed on ice for 3 minutes, addition of 250 µL LB medium, recovered for 45-100 min in a 37 °C air incubator at 200 rpm, plated at LB+amp plates and grown overnight at 37 °C. Transformants were isolated and plasmids extracted. The plasmids were verified by restriction digestion and sequencing (Eurofins Genomics) before transformation into *Y. lipolytica*.

3.6.5 Yeast transformation

The verified plasmids were transformed to *Y. lipolytica* FKP355 by electroporation according to the protocol in section B.3. Briefly, overnight cultures in YPD was diluted 10X in YPD, grown for a few hours and made competent by lithium acetate (LiOAc) and dithiothreitol (DTT) treatment. Electroporation was carried out at 1.5 kV with a pulse duration between 4-6 ms. The protocol was adapted to *Y. lipolytica* at the last step: the cells was resuspended in 200 µL after centrifugation, half plated on SC -LEU plates and the rest inoculated into 14 mL cultivation tubes containing 2 mL of SC -LEU liquid medium. The liquid medium was allowed 4 days of outgrowth before plating on SC -LEU plates and colony isolation.

Genomic DNA for colony PCR was extracted by resuspending a colony in 25 µL of 20 mM NaOH supplemented with a small spoon of 425-600 µm glass beads (Sigma G-8772), heated at 98 °C for 20 minutes, vortexed 15 seconds and cell debris spun down for 1 min with a table top centrifuge. Colony PCR was performed with DreamTaq polymerase with corresponding cPCR primers in Table A.1. Verified colonies were inoculated in YPD medium for 1 day outgrowth to remove the plasmid, allowing recycling of the marker for later genetic modifications. Outgrowth culture was plated on YPD plates, colonies isolated and verified for plasmid removal by streak out on SC -LEU plates, where they should not be able to grow without the plasmid.

4

Results

4.1 Pilot cultivation

Pilot cultivations were performed to test the performance of *Y. lipolytica* and decide which starting level of NaCl to use in the adaptive evolution experiment. The results of the pilot cultivation of *Y. lipolytica* FKP391 in Delft medium with different concentrations of NaCl are shown in Figure 4.1 and 4.2. Trials with 0.6, 0.7 and 0.8 M NaCl did not show strong effect on the growth: the final OD_{600} was 9.1, 7.7 and 5.9 respectively. The final OD_{600} for 1 and 1.2 M NaCl was 3.3 and 2.4 respectively. The higher concentrations resulted in a final OD_{600} of 2 (1.4 M), 1.5 (1.6 M), 0.82 (1.8 M) and 0.2 (2 M). Cultivation without stress reached a final OD_{600} of 10.6. The starting level of the evolution was chosen to be 1.2 M NaCl, as higher concentrations were thought to result in a too low growth rate for daily dilutions. After a few transfers (10 generations), the selective pressure of 1.2 M NaCl was decided to be too low and was increased to 1.4 M.

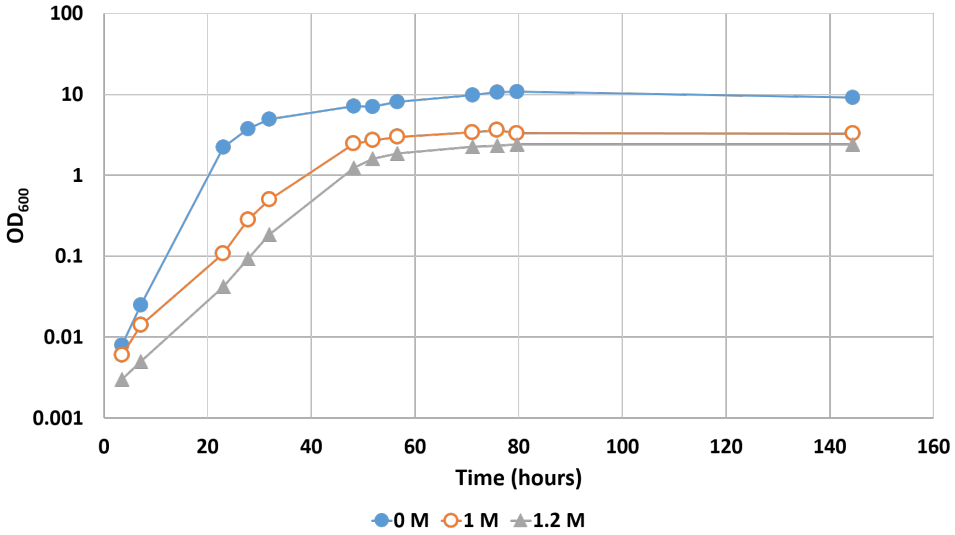

Figure 4.1: Growth of *Y. lipolytica* FKP391 in Delft medium supplied with 0, 1.0 and 1.2 M NaCl

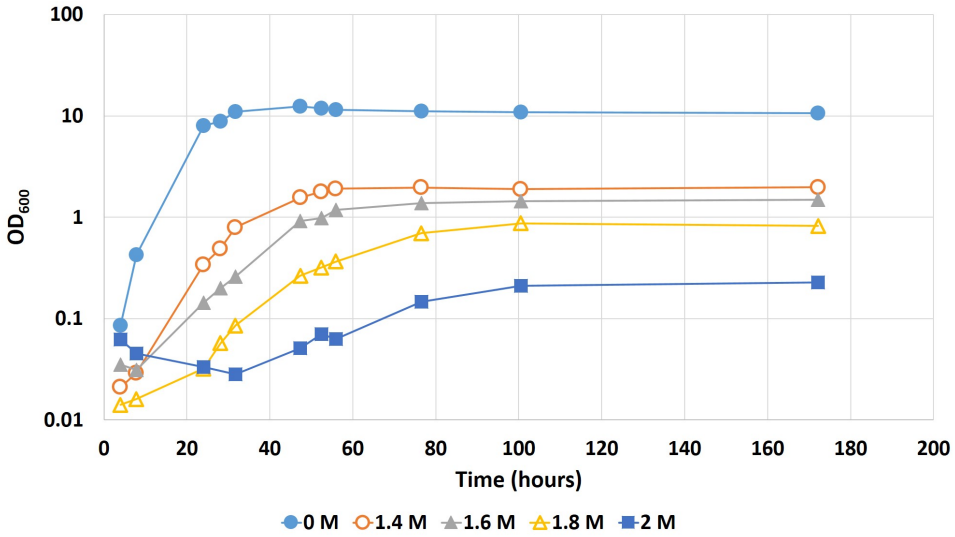

Figure 4.2: Growth of *Y. lipolytica* FKP391 in Delft medium supplied with 0, 1.4, 1.6, 1.8 or 2 M NaCl

4.2 Adaptive laboratory evolution

The course of the evolution in 1.4 M NaCl can be followed in Figure 4.3, which shows the changes in average growth rate as number of generations accumulates.

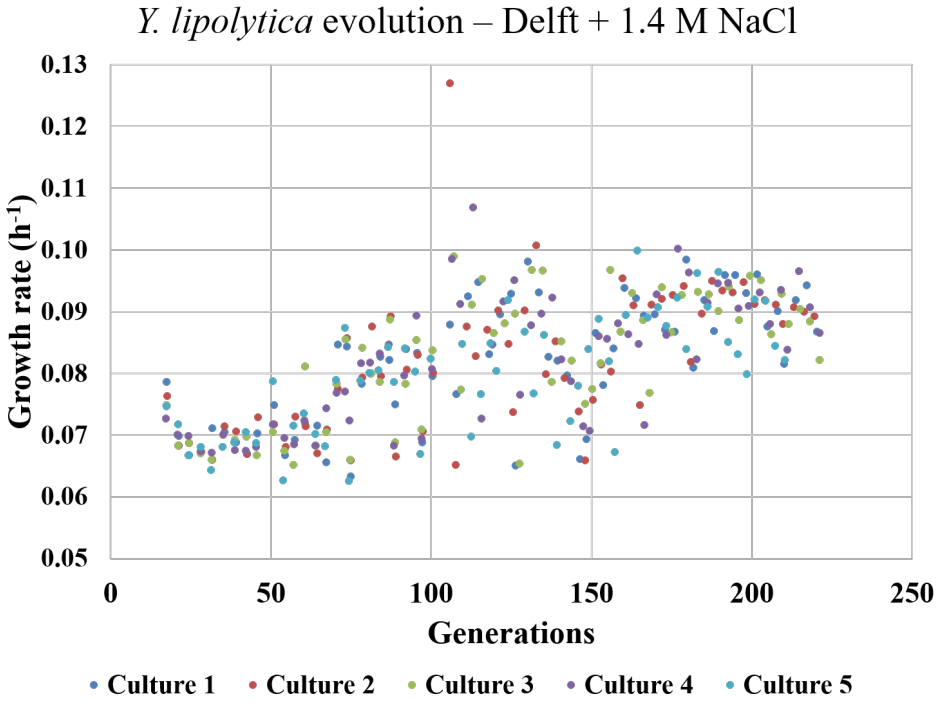

Figure 4.3: Changes in average growth rate for all five cultures during the adaptive evolution experiment in Delft + 1.4 M NaCl

The total improvements are summarized in Figure 4.4, which shows the increase of average growth rate for each of the five populations.

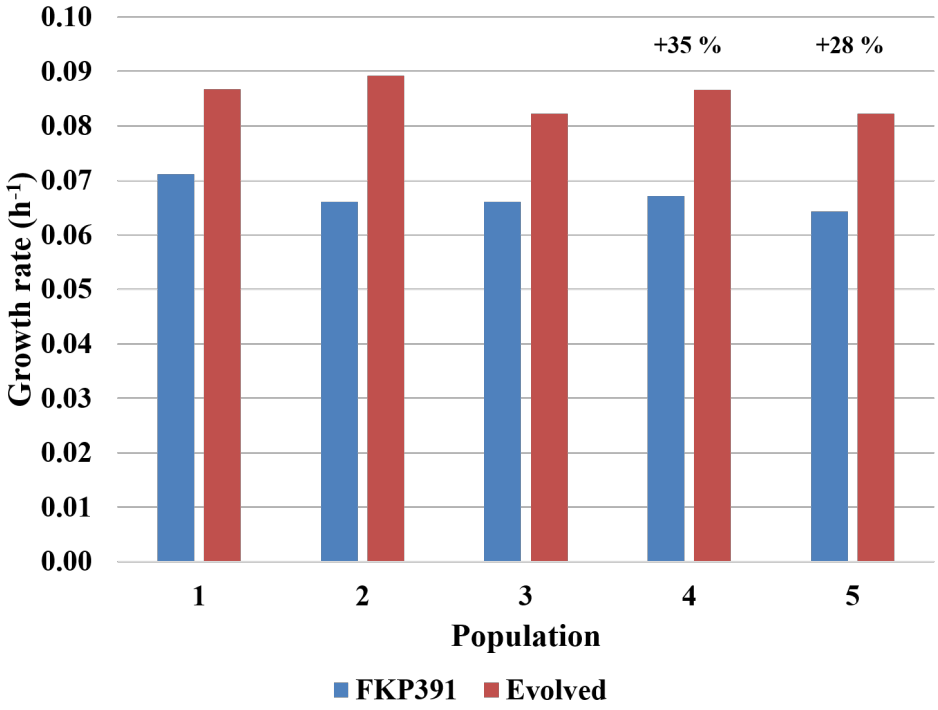

Figure 4.4: Improvement of average growth rate of each of the five populations after 220 generations of evolution in Delft + 1.4 M NaCl

4.3 Bioscreen

Ten colonies from each population were isolated and screened for performance in the Bioscreen to select the best performing ones for whole genome sequencing and further characterization. The results of the two rounds of Bioscreen are summarized in this section. The nomenclature of the strains in this report is the following: S.1.1 stands for saline evolved strain, population 1, colony 1. The strains were chosen both based on growth rate and final OD_{600}.

4.3.1 Round one

The results from the first round of bioscreen are summarized in Figure 4.5, where the final OD_{600} of the 10 strains from each population are shown. The highlighted strains were chosen for the second round of screening.

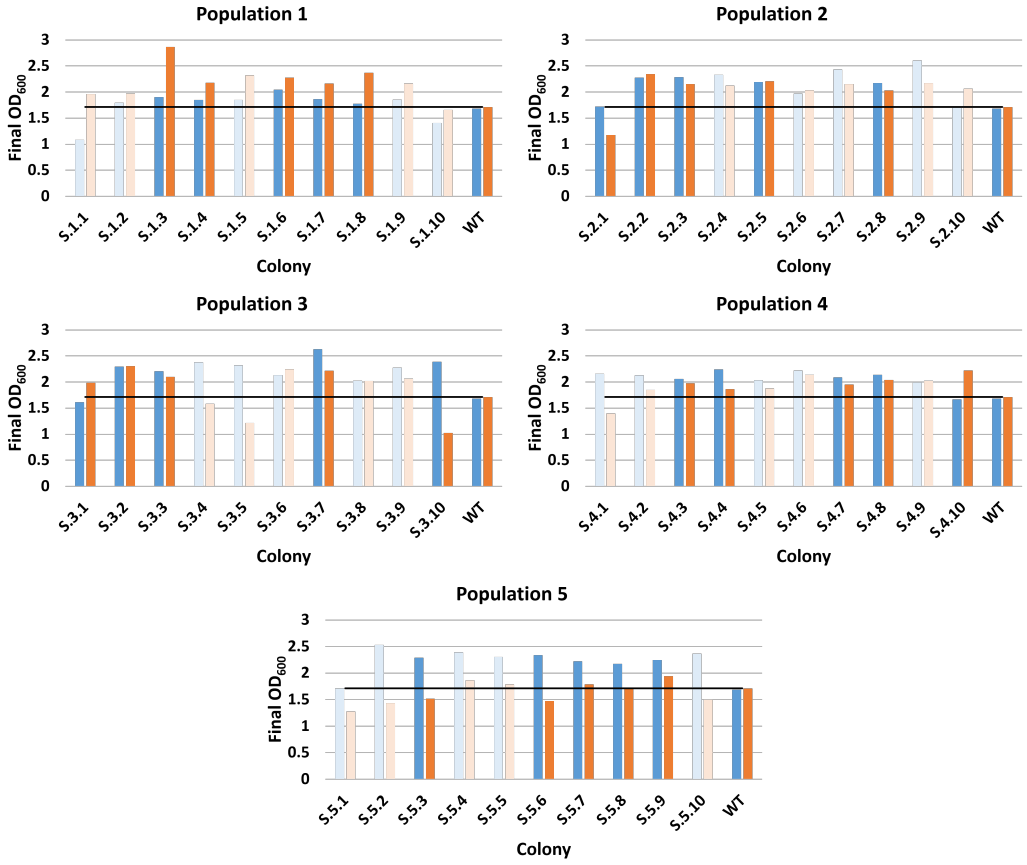

Figure 4.5: Final OD_{600} of each evolved strain in the first round of bioscreen, with two replicates (blue and orange). Highlighted bars correspond to colonies chosen for the second round of bioscreen. The final OD_{600} of the best wild type is marked with a line

The inconsistency of some of the biological replicates could be due to that the outer wells were used in the multiwell plate. Apparently, these wells evaporate a lot of liquid, which the inner wells do not because they are surrounded by wells containing liquid. This results in a higher concentration of NaCl in the outer well and thus higher inhibition of growth. By comparing the low performing strains with their location in the plate (Table A.3), one can see that most of them are located in the outer wells. This problem was uncovered long after the two bioscreen runs and there is a risk that a few high performing colonies located in the outer wells were discarded because of this.

4.3.2 Round two

The growth of some of the top performing strains in the second round of bioscreen are shown in Figure 4.6, with precultures done in Delft medium containing 1.4 M NaCl, with much better performance than the wild type. The upper and lower rows of wells yielded lower OD values than corresponding technical replicates. These wells were removed from the data analyzis, resulting in four technical replicates instead of five.

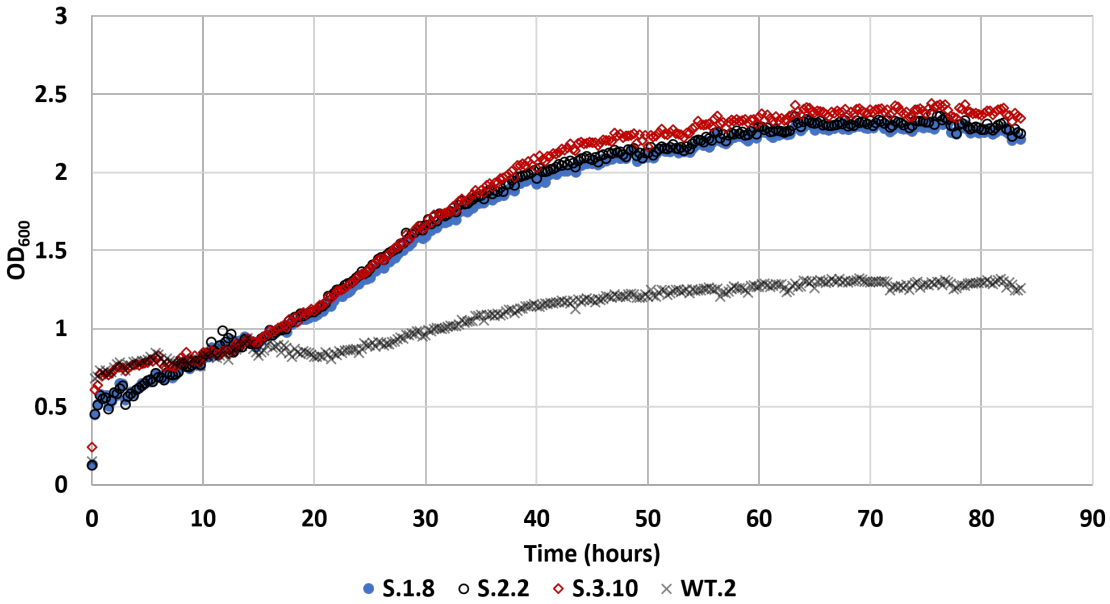

Figure 4.6: Growth of S.1.8, S.2.2, S.3.10 and wild type FKP391 in Delft + 1.4 M NaCl during the second round of bioscreen, using precultures in the same medium

The final OD_{600} for all 25 evolved strains and wild type FKP391 are shown in Figure 4.7. All tested strains performed better than the wild type. The best three colonies of each population were chosen for whole genome sequencing.

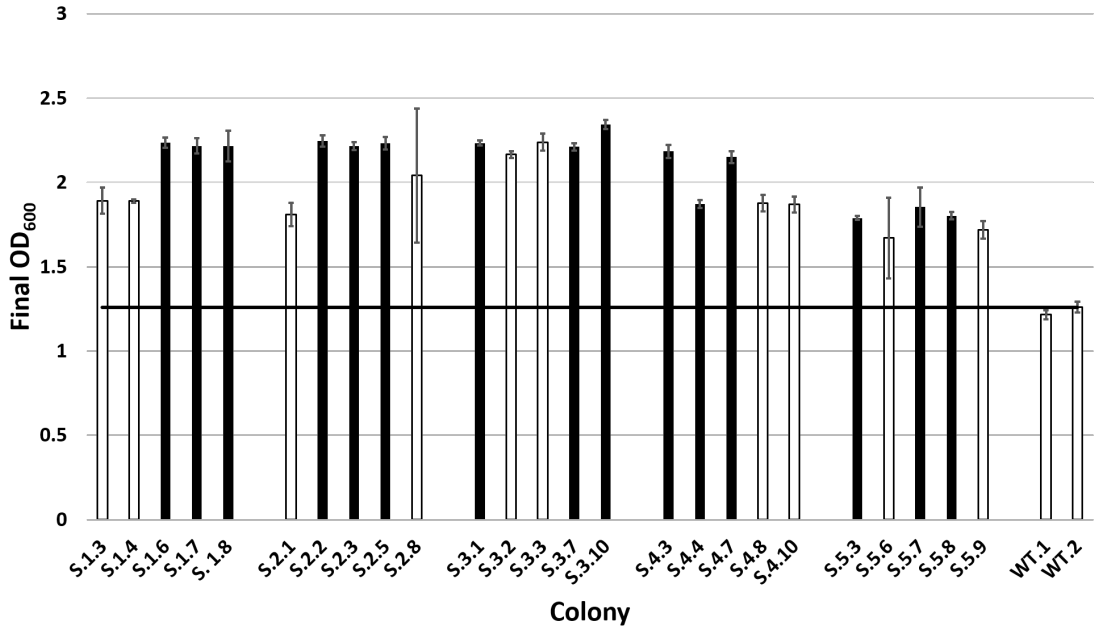

Figure 4.7: Final OD_{600} of each evolved strain in the second round bioscreen. Strains with black bars were chosen for genomic sequencing. The final OD_{600} of the best wild type is shown with a line. Error bars represent standard deviation from four technical replicates

4.4 Shake flasks characterization

In this section, the results from the shake flasks characterization in five different stress conditions (NaCl, sorbitol, low pH, lithium or high temperature) are summarized. The last four are tested to check for cross-tolerance, which could have evolved during the ALE. Shake flask cultivation results in a more representative characterization of the performance of the strains, as the drawbacks of the Bioscreen are sedimentation and reduced aeration.

4.4.1 NaCl

The growth of S.1.8 and wild type FKP391 can be followed in Figure 4.8 after precultured in Delft medium. There is no clear difference between the strains during the first 24 hours in stressful medium, however, the growth of the wild type is slowed after 40 hours, while the evolved strain continues to grow. In the end, the evolved strain obtained a 47 % higher OD_{600} than the wild type.

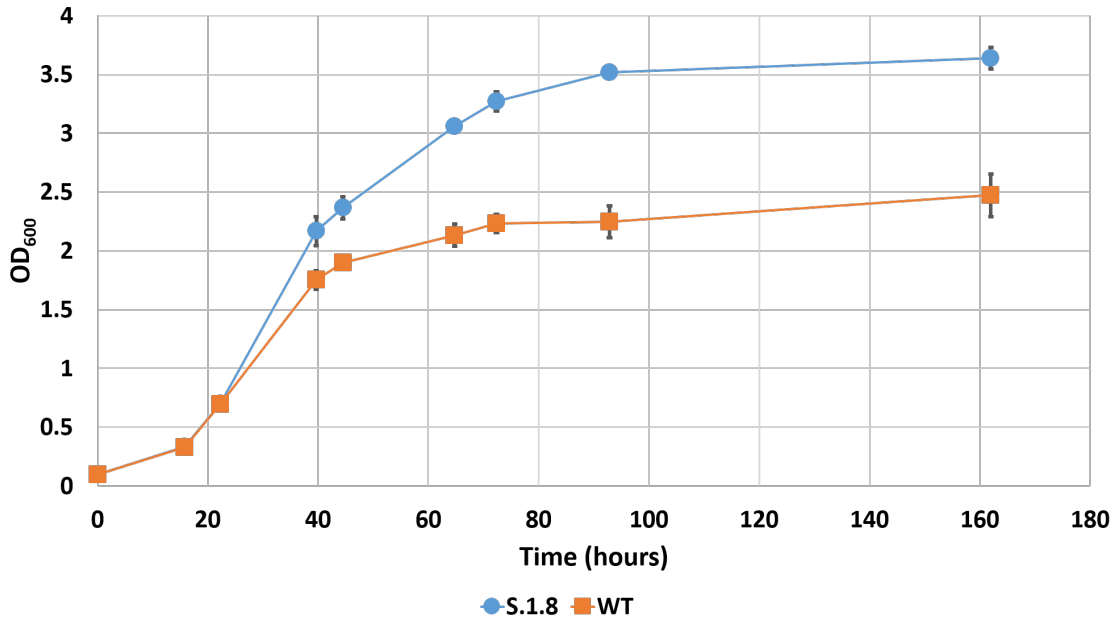

Figure 4.8: Growth S.1.8 and wild type FKP391 in Delft + 1.4 M NaCl. Error bars represent standard deviation from three biological replicates

The difference in morphology of the two strains in this trial are shown in Figure 4.9. S.1.8 shows more pseudohyphae formation, while wild type FKP391 shows hyphae and pseudohyphae formation.

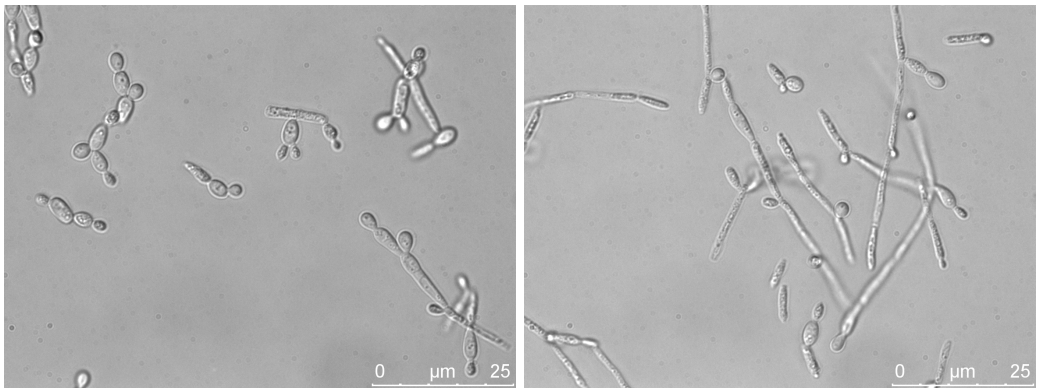

Figure 4.9: Microscope pictures of S.1.8 (left) and wild type FKP391 (right) after 3 days of growth in Delft + 1.4 M NaCl

The trial was repeated with two more evolved strains included. Furthermore, the Delft adapted strain T67C was included to exclude the possibility that the improved phenotype is due to Delft adaptation and not NaCl adaption. The results, shown in Figure 4.10, clearly indicates that the improved performance of the evolved strains is not only due to Delft adaptation, as the Delft adapted strain performed worse than the wild type.

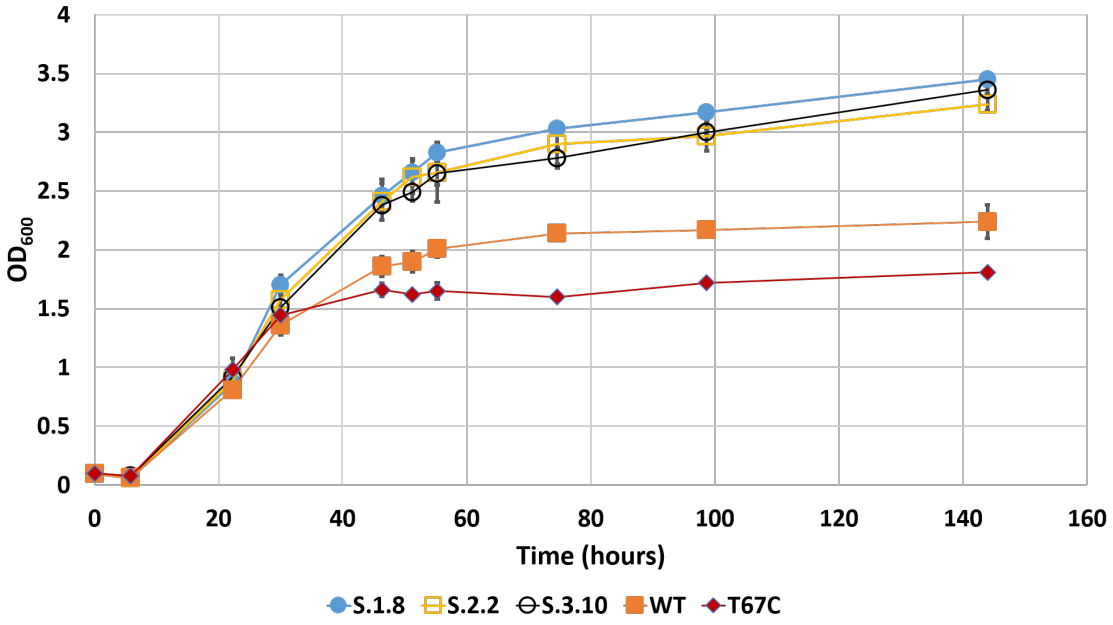

Figure 4.10: Growth of S.1.8, S.2.2, S.3.10, wild type FKP391 and Delft adapted T67C in Delft + 1.4 M NaCl. Error bars represent standard deviation from two technical replicates

The morphology of the strains in this trial can be seen in Figure 4.11, where all saline evolved strains show similar pseudohyphae morphology. The Delft adapted strain shows a more yeast-like phenotype.

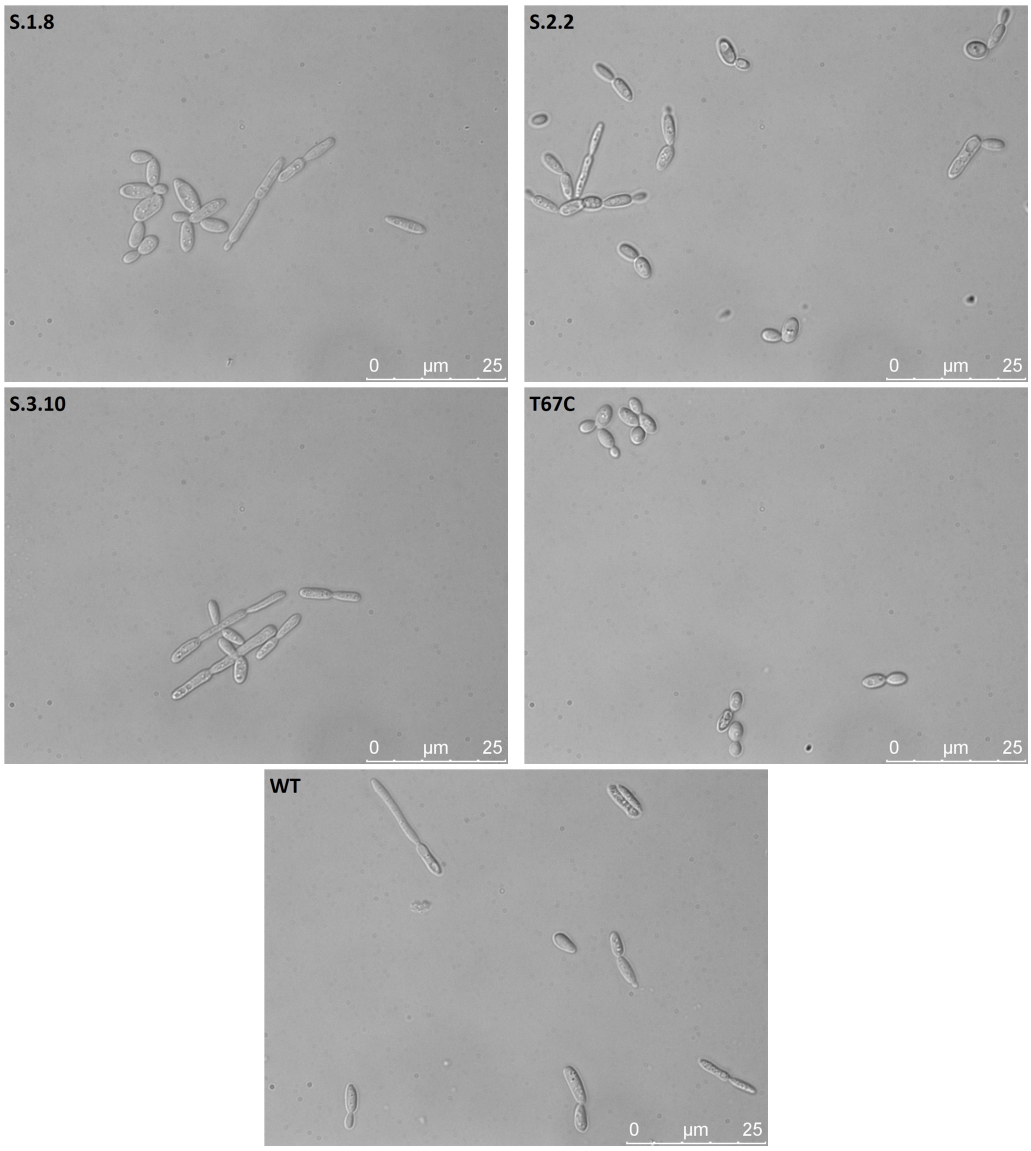

Figure 4.11: Microscope pictures of S.1.8, S.2.2, S.3.10, T67C and WT after 1 day of growth in Delft + 1.4 M NaCl

4.4.2 Sorbitol

NaCl induces both osmotic and ionic stress [19]. Therefore, it would be interesting to test the performance of the evolved strains in pure osmotic stress medium, with sorbitol as the stressor. The growth of three evolved strains and wild type FKP391 can be followed in Figure 4.12, after preculutred in Delft medium. It was clear that the strains evolved in 1.4 M NaCl did not show an improved phenotype towards pure osmotic stress compared to wild type.

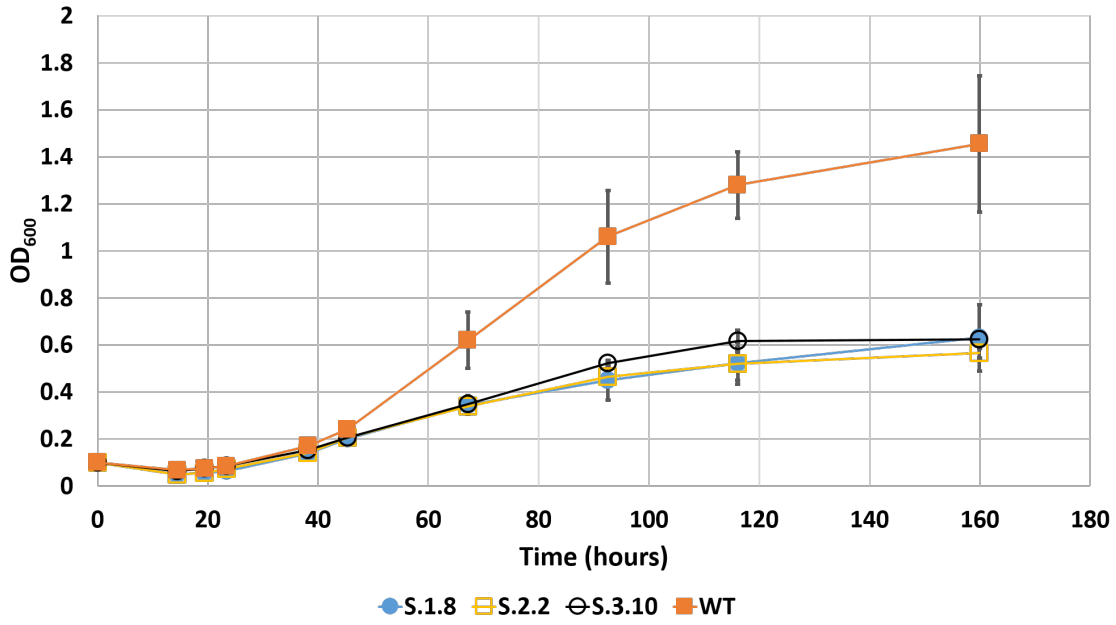

Figure 4.12: Growth of S.1.8, S.2.2, S.3.10 and wild type FKP391 in Delft + 2.5 M sorbitol. Error bars represent standard deviation from two technical replicates

4.4.3 Low pH

The next cross-tolerance tested of the evolved strain was tolerance towards low pH, which could have arisen during the evolution. Figure 4.13 shows the growth of three salt tolerant evolved strains and wild type FKP391 in low pH Delft medium provided by Xiaojun Ji. Interestingly, strain S.3.10 shows good performance at low pH. The purity of the cultures was verified by colony PCR with ITS primers.

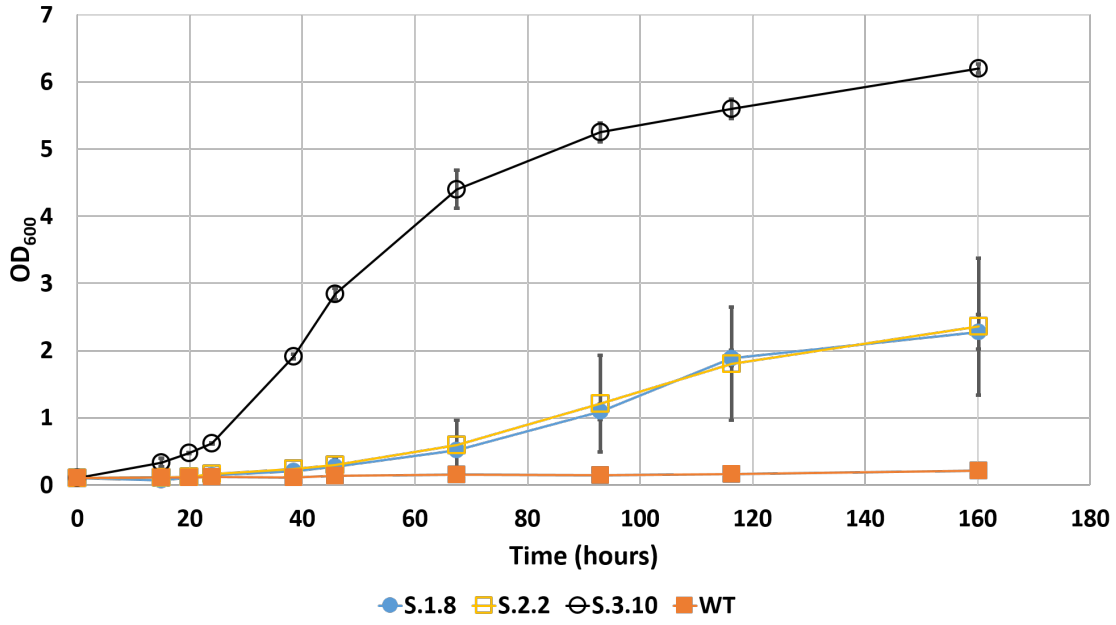

Figure 4.13: Growth of S.1.8, S.2.2, S.3.10 and wild type FKP391 in Delft (pH = 2.27). Error bars represent standard deviation from two technical replicates

The morphology of the strains during this trial is shown in Figure 4.14. All evolved strains formed clumps (some of them insoluble during OD_{600} measurements), which also was visible in the shake flasks during the cultivation. Some of the cells were stuck to the wall when the shake flasks were taken from the shaker, but they could be detached by rotary shaking of the flasks by hand.

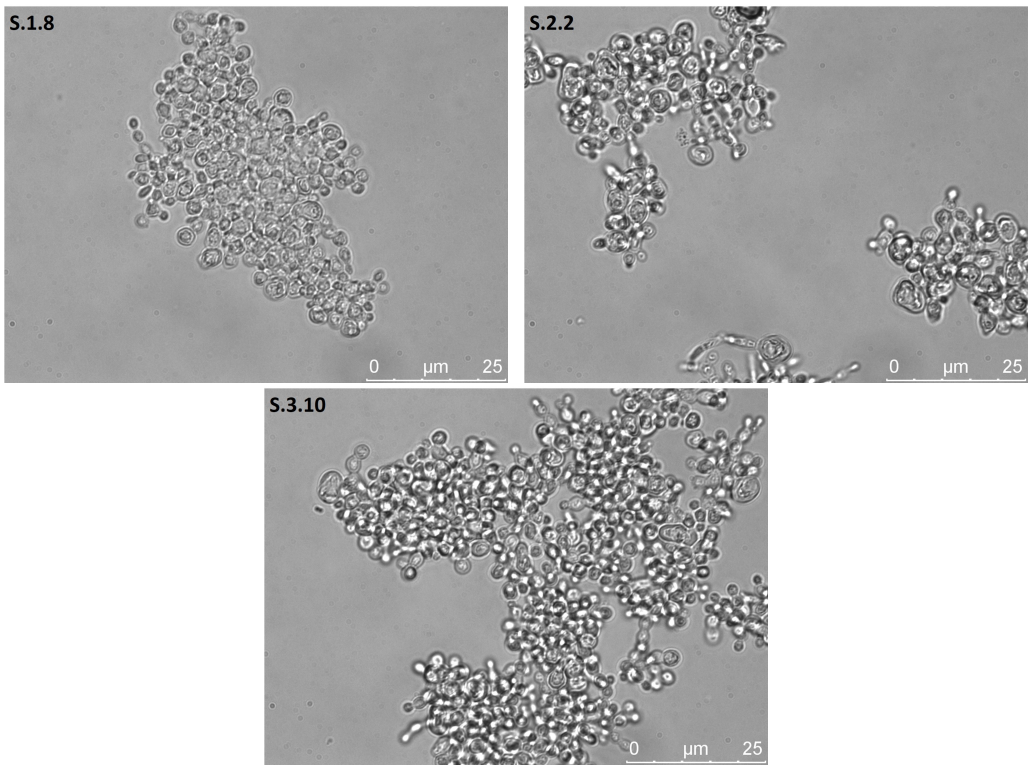

Figure 4.14: Microscope pictures of S.1.8, S.2.2 and S.3.10 in Delft (pH = 2.27) after 160 hours of cultivation

New medium was made and the trial repeated to verify the observed phenotype. The result of this growth trial can be seen in Figure 4.15. All strains showed lower performance, which could be due to the lower pH of this trial compared to the previous one, even though the difference of the pH is very small. The cultures formed even more insoluble clumps in this trial, which made OD measurements hard, resulting in the high standard deviation between the replicates. However, strain S.3.10 still shows a clear improved phenotype compared to the other strains in low pH medium.

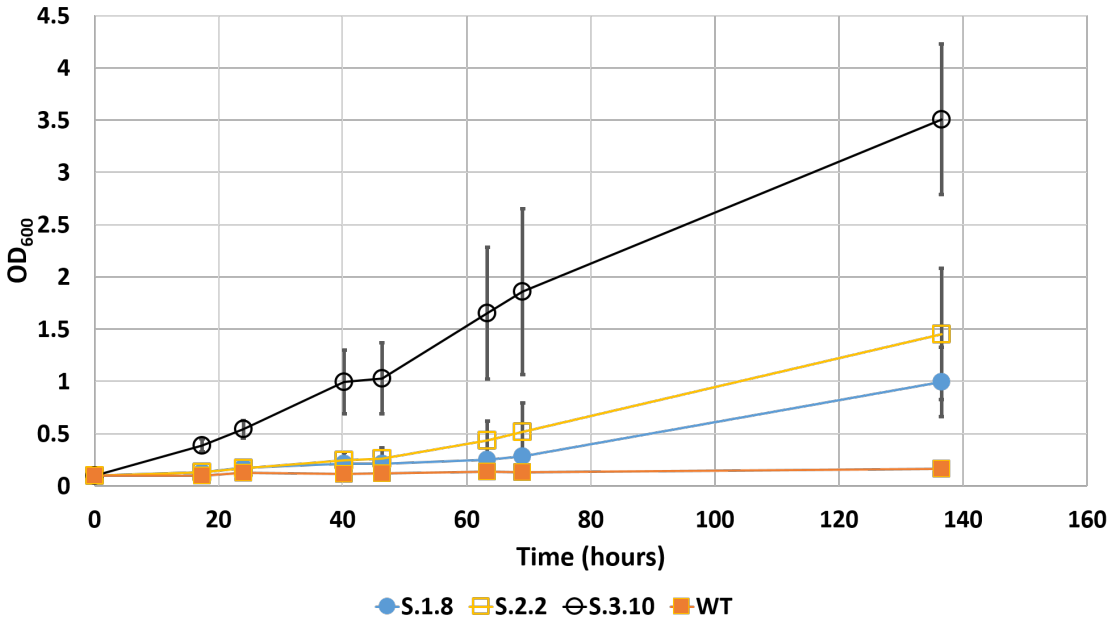

Figure 4.15: Growth of S.1.8, S.2.2, S.3.10 and wild type FKP391 in Delft (pH = 2.15). Error bars represent standard deviation from two technical replicates

The morphology of the tested strains is shown in Figure 4.16. Some of the S.3.10 clumps covered the entire screen, while wild type FKP391 did not seem to form any clumps at all (maybe due to very low number of cells).

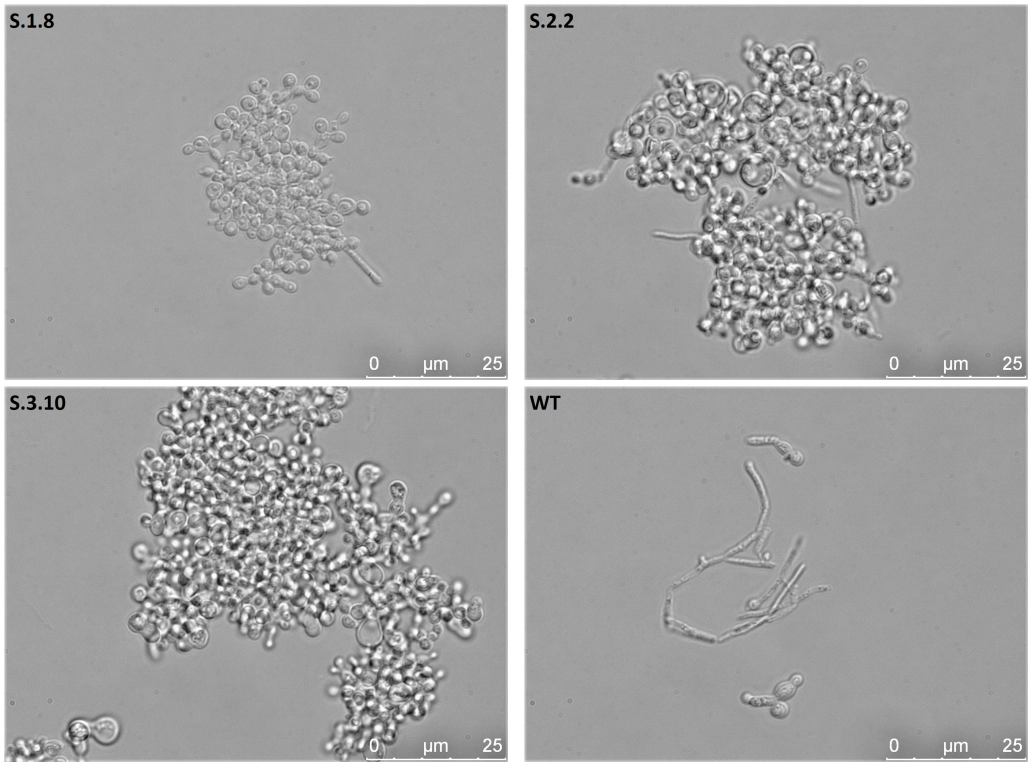

Figure 4.16: Microscope pictures of S.1.8, S.2.2, S.3.10 and wild type FKP391 in Delft (pH = 2.15) after 140 hours of cultivation

4.4.4 Lithium and high temperature

Cross-tolerance towards Li^+ was tested to investigate if evolved strains showed increased tolerance towards pure ionic stress. However, the trials with 0.2 M and 50 mM LiOAc did not result in any growth at all for S.1.8, S.2.2, S.3.10 or the wild type FKP391, as the concentration was too high. The trials were discontinued after one week each. Further experiments are needed to evaluate the ion specific stress resistance of the evolved strains. Cross-tolerance towards high temperature was also tested, but the high temperature trial did not show any growth for the same evaluated strains, indicating no evolved cross tolerance towards elevated temperatures.

4.5 Lipid accumulation

The lipid accumulation of the evolved strains was tested to make sure that the improved performance towards high concentrations of NaCl did not evolve at the cost of lipid accumulation ability. Microscope pictures from the lipid accumulation trial are shown in Figure 4.17, where the cells are stained with Nile red after cultivation in nitrogen limiting medium. All strains show lipid accumulation capabilities after evolution. It looks like S.3.10, S.4.4 and S.5.7 give a weak signal from the Nile red, but if looking closely one can see that there are droplets inside those cells, probably lipid droplets, which somehow have not been stained by Nile red, indicating lipid accumulation capabilities.

It can be noted that some of the strains differ in morphology, even within a population. For example, both S.3.10 and S.4.4 show a more yeast like phenotype, while the other two strains from the same population show a more pseudohyphae and hyphae phenotype.

Lipid accumulation ability was also checked in N-limiting Delft medium supplemented with 1.4 M NaCl for S.1.8, S.2.2 and S.3.10. Nile red staining showed that all 3 evolved strains accumulate lipids at salt stressed conditions as well, see Figure 4.18.

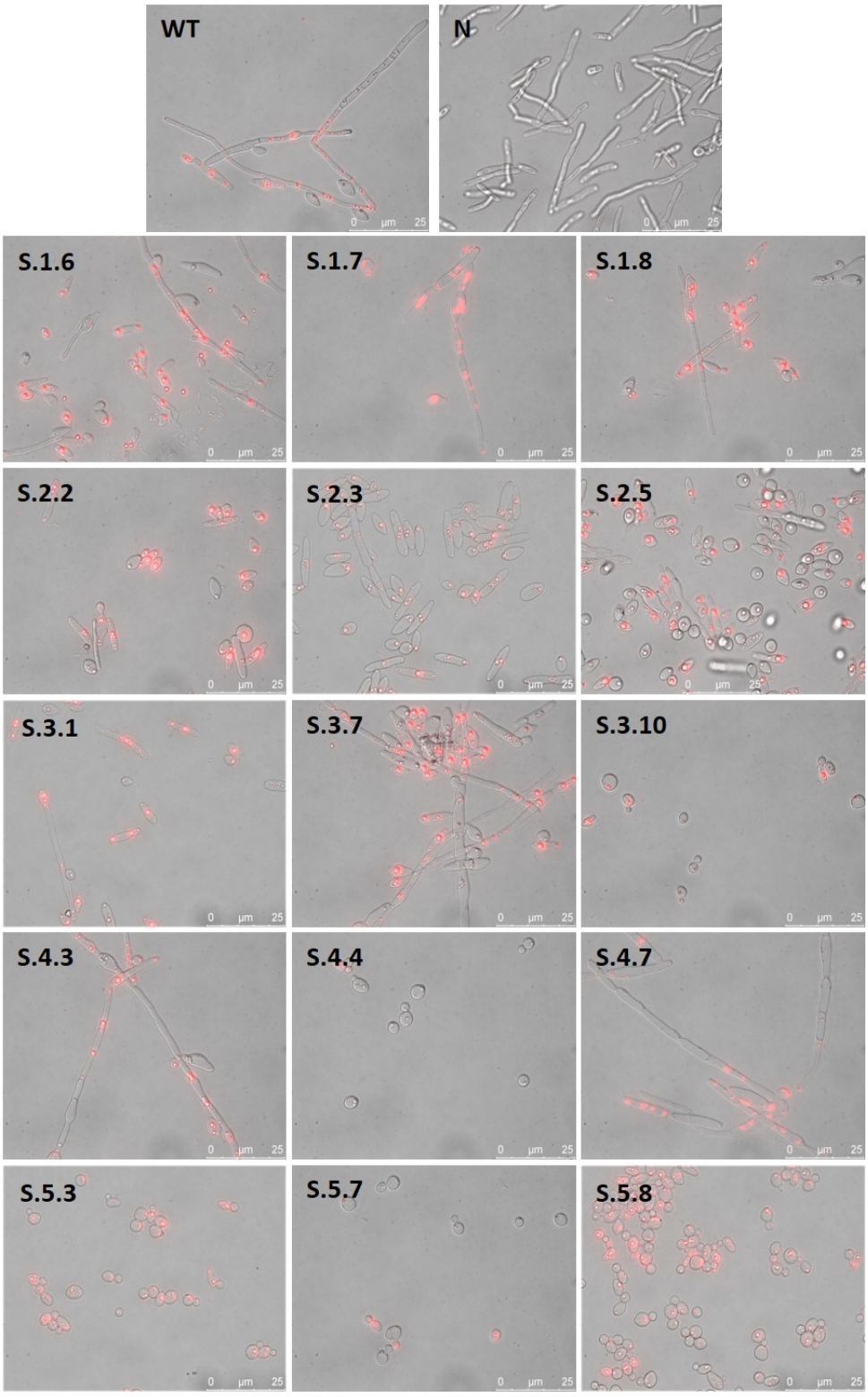

Figure 4.17: Fluorescent microscope pictures of evolved strains stained with Nile red after cultivation in N-limiting Delft medium for 3 days. The negative control (N) is wild type FKP391 cultivated in normal Delft medium

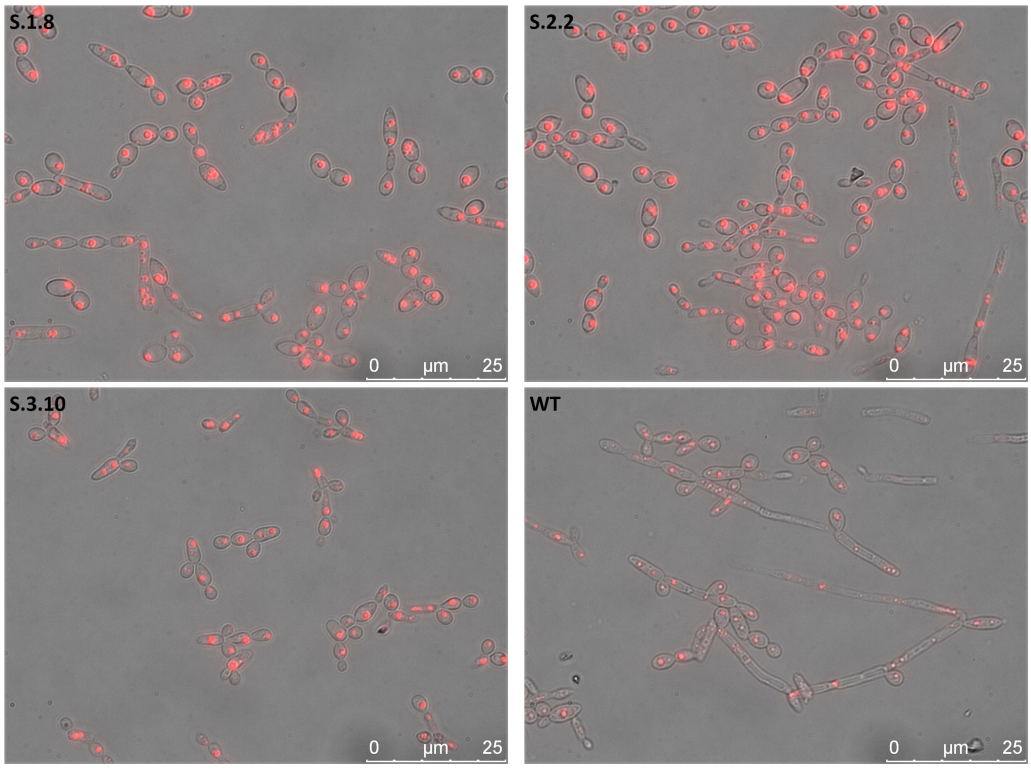

Figure 4.18: Fluorescent microscope pictures of evolved strains stained with Nile red after cultivation in N-limiting Delft medium supplemented with 1.4 M NaCl for 3 days

4.6 Whole genome sequencing

The best three strains from each population (section 4.3.2) were selected for whole genomic sequencing to connect the improved phenotype with changes in the genome. Common changes between the population are potential candidate mutations, which later can be reconstructed using the CRISPR/Cas9 technique.

4.6.1 Genomic extractions

Before whole genome sequencing, the genomic DNA needs to be extracted and purified. The results of the quality control of extracted genomic DNA are summarized in Table 4.1 and Figure 4.19. $A_{260/280}$ should be between 1.6 - 2.0 and $A_{260/230}$ should be between 1.8 - 2.2. Lower ratios indicate presence of contaminants which absorbs at 280 or 230 nm. Furthermore, the genomic DNA should be of high molecular weight and at least 500 ng in amounts. The results from the whole genome sequencing are expected to be received in late 2017.

Table 4.1: Concentrations and purity for the genomic extractions done with phenol:chloroform extraction (P) or with Amresco kit (A).

	NanoDrop 1^{st} read			NanoDrop 2^{nd} read			Qubit	
Sample	**ng/µL**	**260/280**	**260/230**	**ng/µL**	**260/280**	**260/230**	**ng/µL**	**µg**
S.1.6 (P)	376.9	1.84	2.02	373.9	1.84	2.02	48.4	2.42
S.1.7 (P)	1909.3	1.88	2.03	1730.9	1.9	2.13	85.2	4.26
S.1.8 (P)	528.3	1.82	1.99	577.1	1.81	2.08	73.4	3.67
S.2.2 (P)	1977.9	1.9	2.01	1868.4	1.9	2.03	89	4.45
S.2.3 (P)	772.9	1.88	2.26	689.5	1.88	2.2	69	3.45
S.2.5 (P)	260.6	1.84	1.95	316.1	1.83	1.79	69.2	3.46
S.3.1 (P)	889.7	1.83	2.08	806.4	1.83	2.15	63.2	3.16
S.3.7 (P)	227	1.8	1.77	227	1.81	1.78	36.8	1.84
S.3.10 (A)	756.4	2.01	2.16	741.5	2	2.18	44.2	2.21
S:4.3 (P)	2295	1.93	2.01	2405.8	1.92	1.97	96	4.8
S.4.4 (A)	1072.7	2.05	2.24	1077.9	2.04	2.23	47.6	2.38
S.4.7 (P)	157.7	1.78	1.72	157.8	1.78	1.71	30.6	1.53
S.5.3 (P)	883.1	1.87	1.86	710.3	1.9	2.02	95.6	4.78
S.5.7 (P)	1314.7	1.91	1.88	1258	1.91	1.86	48.4	2.42
S.5.8 (P)	948.2	1.88	1.92	951.7	1.87	1.87	97.6	4.88
FKP391 (P)	1488.6	1.99	2.2	1487.6	1.97	2.2	57	2.85
Population 1 (A)	1259.9	1.93	2.25	1085.8	1.92	2.25	40.4	2.02
Population 2 (A)	753	1.94	2.26	764.6	1.95	2.27	30.8	1.54
Population 3 (A)	785	1.9	2.18	722.6	1.91	2.15	54.4	2.72
Population 4 (A)	632.3	1.88	2.12	622.7	1.88	2.11	32.6	1.63
Population 5 (A)	695.9	1.93	2.05	676.1	1.91	2.03	62.4	3.12

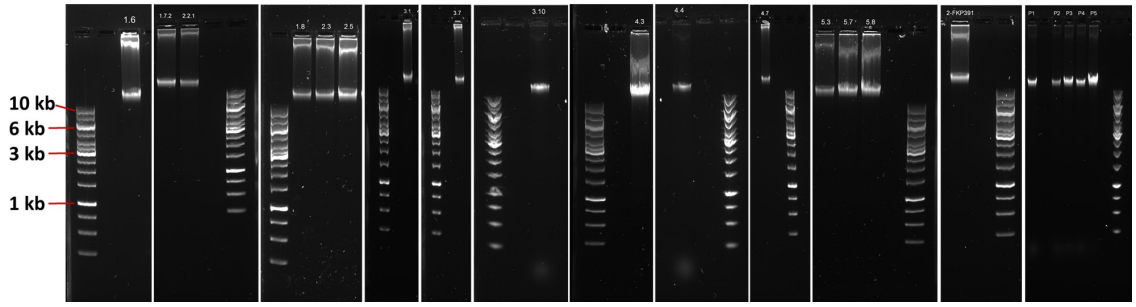

Figure 4.19: 0.7 % agarose gels of extracted genomic DNA run at 80 V for 1 hour and 20 minutes

4.7 Reverse engineering

As no sequencing data could be obtained before the end of this thesis, no targets for reverse engineering were identified. However, the CRISPR/Cas9 system was set up and can be used when genome data arrives.

4.7.1 pCAS2yl-erg3

The first attempt for setting up the CRISPR/Cas9 system was not successful; none of the 8 screened colonies showed evidence of a knock-out event for *ERG3* when using the Gao et al. (2016) strategy [38], even after 4 days of outgrowth. According to Schwartz et al. (2016) the polymerase II design with ribozymes results in low HR frequencies, while the polymerase III design with tRNA hybrid promoters yields much higher HR frequencies (>90 % after 4 days of outgrowth) [44]. Therefore, the CRISPR/Cas9 design was adapted to utilize the design of Schwartz et al (2016) for the expression of the sgRNA.

4.7.2 pCAS4yl-hph

The modified design with polymerase III promoter for the sgRNA instead of polymerase II (pCAS2yl-erg3) proved to be more efficient in *Y. lipolytica*. The map and sequence of the final constructed plasmid for the hph knock-out (pCAS4yl-hph) can be seen in section A.3.1. The results of the colony PCR of *Y. lipolytica* FKP355 transformed with pCAS4yl-hph are shown in Figure 4.20. Successful knock-outs should yield a PCR product of 2453 bp, while colonies with intact hph casette should yield a PCR product of 3890 bp. This matches the results quite well. This strain was cured of the plasmid by YPD outgrowth, named JHYl001 and cryopreservered in 25 % glycerol at -80 °C.

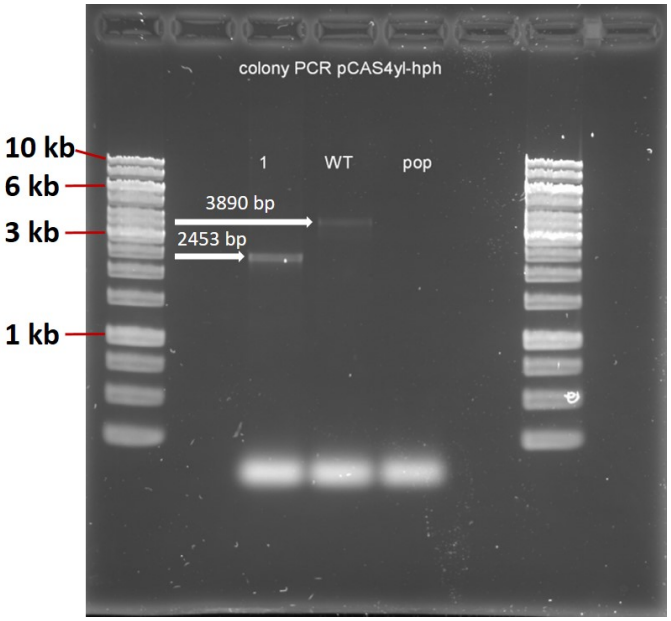

Figure 4.20: Gel electrophoresis results of colony PCR products of colony 1 (transformed with pCAS4yl-hph and isolated after outgrowth) and wild type FKP355. Knock-out event should yield a 2453 bp product, while intact *HPH* should yield a 3890 bp product

This means that the single plasmid based design for markerless genome editing is set up and ready for later reverse engineering based on future analysis of genome data of the evolved strains.

5
Discussion

Adaptive laboratory evolution is a useful technique to understand how cells adapt to different stresses, and to use that information to construct new, more tolerant strains. During the evolution in high salt medium, the cells spend a lot of energy to pump out excess Na^+ to avoid toxic levels [27]. A mutant with more effluent pumps, higher energy production or some other adaptation that lowers the toxicity of excess Na^+ will be able to spend more energy on growing instead of surviving, resulting in faster growth rate and/or higher biomass yield.

After this adaptive evolution experiment, strains from every population showed increased performance in high salt medium, both in microcultivation settings and shake flask cultivation. The growth rate of the evolved strains did not differ from the wild type during the first 24 hours of cultivation, but the evolved strains showed an increase of roughly 50 % in final OD_{600} (Figure 4.11). This could be due to that the evolved strains somehow handle the Na^+ stress more efficiently, resulting in a higher biomass yield. The mechanism of adaption can only be speculated based on the observed phenotype, as the genome data are yet to be obtained.

None of the tested evolved strains showed improved tolerance in LiOAc medium or high sorbitol medium. The concentration of lithium might have been too high in the Li^+ trial, as *S. cerevisiae* strains have been shown to tolerate up to 20 mM LiCl [28]. A screening for Li^+ tolerance should have been made for *Y. lipolytica* FKP391 prior the testing of the evolved strains. The reason for no improvement in sorbitol medium needs further investigations. The high concentration of sorbitol (2.5 M) made the medium very viscous, which might introduce other factors, such as oxygen transfer limits, and thus affect the results. However, it could be possible that some of the other isolated colonies show increased tolerance towards sorbitol and Li^+.

Both biological replicates of the wild type grew very poorly in the second bioscreen round, which is not consistent with the first round and the later shake flasks cultivations. The difference in the setup of the second round of bioscreen was that the preadaptation was done in precultures containing full stress medium (Delft medium containing 1.4 M NaCl), where the strains in the first round was preadapted on YPD plates containing 1.2 M NaCl prior inoculation into the multiwell plate. Even the evolved strains obtained a lower final OD_{600} in the second round (see Figure 4.5 and Figure 4.7), even if the difference was small compared to the reduction of the wild type. Maybe the difference of the evolved strains is mainly caused by the error between screening runs and since they were precultured in full stress medium. The reason why the wild type grew poorly during the second round of bioscreen needs to be further investigated.

Interestingly, all 3 tested strains showed improved performance at low pH medium, especially S.3.10 (Figure 4.13 and 4.15), compared to wild type. A evolved cross-tolerance towards low pH and not sorbitol could indicate adaptation towards ionic stress rather than osmotic stress. This phenotype could potentially be connected to improved efficiency of the Nha1 antiporter, which is responsible for most of the Na^+ efflux at low pH [27]. If the efficiency of Nha1 has been improved, the cell needs to be able to cope with excess protons that are taken up during Na^+ export through the Nha1 antiporter. However, this needs genome data and additional experiments to be confirmed.

Similar formation of multicellular clusters, which was seen in the low pH growth trials (Figure 4.14 and 4.16), was found in lactic acid evolved strains in the study of Fletcher et al. (2017). However, if *Y. lipolytica* had acquired the same mutation (defect *ACE2*), the multicellular phenotype would have occurred at all growth conditions, which it did not. Formation of flocs have previously been shown to increase the tolerance to higher ethanol concentrations [51] and could be a way for cells to cope with some external stresses. The evolved strain S.3.10 have shown a more yeast-like phenotype in medium without NaCl (Figure 4.17), which might aid the formation of multicellular clusters and improve tolerance towards low pH.

In hindsight, it might have been better to use the BioLector (m2p-labs) for the colony screening, even though the throughput is lower (48 wells), as it has a controlled environment inside. This could reduce the problems with evaporation and make sure that no oxygen limit occurs, which is important as *Y. lipolytica* requires oxygen to grow. Indications of oxygen limitations in the bioscreen have been previously reported [50].

Even though the CRISPR tool is ready for reverse engineering (see section 4.7.2), it could still use some further improvements. One of the improvements could be to add multiplexing to the CRISPR design. A potential problem with the current design is that the plasmid might become too large if several genes are going to be knocked out at the same time, as each knock out requires 2000 bp in repair fragments and 547 bp for the sgRNA casette. This could be avoided by integrating Cas9 into the pBR docking platform [52] in the genome of *Y. lipolytica*, using hygromycin B for selection. This integration would remove the need to transform Cas9 every time genetic changes are going to be made and thus reducing the size of the plasmid by almost 5000 bp. Furthermore, the design can be improved by expressing multiple guides from one transcript and process them using the tRNA machinery, shown to be able to process at least eight sgRNAs from one transcript [48].

Future work involves analysis of genome data and further characterization of the evolved strains in high saline medium in controlled settings using bioreactors. The further characterization would include sampling for substrate uptake rate, compatible solute formation and excretion of metabolites. It would be interesting to obtain and analyze the transcription profiles of the evolved strains during high saline cultivation and compare them to the wild type. The information collected from this evolution could then hopefully be transferred to an industrial strain with the aim for improved tolerance and performance in lignocellulose-derived medium for sustainable production of biodiesel.

6
Conclusion

After 220 generations of the adaptive evolution, an improved phenotype could be observed for all isolated strains of *Y. lipolytica* in Delft medium containing 1.4 M NaCl. Furthermore, the strains showed increased performance at low pH, indicating evolved cross protection. Samples are sent for genomic sequencing, which is expected to be finished after the summer. A single plasmid design has been optimized for later reverse engineering in *Y. lipolytica* when genome data arrives.

Bibliography

[1] Li Q, Du W, Liu D. Perspectives of microbial oils for biodiesel production. Applied Microbiology and Biotechnology. 2008;80(5):749–756.

[2] Luo L, van der Voet E, Huppes G. Biorefining of lignocellulosic feedstock - Technical, economic and environmental considerations. Bioresource Technology. 2010;101(13):5023–5032.

[3] Jönsson LJ, Alriksson B, Nilvebrant NO. Bioconversion of lignocellulose: Inhibitors and detoxification. Biotechnology for Biofuels. 2013;6(1).

[4] Friedlander J, Tsakraklides V, Kamineni A, Greenhagen EH, Consiglio AL, MacEwen K, et al. Engineering of a high lipid producing *Yarrowia lipolytica* strain. Biotechnology for Biofuels. 2016;9(1).

[5] Cavka A, Jornsson LJ. Comparison of the growth of filamentous fungi and yeasts in lignocellulose-derived media. Biocatalysis and Agricultural Biotechnology. 2014;3(4):197–204.

[6] Braga A, Mesquita DP, Amaral AL, Ferreira EC, Belo I. Quantitative image analysis as a tool for *Yarrowia lipolytica* dimorphic growth evaluation in different culture media. Journal of Biotechnology. 2016;217:22–30.

[7] Ledesma-Amaro R, Nicaud JM. Metabolic Engineering for Expanding the Substrate Range of *Yarrowia lipolytica*. Trends in Biotechnology. 2016;34(10):798–809.

[8] Fickers P, Benetti PH, Waché Y, Marty A, Mauersberger S, Smit MS, et al. Hydrophobic substrate utilisation by the yeast *Yarrowia lipolytica*, and its potential applications. FEMS Yeast Research. 2005;5(6-7):527–543.

[9] Workman M, Holt P, Thykaer J. Comparing cellular performance of *Yarrowia lipolytica* during growth on glucose and glycerol in submerged cultivations. AMB Express. 2013;3:1–9.

[10] Bankar AV, Kumar AR, Zinjarde SS. Environmental and industrial applications of *Yarrowia lipolytica*. Applied Microbiology and Biotechnology. 2009;84(5):847–865.

[11] Madzak C. *Yarrowia lipolytica*: recent achievements in heterologous protein expression and pathway engineering. Applied Microbiology and Biotechnology. 2015;99(11):4559–4577.

[12] Kavšček M, Bhutada G, Madl T, Natter K. Optimization of lipid production with a genome-scale model of *Yarrowia lipolytica*. BMC Systems Biology. 2015;9(1).

[13] Wolf K. Nonconventional Yeasts in Biotechnology: A Handbook. 1st ed. Springer Berlin Heidelberg; 1996.

[14] Kerscher S, Dröse S, Zwicker K, Zickermann V, Brandt U. Yarrowia lipolytica, a yeast genetic system to study mitochondrial complex I. Biochimica et Biophysica Acta - Bioenergetics. 2002;1555(1-3):83–91.

[15] Dulermo T, Nicaud JM. Involvement of the G3P shuttle and β-oxidation pathway in the control of TAG synthesis and lipid accumulation in *Yarrowia lipolytica*. Metabolic Engineering. 2011;13(5):482–491.

[16] Beopoulos A, Cescut J, Haddouche R, Uribelarrea JL, Molina-Jouve C, Nicaud JM. *Yarrowia lipolytica* as a model for bio-oil production. Progress in Lipid Research. 2009;48(6):375–387.

[17] Pomraning KR, Kim YM, Nicora CD, Chu RK, Bredeweg EL, Purvine SO, et al. Multi-omics analysis reveals regulators of the response to nitrogen limitation in *Yarrowia lipolytica*. BMC Genomics. 2016;17(1).

[18] Li H, Wu M, Xu L, Hou J, Guo T, Bao X, et al. Evaluation of industrial Saccharomyces cerevisiae strains as the chassis cell for second-generation bioethanol production. Microbial Biotechnology. 2015;8(2):266–274.

[19] Hohmann S. Osmotic stress signaling and osmoadaptation in yeasts. Microbiology and Molecular Biology Reviews. 2002;66(2):300–372.

[20] Babazadeh R, Adiels CB, Smedh M, Petelenz-Kurdziel E, Goksör M, Hohmann S. Osmostress-induced cell volume loss delays yeast Hog1 signaling by limiting diffusion processes and by Hog1-specific effects. PLoS ONE. 2013;8(11).

[21] Yancey PH. Organic osmolytes as compatible, metabolic and counteracting cytoprotectants in high osmolarity and other stresses. Journal of Experimental Biology. 2005;208(15):2819–2830.

[22] Hohmann S. An integrated view on a eukaryotic osmoregulation system. Current Genetics. 2015;61(3):373–382.

[23] Babazadeh R, Lahtvee PJ, Adiels CB, Goksör M, Nielsen JB, Hohmann S. The yeast osmostress response is carbon source dependent. Scientific Reports. 2017;7(1).

[24] Yang LB, Dai XM, Zheng ZY, Zhu L, Zhan XB, Lin CC. Proteomic analysis of erythritol-producing *Yarrowia lipolytica* from glycerol in response to osmotic pressure. Journal of Microbiology and Biotechnology. 2015;25(7):1056–1069.

[25] MAPK signaling pathway - yeast - *Yarrowia lipolytica* [Internet]. Kyoto Encyclopedia of Genes and Genomes (KEGG); [updated 2016 Feb 16; cited 2017 Jun 1]. Available from: `http://www.genome.jp/kegg-bin/show_pathway?org_name=yli&mapno=04011&mapscale=&show_description=hide`.

[26] Cervantes-Chávez JA, Ruiz-Herrera J. STE11 disruption reveals the central role of a MAPK pathway in dimorphism and mating in *Yarrowia lipolytica*. FEMS Yeast Research. 2006;6(5):801–815.

[27] Ariño J, Ramos J, Sychrová H. Alkali metal cation transport and homeostasis in yeasts. Microbiology and Molecular Biology Reviews. 2010;74(1):95–120.

[28] Daran-Lapujade P, Daran JM, Luttik MAH, Almering MJH, Pronk JT, Kötter P. An atypical *PMR2* locus is responsible for hypersensitivity to sodium and lithium cations in the laboratory strain *Saccharomyces cerevisiae* CEN.PK113-7D. FEMS Yeast Research. 2009;9(5):789–792.

[29] Ramos J, Ariño J, Sychrova H. Alkali-metal-cation influx and efflux systems in nonconventional yeast species. FEMS Microbiology Letters. 2011;317(1):1–8.

[30] Papouskova K, Sychrova H. *Yarrowia lipolytica* possesses two plasma membrane alkali metal cation/H+ antiporters with different functions in cell physiology. FEBS Letters. 2006;580(8):1971–1976.

[31] Casaregola S, Neuvéglise C, Lépingle A, Bon E, Feynerol C, Artiguenave F, et al. Genomic exploration of the hemiascomycetous yeasts: 17. *Yarrowia lipolytica*. FEBS Letters. 2000;487(1):95–100.

[32] Gasch AP, Werner-Washburne M. The genomics of yeast responses to environmental stress and starvation. Functional and Integrative Genomics. 2002;2(4-5):181–192.

[33] Dragosits M, Mattanovich D. Adaptive laboratory evolution - principles and applications for biotechnology. Microbial Cell Factories. 2013;12(1).

[34] Caspeta L, Chen Y, Ghiaci P, Feizi A, Baskov S, Hallström BM, et al. Altered sterol composition renders yeast thermotolerant. Science. 2014;346(6205):75–78.

[35] Fletcher E, Feizi A, Bisschops MMM, Hallström BM, Khoomrung S, Siewers V, et al. Evolutionary engineering reveals divergent paths when yeast is adapted to different acidic environments. Metabolic Engineering. 2017;39:19–28.

[36] Dhar R, Sägesser R, Weikert C, Yuan J, Wagner A. Adaptation of *Saccharomyces cerevisiae* to saline stress through laboratory evolution. Journal of Evolutionary Biology. 2011;24(5):1135–1153.

[37] Hussain MS, Rodriguez GM, Gao D, Spagnuolo M, Gambill L, Blenner M. Recent advances in bioengineering of the oleaginous yeast *Yarrowia lipolytica*. AIMS Bioengineering. 2016;3(4):493–514.

[38] Gao S, Tong Y, Wen Z, Zhu L, Ge M, Chen D, et al. Multiplex gene editing of the *Yarrowia lipolytica* genome using the CRISPR-Cas9 system. Journal of Industrial Microbiology and Biotechnology. 2016;43(8):1085–1093.

[39] Kretzschmar A, Otto C, Holz M, Werner S, Hübner L, Barth G. Increased homologous integration frequency in *Yarrowia lipolytica* strains defective in non-homologous end-joining. Current genetics. 2013;59(1-2):63–72.

[40] Verbeke J, Beopoulos A, Nicaud JM. Efficient homologous recombination with short length flanking fragments in Ku70 deficient *Yarrowia lipolytica* strains. Biotechnology Letters. 2013;35(4):571–576.

[41] Näätsaari L, Mistlberger B, Ruth C, Hajek T, Hartner FS, Glieder A. Deletion of the pichia pastoris ku70 homologue facilitates platform strain generation for gene expression and synthetic biology. PLoS ONE. 2012;7(6).

[42] Bredeweg EL, Pomraning KR, Dai Z, Nielsen J, Kerkhoven EJ, Baker SE. A molecular genetic toolbox for *Yarrowia lipolytica*. Biotechnology for Biofuels. 2017;10(1).

[43] Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science. 2012;337(6096):816–821.

[44] Schwartz CM, Hussain MS, Blenner M, Wheeldon I. Synthetic RNA Polymerase III Promoters Facilitate High-Efficiency CRISPR-Cas9-Mediated Genome Editing in *Yarrowia lipolytica*. ACS Synthetic Biology. 2016;5(4):356–359.

[45] Ryan OW, Skerker JM, Maurer MJ, Li X, Tsai JC, Poddar S, et al. Selection of chromosomal DNA libraries using a multiplex CRISPR system. eLife. 2014;3(August2014):1–15.

[46] Yao L, Cengic I, Anfelt J, Hudson EP. Multiple Gene Repression in Cyanobacteria Using CRISPRi. ACS Synthetic Biology. 2016;5(3):207–212.

[47] Gao Y, Zhao Y. Self-processing of ribozyme-flanked RNAs into guide RNAs in vitro and in vivo for CRISPR-mediated genome editing. Journal of Integrative Plant Biology. 2014;56(4):343–349.

[48] Xie K, Minkenberg B, Yang Y. Boosting CRISPR/Cas9 multiplex editing capability with the endogenous tRNA-processing system. Proceedings of the National Academy of Sciences of the United States of America. 2015;112(11):3570–3575.

[49] Fujita SI, Senda Y, Nakaguchi S, Hashimoto T. Multiplex PCR using internal transcribed spacer 1 and 2 regions for rapid detection and identification of yeast strains. Journal of Clinical Microbiology. 2001;39(10):3617–3622.

[50] Warringer J, Blomberg A. Automated screening in environmental arrays allows analysis of quantitative phenotypic profiles in *Saccharomyces cerevisiae*. Yeast. 2003;20(1):53–67.

[51] Lei J, Zhao X, Ge X, Bai F. Ethanol tolerance and the variation of plasma membrane composition of yeast floc populations with different size distribution. Journal of Biotechnology. 2007;131(3):270–275.

[52] Madzak C, Tréton B, Blanchin-Roland S. Strong hybrid promoters and integrative expression/secretion vectors for quasi-constitutive expression of heterologous proteins in the yeast *Yarrowia lipolytica*. Journal of Molecular Microbiology and Biotechnology. 2000;2(2):207–216.

A

Appendix 1

A.1 Primers

Table A.1: Primers used in this study. Uppercase characters hybridize to the template and lowercase characters contain the overlapping overhang

Primer	**Sequence (5′ → 3′)**
ITS1	TCCGTAGGTGAACCTGCGG
ITS4	TCCTCCGCTTATTGATATGC
erg3-up FW	ttcatggtcaaagatGAGCGTTACAGTGGTAGTCC
erg3-up RV	caacccggtctctgtGTCTTGTATGGAGTCATTAATGG
erg3-dw FW	tgaccatgattacgccaagcttgtttgtttTGTAGTCATCGAGCTGTTGG
erg3-dw RV	accactgtaacgctcATCTTTGACCATGAAGTGAAGG
gRNA-erg3-up FW	gactccatacaagacACAGAGACCGGGTTGGC
gRNA-erg3-up RV	gctctaaaacacgtagactcgatcgtcagcGACGAGCTTACTCGTTTCG
gRNA-erg3-dw FW	gctgacgatcgagtctacgtGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGC
gRNA-erg3-dw RV	gacacaaatacgccgccaacccggtctctgCAAACCCAAAAGGGCCGA
cPCR-erg3 FW	ATCCCAGAGTGGTTTCGCTG
cPCR-erg3 RV	GATTCTACACCGGTCTCGCC
Pol3 promoter FW	atgattacgccaagcttgtttgtttaaacCCCCAGTTGCAAAAGTTGAC
Pol3 promoter RV	aaatacgccgccaacccggtctctgtcgacAAAAAAAAAGCACCGACTCG
hph-up FW	tgaccatgattacgccaagcttgtttgtttaaacTAAAGGAGGCCGCAAAGGTT
hph-up RV	aagtggatcccggtcGTCGACTTCCAGATCAGAACATG
hph-dw FW	gatctggaagtcgacGACCGGGATCCACTTAACG
hph-dw RV	ttgcaactgggggcatgcGAGTGTCCAGATCCGCAAGA
gRNA-hph-up FW	atctggacactcgcatgcCCCCAGTTGCAAAAGTTGAC
gRNA-hph-up RV	gctctaaaactggttggcttgtatggagcaacGTCAACCTGCGCCGACC
gRNA-hph-dw FW	tgctccatacaagccaaccaGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGG
gRNA-hph-dw RV	ATGGGTTACCCGGATGGC
cPCR-hph FW	GGCCCACATTCAGACCCTT
cPCR-hph RV	CTCGTACTCTCCCACGGATGC

A.2 CRISPR guides

Table A.2: CRISPR guides

Guide	**Sequence (5′ → 3′)**	**Efficiency score**	**Specificity Score**
ERG3	GCTGACGATCGAGTCTACGT	75.04	99.89
HPH	TGCTCCATACAAGCCAACCA	74.29	100

A.3 Plasmids

A.3.1 pCAS4yl-hph

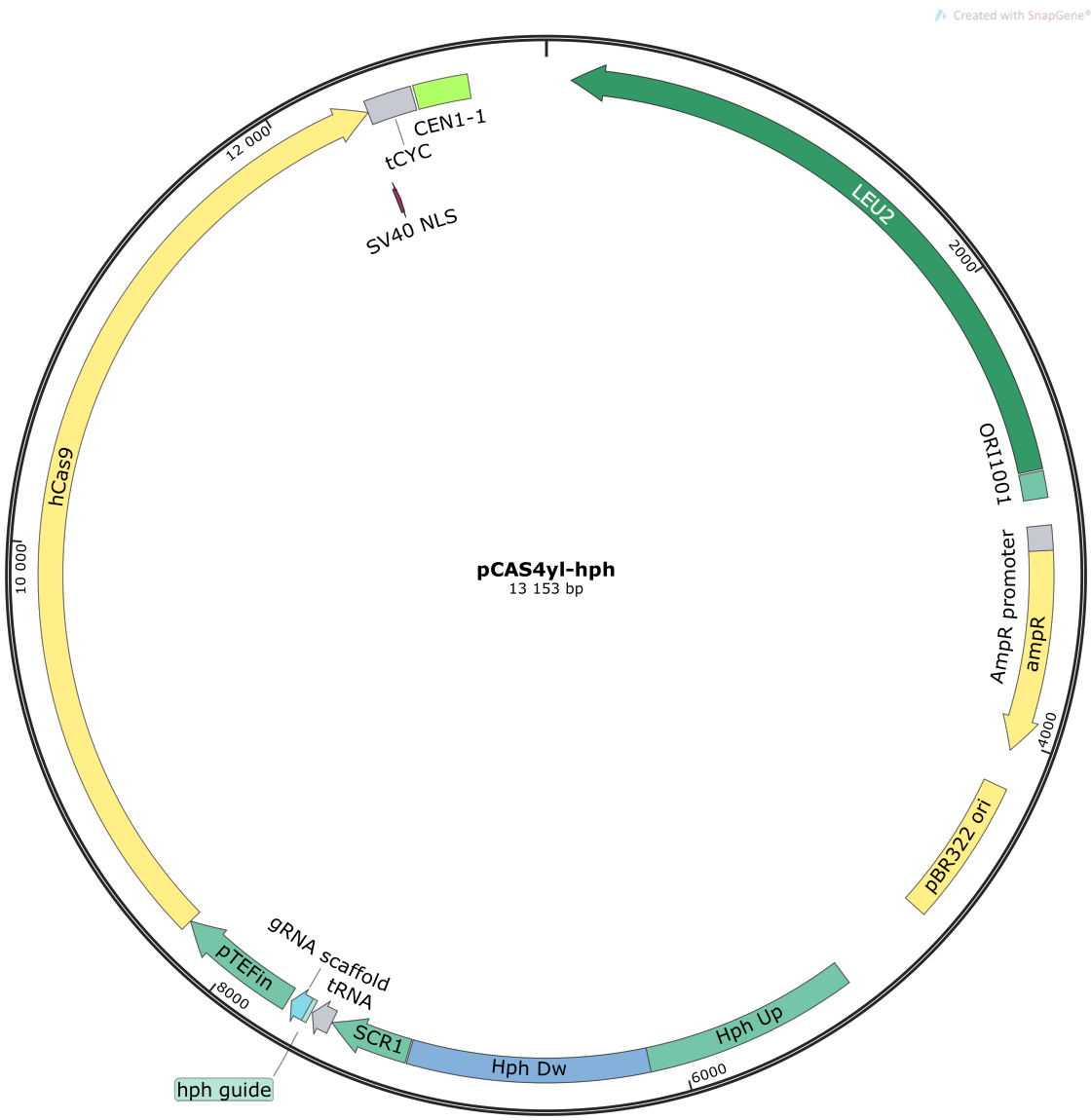

Figure A.1: Plasmid map pCAS4yl-hph

Sequence:

```
LOCUS       Exported                13153 bp ds DNA     circular SYN 22 MAY 2017
DEFINITION  synthetic circular DNA
ACCESSION   .
VERSION     .
KEYWORDS    pCAS4yl hph
SOURCE      synthetic DNA construct
  ORGANISM  synthetic DNA construct
REFERENCE   1  (bases 1 to 13153)
  AUTHORS   .
  TITLE     Direct Submission
  JOURNAL   Exported Monday, May 22, 2017 from SnapGene Viewer 3.3.4
            http://www.snapgene.com
FEATURES             Location/Qualifiers
     source          1..13153
                     /organism="synthetic DNA construct"
                     /mol_type="other DNA"
```

```
     misc_feature    complement(107..2845)
                     /label=LEU2
     misc_feature    2854..2966
                     /label=ORI1001
     promoter        3081..3185
                     /label=AmpR promoter
     CDS             3186..4043
                     /label=ampR
     misc_feature    4201..4820
                     /label=pBR322_ori
     misc_feature    5238..6141
                     /label=hph Up
     misc_feature    6142..7159
                     /label=Hph Dw
     misc_feature    7166..7511
                     /label=SCR1
     misc_feature    7512..7603
                     /label=tRNA
     misc_feature    7609..7628
                     /label=hph guide
     misc_RNA        7629..7704
                     /label=gRNA scaffold
     misc_feature    7705..7712
                     /label=polyT
     promoter        7719..8249
                     /label=pTEFin
     CDS             8250..12386
                     /label=hCas9
     CDS             12363..12383
                     /label=SV40 NLS
     terminator      12387..12587
                     /label=tCYC
     misc_feature    12599..12828
                     /label=CEN1 1
ORIGIN
        1 ctcccggcat ccgcttacag acaagctgtg accgtctccg ggagctgcat gtgtcagagg
       61 ttttcaccgt catcaccgaa acgcgcgaga cgaaagggcc agatctgttc ggaaatcaac
      121 ggatgctcaa ccgatttcga cagtaataat ttgaatcgaa tcggagccta aaatgaaccc
      181 gagtatatct cataaaattc tcggtgagag gtctgtgact gtcagtacaa ggtgccttca
      241 ttatgccctc aaccttacca tacctcactg aatgtagtgt acctctaaaa atgaaataca
      301 gtgccaaaag ccatggcact gagctcgtct aacggacttg atatacaacc aattaaaaca
      361 aatgaaaaga aatacagttc tttgtatcat ttgtaacaat taccctgtac aaactaaggt
      421 attgaaatcc cacaatattc ccaaagtcca ccccttccca aattgtcatg cctacaactc
      481 atataccaag cactaaccta ccaaacacca ctaaaacccc acaaaatata tcttaccgaa
      541 tatacagtaa caagctacca ccacactcgt tgggtgcagt cgccagctta aagatatcta
      601 tccacatcag ccacaactcc cttcctttaa taaaccgact acacccttgg ctattgaggt
      661 tatgagtgaa tatactgtag acaagacact ttcaagaaga ctgtttccaa aacgtaccac
      721 tgtcctccac tacaaacaca cccaatctgc ttcttctagt caaggttgct acaccggtaa
      781 attataaatc atcatttcat tagcagggca gggccctttt tatagagtct tatacactag
      841 cggaccctgc cggtagacca acccgcaggc gcgtcagttt gctccttcca tcaatgcgtc
      901 gtagaaacga cttactcctt cttgagcagc tccttgacct tgttggcaac aagtctccga
      961 cctcggaggt ggaggaagag cctccgatat cggcggtagt gataccagcc tcgacggact
     1021 ccttgacggc agcctcaaca gcgtcaccgg cgggcttcat gttaagagag aacttgagca
     1081 tcatggcggc agacagaatg gtggcaatgg ggttgacctt ctgcttgccg agatcggggg
     1141 cagatccgtg acagggctcg tacagaccga acgcctcgtt ggtgtcgggc agagaagcca
     1201 gagaggcgga gggcagcaga cccagagaac cggggatgac ggaggcctcg tcggagatga
     1261 tatcgccaaa catgttggtg gtgatgatga taccattcat cttggagggc tgcttgatga
     1321 ggatcatggc ggccgagtcg atcagctggt ggttgagctc gagctggggg aattcgtcct
     1381 tgaggactcg agtgacagtc tttcgccaaa gtcgagagga ggccagcacg ttggcccttgt
     1441 caagagacca cacgggaaga gggggttgt gctgaagggc caggaaggcg gccattcggg
     1501 caattcgctc aacctcagga acggagtagg tctcggtgtc ggaagcgacg ccagatccgt
     1561 catcctcctt tcgctctcca aagtagatac ctccgacgag ctctcggaca atgatgaagt
     1621 cggtgccctc aacgtttcgg atgggggaga gatcggcgag cttgggcgac agcagctggc
     1681 agggtcgcag gttggcgtac aggttcaggt cctttcgcag cttgaggaga ccctgctcgg
     1741 gtcgcacgtc ggttcgtccg tcgggagtgg tccatacggt gttggcagcg cctccgacag
     1801 caccgagcat aatagagtca gcctttcggc agatgtcgag agtagcgtcg gtgatgggct
     1861 cgccctcctt ctcaatggca gctcctccaa tgagtcggtc ctcaaacaca aactcggtgc
     1921 cggaggcctc agcaacagac ttgagcacct tgacggcctc ggcaatcacc tcggggccac
     1981 agaagtcgcc gccgagaaga acaatcttct tggagtcagt cttggtcttc ttagtttcgg
     2041 gttccattgt ggatgtgtgt ggttgtatgt gtgatgtggt gtgtggagtg aaaatctgtg
     2101 gctggcaaac gctcttgtat atatacgcac ttttgcccgt gctatgtgga agactaaacc
     2161 tccgaagatt gtgactcagg tagtgcggta tcggctaggg acccaaacct tgtcgatgcc
     2221 gatagcgcta tcgaacgtac cccagccggc cgggagtatg tcggagggga catacgagat
     2281 cgtcaagggt ttgtggccaa ctggtaaata aatgatgact caggcgacga cggaattcga
     2341 cagcaactac tcctttcacc aaccatgtgc attttagctc gaataacatt cacaggcttg
     2401 gtgatctaca tccatggtgt ctggccgatt accgtggtgt tttggcagta acgagaatat
     2461 tgagtgaact cttcccatca ccaataaaga ctcatactac aatcacgagc gcttcagctg
     2521 ccactatagt gttggtgaca caatacccct cgatgctggg cattactgta gcaagagata
     2581 ttcatttcat ggcgcatttt ccagtctacc tgacttttta gtgccgattt cttctccaca
     2641 ttttacgctc agtgtgaaaa gttggagtgc acacttaatt atcgccggtt ttcggaaaag
     2701 tactatgtgc tcaaggttgc accccacgtt acgtatgcag cacattgagc agcctttgga
     2761 ccgtggagat aacggtgtgg agatagcaac gggtagtctt cgtaataagc aatgcattgt
     2821 tagttttata tgatatggtg tcgaagcggc cgcatactac tgtatattca agcaagtata
     2881 tccgtgggtg cgggtgattt ggatctaagg ttcgtactca acactcacga gcagcttgcc
     2941 tatgttacat ccttttatca gacatagcgg ccgcttcgac tctagaggat ctgggcctcg
     3001 tgatacgcct atttttatag gttaatgtca tgataataat ggtttcttag acgtcaggtg
     3061 gcacttttcg gggaaatgtg cgcggaaccc ctatttgttt atttttctaa atacattcaa
     3121 atatgtatcc gctcatgaga caataaccct gataaatgct tcaataatat tgaaaaagga
     3181 agagtatgag tattcaacat ttccgtgtcg cccttattcc ctttttgcg gcattttgcc
     3241 ttcctgtttt tgctcaccca gaaacgctgg tgaaagtaaa agatgctgaa gatcagttgg
     3301 gtgcacgagt gggttacatc gaactggatc tcaacagcgg taagatcctt gagagttttc
     3361 gccccgaaga acgttttcca atgatgagca cttttaaagt tctgctatgt ggcgcggtat
     3421 tatcccgtat tgacgccggg caagagcaac tcggtcgccg catacactat tctcagaatg
     3481 acttggttga gtactcacca gtcacagaaa agcatcttac ggatggcatg acagtaagag
     3541 aattatgcag tgctgccata accatgagtg ataacactgc ggccaactta cttctgacaa
```

```
3601 cgatcggagg accgaaggag ctaaccgctt ttttgcacaa catgggggat catgtaactc
3661 gccttgatcg ttgggaaccg gagctgaatg aagccatacc aaacgacgag cgtgacacca
3721 cgatgcctgt agcaatggca acaacgttgc gcaaactatt aactggcgaa ctacttactc
3781 tagcttcccg gcaacaatta atagactgga tggaggcgga taaagttgca ggaccacttc
3841 tgcgctcggc ccttccggct ggctggttta ttgctgataa atctggagcc ggtgagcgtg
3901 ggtctcgcgg tatcattgca gcactggggc cagatggtaa gccctcccgt atcgtagtta
3961 tctacacgac ggggagtcag gcaactatgg atgaacgaaa tagacagatc gctgagatag
4021 gtgcctcact gattaagcat tggtaactgt cagaccaagt ttactcatat atactttaga
4081 ttgatttaaa acttcatttt taatttaaaa ggatctaggt gaagatcctt tttgataatc
4141 tcatgaccaa aatcccttaa cgtgagtttt cgttccactg agcgtcagac cccgtagaaa
4201 agatcaaagg atcttcttga gatccttttt ttctgcgcgt aatctgctgc ttgcaaacaa
4261 aaaaaccacc gctaccagcg gtggtttgtt tgccggatca agagctacca actcttttte
4321 cgaaggtaac tggcttcagc agagcgcaga taccaaatac tgtccttcta gtgtagccgt
4381 agttaggcca ccacttcaag aactctgtag caccgcctac atacctcgct ctgctaatcc
4441 tgttaccagt ggctgctgcc agtggcgata agtcgtgtct taccgggttg gactcaagac
4501 gatagttacc ggataaggcg cagcggtcgg gctgaacggg gggttcgtgc acacagccca
4561 gcttggagcg aacgacctac accgaactga gatacctaca gcgtgagcat tgagaaagcg
4621 ccacgcttcc cgaagggaga aaggcggaca ggtatccggt aagcggcagg gtcggaacag
4681 gagagcgcac gagggagctt ccagggggaa acgcctggta tctttatagt cctgtcgggt
4741 ttcgccacct ctgacttgag cgtcgatttt tgtgatgctc gtcaggggggg cggagcctat
4801 ggaaaaacgc cagcaacgcg gccttttttac ggttcctggc cttttgctgg ccttttgctc
4861 acatgttctt tcctgcgtta tcccctgatt ctgtggataa ccgtattacc gcctttgagt
4921 gagctgatac cgctcgccgc agccgaacga ccgagcgcag cgagtcagtg agcgaggaag
4981 cggaagagcg cccaatacgc aaaccgcctc tccccgcgcg ttggccgatt cattaatgca
5041 gctggcacga caggtttccc gactggaaag cgggcagtga gcgcaacgca attaatgtga
5101 gttagctcac tcattaggca ccccaggctt tacactttat gcttccggct cgtatgttgt
5161 gtggaattgt gagcggataa caatttcaca caggaaacag ctatgaccat gattacgcca
5221 agcttgtttg tttaaactaa aggaggccgc aaaggttcta ttggcttagt gccgagtttg
5281 gtggtccatt gttcacttga tggtatgtaa tgaactgtag ctacaataga gcagttgacg
5341 atatagccct ttgtagtgct ctatagacta ttgtgtgtcg taacattgga gtgccaacct
5401 acctctagca acacactcat tcacaagaga gaaacgacac agaagctgtt gttaatcagg
5461 cagcaggtat tccattgccc tttgctctcg ttaagtgatc tgatttcact cgctacttct
5521 tactgagcaa atacaaccct cttcatcggt acattatata aagtcatcca ccactagcca
5581 ctactttcgc caccaagatg ggcaagaaga accgcttcaa aaaggtgagt cgagaacttc
5641 attttggccg caagttggag caatcaccta ctctctctct cgacggagag gctcaattgg
5701 agcggtttgc tcgaaaccag tggtaccgct actttgtgca gattgaggcc aagtaccgaa
5761 tggtatccaa cgacgatgag atcacggctg cccaccaaac catgatgttg cttggacaac
5821 ctgttggcat ttccctggac tccgagctcg gtcagaagat tgaggctgag ctgcagggac
5881 aacgacagga acaggaggcg gacgctcatc cagtggaaac tgtcaacccg cccagaagac
5941 catggagacg gtcaagaaag gcgtagctac ggtgtaactt gggccaaacc acgtgacagt
6001 cacatgattc accgctcata cgacatcgca cagacaccag tatatagata atatatttgt
6061 actctatctt aatctgaaag aggactcatc ctcatatcgt ttggaagcag cagtcataca
6121 tgttctgatc tggaagtcga cgaccgggat ccacttaacg ttactgaaat catcaaacag
6181 cttgacgaat ctggatataa gatcgttggt gtcgatgtca gctccggagt tgagacaaat
6241 ggtgttcagg atctcgataa gatacgttca tttgtccaag cagcaaagag tgccttctag
6301 tgatttaata gctccatgtc aacaagaata aaacgcgttt cgggtttacc tcttccagat
6361 acagctcatc tgcaatgcat taatgcattg gacctcgcaa ccctagtacg cccttcaggc
6421 tccggcgaag cagaagaata gcttagcaga gtctattttc attttcggga gacgagatca
6481 agcagatcaa cggtcgtaga cagacgacga tgacgatggt ggaagaatac aggaaaggga
6541 ttgtgtagta ttgttggact ttgctctcaa atctgatctg aattttcctt gaatttcctc
6601 aataatctca atattcacga gtttgcagta ggcggttctt gacgtgctct tccccatgta
6661 caaatccaat tcatcaagat caaatgcacg cacggaaatc gaactcccaa ccacacaccg
6721 cccatctgca taaactctaa atgcggtgaa atttgtaact tctgaaagac acgattttac
6781 acactcacgc aaccacctga aacgccatgt tctacaaagt cacacaggtc cgatctatcc
6841 tgcgaatgcc gtggagaaca cgtgacgtgc tcaagaccct cggtctgggc aaggtcggct
6901 cccagaagta ctacaaggtg tctcctggta tcgccggtca gctctacaag gtcaaggagc
6961 tggtcaacgt ggagatggcc gaccagtaca agtccaaggc ccagatcaag gccgaccgaa
7021 aaaccgagcc cggattctcc gttgtcggtt ctaagcgagc ctaagagtga gacagaatga
7081 tctcaaacga cacaaaacga caggctgatt tgaaacaaca caacacaagt gcgagccgat
7141 cttgcggatc tggacactcg catgccccca gttgcaaaag ttgacacaac tctagatctg
7201 cttccaaata tagaatcata acaagggtta gggtgtgatt atataatatt ggtcttaatt
7261 gatgtgctag ggctttaaaa gttggttaaa ataacgctct aatgcctttt taatatattg
7321 tctttttcaa aatctcaaat cggacacttc ttcgtgtatg agactccatt ttttggctcc
7381 gtcacgtgat atgtattatc agctatagtg gtgtaaacaa agtttttac tagctgtaat
7441 ggcatttttgt cggagtggta aatcgccttc ttgttgtgcg ttcgagttct ggactctgca
7501 ctgggctact ttgaaaaata cctctaatgc gccgatggtt tagtggtaaa atccatcgtt
7561 gccatcgatg ggcccccggt tcgattccgg gtcggcgcag gttgacgttg ctccatacaa
7621 gccaaccagt tttagagcta gaaatagcaa gttaaaataa ggctagtccg ttatcaactt
7681 gaaaaagtgg caccgagtcg gtgctttttt ttgtcgacag agaccgggtt ggcggcgtat
7741 ttgtgtccca aaaaacagcc ccaattgccc caattgaccc caaattgacc cagtagcggg
7801 cccaaccccg gcgagagccc ccttcacccc acatatcaaa cctcccccgg ttcccacact
7861 tgccgttaag ggcgtagggt actgcagtct ggaatctacg cttgttcaga ctttgtacta
7921 gtttctttgt ctggccatcc gggtaaccca tgccggacgc aaaatagact actgaaaatt
7981 tttttgcttt gtggttggga ctttagccaa gggtataaaa gaccaccgtc cccgaattac
8041 ctttcctctt cttttctctc tctccttgtc aactcacacc cgaaatcgtt aagcatttcc
8101 ttctgagtat aagaatcatt caaaatggtg agtttcagag gcagcagcaa ttgccacggg
8161 ctttgagcac acggccgggt gtggtcccat tcccatcgac acaagacgcc acgtcatccg
8221 accagcactt tttgcagtac taaccgcagg acaagaagta ctccattggg ctcgatatcg
8281 gcacaaacag cgtcggctgg gccgtcatta cggacgagta caaggtgccg agcaaaaaat
8341 tcaaagttct gggcaatacc gatcgccaca gcataaagaa gaacctcatt ggcgccctcc
8401 tgttcgactc cggggagacg gccgaagcca cgcggctcaa aagaacagca cggcgcagat
8461 atacccgcag aaagaatcgg atctgctacc tgcaggagat ctttagtaat gagatggcta
8521 aggtggatga ctctttcttc cataggctgg aggagtcctt tttggtggag gaggataaaa
8581 agcacgagcg ccacccaatc tttggcaata tcgtggacga ggtggcgtac catgaaaagt
8641 acccaaccat atatcatctg aggaagaagc ttgtagacag tactgataag gctgacttgc
8701 ggttgatcta tctcgcgctg gcgcatatga tcaaatttcg gggacacttc ctcatcgagg
8761 gggacctgaa cccagacaac agcgatgttg acaaactctt tatccaactg gttcagactt
8821 acaatcagct tttcgaagag aacccgatca acgcatccgg agttgacgcc aaagcaatcc
8881 tgagcgctag gctgtccaaa tcccggcggc tcgaaaacct catcgcacag ctccctgggg
8941 agaagaagaa cggcctgttt ggtaatctta tcgccctgtc actcgggctg acccccaact
9001 ttaaatctaa cttcgacctg gccgaagatg ccaagcttca actgagcaaa gacacctacg
9061 atgatgatct cgacaatctg ctggcccaga tcggcgacca gtacgcagac ctttttttgg
9121 cggcaaagaa cctgtcagac gccattctgc tgagtgatat tctgcgagtg aacacggaga
9181 tcaccaaagc tccgctgagc gctagtatga tcaagcgcta tgatgagcac caccaagact
9241 tgactttgct gaaggccctt gtcagacagc aactgcctga gaagtacaag gaaatttttct
```

```
     9301 tcgatcagtc taaaaatggc tacgccggat acattgacgg cggagcaagc caggaggaat
     9361 tttacaaatt tattaagccc atcttggaaa aaatggacgg caccgaggag ctgctggtaa
     9421 agcttaacag agaagatctg ttgcgcaaac agcgcacttt cgacaatgga agcatccccc
     9481 accagattca cctgggcgaa ctgcacgcta tcctcaggcg gcaagaggat ttctaccccc
     9541 ttttgaaaga taacagggaa aagattgaga aaatcctcac atttcggata ccctactatg
     9601 taggcccccc cgcccgggga aattccagat tcgcgtggat gactcgcaaa tcagaagaga
     9661 ccatcactcc ctggaacttc gaggaagtcg tggataaggg ggcctctgcc cagtccttca
     9721 tcgaaaggat gactaacttt gataaaaatc tgcctaacga aaaggtgctt cctaaacact
     9781 ctctgctgta cgagtacttc acagtttata acgagctcac caaggtcaaa tacgtcacag
     9841 aagggatgag aaagccagca ttcctgtctg gagagcagaa gaaagctatc gtggacctcc
     9901 tcttcaagac gaaccggaaa gttaccgtga aacagctcaa agaagactat ttcaaaaaga
     9961 ttgaatgttt cgactctgtt gaaatcagcg gagtggagga tcgcttcaac gcatccctgg
    10021 gaacgtatca cgatctcctg aaaatcatta aagacaagga cttcctggac aatgaggaga
    10081 acgaggacat tcttgaggac attgtcctca cccttacgtt gtttgaagat agggagatga
    10141 ttgaagaacg cttgaaaact tacgctcatc tcttcgacga caaagtcatg aaacagctca
    10201 agaggcgccg atatacagga tgggggcggc tgtcaagaaa actgatcaat gggatccgag
    10261 acaagcagag tggaaagaca atcctggatt ttcttaagtc cgatggattt gccaaccgga
    10321 acttcatgca gttgatccat gatgactctc tcacctttaa ggaggacatc cagaaagcac
    10381 aagtttctgg ccaggagggac agtcttcacg agcacatcgc taatcttgca ggtagcccag
    10441 ctatcaaaaa gggaatactg cagaccgtta aggtcgtgga tgaactcgtc aaagtaatgg
    10501 gaaggcataa gcccgagaat atcgttatcg agatggcccg agagaaccaa actacccaga
    10561 agggacagaa gaacagtagg gaaaggatga agaggattga agagggtata aaagaactgg
    10621 ggtcccaaat ccttaaggaa cacccagttg aaaacaccca gcttcagaat gagaagctct
    10681 acctgtacta cctgcagaac ggcagggaca tgtacgtgga tcaggaactg gacatcaatc
    10741 ggctctccga ctacgacgtg gatcatatcg tgccccagtc ttttctcaaa gatgattcta
    10801 ttgataataa agtgttgaca agatccgata aaaatagagg gaagagtgat aacgtcccct
    10861 cagaagaagt tgtcaagaaa atgaaaaatt attggcggca gctgctgaac gccaaactga
    10921 tcacacaacg gaagttcgat aatctgacta aggctgaacg aggtggcctg tctgagttgg
    10981 ataaagccgg cttcatcaaa aggcagcttg ttgagacacg ccagatcacc aagcacgtgg
    11041 cccaaattct cgattcacgc atgaacacca agtacgatga aaatgacaaa ctgattcgag
    11101 aggtgaaagt tattactctg aagtctaagc tggtctcaga tttcagaaag gactttcagt
    11161 tttataaggt gagagagatc aacaattacc accatgcgca tgatgcctac ctgaatgcag
    11221 tggtaggcac tgcacttatc aaaaaatatc ccaagcttga atctgaattt gtttacggag
    11281 actataaagt gtacgatgtt aggaaaatga tcgcaaagtc tgagcaggaa ataggcaagg
    11341 ccaccgctaa gtacttcttt tacagcaata ttatgaattt tttcaagacc gagattacac
    11401 tggccaatgg agagattcgg aagcgaccac ttatcgaaac aaacggagaa acaggagaaa
    11461 tcgtgtggga caagggtagg gatttcgcga cagtccggaa ggtcctgtcc atgccgcagg
    11521 tgaacatcgt taaaaagacc gaagtacaga ccggaggctt ctccaaggaa agtatcctcc
    11581 cgaaaaggaa cagcgacaag ctgatcgcac gcaaaaaaga ttgggacccc aagaaatacg
    11641 gcggattcga ttctcctaca gtcgcttaca gtgtactggt tgtggccaaa gtggagaaag
    11701 ggaagtctaa aaaactcaaa agcgtcaagg aactgctggg catcacaatc atggagcgat
    11761 caagcttcga aaaaaacccc atcgactttc tcgaggcgaa aggatataaa gaggtcaaaa
    11821 aagacctcat cattaagctt cccaagtact ctctctttga gcttgaaaac ggccggaaac
    11881 gaatgctcgc tagtgcgggc gagctgcaga aaggtaacga gctggcactg ccctctaaat
    11941 acgttaattt cttgtatctg gccagccact atgaaaagct caaagggtct cccgaagata
    12001 atgagcagaa gcagctgttc gtggaacaac acaaacacta ccttgatgag atcatcgagc
    12061 aaataagcga attctccaaa agagtgatcc tcgccgacgc taacctcgat aaggtgcttt
    12121 ctgcttacaa taagcacagg gataagccca tcagggagca ggcagaaaac attatccact
    12181 tgtttactct gaccaacttg ggcgcgcctg cagccttcaa gtacttcgac accaccatag
    12241 acagaaagcg gtacacctct acaaaggagg tcctggacgc cacactgatt catcagtcaa
    12301 ttacggggct ctatgaaaca agaatcgacc tctctcagct cggtggagac agcagggctg
    12361 accccaagaa gaagaggaag gtgtgatttt ggacctcgag tcattggacc tcgagtcatg
    12421 taattagtta tgtcacgctt acattcacgc cctcccccca catccgctct aaccgaaaag
    12481 gaaggagtta gacaacctga agtctaggtc cctatttatt tttttatagt tatgttagta
    12541 ttaagaacgt tatttatatt tcaaattttt ctttttttc tgtacagacg cgtgatcccc
    12601 ctttcatcaa atttagggat gccatcaact ttcagttcat aattaatatc ttaccaaatt
    12661 aggtaatctg caaaagttca gactgtgaaa tgtaacattt tatatatcaa gctctattta
    12721 acgcctcaca gtagttaaca taaagagata cagaattgtc gtgtcagtgt atactatcca
    12781 tgtgtatact ctggatatcc atttgtattc cattatctac gaaaagcggg taccgagctc
    12841 gaattcactg gccgtcgttt tacaacgtcg tgactgggaa aaccctggcg ttacccaact
    12901 taatcgcctt gcagcacatc ccccttcgc cagctggcgt aatagcgaag aggcccgcac
    12961 cgatcgccct tcccaacagt tgcgcagcct gaatggcgaa tggcgcctga tgcggtattt
    13021 tctccttacg catctgtgcg gtatttcaca ccgcatatgg tgcactctca gtacaatctg
    13081 ctctgatgcc gcatagttaa gccagccccg acacccgcca acacccgctg acgcgccctg
    13141 acgggcttgt ctg
//
```

A.4 Bioscreen sample outline

Table A.3: Sample outline of well 1-100 in bioscreen round 1

1 - 10	11 - 20	21 - 30	31 - 40	41 - 50	51 - 60	61 - 70	71 - 80	81- 90	91 - 100
S.1.1	S.2.1	S.3.1	S.4.1	S.5.1	S.1.1	S.2.1	S.3.1	S.4.1	S.5.1
S.1.2	S.2.2	S.3.2	S.4.2	S.5.2	S.1.2	S.2.2	S.3.2	S.4.2	S.5.2
S.1.3	S.2.3	S.3.3	S.4.3	S.5.3	S.1.3	S.2.3	S.3.3	S.4.3	S.5.3
S.1.4	S.2.4	S.3.4	S.4.4	S.5.4	S.1.4	S.2.4	S.3.4	S.4.4	S.5.4
S.1.5	S.2.5	S.3.5	S.4.5	S.5.5	S.1.5	S.2.5	S.3.5	S.4.5	S.5.5
S.1.6	S.2.6	S.3.6	S.4.6	S.5.6	S.1.6	S.2.6	S.3.6	S.4.6	S.5.6
S.1.7	S.2.7	S.3.7	S.4.7	S.5.7	S.1.7	S.2.7	S.3.7	S.4.7	S.5.7
S.1.8	S.2.8	S.3.8	S.4.8	S.5.8	S.1.8	S.2.8	S.3.8	S.4.8	S.5.8
S.1.9	S.2.9	S.3.9	S.4.9	S.5.9	S.1.9	S.2.9	S.3.9	S.4.9	S.5.9
S.1.10	S.2.10	S.3.10	S.4.10	S.5.10	S.1.10	S.2.10	S.3.10	S.4.10	S.5.10

Table A.4: Sample outline of well 181-185 and 191-195 in bioscreen round 1. The rest of the wells were used by Xiaojun Ji

181 - 185	191 -195
WT	Blank
WT	Blank
WT	Blank
WT	Blank
WT	Blank

Table A.5: Sample outline of well 1-100 in bioscreen round 2

1 - 10	11 - 20	21 - 30	31 - 40	41 - 50	51 - 60	61 - 70	71 - 80	81- 90	91 - 100
S.1.3	S.1.6	S.1.8	S.2.2	S.2.5	S.3.1	S.3.3	S.3.10	S.4.4	S.4.8
S.1.3	S.1.6	S.1.8	S.2.2	S.2.5	S.3.1	S.3.3	S.3.10	S.4.4	S.4.8
S.1.3	S.1.6	S.1.8	S.2.2	S.2.5	S.3.1	S.3.3	S.3.10	S.4.4	S.4.8
S.1.3	S.1.6	S.1.8	S.2.2	S.2.5	S.3.1	S.3.3	S.3.10	S.4.4	S.4.8
S.1.3	S.1.6	S.1.8	S.2.2	S.2.5	S.3.1	S.3.3	S.3.10	S.4.4	S.4.8
S.1.4	S.1.7	S.2.1	S.2.3	S.2.8	S.3.2	S.3.7	S.4.3	S.4.7	S.4.10
S.1.4	S.1.7	S.2.1	S.2.3	S.2.8	S.3.2	S.3.7	S.4.3	S.4.7	S.4.10
S.1.4	S.1.7	S.2.1	S.2.3	S.2.8	S.3.2	S.3.7	S.4.3	S.4.7	S.4.10
S.1.4	S.1.7	S.2.1	S.2.3	S.2.8	S.3.2	S.3.7	S.4.3	S.4.7	S.4.10
S.1.4	S.1.7	S.2.1	S.2.3	S.2.8	S.3.2	S.3.7	S.4.3	S.4.7	S.4.10

Table A.6: Sample outline of well 101-200 in bioscreen round 2

101 - 110	111 - 120	121 -130	171 - 180	181 - 190	191 -200
S.5.3	S.5.7	S.5.9	WT.1	WT.2	Blank
S.5.3	S.5.7	S.5.9	WT.1	WT.2	Blank
S.5.3	S.5.7	S.5.9	WT.1	WT.2	Blank
S.5.3	S.5.7	S.5.9	WT.1	WT.2	Blank
S.5.3	S.5.7	S.5.9	WT.1	WT.2	Blank
S.5.6	S.5.8				Blank
S.5.6	S.5.8				Blank
S.5.6	S.5.8				Blank
S.5.6	S.5.8				Blank
S.5.6	S.5.8				Blank

B

Appendix 2

Protocols used in this study

B.1 Genomic extraction for colony PCR

- Prepare eppendorf tubes with 25 µL of 20 mM NaOH for each colony you want to test.
- Add few glass beads (diameter
- With a sterile pipetting tip transfer a 5 µL of culture to the tube and resuspend the yeast cells in the solution.
- Boil the sample for 15 min (heating block at 95-100 °C)
- Vortex 15 sec
- Spin for 1 min a max. speed in a tabletop centrifuge.
- Transfer 0.5 µL of the solution to a 25 µL PCR reaction.

B.2 Phenol:chloroform genomic extraction

1. Inoculate 1 colony in 5 mL YPD. Grow overnight
2. Spin down cells 4000 g, 5 min, +4 °C
3. Discard supernatant and remove remaining traces with pipette.
4. Resuspend pellet in 2 mL TE buffer and transfer to 2 mL Eppendorf tube
5. Spin down cells 4000 g, 5 min, +4 °C
6. Discard supernatant and remove remaining traces with pipette.
7. Resuspend pellet in 1 mL buffer Y1 (1 M sorbitol, 100 mM EDTA, 14 mM 2-Mercaptoethanol (toxic))
8. Add 100 units of Zymolyase or lyticase
9. Incubate 30 °C with gentle shaking (ex post-stain shaker or normal shaker with 70 rpm) for 1-2 hours
10. Check spheroplasting by measuring OD of in 1 M sorbitol and in 5 % SDS (ex 50 µL sample + 950 µL sorbitol or SDS)
    (a) SDS sample should lyse, yielding high OD reduction
    (b) Handle spheroplasts gently
11. Pellet spheroplasts at 500 g, 10 min, +4 °C
12. Remove supernatant and dissolve pellet in 600 µL lysis buffer (1 % SDS, 100 mM NaCl, 2 % TritonX-100, 1x TE (10 mM Tris-Cl, 1 mM EDTA, pH 8))
13. Add 12 µL RNase A
14. Incubate 50 °C for 1 hour with occasional inversions
15. Add 1 volume of phenol:chloroform:isoamyl and vortex 10-20 sec
16. Centrifuge 5 min 16 000 g room temperature and transfer upper phase to new tube.
    (a) Protein will form a visible white precipitate at the interface

17. Repeat step 15-16 until no protein is visible at the interface
18. Add 1 volume chloroform
19. Vortex 10-20 sec
20. Centrifuge 5 min 16 000g RT
21. Transfer upper phase to new tube
22. Add 2.5 volumes of 100 % ethanol.
23. Mix and spin down sample.
24. Place at -80 °C for 60 min (or -20 °C overnight).
25. Centrifuge 16 000 g, 20 min, +4 °C with hinges outward
26. Carefully remove supernatant without disturbing the pellet.
27. Wash pellet with 70 % ethanol (cold).
28. Centrifuge 16 000 g, 5 min, +4 °C with hinges outward
29. Wash again with 70 % ethanol (cold).
30. Pull off all ethanol with pipet tip (use table top centrifuge to collect all ethanol at the bottom of the tube).
31. Air dry pellet under flame for 5 min
32. Resuspend pellet in 50 µl of 1/10 TE (10 mM Tris, pH 7.5-8.0, 0.1 mM EDTA) at 55 °C. 1-2 hours.
33. Measure purity with nanodrop, concentration with Qubit and run 30-60 ng on 0.7 % agarose gel 80 V 1h 20 min
34. Store DNA @ -80 °C or -20 °C.

B.3 Yeast electroporation

10×TE buffer (pH 7.5): 100 mM Tris-HCl, 10 mM EDTA. Filter-sterilized
10× LiAc: 1 M LiAc, pH 7.5 (adjusted using HAc). Filter-sterilized
1 M DTT: stored at -20 °C. Filter-sterilized
Day 1
Pick up single colonies on plates to 5 ml YPD medium. Culture at 30 °C for 12-16 h.
Day 2

1. Inoculate into 50 ml YPD medium in flask. Culture at 30 °C for 6-9 h.
   ««Put sterilized water on ice »»
2. When OD is 0.5-1.2, transfer cell culture into 50 ml cap tube (sterilized).
   «« From here cells should be always on ice »»
3. Collect cells by centrifugation (1100 g, 4 °C, 5 min). Decant supernatant.
4. Re-suspend cells with 20 ml of sterilized H2O (ice-cold). Mix with pipetman. Centrifuge and decant supernatant.
5. Treat cells with 20 ml of 0.1 M LiAc (16 ml 1 M sorbitol plus 2 ml 10×TE buffer plus 2 ml 1 M LiAc) at 30° C for 30 min. Add 0.2 ml 1 M DTT and keep cells at 30 °C for 15 min. Centrifuge and decant supernatant.
   ««Put 1 M sorbitol on ice »»
6. Wash cells twice with 20 ml of 1 M sorbitol (ice-cold). Centrifuge and decant supernatant.
7. Re-suspend cells with 100-200 µl of sterilized ice-cold 1 M sorbitol (final OD=100-200).
8. Take 50 µl of suspended cells into a new 1.5 ml tube on ice.

9. Add 5 µl fragment DNA (> 200 ng/µl). Mix with pipetman and keep on ice for 15 min. Transfer all to a sterilized cuvette (green cap). Add 1 ml of cold 1 M sorbitol to new labeled tubes (used later).
10. Set the cuvette in the holder of Micro Pulser Electroporator. Chose "Manual", and set voltage at 1.5 kV. Push the pulse button. Read "Time / ms", if it is between 4.0-6.0, this process is successful.
11. Add 1 ml of cold 1 M sorbitol, immediately after the pulse. Mix well by pipetting up and down. (After this, it is OK to be at RT.) Transfer all to the sorbitol tube ASAP.
12. Incubate at 30 °C for 1-3 h. Centrifuge (3000 g, 1 min) to 150 µl and then spread cells on selection plates. Make a negative control plate. Place the plates at 30 °C air incubator.