TECHNISCHE UNIVERSITÄT WIEN
Vienna University of Technology

Diplomarbeit

Characterization of the phenotypic diversity of *Komagataella* strains for biotechnological applications

Durchgeführt am

Institut für Verfahrenstechnik, Umwelttechnik und technische Biowissenschaften

der Technischen Universität Wien

Zur Erlangung des akademischen Grades

Diplom Ingenieur

Unter der Anleitung von

Univ.Prof. Mag. Dr.rer.nat. Robert Mach

Univ.Prof. Dipl.-Ing. Dr.nat.techn. Diethard Mattanovich

und unter der Betreuung von

Dipl.-Ing. Lina Heistinger, Ph.D.

Vorgelegt von

Daniel Koizar, 001125852

Wien, 29.04.2021

Statutory Declaration

I declare in lieu of an oath that I have written this master thesis independently and performed the associated research by myself using no other source than the declared literature.

I confirm that this work has not been submitted elsewhere for examination purposes, neither in this nor in any other country.

Date: _____ Signature:_____

Acknowledgments

First, I would like to thank my supervisors Prof. Diethard Mattanovich at the University of Natural Resources and Life Sciences, and Prof. Robert Mach at the University of Technology in Vienna, for guiding me through this thesis and the support in every situation.

Special thanks to Lina Heistinger, who was always in the lab, ready to help and answer every question. Her massive support and knowledge made it easy for me to enjoy the exciting work in this research group.
I also want to thank the whole Pichia group for the great atmosphere and fun discussions throughout my time at BOKU. Especially Petra Cazzanelli, who introduced me and helped me in my daily lab work.

Finally, I want to thank my family, girlfriend, and friends who supported me all the years and motivated me during challenging times. Not to forget my Dad, who would be very proud that I made it.

ABSTRACT

The methylotrophic yeast *Pichia pastoris* is an important model organism for heterologous protein production. It belongs to the genus *Komagataella*, which is now divided into seven different *Komagataella* species. The main focus of this thesis was the characterization of growth and mating behavior of 25 different *Komagataella* strains. The growth characterization was realized with spotting assays using various carbon sources and stressors at two different temperatures, which gives a good selection of strains with useful properties for further analysis. Not only growth behavior but also mating between different strains and species is not well studied for *Komagataella*. For the mating experiments, strains were transformed with a vector containing antibiotic resistance and a reporter protein, either geneticin/eGFP or nourseothricin/mCherry. Characterization of mating and the ability to undergo meiosis is essential for strain characterization, especially in terms of species definition but also for biotechnological applications. Based on the obtained results, two strains with differences in xylose utilization were mated, and growth curves of the F1 generation were recorded. This small-scale experiment showed the possibility to create descendants harboring the benefits of both parents as well as the analysis of complex traits and their genetic background.

CONTENTS

1. INTRODUCTION

1.1. *Komagataella phaffii* or *Pichia pastoris*, resp.

In 1920, Alexandre Guilliermond isolated the yeast *Zygosaccharomyces* (Guilliermond A. 1920), which was renamed by Herman Phaff to *Pichia pastoris* in 1956 (Phaff et al. 1956). After decades, *P. pastoris* was reallocated to the genus *Komagataella* based on 18S and 26S rRNAs sequences (Yamada et al. 1995). Today, this genus consists of seven different species (Naumov et al. 2018). Industrial relevant strains belong to *K. phaffii* or *K. pastoris*, and besides *Saccharomyces cerevisiae*, these species are important host organisms for heterologous protein production and are well studied in this context. However, so far little research has been done with strains of the other *Komagataella* species. As a methylotrophic yeast, *K. phaffii* can use methanol as a carbon and energy source while using an aerobic metabolism. Its high-density growth on substrates like glucose, glycerol or methanol and the ability to perform post-translational modifications (PTMs), as well as the low risk of pathogenic virus contamination, have significant advantages for industry. The availability of a stable expression system, a vast range of genetic modification methods, and well-annotated genome sequences make *Komagataella* also suitable for scientific research. The most commonly used promotor system is derived from the alcohol oxidase 1 (*AOX1*) gene. The *AOX1* promotor is strongly induced on methanol, which makes a growth decoupled protein production possible (Cregg et al. 1985).

Nevertheless, strain improvement is one of the big challenges for process engineering, and there are multiple strategies to do so. Industrially used strains represent only a small fraction of all naturally available species, which harbor a massive potential in their genetic diversity (Steensels et al. 2014).

1.2. The environment of *Komagataella* strains

Most of the *Komagataella* species were isolated from tree exudates and decomposing wood, like the first strain of *Komagataella* isolated by Guillermond from a chestnut tree near Lyon, France. Such habitats provide a large supply of sugars, amino acids, and other organic compounds that serve as a perfect environment for yeast growth.

While most of the species of *K. pastoris* and *K. pseudopastoris* were isolated in European woods in Hungary, the other species, *K. phaffii* (Ontario and Arizona), *K. populi* (Illinois), *K. ulmi* (Ontario), *K. kurtzmanii* (Arizona), and *K. mondaviorum* (California) were isolated from different trees in the USA and Canada (Sibirny 2019). One theory behind the limited area of distribution is that some yeasts are associated with specific insects. These insects are responsible for the dissemination of the yeasts by introducing them into tree fluxes. The limited habitat of the insects results in a regional distribution, as is the case with *Komagataella spp.* (Mrak et al. 1978).

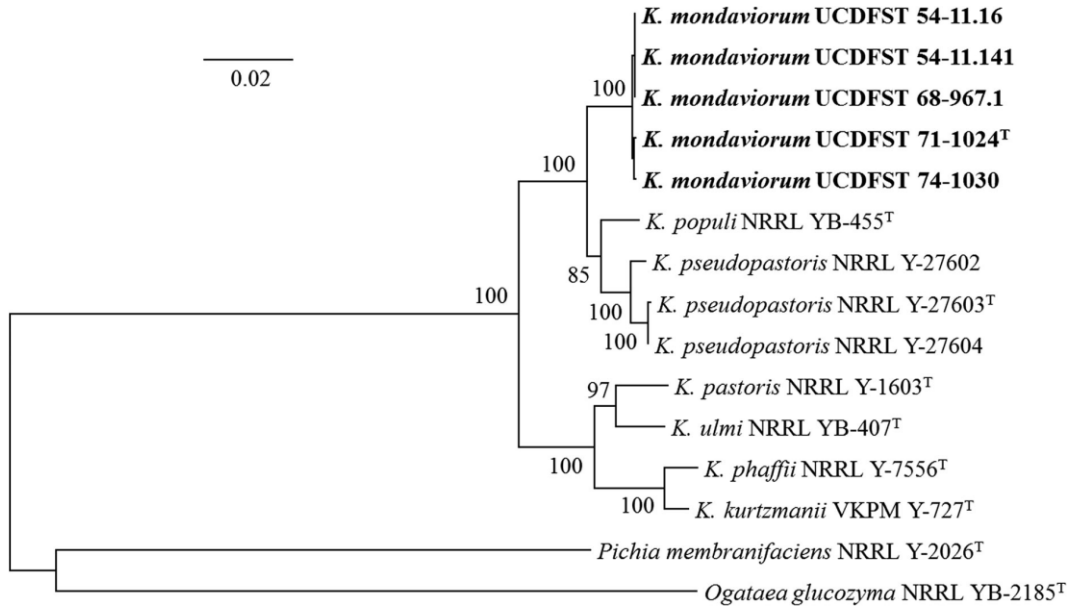


Figure 1: Neighbour-joining tree of the genus Komagataella based on combined sequences for D1/D2 and ITS1-5.8S-ITS2 rDNA, translation elongation factor-1a, and RNA polymerase II (subunit RPB1). Figure taken from (Naumov et al. 2018).

The phylogenetic tree in Figure 1 shows the genetic relationship between the seven *Komagataella* species based on conserved marker gene sequences, analyzed with the neighbor-joining method. It splits them into two major clades, one with *K. mondaviorum*, *K. populi*, *K. pseudopastoris,* and the other clade with *K. pastoris*, *K. ulmi*, *K. phaffii*, *K. kurtzmanii*. Even though *K. pastoris* and *K. pseudopastoris* were both isolated in Europe, they are found in different clades. No *Komagataella* isolates have been found in Asia or South America so far.

The taxonomic classification changes with the use of different species concepts applied to yeasts. There are several strategies, including morphology, physiology, genetics, biochemistry, ecology, and molecular genetics for taxonomic evaluation (taxonomy | Definition, Examples, Levels, & Classification 2020). There are three major classification concepts to represent the most likely relationship of a specific group of organisms.

Especially in early studies, the phenetic species concept, including physiological and morphological attributes, was used. Unfortunately, the used growth tests lack compounds that are complicated to metabolize, which would increase the significance of such classifications. The second method is based on the ability to breed with other individuals, which indicates the same species. However, closely related yeast species are often also able to mate, so the progenies' viability is a requirement for a well-founded statement about taxonomy. This biological species concept only works for sexual yeast species and often conflicts with other concepts, like the phylogenetic species concept. The phylogenetic species concept is based solely on evolutionary parentage and is, therefore, less strict in the context of interspecies mating. Due to the lack of archaeological data, the application is primarily dependent on genetic data, which mostly comes from sequenced ribosomal RNA and DNA. The usual threshold obtained for different species is about 2-3 % difference in their base composition, while no difference in the base composition necessarily implicates the same species. While interspecies mating seems to correlate with DNA similarity, even small percentages of similarity are not a guarantee for preventing genetic exchange (Kurtzman et al. 2010) (Boekhout and Kurtzman 1996).

1.3. Strain diversity

Overall, most microorganisms in the environment are not cultivable in laboratory settings. However, the diversity in their characteristics could be relevant for industrial use. Standard methods to increase the diversity of laboratory strains, complying with the non-genetically modified (non-GMO) directives, are varying forms of mutagenesis, sexual hybridization, protoplast fusion, cytoduction, and evolutionary engineering (Figure 2).

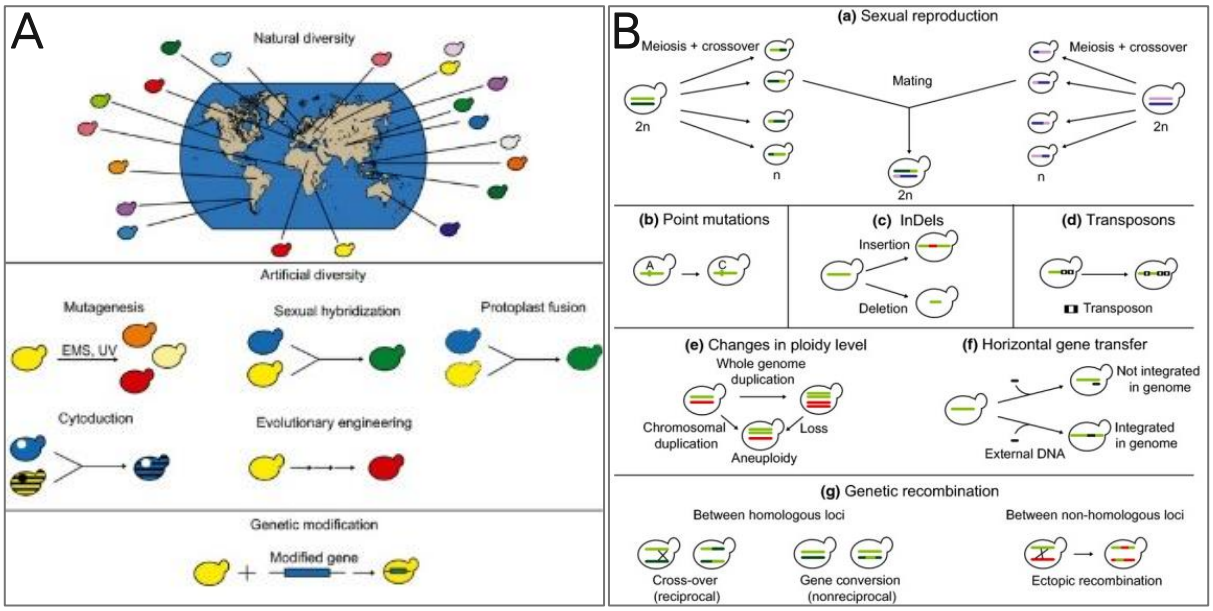


Figure 2: (A) Schematic figure about the different strategies to increase strain diversity. (B) Schematic figure about different genetic modification methods. Figures taken from (Steensels et al. 2014).

With these methods, the genetic variability follows natural processes for genomic modifications like sexual reproduction, point mutations, or gene loss over generations.

Yeast cells usually propagate by mitotic cell division. Under non-favorable conditions, many species can undergo sexual reproduction as a way to generate genetic diversity, as it is observed in higher eukaryotes. Sexual reproduction (or mating) can be used to breed strains with beneficial phenotypes for industrial applications. Also, crossing of strains with differences in a certain trait can be exploited to determine the genetic loci responsible for a specific quantitative trait with a method called quantitative trait loci (QTL) mapping. The knowledge on the genetic determinants of a phenotype enables targeted genetic engineering to create new strains with beneficial features and genetic diversity for industrial applications (Swinnen et al. 2012). Furthermore, it also helps to understand the mechanisms of genetic recombination during mating.

1.4. Mating

As mentioned above, yeasts like *S. cerevisiae* and *K. phaffii* reproduce almost exclusively by budding, an asexual process where the DNA is doubled, and cell division occurs. Spontaneous mutations drive the evolutionary trial of genetic diversification throughout the genome. Sexual reproduction is far less frequent and mostly induced by starvation conditions. For example, *Saccharomyces paradoxus* cells undergo meiosis with a probability of one meiotic reproduction for every 1000 mitotic cycles (Tsai et al. 2008). Sexual reproduction requires two cells with opposite mating types, which are called a and α for most yeast species. Each cell type secretes a specific mating factor pheromone, which binds to the receptors on the cell surface of the opposite mating type, initiating the formation of diploid cells. In *K. phaffii*, mating pheromone and receptor gene expression are regulated by specific transcription factors encoded in two *MAT* loci on chromosome 4, harboring two *MATa* and *MATα* genes, respectively. While the locus close to the telomeric region is downregulated, the *MAT* genes in the second locus are transcribed and determine the mating type (Heistinger et al. 2018). As a homothallic species, *K. phaffii* can also switch its mating type by exchanging the two genes in the *MAT* loci in a homologous recombination process (Hanson et al. 2014). It has been shown that under nitrogen starvation, the methylotrophic yeast *Ogataea polymorpha* activates mating-type switching through the activation of the transcription factor Efg1 (Hanson et al. 2017). The same nitrogen-limiting environment also forces the diploid cells to undergo meiosis. Unlike *S. cerevisiae*, diploid cells of *K. phaffii* are not stable and immediately undergo mitosis and sporulate by forming an ascus containing up to 4 spores. The spore wall, containing dityrosine and chitosan layers, makes the cells resistant to harsh conditions, a primary argument for the occurrence of a meiotic cell cycle (Briza et al. 1988).

1.5. Homologous recombination during meiosis

After mating and fusion of the nuclei, the diploid cell can either divide mitotically or undergo meiotic recombination. During the first phase of meiosis, pairs of homologous chromosomes align, and the DNA is exchanged by homologous recombination. It is reported that these recombination events do not only happen at random sites but hot spots related to double-strand breaks (Cao et al. 1990). There, reciprocal crossing over or non-reciprocal gene conversion occurs, and parts of the chromosome change their position after the structures are dissolved by double-strand break repair or synthesis-dependent-strand-annealing (Figure 3) (Duret and Galtier 2009).

Which parts of the chromosome are exchanged and how the cells control the crossing overs are still unknown for *Komagataella spp.* It has only been shown that meiotic recombination is approximately 3.5 times lower than for *S. cerevisiae*. Meiotic recombination around the centromeres is suppressed (~300 kb), which results in low sequence diversity around these areas (Braun-Galleani et al. 2019).

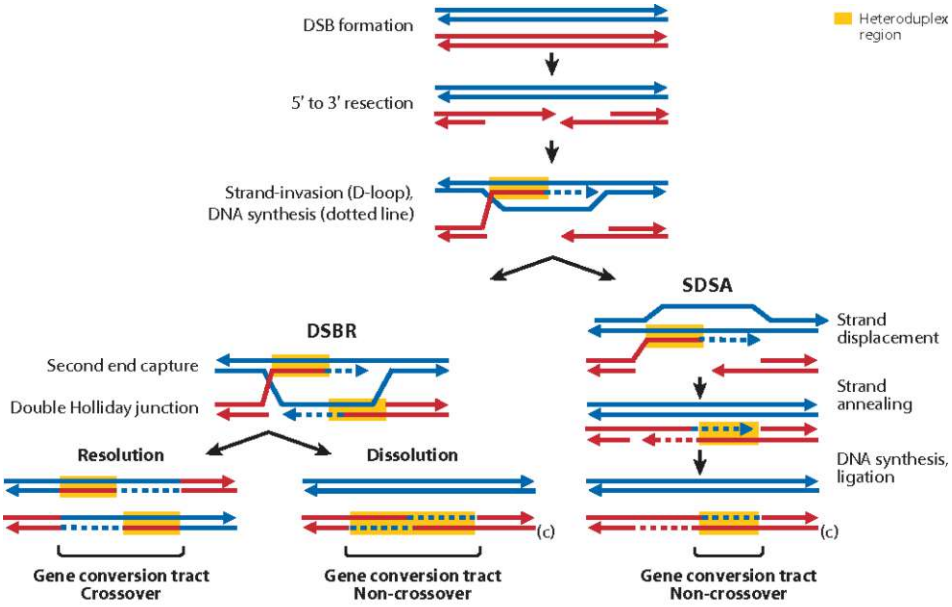


Figure 3: Homologous recombination at double-stranded recombination hotspots, dissolved by double-strand-break repair or synthesis-dependent-strand-annealing. Figure taken from (Duret and Galtier 2009).

After meiosis 1 ends and the chromosomes are separated by the spindle apparatus, meiosis 2 starts with the separation of the sister chromatids (Neiman 2005). Then, the pro-spore membrane is formed, and as meiosis 2 proceeds, the membrane expands and encloses the four haploid nuclei. The spore wall is formed around the immature spores inside the ascus's cytoplasm and after completion, the mother cell, still encapsulating the spores, collapses (Neiman 1998). Usually, the spores are then released from the ascus by bursting or self-digestion. The released haploid spores can now again divide or mate, depending on the surrounding conditions.

1.6. Aim of the thesis

The yeast genus *Komagataella* includes seven closely related species with different characteristics, which are potentially interesting for industrial use. In the first part, growth tests on different carbon sources as well as stressors were performed at different incubation temperatures to get an overview of the characteristics of 25 different *Komagataella* strains.

The second goal was to mate the most relevant strains, transformed with fluorescent markers and antibiotic resistance genes to test the compatibility within the different strains and species, as meiosis and the formation of viable spores are often impaired when crossing strains of different species. DNA content measurements by flow cytometry were used to analyze the mating results.

According to the observations during the first two series of experiments, strains with interesting properties were tested in more detail regarding their growth behavior and mating ability. By crossing species with different traits, it is possible to create progeny carrying diverse characteristics due to the genetic variation occurring during meiosis. This is not only of interest for optimizing industrially relevant characteristics but also to get an insight into the complex process of meiosis.

2. MATERIALS & METHODS

2.1. Strains

All yeast strains used in this master thesis are listed in table 1. For Golden Gate Cloning and plasmid amplification, the *Escherichia coli* strain DH10B was used.

Table 1: Komagataella strains

species name	strain code	place of origin	location
Komagataella phaffii	CBS 2612	*Quercus kelloggii*	California, USA
Komagataella phaffii	CBS 7435	unknown	unknown
Komagataella phaffii	85- 263.1	*Quercus emoryi*	Tucson, Arizona, USA
Komagataella phaffii	03- 318t1	*Quercus rubra*	Long Point, Ontario
Komagataella phaffii	03- 328y3	*Quercus rubra*	Long Point, Ontario
Komagataella phaffii	81- 86	*Quercus rubra*	unknown
Komagataella phaffii	81- 18	unknown	unknown
Komagataella phaffii	82- 16	*Quercus rubra*	Pinery Ontario
Komagataella phaffii	85- 348.1	*Quercus emoryi*	Tucson, Arizona, USA
Komagataella phaffii	85- 926.1	*Quercus emoryi*	Tucson, Arizona, USA
Komagataella phaffii	91- 119.3	*Quercus rubra*	Wilderness Area Pinery, Ontario
Komagataella phaffii	91- 132.2	*Quercus rubra*	South of Road Backus Woods, Ontario
Komagataella (species unknown)	HA71	unknown	unknown
Komagataella pastoris	DSMZ 70382 (CBS 704)	*Chestnut tree*	Lyon, France
Komagataella pastoris	DSMZ 70877	*Beech*	Berlin, Germany
Komagataella pastoris	CBS 9178	thick mass in the cavity of *Quercus sp.*	Pilis Mountain, Hungary
Komagataella pastoris	CBS 9173	exudate of *Acer platanoides*	Hungary
Komagataella pastoris	CBS 9185	exudate of *Fagus sylvatica*	Hungary
Komagataella ulmi	CBS 12361	elm tree exudate (*Ulmus americana*)	Peoria, Illinois, USA
Komagataella ulmi	91- 206.2	*Quercus rubra*	Long Point, Ontario
Komagataella ulmi	03- 338t1	*Quercus rubra*	Long Point, Ontario
Komagataella pseudopastoris	CBS 9187	rotten wood (cavity of *Salix alba*)	Dorog, Hungary
Komagataella pseudopastoris	CBS 9189	rotten wood (*Salix alba*)	Budapest, Hungary
Komagataella kurtzmanii	CBS 12817	*fir flux*	Catalina Mountains, Arizona, USA
Komagataella populi	CBS 12362	cottonwood tree exudate (*Populus deltoides*)	Peoria, Illinois, USA

2.2. Media

All media, except 50x TAE, 1.5 % agarose gels, and 500x Biotin, were sterilized by autoclaving at 121 °C for 20 min. Media containing agar-agar was melted in the microwave and cooled down to 60 °C before pouring the plates.

LB-Medium and LB-Agar

- 10 g soy peptone
- 5 g yeast extract
- 5 g NaCl
- Filled up to 900 g with AD
- pH set between 7.4-7.6 with NaOH
- 20 g agar-agar was added for LB-agar

YP-Medium and YP-Agar

- 20 g soy peptone
- 10 g yeast extract
- Filled up to 900 g with AD
- pH set between 7.4-7.6 with NaOH
- 20 g agar-agar was added for YP-agar
- 100 mL glucose (20%) or 50 mL xylose (20 %) was added after autoclaving

Acetate-Agar

- 5 g $C_2H_3NaO_2$
- 10 g KCl
- 20 g agar-agar
- Filled up to 900 g with AD
- 50 mL glucose (20%) was added after autoclaving

Agar for spotting Assays

- 5 g agar-agar
- 25 g potassium-phosphate buffer (1 M, pH 6)
- Filled up to 200 g with AD
- pH set to 6.5 with NaOH
- Prior use addition of
    - 25 mL YNB
    - 500 µL Biotin 0,2 g/L
    - 25 mL carbon source
        - Glucose 20 %
        - Xylose 20 %
        - Glycerol 10 %
    - 250 µL carbon source distributed on the surface of the already poured plate
        - Methanol 100 %
        - Ethanol 100 %

2.3. Stock solutions

10x Glucose 20 %

- 220 g D(+)-glucose monohydrate
- Filled up to 1000 g with AD

10x Xylose 20 %

- 200 g D-xylose
- Filled up to 1000 g with AD

10x Glycerol 20 %

- 126 g glycerol 100 %
- Filled up to 1000 g with AD

Sorbitol 1 M

- 182.2 g D-sorbitol
- Filled up to 1000 g with AD

10x PBS

- 2.4 g KH_2PO_4
- 18 g Na_2HPO_4*2 H_2O
- 2 g KCl
- 80 g NaCl
- Filled up to 1000 g with AD

500x Biotin 0.2 g/L

- 200 mg D-biotin
- Filled up to 1000 g with AD
- Sterile filtered and stored at 4 °C

YNB

- 34 g yeast nitrogen base w/o amino acids and ammonium sulfate
- 100 g NH_4SO_4
- Filled up to 1000 g with AD
- Sterile filtered and stored at 4 °C protected from light

50x TAE Buffer

- 242 g tris base
- 57.1 mL glacial acetic acid
- 100 mL EDTA solution pH 8
- Filled up to 1000 g with AD

1.5 % Agarose Gel

- 15 g agarose
- 20 mL 50x TAE
- 5 µL Midori Green (analytical gels) or 50 µL SYBR Safe (preparative gels)
- Filled up to 1000 mL with AD

TFB 1 solution

- 30 mM potassium acetate
- 10 mM $CaCl_2$
- 50 mM $MnCl_2$
- 100 mM RbCl
- pH set to 5.1 using acetic acid
- Sterile filtered and stored at 4 °C

TFB 2 solution

- 100 mM MOPS
- 75 mM $CaCl_2$
- 10 mM RbCl
- 15 % glycerol
- pH set to 6.5 using KOH
- Sterile filtered and stored at 4 °C

2.4. Antibiotic selection markers

The working concentrations for the required antibiotics are listed in the table below.

Table 2: Antibiotic concentrations

antibiotic	working concentration *E. coli* [µg/mL]	working concentration *Komagataella* [µg/mL]
Kanamycin	50	-
Geneticin	-	500
Nourseothricin	50	100

2.5. Golden Gate Cloning

Golden Gate assembly was used for generation of the vectors for integration into the different *Komagataella* strains. Golden Gate assembly uses type IIs restriction enzymes like *BsaI* and *BpiI*, cutting DNA outside their recognition sites and creating four base pair long overhangs. The T4 ligase fuses these overhangs (specific fusion sites) and thereby creates a scarless vector after ligation.

The Golden Gate library available in the group consists of three levels of backbone cassettes for modular cloning. The first level, the BB1, contains basic modules like promotors and terminators. The second level contains already assembled expression units with a certain promotor-CDS-terminator combination. Several of these constructs can be combined into a BB3 vector, representing the last level for transformation into yeast.

The final BB3 vectors used in this study were assembled directly from the BB1 vectors of the promotor, terminator and coding sequence (CDS) carrying *BpiI* recognition sites into a special BB3 with compatible fusion sites. The composition of the Golden Gate reaction mix and the parameters for the thermocycler are listed below and in Table 3.

Golden Gate reaction mix 20 µL:

- 2 µL restriction enzyme (20 U)
- 1 µL T4 ligase (1:10 diluted) (40 U)
- 2 µL CutSmart buffer (10x)
- 2 µL ATP (10x)
- 1 µL of the target vector (40 nM)
- 1 µL of each fragment (40 nM)
- Filled up to 20 µL with AD

Table 3: Golden Gate program

step	temperature	time	cycles
Cleavage	37 °C	5 min	20 x
Ligation	16 °C	5 min	
Cleavage	37 °C	10 min	
Inactivation of T4 ligase	55 °C	30 min	
Inactivation of restriction enzymes	80 °C	10 min	

2.6. *E. coli* transformation

2.6.1. Preparation of chemically competent E. coli

Materials and Procedure

- LB-agar
- LB-medium
- TFB 1 solution
- TFB 2 solution

A single colony *E. coli* DH10B streaked out on LB-agar was inoculated in 10 mL LB-medium overnight at 37 °C and 180 rpm. The next day, 2 mL were transferred into 200 mL LB-medium and incubated for two hours until OD_{600} reached 0.6.

After the cells cooled down on ice for 10 min and centrifuged for 5 min (4500 rpm) at 4 °C, the cells were resuspended in 0.4 volumes ice-cold TFB1. The centrifugation step was repeated, and the pellet resuspended in 1/25 volume of ice-cold TFB2. Aliquots of 100 µL were stored at -80 °C.

2.6.2. *E. coli* transformation

Frozen *E. coli* aliquots were thawed on ice, and after addition of the DNA for transformation, incubated for 30 min. After 90 s of heat shock at 42 °C, the mix was recovered on ice for another 5 min. The cells were mixed with 1 mL LB-medium and incubated at 37 °C for 45-60 min. The cells were plated on selective LB-agar and incubated overnight at 37 °C.

2.7. *Komagataella* transformation

2.7.1. Preparation of electro- competent *Komagataella spp.*

Materials and procedure

- YPD-agar
- YPD-medium
- 1 M HEPES, pH 8
- 1 M DTT
- 1 M Sorbitol
- AD

A single colony of a *Komagataella* strain streaked out on YPD-agar was inoculated in 10 mL YPD-medium overnight at 25 °C and 180 rpm. The amount of pre-culture required for the main culture was calculated according to the equation below and inoculated into 100 mL YPD-medium for cultivation at 25 °C and 180 rpm overnight.

$$V_{inoc}[\mu L] = \frac{OD_m * V_m}{e^{\mu * t}} * \frac{1000}{OD_{pre}}$$

Equation 1: Calculation of the inoculation volume for the main culture

OD_m.... OD_{600} of the main culture after time t (OD_{600} 3 is used for calculation)

V_m..... volume of the main culture [mL]

t..... incubation time of the main culture [h] (at least 15)

µ..... 0.28 h^{-1} for *Komagataella phaffii* wild type in YPD at 25°C

OD_{pre} .. OD_{600} of the pre-culture

After the desired OD_{600} was reached, the main culture was split into two falcon tubes (2x45 mL) and centrifuged for 5 min at 1500 x g and 4 °C. The pellet was resuspended in pre-warmed YPD/HEPES/DTT (20 mL YPD + 400 µL HEPES + 500 µL DTT) and incubated for 30 min at 25 °C and 180 rpm. Each falcon tube was mixed with 40 mL ice-cold AD and centrifuged for 5 min at 1500 x g and 4 °C. The pellets were combined, resuspended in 45 mL ice-cold HEPES (45 mL AD + 45 µL HEPES), and again centrifuged for 5 min at 1500 x g and 4 °C. 45 mL of ice-cold sorbitol was added to the pellet, centrifuged for 5 min at 1500 x g and 4 °C, and resuspended in 500 µL ice-cold sorbitol. Aliquots of 80 µL were stored at -80 °C.

2.7.2. Transformation

Purified and linearized plasmids were added to the electro-competent cells and incubated for 5 min on ice. The mixture was transferred into a pre-cooled electroporation cuvette, where the electroporation was done using 2 kV and 4 ms time. After electroporation, ice-cold YPD-medium was added to the cells and the culture was transferred into a microcentrifuge tube where the cells were regenerated at 28 °C for 1.5-3 h. The cells were plated on selective YPD-agar and incubated for 2 days at 30 °C or 3 days at 25°C.

2.8. Polymerase Chain Reaction (PCR)

The polymerase chain reaction uses a thermostable DNA polymerase (Q5 or Taq) and a forward and reverse primer binding on the template DNA to amplify a specific region of the DNA. In the first step, the DNA is denatured at 98 °C where the hydrogen bonds between the two strains are broken up. The second step is for primer annealing at the specific temperature, 5-10 °C lower than the melting temperature (T_m), according to the primers' base composition. The forward primer binds to the non-coding strand, and the reverse primer binds to the coding strand or 3´end of a gene. From these starting points on the third step, the elongation is initiated, and the polymerase amplifies the strands in 5´-3´ direction.

For complex templates, a touch-down (TD) PCR is preferred. It starts at higher annealing temperatures to avoid unspecific primer binding, and the temperature is lowered step by step until the optimal annealing temperature is reached. The composition of the PCR reaction mix and the thermocycler parameters are listed below and in table 4 and table 5.

PCR reaction mix 20 µL:

- 4 µL Q5 reaction buffer (5x)
- 4 µL GC-Enhancer (5x)
- 0.4 µL dNTPs (10 mM)
- 1 µL Primer fwd. (10 µM)
- 1 µL Primer rev. (10 µM)
- 0.2 µL Q5 polymerase
- 8.4 µL nuclease-free water
- 1 µL template DNA

Table 4: PCR program

step	temperature	time	cycles
First denaturation	98 °C	60 s	
Denaturation	98 °C	10 s	25-35x
Annealing	calculated[1]	20 s	
elongation	72 °C	30 s/kb	
Final elongation	72 °C	120 s	

Table 5: Touch-down PCR program

step	temperature	time	cycles
First denaturation	95 °C	30 s	
Annealing	calculated[1]; decreased by 1 °C after each cycle	30 s	10x
Elongation	72 °C	30 s/kb	
Denaturation	95 °C	30 s	
Annealing	calculated[1]	30 s	25x
Elongation	72 °C	30 s/kb	
Final elongation	72 °C	10 s	

2.8.1. Colony PCR (cPCR)

A single colony of a *Komagataella* strain was picked and suspended in 10 µL of 0.02 M NaOH. The mixture was heated to 99 °C for 10 min, and 1 µL was used as a TD-PCR template.

2.9. Gel electrophoresis

Materials and procedure

- Purple Loading Dye 6x
- 2-log DNA Ladder
- TAE Buffer 50x

[1] The annealing temperature is calculated on the website tmcalculator.neb.com (Biolabs 2020).

Gel electrophoresis is used to separate DNA fragments according to their size. By applying an electric field, the negatively charged DNA passes through the agarose gel and is divided by the filter effect into bigger and smaller molecules. The gels were cut to the specified pocket number required and placed in TAE buffer in the electrophoresis chamber. For analytical gels 130 V and preparative gels, 90 V were applied for 30-50 min. Before loading, the samples need to be mixed with a loading dye for visualization and to increase the density to ensure the sample sinks into the well. For size determination of the bands, 8 µL of DNA-ladder was also loaded. The specific pattern of the DNA-ladder is shown in figure 4 (Biolabs 2020b).

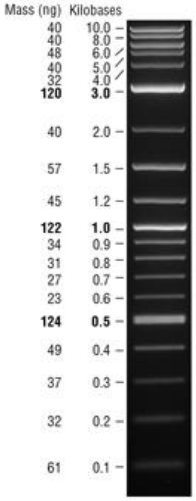


Figure 1: neb Quick-Load 1 kb Plus DNA Ladder

2.9.1. Extraction of DNA bands

The desired bands of the preparative gels were cut out under UV-light and purified with the "innuPREP DOUBLEpure" kit (Analytic Jena), according to the manufacturer´s instructions. The elution of the DNA was done with 30 µL RO-water.

2.10. Plasmid purification

For small scale plasmid preparation, 2 mL of LB-medium were inoculated with a single colony and incubated overnight at 37 °C and 180 rpm. The "Hi Yield Plasmid Mini Kit" (SLG) was used for purification following the manufacturer´s instruction, and the elution step was done with 50 µL RO-water.

Plasmid midi-preparation was started by inoculation of 50 mL LB-medium with a single colony and incubated overnight at 37 °C and 180 rpm. The "NucleoBond Xtra Midi Kit" (Machery-Nagel) was used for purification following the manufacturer´s instruction, and the rehydration step was done with 500 µL RO-water.

2.11. Restriction digest

To verify plasmid DNA after cloning, a restriction digest with one or more restriction enzymes is used to cut the plasmid at specific recognition sites. The cleaved DNA is then separated by gel electrophoresis showing a plasmid-specific pattern.

The reaction mix describes the procedure with one restriction enzyme. If more enzymes were used, the amount of RO-water was decreased by the additional amount of enzyme. The mixture was incubated at the recommended temperature of the restriction enzymes for 1 h.

Restriction reaction mix 10 µL:

- 2 µL plasmid DNA
- 0.25 µL restriction enzyme
- 1 µL appropriate buffer 10x
- 6.5 µL AD-water

2.12. Sequencing

For verification of plasmid sequences, a mix of 12 µL (80 ng/µL) per sample with 3 µL (10 µM) of the appropriate primer was sent to Microsynth AG for Sanger sequencing. Results were analyzed with CLC Main Workbench 8 (Qiagen).

2.13. Spotting assays

A single colony of a *Komagataella* strain streaked out on YPD-agar was picked and suspended in 1 mL of PBS. The suspension was diluted to reach an OD_{600} of 0.3, and 200 µL of this suspension was applied to the first row of a sterile 96 deep-well plate. For the dilution series, the next four rows were filled with 180 µL of PBS, and 20 µL were taken from the first row to the second, always diluting it 1:10 until the 5th row was reached.

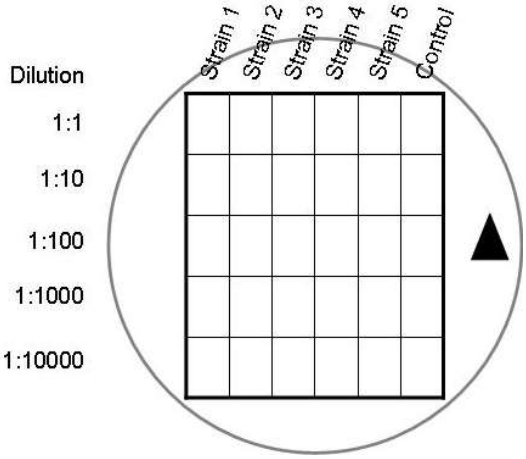


Figure 2: Pipetting scheme of the spotting assays

Using a multichannel pipette, it was possible to pipette 5 strains plus the control strain side-by-side onto a minimal YNB-agar plate with the respective carbon source. For pipetting, sealed filter tips were used to spot 4 µL of each dilution on the agar, according to figure 5. After the drops were dried, the plates were incubated at the appropriate conditions (30 °C and 37 °C) for at least two days. The growth was analyzed by scanning the plates and visual comparison between the spots.

2.14. Mating

For mating, the strains were streaked out on YPD-agar in parallel lines and incubated at 30 °C overnight. On day two, the cells were transferred onto acetate-agar plates with an autoclaved cloth wrapped around a stamp. The strains were crossed by turning the stamp by 90 ° after the first contact with the agar. After 3 days of incubation at 25 °C, the cells on acetate-agar plates were replicated on selective (2x G418 + NTC) YPD-agar plates and incubated for 3 days at 25 °C. On day 8, positive clones were picked from the crossing points, streaked out again on selective agar plates, and incubated at 30 °C for two days. For further analysis, the diploid cells were either replicated on acetate-agar for sporulation or inoculated in selective YPD medium in a 24-well plate for DNA content staining and flow cytometry measurement.

2.15. Random spore analysis

Materials and Procedure

- Diethyl ether ≥ 99.5 %

The acetate-agar plates from the mating step were incubated for 3 days at 25 °C, and then a small number of spores were scratched from the plate and resuspended in 700 µL RO-water. For spore extraction, 700 µL of diethyl ether was added and vortexed for 5 min. After a 2 min centrifugation step at max. speed, the diethyl ether phase was removed, followed by eliminating the water leaving ~ 50 µL in the tube. The cells were resuspended in the rest of the water and diluted 1:100. This suspension was sonicated two times (Amplitude: 85 %, Cycle: 1, Duration: 10 s), and 20 µL were plated on YPD-agar. After incubation for two days at 30 °C, clones were picked, streaked out on selective agar (YPD + G418 and YPD + NTC), and incubated for two days at 30 °C. These F1 generation clones were further analyzed by flow cytometry.

2.16. DNA content staining

Materials and Procedure

- Ethanol 70 %
- PBS
- Tween 20
- RNase 10 mg/mL
- Propidium iodide 2 mM

For determining the DNA content of cells, staining with propidium iodide (PI) solution was used. PI penetrates dead cells and intercalates between the DNA bases. The excitation at 488 nm causes a fluorescence emission maximum at 640 nm, which can be measured by flow cytometry.

A single colony of a *Komagataella* strain was picked and grown overnight at 25 °C in 2 mL (selective) YPD-medium. The cell density was measured on a lab photometer, and the volume required for 10^7 cells was calculated using the assumption that OD_{600}=1 corresponds to approximately 10^7 cells per mL. This calculated volume of cell suspension was pipetted into a fresh microcentrifuge tube, and 1 mL of 70 % Ethanol was added slowly under continuous mixing for fixation of the cells. 300 µL of this suspension was centrifuged at 10000 g and 4 °C for 5 min, washed in 1 mL PBST (PBS + 1:100 Tween 20), and again centrifuged as described. The cell pellet was resuspended in 990 µL PBS + 10 µL RNase (10 mg/mL) and incubated at 37 °C and 800 rpm for 1-2 h. After centrifugation at 4 °C, 10000 g, and 2 min, the supernatant was discarded, and the pellet was resuspended in 1 mL PBS. After a second centrifugation step at the same conditions, the pellet was again resuspended in 1 mL PBS, and the suspension was sonicated twice (Amplitude: 85 %, Cycle: 1, Duration: 10 s). 150 µL of the treated cell suspension was mixed with 150 µL PI-solution (1:100 in PBS) in a 96-well plate and analyzed by flow cytometry.

2.17. Flow Cytometry

Materials and Procedure

- Gallios Flow Cytometer
- CytoFLEX Flow Cytometer
- PBS

For vector integration analysis, cells were cultivated overnight in 2 mL selective YPD-medium at 25 °C in a 24-well plate. The samples were diluted to OD_{600}= 0.4 in 500 µL PBS into chilled FACS tubes. For all samples, forward scatter, side scatter, eGFP (excitation 488 nm, emission 525 nm), and mCherry (excitation 587 nm, emission 610 nm) fluorescence were measured. For DNA content measurements, forward scatter, side scatter, and PI fluorescence (excitation 488 nm, emission 640 nm) were measured. For all samples, 10,000 events were detected.

2.18. Growth curves

Materials and procedure

- Tecan sunrise microplate reader
- YP-media with xylose

Single colonies, streaked out on YPD-agar plates, were picked and grown overnight in 2 mL YPD at 25 °C and 280 rpm in 24-well plates. The next day the plate was centrifuged for 5 min and 1500 g, the pellets resuspended in 1 mL PBS and centrifuged again for 5 min and 1500 g. This washing step was repeated, and the pellets were again resuspended in 1 mL PBS. For OD_{600} determination, the culture was diluted 1:11 and 1:121 by pipetting 30 µL of the culture into a microtiter plate filled with 300 µL of RO-water. 30 µL of this first dilution was transferred into the next well for the 1:121 dilution. According to the OD_{600}, the culture was diluted to OD_{600}= 1.6 into a sterile 96-well plate. For a simultaneous starting time, the culture was diluted 1:20 into a new sterile 96-well plate containing growth medium. To keep the evaporation of the medium low, the outer rows were filled with 200 µL of water to saturate the air with water vapor. One row was filled only with 200 µL medium as a reference. The plate was covered with a lid and sealed with parafilm for measurement. The table below shows the parameters chosen for the measurement.

Table 6: Settings for 24 h growth curve recording

	setpoint
Wavelength	600 nm
Run time	24 h
Numbers of cycles	99
Temperature range	29.5 - 30.5 °C
Shaking between kinetic cycles	870 s

3. RESULTS & DISCUSSION

3.1. Vector cloning

To obtain the strains for the mating experiments, the *AOX1* terminator locus was chosen for integration of one of two designed vectors BB3aN_P_{GAP}_mCherry_RPS3tt or BB3aK_P_{GAP}_eGFP_RPS3tt, which were assembled into a Golden*PI*CS BB3 vector. This was done by using already available BB1 vectors for the promoter, coding sequences and terminator, which are listed in table 7 (Prielhofer et al. 2017).

Table 7: BB3 assembly with the BB1 vectors for promotor, coding sequence and terminator for each vector

BB1	cloned into 001_BB3aK_FS1_FS4_empty vector	cloned into 109_BB3aN_FS1_FS4_empty vector
Promotor	023_BB1_12_P_{GAP}	023_BB1_12_P_{GAP}
Coding sequence	050_BB1_23_eGFP	124_BB1_23_mCherry
Terminator	018_BB1_34_RPS3tt	018_BB1_34_RPS3tt

The vectors carry the NTC or Kan/G418 resistance gene and the gene for the mCherry or eGFP fluorescence reporter for easy detection by flow cytometry or fluorescence microscopy.

3.2. Vector integration

After the Golden Gate assembly, the two vectors were separately transformed into *E. coli* cells for plasmid amplification, followed by plasmid purification and linearization. The two constructs were then integrated into the *AOX1* terminator region of all *Komagataella* strains via homologous recombination, as is shown for BB3aK_P_{GAP}_eGFP_RPS3tt in figure 6.

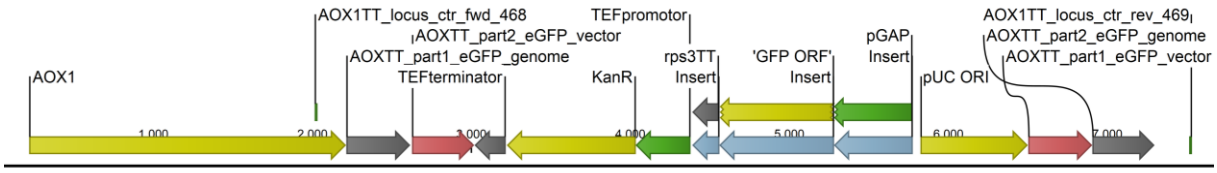


Figure 6: AOX1 terminator locus of K. phaffii CBS 7435, after integration of BB3aK_P_{GAP}_eGFP_RPS3tt with the binding positions of primers AOX1TT_locus_ctr_fwd_468 and AOX1TT_locus_ctr_rev_469

In the end, transformed strains were streaked out on selective YPD+NTC/G418 to obtain in total 25 strains carrying the NTC/mCherry and 25 strains carrying the G418/eGFP construct.

3.3. Integration control

After transformation and vector integration, the expression of the fluorescence proteins, mCherry or eGFP, was controlled via flow cytometry. Positive clones were further analyzed for plasmid integration into the correct locus using colony PCR. Control primers were designed to bind outside of the *AOX1* terminator locus. The size of the obtained fragments, indicating the right integration, was 5518 bp for the eGFP- vector and 5272 bp for the mCherry- vector. figure 7 shows the results of the locus PCR.

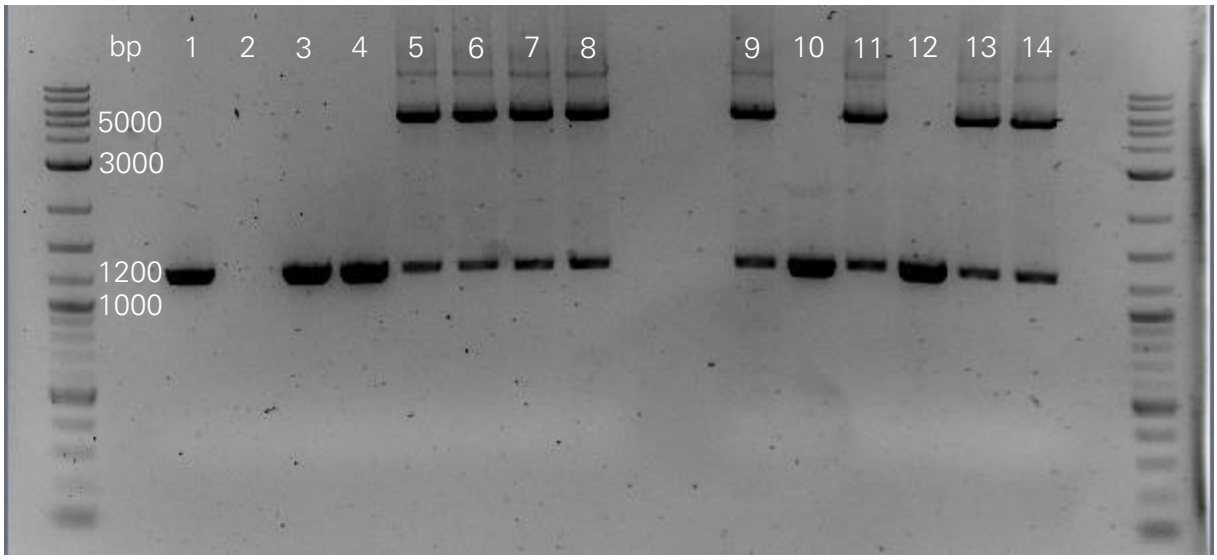


Figure 7: Agarose gel with Quick-Load 1 kb Plus DNA Ladder in the first and last lane of the gel. Lane 1 contains the wildtype control with the expected band at 1227 bp and lane 2 the negative control. The mCherry clones were loaded on lanes 3-8 (3/4: CBS2 612; 5/6: CBS 7435; 7/8: 03-328y3), eGFP clones were loaded on lanes 9-14 (9/10: CBS 2612; 11/12: CBS 7435; 13/14: 03-328y3). PCR was done with primers 468 and 469, as shown in Figure 6.

Although the integration method of homologous recombination is well established in some *Komagataella* species e.g. *K. phaffii*, it seemed that the integration was less effective in some other species. In figure 7, lane 1 shows the wildtype with the expected band at 1227 bp and lane 2 as the negative control. Lane 3 and 4 are clones of the strain CBS 2612 with the mCherry-vector signaling no vector at the specific site. The lanes 5, 6, 7, 8, 9, 11, 13, and 14 show two bands, one at the right size of the vector and one at the wildtype size. All colony PCRs always led to either wildtype bands, both bands, or no signal at all. Therefore, the transformations were repeated, and more clones were analyzed, but no clones showing only the correct band could be obtained. Also, new primers with different binding sites and the use of primers binding inside of the vector, amplifying towards the outside binding primer, did not clarify the results. During vector integration, the homologous parts of the *AOX1* terminator locus are duplicated in the genome.

The PCR results indicated that with the vectors used, the duplicated region was larger and included the primer binding sites used to check for correct integration. Therefore, all clones showing fluorescence in flow cytometry and a band at the expected size were classified as correct. Sequencing of the whole locus after the integration would be necessary to confirm the actual organization of the genomic region.

For some of the strains it was not possible to obtain strains with the correctly integrated plasmid, even though the integration was tried several times. The fact that the integration efficiency was high for the *K. phaffi* strains than strains of the other species in the collection, indicated that the sequence of the homologous region for integration could be different throughout the species. A comparison of the *AOX1* terminator sequence of the strains *K. phaffii* CBS 7435, *K. phaffii* CBS 2612, and *K. pastoris* DSMZ 70382 showed a high sequence identity of around 90 % at the integration locus. Differences in the non-coding integration sequence could be a reason for the integration difficulties in some of the strains. However, the genome sequences of the other species were not available for analysis. Overall, the eGFP vector was integrated correctly in 52% of the clones, while the mCherry vector was integrated into the correct locus with an efficiency of 72 %. It is unclear why the efficiency was different as the vectors shared the same homologous regions, the same backbone and had a similar size.

For the characterization of the strains, it was not necessary to have the vector with the selection marker integrated into the correct locus. Still, for further mating experiments it would be an advantage as integration into the same genomic locus prevents recombination of the markers, which could lead to marker loss or clones with both markers in one haploid genome.

3.4. Spotting Assay

To get an overview of the growth characteristics of the 25 strains, spotting assays were carried out on several different media under specific growth conditions. The strains were tested on conditions summarized in table 8. The growth was controlled visually, and the Petri dishes were photographed after two and/or seven days.

Table 8: Testing and media conditions for spotting assays

carbon sources	stressors	temperature
Glucose	NaCl c= 1 M	30 °C
Xylose	pH 9	37 °C
Glycerol	pH 4	
MeOH	H_2O_2 c= 0.5 mM	
EtOH	H_2O_2 c= 1 mM	
Agar without carbon source		

For better comparison of the different plates and conditions, the reference *K. phaffii* CBS 2612 was always spotted on every plate on the right outside. Cell growth and spot density was analyzed visually and compared to the growth of the reference *K. phaffii* CBS 2612. The summarized results for the 25 strains are listed in the Appendix (5.3).

3.4.1. Carbon sources

When testing different carbon sources, it turned out that most of the strains did not show significant growth differences. Especially at 30 °C, the growth of the different strains was similar to the reference for most carbon sources. However, the results for xylose showed some interesting differences in the growth pattern (Figure 8). After two days on minimal medium with xylose, *K. populi* CBS 12362 and *K. pastoris* DSMZ 70382 did show better growth than the reference, and also denser spots than the rest of the strains. The difference was even more pronounced at 37 °C where *K. populi* CBS 12362 was the strain with the best growth while *K. kurtzmanii* CBS 12817 was the worst and did not grow at all.

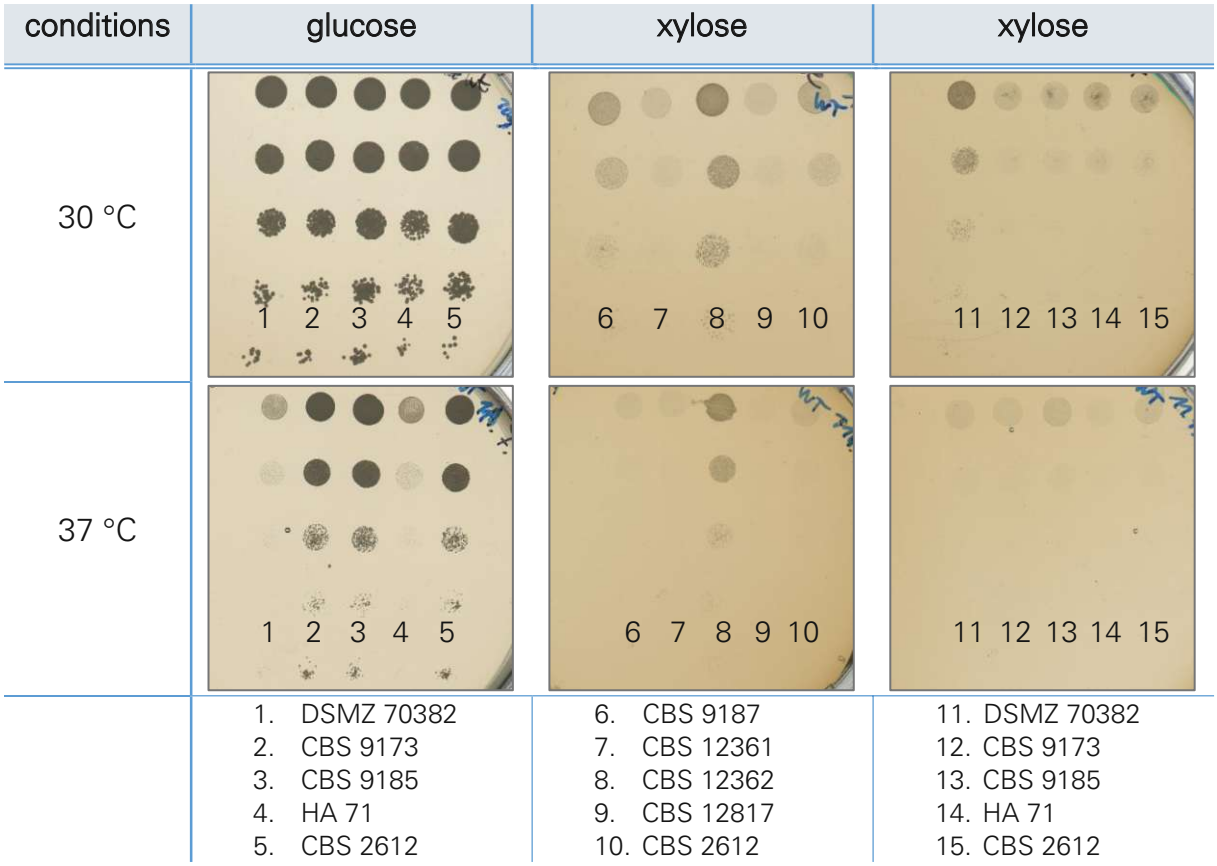


Figure 8: Spotting assays representing the difference between glucose and xylose at different temperatures

Apart from xylose, *K. kurtzmanii* CBS 12817 also showed the worst growth on all other carbon sources. It should be mentioned that the comparably good growth of *K. populi* CBS 12362 and *K. pastoris* DSMZ 70382 on xylose was still very weak compared to growth on glucose. Overall, higher variation was observed at 37 °C, where most of the *K. phaffii* strains showed slightly better growth, *K. pastoris* grew similar to the reference and *K. ulmi*, *K. pseudopastoris,* and *K. kurtzmanii* CBS 12817 grew generally worse than the reference. If a strain did not grow well on glucose at 37 °C, it did not show good growth on glycerol, methanol or ethanol either, although this observation was not true for xylose utilization. As expected, growth at 30 °C was better than at 37°C for all carbon sources tested.

3.4.2. Stressors

3.4.2.1. After 2 days

Incubation at 30 °C

When testing the effect of different stressors on growth, the spotting assays showed only small differences between the reference and the different strains (Figure 9). Compared to growth on glucose at 30 °C, the stressors inhibited the growth of every strain, which resulted in smaller spots and less density of the colonies. While the spots at pH 4 are still comparable to those without stressors, growth was strongly inhibited by NaCl, pH 9, and H_2O_2 1 M, where the spots are mostly very sparse, small, and only recognizable in the first dilutions. It seemed that a higher salt concentration of NaCl and a high pH had a stronger impact on growth than a low pH or oxidative stress in a low concentration. It was previously reported that *Komagataella phaffii* grows in a pH range of 3-7, so the results fit the observation that in general, *Komagataella* strains grew better in an acidic environment (Cregg et al. 1993).

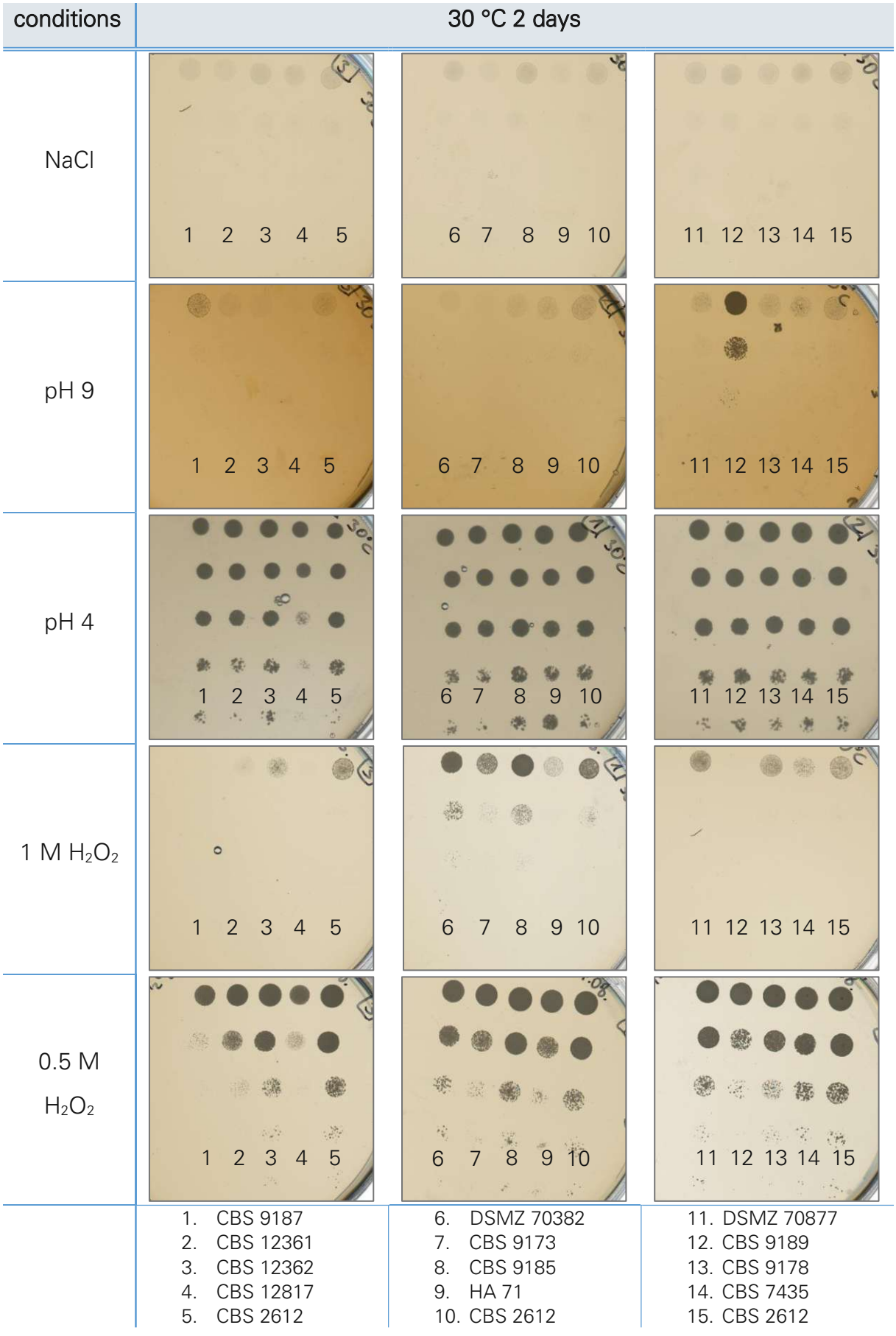


Figure 9: Spotting assays representing the difference under stressors at 30 °C after two days

Only *K. pseudopastoris* CBS 9189 was able to grow to a relatively dense spot on pH 9, and again *K. kurtzmanii* CBS 12817 was the strain with the weakest growth at 30 °C as well as 37 °C.

Incubation at 37 °C

The density and intensity of the spots decreased significantly when the plates were incubated at 37 °C. Only the plates with pH 4 and 0.5 mM H_2O_2 showed some sparse growth in the first two dilutions. Due to the slow growth after two days, the spotting assays at 37°C were incubated for another five days to see if the strains could overcome the stressor's effect.

3.4.2.2. After 7 days

Incubation at 30 °C

After seven days of incubation, most of the spots were as dense as the ones on 30 °C with glucose after two days. The strains incubated at 30 °C did grow well on H_2O_2 in both concentrations, and also the cells spotted on pH 4 were able to grow in every dilution, although the size of the spots was smaller than at 30 °C glucose. Again, the higher salt concentration and the high pH resulted in less growth. The NaCl assays showed smaller spots but mostly grown to the highest dilution, while pH 9 offered more significant differences. Both *K. pseudopastoris* CBS 9189 and CBS 9187 showed good and dense growth, and after seven days at 30 °C even *K. kurtzmanii* CBS 12817 showed the same pattern as the two *K. pseudopastoris* strains, though the spots were barely visible after two days. Yeasts can adapt their organism to environmental changes especially described for stressors by (Gasch et al. 2000), which could also have taken place in, e.g. *K. kurtzmanii* on pH 9, where the organism needed a longer time to adapt to these conditions. Once the agar was covered with a thin layer of cells, they could grow on that basis into dense colonies.

Incubation at 37 °C

The spotting assays incubated at 37 °C over seven days are shown in figure 10.

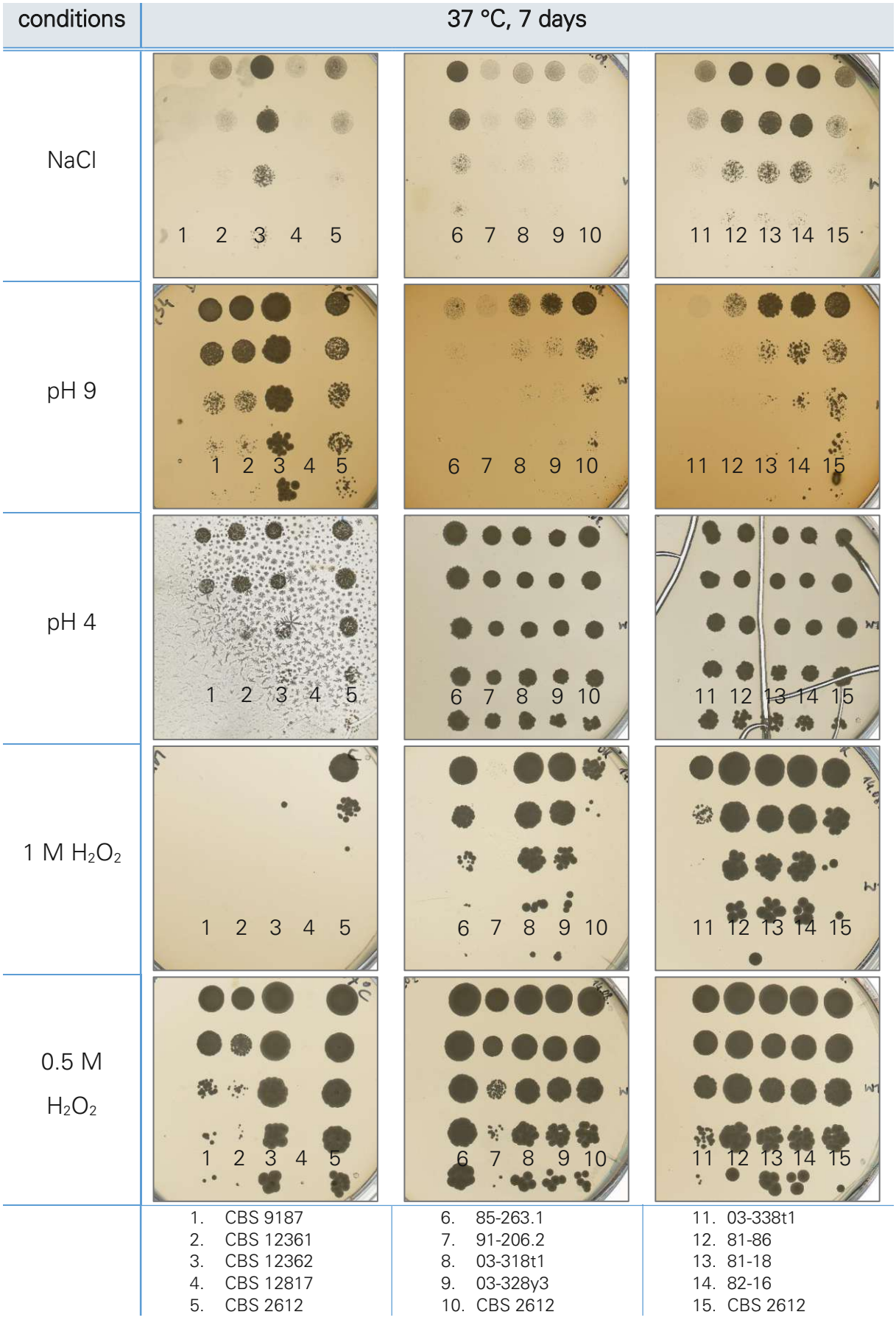


Figure 10: Spotting assays representing the difference under stressors at 37 °C after seven days

The best growth was observed in the presence of 0.5 mM H_2O_2, which was comparable to the one at 30 °C glucose. On plates with 1 mM H_2O_2, most of the *K. phaffii* strains, like the 03-318t1 and 03-328y3, did grow in dense colonies in several dilutions. The spots on pH 4 were smaller, and the faster growth at the beginning of the incubation slowed down during longer incubation times. The reference strain's behavior at pH 4 was similar to that of the other strains, which also showed a decline in growth. The basic condition at pH 9 was again a challenge for the yeasts, and the growth was suppressed for every strain. Only the most tolerant strains like the two *K. pseudopastoris* CBS 9189 and CBS 9187 strains, *K. populi* CBS 12362 and the reference *K. phaffii* CBS 2612 were able to form a dense spot in the first dilutions. Under high salt concentration, the cells grew to a maximum of three dilutions and did not reach the same result as on 1 M NaCl at 30 °C. Unlike the results at 30 °C, where *K. kurtzmanii* showed some dense spots, especially at pH 4, it did not grow at all at 37 °C, independent of the type of stressor tested.

Based on the results of the spotting assays, five strains were selected for further investigation. The idea was to characterize different strains in terms of their mating ability and growth phenotypes as a basis for crossing to create hybrid descendants with improved properties. The selected strains are listed in table 9.

Table 9: Five strains which were selected for further mating experiments

Yeast organsism	Code	Clone No	
		eGFP_G418	mCherry_NTC
K. ulmi	91-206.2	#2	#1
K. phaffii	81- 86	#1	#2
K. pastoris	DSMZ 70382 (CBS 704)	#2	#3
K. populi	CBS 12362	#1	#2
K. kurtzmanii	CBS 12817	#2	#1

K. phaffii 81-86 was chosen because of the good growth at 37 °C and to include another *K. phaffii* strain besides the reference for testing the mating ability. All three of the *K. ulmi* strains showed similar properties during the spotting assays, but 91-206.2 was chosen because of the successful integration of the NTC/mCherry constructs into the correct genomic locus. *K. pastoris* DSMZ 70382 is the type strain of *K. pastoris,* and its performance differed most from the reference. *K. populi* CBS 12362 and *K. kurtzmanii* CBS 12817 are also the type strains of their species, and while the *K. populi* strain showed the best growth on xylose compared to the reference, *K. kurtzmanii* was growing the worst in nearly every condition tested. It was important to choose strains from different species but also with diverse characteristics for the mating experiments (section 3.5).

Additional to the mating experiments, *K. populi* CBS 12362 was selected for further investigation of the improved growth on xylose (section 3.6).

3.5. Mating

Mating experiments can give an insight into the compatibility between different strains and species. In addition to the genomic compatibility during meiosis, cell recognition among each other is the basic requirement for successful sexual recombination. This means that the secreted mating pheromones need to be recognized by the pheromone receptors of the opposite mating partner, which is not always the case for more distantly related species (Merlini et al. 2013).

The DNA content of the cells after mating and sporulation was determined by PI staining and flow cytometry measurements. All the samples were compared to a haploid reference to determine the ploidy of the cells. In figure 11 (A-B), the distribution of a haploid chromosome pattern of the control strain *K. phaffii* CBS 2612 can be seen. The cell population highlighted in figure 11A was gated in the diagram of side-scatter vs. forward-scatter. All yeast cells within the gate were used to generate the histogram in figure 11B showing the two typical DNA peaks. The first peak represents the cell population in the G1 phase of mitosis with one copy of each chromosome. The second peak represents the cell population in the G2 phase, where the amount of DNA has already doubled. Figure 11 (C-D) shows the cell population and DNA content of diploid cells. The first peak in figure 11 D again shows the cells in G1 phase, whereas the shift of the peak to higher fluorescence is due to the doubled chromosomes after mating. The second peak again represents the G2 phase of diploid cells containing four chromosomal copies. Comparison of the two histograms can be used to make a statement about the DNA content and thus the ploidy of the cells.

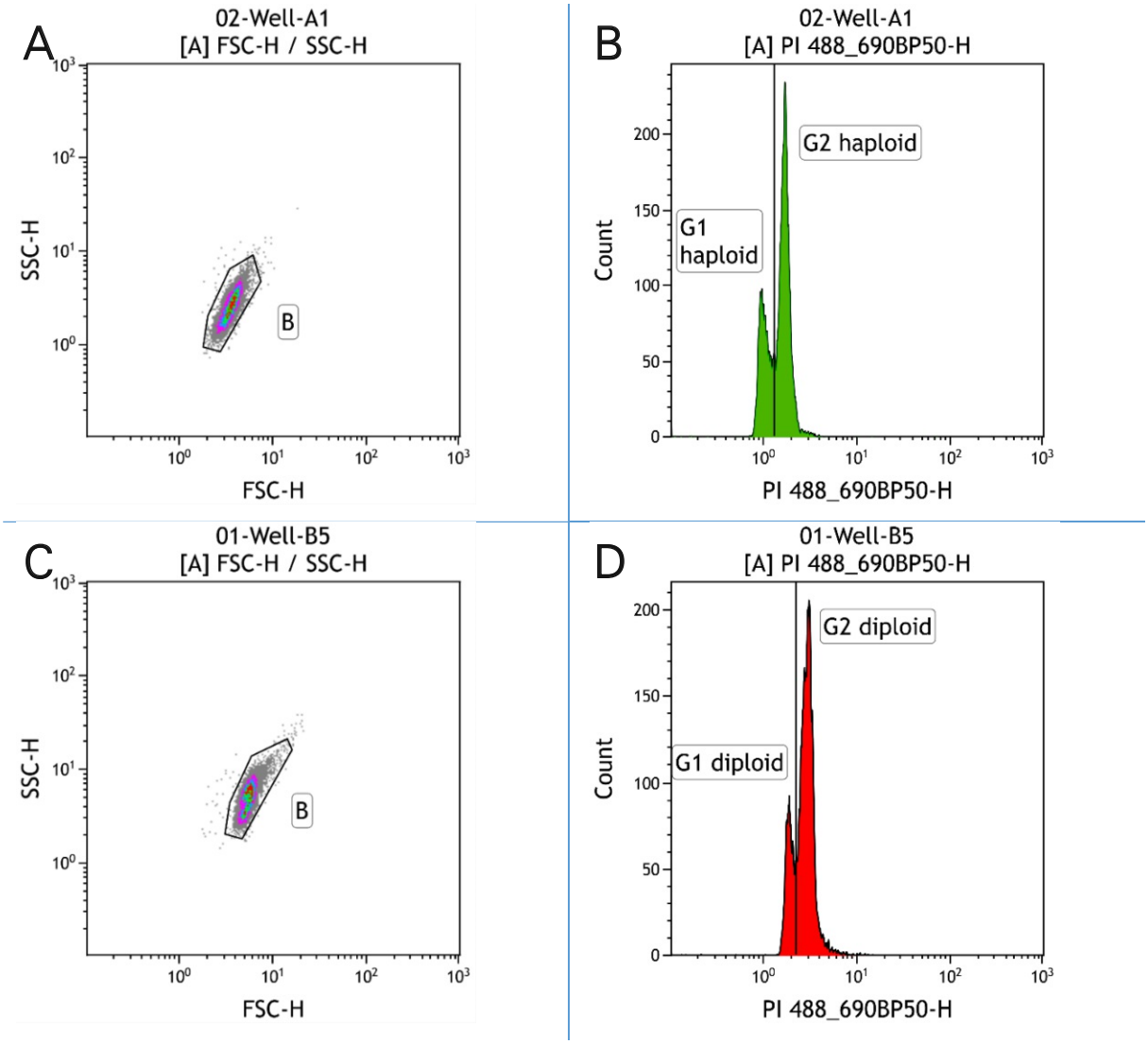


Figure 11: (A) Forward-side scatter plot of the haploid control CBS 2612 with gate B around the population. (B) Fluorescent intensity histogram of the gated cells of figure (A). The second vertical gate is for separation of G1 and G2 phase cells. (C) Forward-side scatter plot of the diploid mating sample between CBS 2612 and CBS 12362 with gate B around the population. (D) Fluorescent intensity histogram of the gated cells of figure (C). The second vertical gate is for separation of G1 and G2 phase cells.

By visual comparison of the two histograms in figure 11, it can be seen that the red mated sample has nearly doubled its fluorescent intensity. For a diploid sample, the G1 phase peak should overlap with the G2 phase peak of the haploid control as the histogram is shifted to double the fluorescent intensity. Atypical shifts can be caused by the loss of single chromosomes or can be the result of failures during the recombination process of meiosis.

3.5.1. Mating with *K. phaffii* CBS 2612

In the first mating experiment, every strain in the collection was mated with the *K. phaffii* CBS 2612 reference strain. For this, all strains with the opposite marker to the reference were tested, even if the locus of the integrated vector could not be verified by PCR.

The mating success was controlled visually. Diploid colonies visible at the crossing points of two strains were picked and cultivated for DNA content analysis by PI staining and flow cytometry. The results are listed in the table below.

Table 10: Mating test between the reference strain K. phaffii CBS 2612 and all other 24 strains. The "X" marks a positive mating result between the opposing transformed strains.

		K. phaffii CBS 2612			
		eGFP/G418 vector		mCherry/NTC vector	
species name	strain code	visible colonies	diploid result at flow cytometry	visible colonies	diploid result at flow cytometry
K. phaffii	CBS 7435	not tested[2]		not tested[2]	
K. phaffii	85- 263.1	X	X	X	X
K. phaffii	03- 318t1	X	X	X	X
K. phaffii	03- 328y3	X	X	X	X
K. phaffii	81- 86	X	X	X	X
K. phaffii	81- 18	X	X	X	X
K. phaffii	82- 16	X	X	X	X
K. phaffii	85- 348.1	X	X	X	X
K. phaffii	85- 926.1	X	X	X	X
K. phaffii	91- 119.3	X	X	X	X
K. phaffii	91- 132.2	X	X	X	X
Komagataella	HA71	very bad	X	X	X
K. pastoris	DSMZ 70382 (CBS 704)	-	-	-	-
K. pastoris	DSMZ 70877	X	X	X	X
K. pastoris	CBS 9178	X	X	X	X
K. pastoris	CBS 9173	X	X	X	X
K. pastoris	CBS 9185	X	X	X	X
K. ulmi	CBS 12361	X	X		
K. ulmi	91- 206.2	X	X	X	X
K. ulmi	03- 338t1	X	X	X	
K. pseudopastoris	CBS 9187	X	X	X	X
K. pseudopastoris	CBS 9189	X	X	X	X
K. kurtzmanii	CBS 12817	-	-	-	-
K. populi	CBS 12362	X	X	very bad	X

The mating experiment showed that 21 strains out of 24 could mate with the CBS 2612 reference strain in at least one vector combination. The only strains where no diploid colonies were observed were *K. pastoris* DSMZ 70382 and *K. kurtzmanii* CBS 12817. It was rather unexpected that the *K. pastoris* type strain could not mate with the *K. phaffii* type strain while all other *K. pastoris* strains did not show any mating difficulties.

[2] This strain was not tested because of previous positive mating experiments carried out in the working group.

Incorrect vector integration should not be the reason for this, as the same integration problem was also observed with all other *K. pastoris* strains.

It has to be noted that this mating experiment was done by a replica plating method, which is an easy and fast method for a high number of samples like in this experiment. Still, there are also some limitations coming with this method. As the mating spots where two strains are combined on the plate are small, the probability of seeing diploid colonies, especially for strains mating with very low efficiency, is reduced. Also, the tissue handling during replica plating can cause blurring and can contaminate the separated lanes of the different strains, which can lead to wrong results.

The ability of two strains to reproduce sexually can only be determined by studying the full process of mating and sporulation. This means that after successful mating, sporulation and the viability of the spores should also be analyzed. However, due to time limitation, this has not been done for this experiment.

3.5.2. Mating between five selected strains

The five strains listed in table 9 were selected for further analysis based on the results of the spotting assays and the first mating experiments with *K. phaffii* CBS 2612. They were tested for their ability to mate with each other, which was done in the same testing procedure as before. The short summary in table 11 shows an overview of the results from flow cytometry data. However, the more detailed description shows that unambiguous statements cannot always be made, as indicated in this table. The presence of colonies after the spore extraction was not a guarantee for a clear result of the DNA content.

Table 11: Mating results between the five selected strains. The "X" marks positive diploid clones after mating or viable clones obtained after sporulation, while the number beside shows the number of clearly haploid clones determined by flow cytometry.

species name	strain code	*K. phaffii* 81- 86		*K. pastoris* DSMZ 70382		*K. populi* CBS 12362		*K. kurtzmanii* CBS 12817	
		diploid cells	*haploid spores*	*diploid cells*	*haploid spores*	*diploid cells*	*haploid spores*	*diploid cells*	*haploid spores*
K. ulmi	91-206.2	X	X (0/4)	X	X (0/4)	X	X (3/4)	X	X (1/4)
K. phaffii	81- 86			-	-	X	X (0/4)	X	X (2/4)
K. pastoris	DSMZ 70382					-	-	-	-
K. populi	CBS 12362							X	X (2/4)
K. kurtzmanii	CBS 12817								

Both *K. ulmi* 91-206.2 clones mated with every other strain and formed diploid cells. After sporulation, the analyzed colonies only showed haploid signals for the crossings with *K. populi* CBS 12362 and *K. kurtzmanii* CBS 12817, where even only one colony was haploid.
The strain *K. phaffii* 81-86 mated and formed diploid cells with *K. ulmi* 91-206.2 and *K. populi* CBS 12362, but all samples gave a diploid signal after the spore extraction.

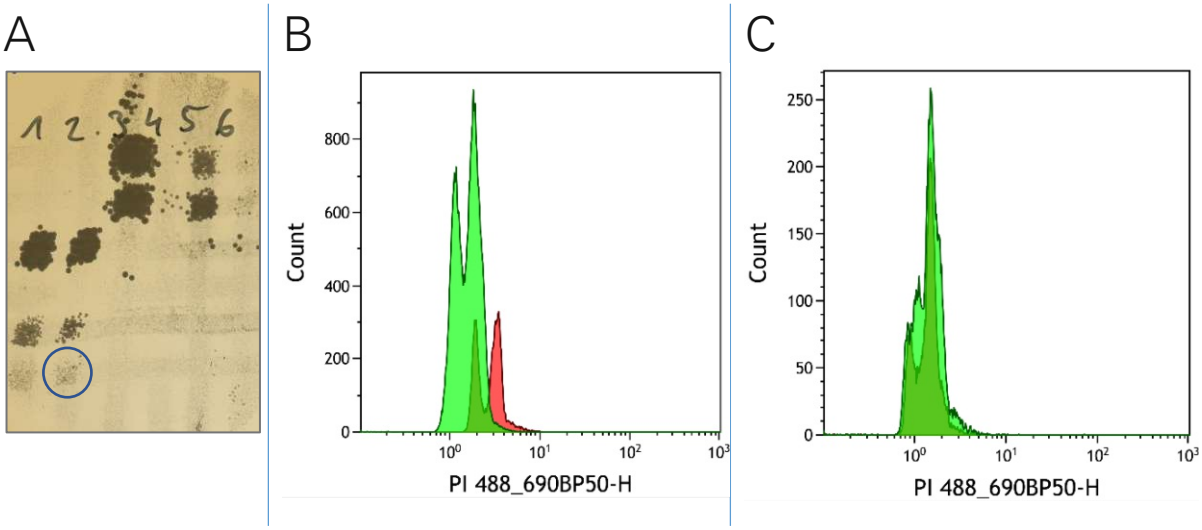


Figure 12: (A) Mating plate on selective agar with the surrounded crossing spot between CBS 12817 and 81-86. (B) The green signal represents the haploid reference population CBS 2612, and the red signal the population of the mated sample from (A). (C) Overlay of the green haploid reference CBS 2612 and the sporulated haploid sample from (A).

For crossings between *K. kurtzmanii* CBS 12817 and *K. phaffii* 81-86 the agar-plate spots always showed a sparse but noticeable growth, making it hard to confirm a positive mating result. The analyzed cells always showed a diploid histogram as it is shown in figure 12 (B), so probably the mating occurred in a very low efficiency, or the spot was contaminated with other cells of the mating plate. In the spore analysis, two samples showed the same pattern as the haploid control in figure 12 (C), while two other samples remained diploid according to the shifted histogram to the higher fluorescence signal.
K. pastoris DSMZ 70382 only mated with *K. ulmi* 91-206.2. *K. populi* CBS 12362 was able to mate with every other strain except *K. pastoris* DSMZ 70382, although the growth spot at the crossing with *K. kurtzmanii* CBS 12817 was not very dense. After mating, the DNA content showed a diploid pattern, as shown in figure 13 (B). After sporulation again, colonies were observed on the acetate-agar plate, but two samples were haploid (shown in figure 13 (C)), and two showed a shifted population indicating diploid cells.

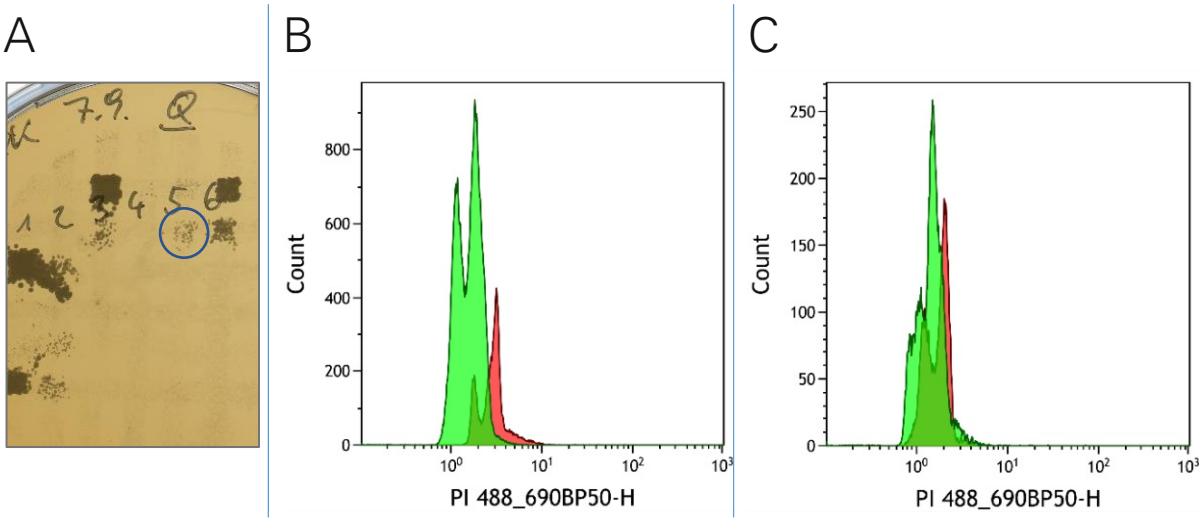


Figure 13: (A) Mating plate on selective agar with the surrounded crossing spot between CBS 12817 and CBS 12362. (B) The green signal represents the population of the haploid reference CBS 2612, and the red signal the population of the mated sample from (A). (C) Overlay of the green haploid reference CBS 2612 and the sporulated haploid sample from (A).

The results confirmed the low mating ability and peculiarity of the strain *K. kurtzmanii* CBS 12817 observed in previous experiments. It was able to mate with all other strains but with low efficiency. Most of the analyzed cells were diploid after mating, but the cells' analysis after sporulation did not give a clear result about their ploidy.

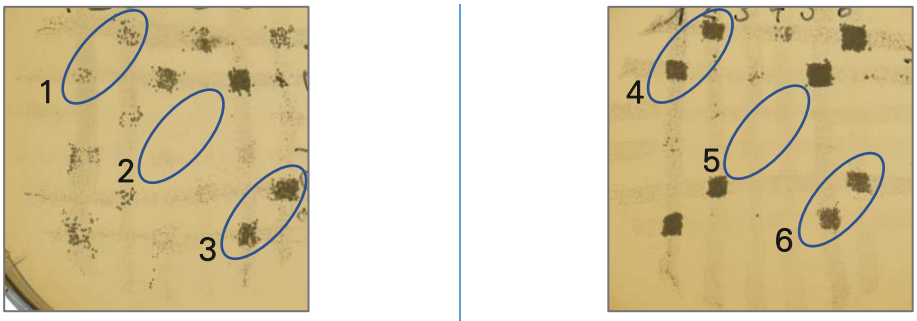
The mating and sporulation experiments showed that it is hard to obtain a clear result for the tested strains. Again, the mating ability of *K. kurtzmanii* CBS 12817 was not very pronounced but still better than for *K. pastoris* DSMZ 70382, which was not able to mate with the reference or any of the other strains except the *K. ulmi* 91-206.2. *K. ulmi* and *K. pastoris* are located in the same branch of the phylogenetic tree, which might be the reason for the observable mating between these species, even though the sporulating colonies could not be confirmed as haploid descendants. This compatibility of *K. pastoris* DSMZ 70382 apparently only relates to the closest relatives, although mating between the different *K. pastoris* strains remains to be tested.

The results also showed that the formation of diploid cells is not the critical step in the sexual reproduction between diverse yeast strains, as hybrids could be formed between almost all strains and species tested. As expected from literature, the critical step is meiosis and the formation of haploid spores. The inability of hybrids to form viable haploid spores leads to reproductive isolation and species separation (Greig 2009). A clear statement on the degree of reproductive isolation between the tested strains would require the quantitative determination of spore viability. Also, it should be kept in mind that the mating efficiency always depends on the respective mating type ratio of the partners.

Since the strains in this study are homothallic and have not been tested for their mating type or the frequency of mating-type switching, no conclusions can be drawn about the mating types of the cells used in the experiments. Poor mating behavior could also be the result of an unequal ratio of mating types.

3.5.3. Self mating

One possibility to show mating-type switching is to test the ability to self-mate within a population originally derived from a single colony. For this, the five strains selected in table 9 were mated by replica plating and the ploidy of the cells was controlled by PI-staining and flow cytometry.



1. *K. populi* CBS 12362
2. *K. kurtzmanii* CBS 12817
3. *K. phaffii* 81-86
4. *K. ulmi* 91-206.2
5. *K. pastoris* DSMZ 70382 (CBS 704)
6. *K. phaffii* CBS 2612

Figure 14: Self-mating spots on selective agar

Figure 14 shows that self-mating can be observed between *K. phaffii* 81-86, *K. ulmi* 91-206.2, and *K. phaffii* CBS 2612. However, flow cytometry analysis showed that only the cells of *K. phaffii* 81-86 and *K. phaffii* CBS 2612 were really diploid. For both strains, haploid spores were detected after sporulation. The result for *K. ulmi* 91-206.2 was not clear to evaluate. While the mating spots showed good diploid cell formation, the flow cytometry data indicated haploid samples. One explanation would be a significantly faster initiation of sporulation together with a higher tolerance to the NTC/G418 selection, which might result in the presence of already germinated spores in the flow cytometry analysis. Nevertheless, the cells were transferred to sporulation agar and analyzed again, which resulted in mainly haploid samples and one showing three peaks indicating cells in an intermediate stage with an aneuploid DNA content, indicating problems during meiosis. This was also the case for *K. populi* CBS 12362, where the mated samples were found to be half haploid and diploid. Also, half of the analyzed spores showed a haploid profile, while half of the histograms again showed three peaks.

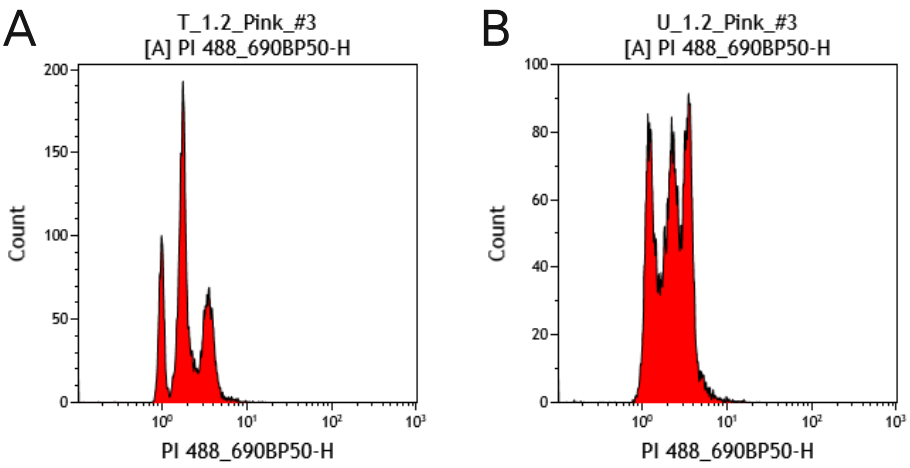


Figure 15: (A) Histogram of the haploid spores from strain 91-206.2 showing a triple peak population. (B) Histogram of the haploid spores from strain CBS 12362 showing a triple peak population.

During meiosis, homologous recombination and its crossing over events can lead to spores with an unevenly distributed set of chromosomes and genetic errors, which is probably an explanation for the observed data with shifted peaks. Samples with three peaks probably represent a mixed population of haploid and diploid or aneuploid cells.

For the strains *K. pastoris* DSMZ 70382 and *K. kurtzmanii* CBS 12817, no positive clones were obtained after self-mating. This could indicate a low rate or absent mating-type switching in these strains.

3.5.4. Mating between *K. populi* CBS 12362 and *K. phaffii* CBS 2612

After mating was tested for every strain, a more detailed analysis was done for *K. populi* CBS 12362, which was also picked for a detailed growth analysis on xylose. So again, the transformed clones of *K. phaffii* CBS 2612 were mated with the ones of *K. populi* CBS 12362, which can be seen in table 12.

Table 12: Flow cytometry results of the mated cells and spores from CBS 2612 and CBS 12362

strain	marker	strain	marker	diploid cells	haploid spores
K. phaffii CBS 2612	mCherry_NTC	*K. phaffii* CBS 2612	eGFP_G418	6/6	6/8
K. phaffii CBS 2612	mCherry_NTC	*K. populi* CBS 12362	eGFP_G418	6/6	1/8
K. phaffii CBS 2612	eGFP_G418	*K. populi* CBS 12362	mCherry_NTC	5/6	5/8
K. populi CBS 12362	mCherry_NTC	*K. populi* CBS 12362	eGFP_G418	4/6	8/8

The results underline the previous experiments' observations that *K. phaffii* CBS 2612 and *K. populi* CBS 12362 could mate and successfully form spores.

The mating between the two different strains efficiently resulted in diploid hybrids, but the formation of haploid spores was not that successful throughout the experiment. All the samples after spore extraction, which were not haploid, showed a diploid pattern. During spore extraction, vegetative cells are killed, so only the germinating spores can grow after plating on rich medium. To avoid clumping of the spores and potential mating within a tetrad, cells were sonicated after spore extraction to separate them. However, the high number of diploid samples is most likely the result of an inefficient spore extraction because, without external pressure like nitrogen starvation, the germinating cells would rather grow by mitotic cell division than by mating. To confirm the formation of spores by the hybrids, cells were also analyzed by microscopy.

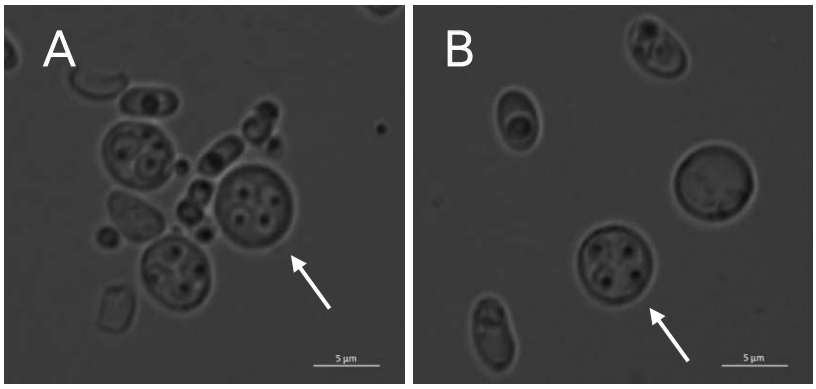


Figure 16: (A, B) cells of the crossing between CBS 2612 and CBS 12362 after sporulation. Arrows indicate tetrads harboring four spores.

In figure 16, a sample of *K. phaffii* CBS 2612 mated with *K. populi* CBS 12362 after sporulation on an acetate agar can be seen. The marked cells harbor four spores enclosed by the spore wall, while the smaller dots represent already released spores. The microscopy pictures confirmed the successful formation of spores by hybrids of the two strains.

3.6. Growth curves

The improved growth on xylose and the fact that *K. populi* CBS 12362 can be crossed with strains of other species made it a promising candidate for more in-depth analysis. To get a better idea of the growth differences on xylose, the growth rates of the reference *K. phaffii* CBS 2612 and *K. populii* CBS 12362 were determined. Also, the growth of a number of F1 generation clones obtained by crossing of the strains was evaluated to see potential variation in the growth phenotype. For this, the strains were cultivated in a 96-well plate for 24 h while the optical cell density was measured continuously. After mating and spore extraction, 48 colonies were picked and analyzed by measuring the OD_{600} of biological triplicates.

Due to the limited space, the 48 samples had to be divided and were measured in three test series on three different days. The first two series showed an expected result as the growth of *K. populi* was slightly better than *K. phaffii,* and the different spore samples produced a variety of different growth patterns. Figure 17 shows the growth curves of the parent strains and two different spore samples.

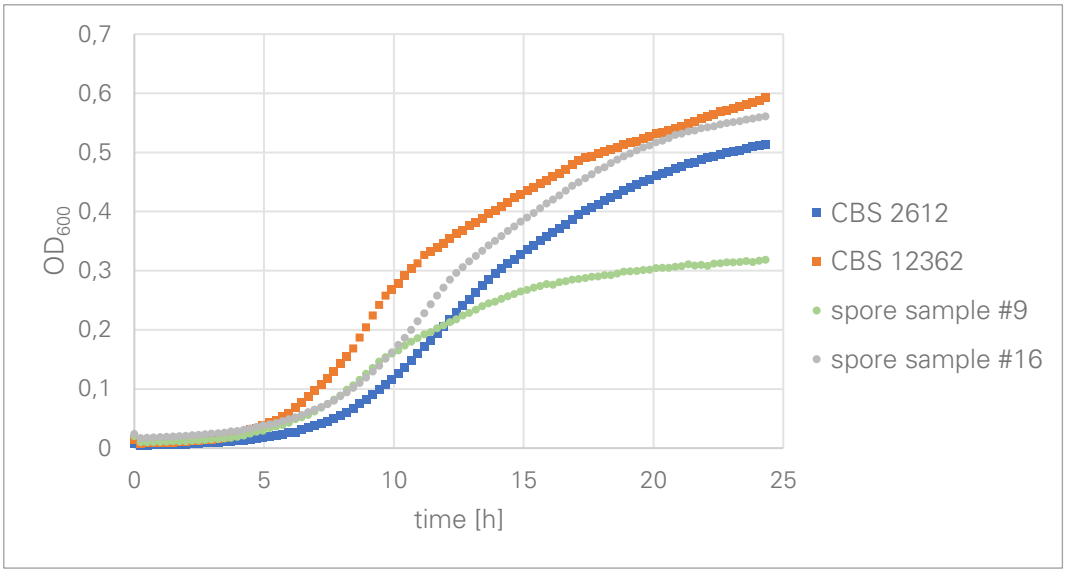


Figure 17: Comparison of the OD_{600} results of the references CBS 2612 and CBS 12362 and the spore samples #9 and #16 of the F1 generation.

While the growth profiles trend fit the expectation that the best strain would be the *K. populi* CBS 12362, the growth profiles were not the classic exponential curves. After a short lag phase in which the cells adapt to their environment, the exponential phase started but soon turned into a more linear growth, which flattened at the end of the 24 h measurement. This short exponential growth phase indicated limited growth due to setup-related problems, which, however, occurred strain independent.

In the second test series, the data show similar behavior. However, the growth curves of the parent clones were different in the third test, as can be seen in figure 18.

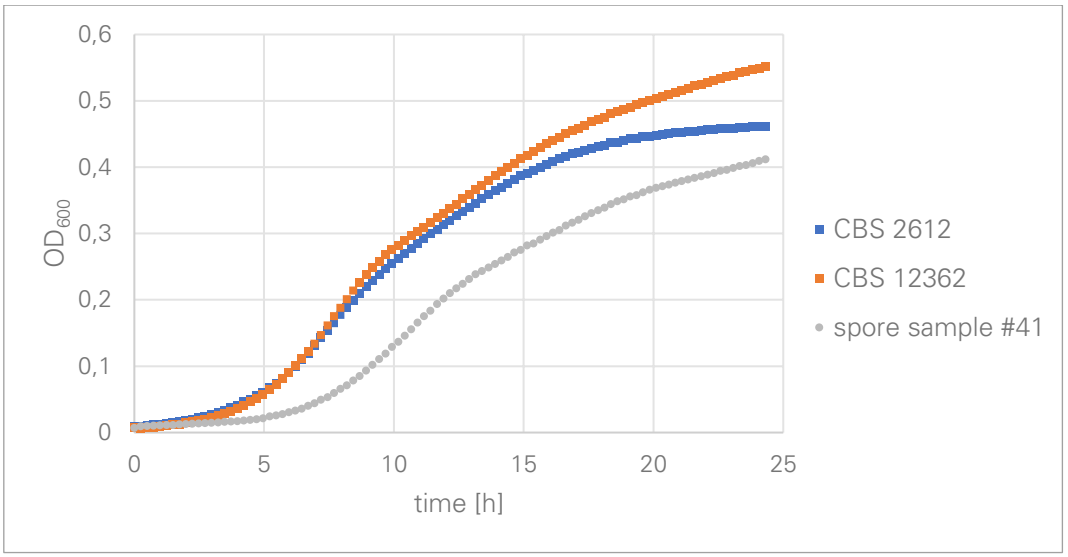


Figure 18: Comparison of the OD_{600} results of the references CBS 2612 and CBS 12362 and the spore sample #41 of the F1 generation. Showing different behavior of the parents in comparison to Figure 17.

The gap between the growth curves of the two parent strains got smaller during the exponential phase, and it seemed that the curve of the CBS 2612 flattens at the end into the stationary phase, where growth slows down. In the end, the two references show again a gap of about $\Delta OD_{600} \approx 0.1$. The different behavior makes it harder to compare the three runs and to confirm the better growth of *K. populi* CBS 12362 on xylose. For a more precise comparison of the parent strains, their growth rates and the respective generation times were plotted in the following diagrams (figure 19 and figure 20).

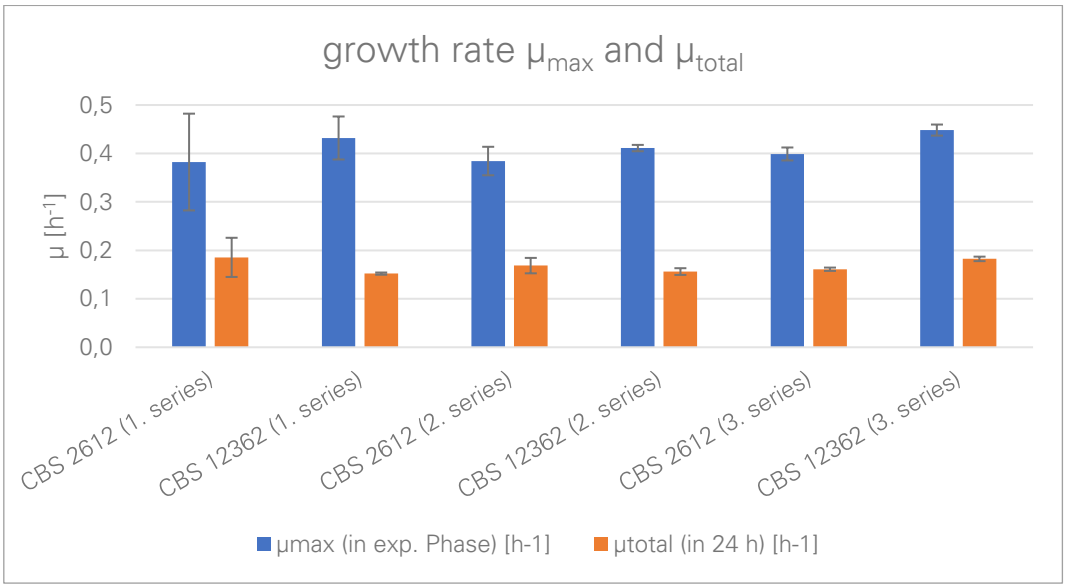


Figure 19: Comparison of the maximal growth rate and the total growth rate. Showing no significant difference between the strains in the first two series, but a significant difference in the 3rd series.

On the first sight neither the maximal growth rate μ_{max} nor the total growth rate μ showed any clear differences between the two strains, which was confirmed by a Student's t-test ($P<0.05$) for the first two test series. Only the third test series showed a significant difference between the growth of CBS 2612 and CBS 12362 which was actually expected for every run. Maybe the 24 h observation period was too short to get a reliable result for growth differences because the plots stop before the stationary phases are reached. Also, the generation times, of course, related to the growth rate, did not show a significant difference for the first two test series (figure 20).

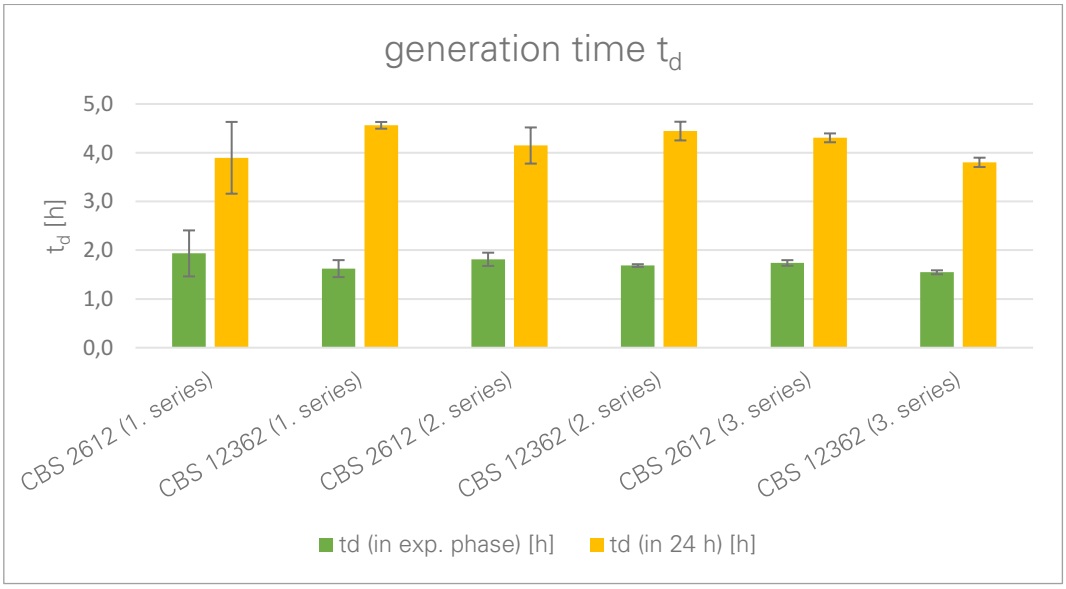


Figure 3: Comparison of the generation time in the exponential phase, and over 24 h. Showing no significant difference between the strains in the first two series, but a significant difference in the 3rd series.

If the growth had been monitored for more than 24 h, the total OD_{600} difference would be larger, which would also influence the total μ. Also, it has to be said that the media consisted of rich YP, which contains amino acids, peptides, vitamins, and carbohydrates. This medium alone contains enough nutrients for some cell growth (at least 24 h) and could thus further reduce the growth difference on xylose as carbon source. In comparison, a defined minimal medium with xylose was used for the spotting assays. Therefore growth rates are not comparable to reported ones on xylose with about 20 times lower growth rates (Li et al. 2015). Another factor that has a huge impact on growth, especially in small volumes inside 96-well plates, is aeration. The linear trend after the short exponential phase could also be the consequence of an insufficient O_2 supply. Further experiments testing longer growth periods with better aeration and controls growing on minimal media would be necessary to get a good overview of the growth difference between the strains as well as the clones of the F1 generation.

4. CONCLUSION

This work aimed to characterize the growth and mating phenotypes of 25 different *Komagataella* strains of the six species *K. phaffii*, *K. pastoris*, *K. ulmi*, *K. pseudopastoris*, *K. kurtzmanii*, and *K. populi.* This data will be the basis to select strains for crossing to combine desired phenotypes and the analysis of their genetic determinants by quantitative trait loci (QTL) mapping. As an example, the improved xylose utilization of *K. populi* CBS 12362 was investigated in more detail.

For the characterization of growth behavior, the strains were tested on different carbon sources and in the presence of different stressors using spotting assays, both cultivated at 30 °C and 37 °C. *K. phaffii* is typically cultivated at a temperature range between 25 °C – 30 °C, but strains with a better growth at higher temperatures might be relevant for industrial applications. While, depending on the media, growth at 30 °C was adequate for all strains, as expected, it was found that at 37 °C, hardly any strain was able to achieve dense growth within two days. Only some *K. phaffii* strains showed a slightly improved growth at 37 °C when compared to the reference. After incubation for another five days, the strains showed a similar growth pattern as the ones at 30 °C (after 2 days).
When looking at growth on different carbon sources, two strains stood out for growth on xylose. *K. populi* CBS 12362 and *K. pastoris* DSMZ 70382 were able to grow in higher dilutions than any other strain at 30 °C and even 37 °C.
The growth behavior under stress conditions, in general, was again much better at 30 °C than 37 °C, where the cells needed more than two days to grow to dense colonies. While strains grew well on pH 4 and 0.5 mM H_2O_2, *K. pseudopastoris* CBS 9189 was the best performing strain under basic conditions at pH 9. Another interesting observation was the extremely poor growth of the *K. kurtzmanii* CBS 12817 under almost all conditions tested.

From all the conditions tested, the most exciting observation was the growth differences on xylose. Together with arabinose, xylose is a central component of hemicellulose in the plant cell wall. The biotechnological conversion of lignocellulosic biomass to biofuels is a big challenge, and the removal of lignin and hemicellulose is an energy-demanding step. When selecting a suitable organism for the process, the ability to utilize pentoses like xylose is a significant advantage. At the same time, this lignocellulosic biomass would provide a cheap and renewable feedstock for protein production in *Komagataella*.

The pathway for xylose utilization is well characterized in *Pichia stipites*, and orthologs of xylose reductase (XR), xylitol dehydrogenase (XDH) and xylulokinase (XK) are found in *K. phaffii* CBS 7435, suggesting that *Komagataella* species are also capable of utilizing xylose (Bengtsson et al. 2009). Pengfei Li and colleagues introduced an additional pathway with xylose isomerase (XI) into *K. phaffii* and showed that XI and XK play a key role in xylose assimilation (Li et al. 2015). This promising first attempts of trying to use xylose as a carbon source also motivated to do a more detailed analysis of the growth behavior of *K. populi* CBS 12362 on xylose.

Prior to the mating experiments, the strains were transformed with two different vectors carrying either the geneticin/eGFP marker or nourseothricin/mCherry marker for diploid cell selection. The efficiency of integration varied throughout the strains and resulted in an average integration into the correct locus of 52 % for the eGFP and 72 % for the mCherry vector, respectively. This could probably be improved using homologous regions that are adapted to the specific species.

Analysis of the mating behavior of the different strains generally showed good compatibility between the strains. Apart from the two strains *K. pastoris* DSMZ 70382 and *K. kurtzmanii* CBS 12817, all strains were able to mate with the reference *K. phaffii* CBS 2612. The five selected strains which were tested in more detail also showed that with the method used, it is hard to draw a clear line between the ability to mate or not, especially if the efficiency is low. Also, in many cases, the DNA content analysis by flow cytometry delivered unclear results with slightly shifted peaks or a vague distribution pattern. After mating, the results were generally clearer than for the spore samples, which indicates problems during meiosis leading to aneuploidy. When mated with other strains, *K. kurtzmanii* CBS 12817 could not form large diploid colonies. Another interesting observation was that *K. pastoris* DSMZ 70382 could only mate with *K. ulmi* 91-206.2.

In addition to the compatibility between the strains, self-mating was analyzed to get an idea about mating and mating-type switching within cells of the same strain. Both strains, *K. kurtzmanii* CBS 12817 and *K. pastoris* DSMZ 70382, which had difficulties during the previous mating experiments, also struggled to mate with themselves. All other strains could mate normally, but the flow cytometry analysis again delivered some unclear results, especially for the analyzed cells after sporulation, which was expected for hybrids but interesting for self-mating. A detailed analysis of the chromosomes by classical karyotyping, could help to explain the results of the mating and sporulation experiments.

Differences in chromosome number and structure are very likely the main reasons preventing sexual reproduction between the different species.

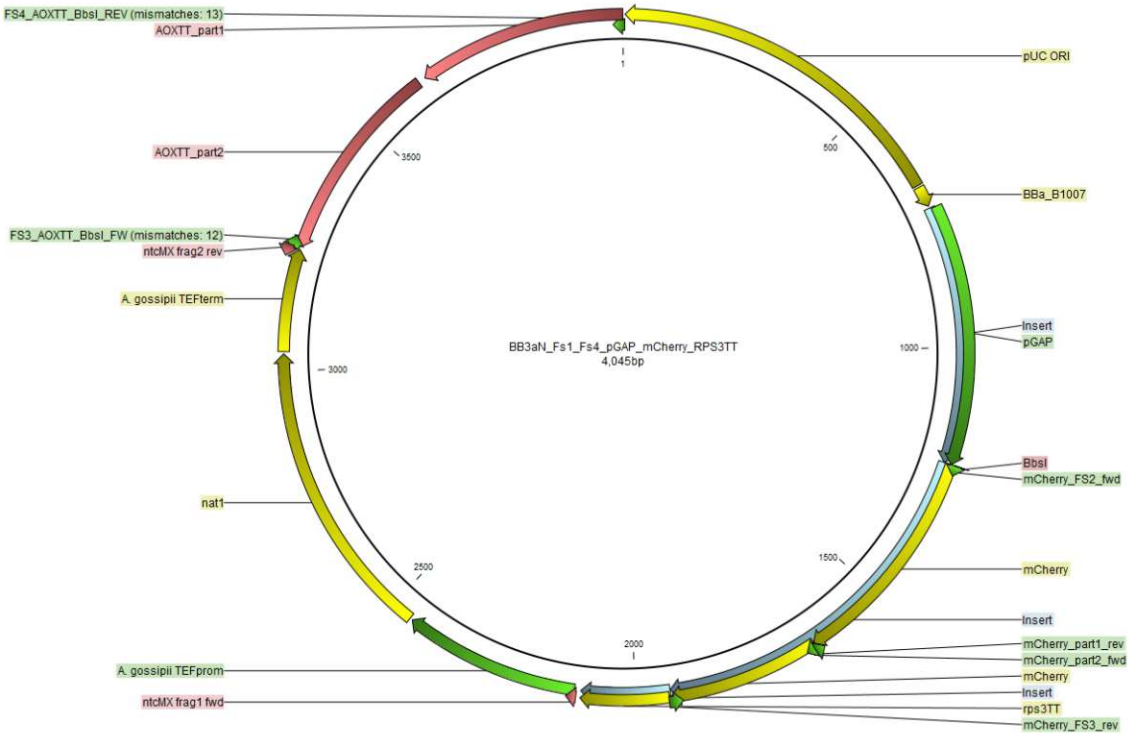
As mentioned above, the improved growth of *K. populi* CBS 12362 on xylose was selected for detailed analysis. Before performing growth experiments in liquid medium, *K. phaffii* CBS 2612 and *K. populi* CBS 12362 were again mated, and spores were extracted and cultivated. This confirmed that the formation of diploid cells is easily possible, but the formation of haploid spores is less efficient. Nevertheless, a high number of spores could be detected under the microscope, which were cultivated for further analysis on xylose.
The direct growth comparison between *K. populi* CBS 12362 and *K. phaffii* CBS 2612 confirmed the results of the spotting assays where a slightly better growth of *K. populi* on xylose could be observed. The 48 clones from the F1 generation after mating and sporulation delivered a diverse growth pattern. The aim to mate strains with different phenotypes and analyze the F1 generation to obtain improved cells could be realized. However, the experimental setup for the growth assay requires optimization in terms of media composition, aeration, and duration.

Overall, it can be said that the characterization of the 25 strains was accomplished. The strains were tested for their growth behavior on different carbon sources and stress conditions, where some differences, e.g., in growth on xylose or at 37 °C, were found. The mating experiments showed general compatibility between the strains and species, and diploid cell formation was possible in most cases. Also, viable spores could be obtained for most strain combinations, and spore formation could be confirmed under the microscope. However, the unclear results of the DNA content analysis indicated problems during meiosis, resulting in the formation of aneuploid progeny. Finally, a 24-hour growth experiment in liquid culture showed that through mating and sporulation, variation in growth on xylose could be achieved in the F1 generation. This opens the opportunity to identify clones with improved growth on xylose and continue with crossings for in-depth analysis of the genes involved in improved xylose utilization by quantitative trait loci (QTL) mapping.

5. APPENDIX

5.1. Plasmids

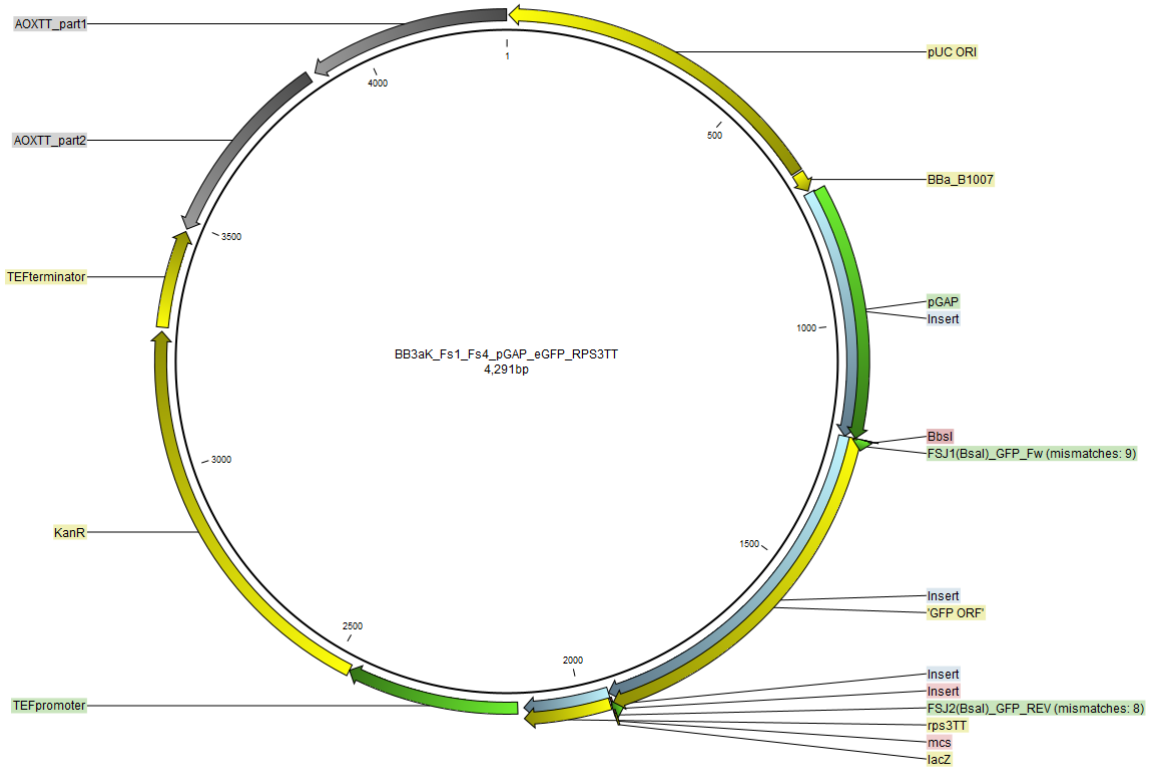
5.1.1. BB3aN_Fs1_pGAP_mCherry_RPS3TT



5´-CGCTGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCG
TTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTG
GCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCT
CTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGG
CGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGG
GCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTG
AGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGC
AGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACAC
TAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGG
TAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCA
GATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTGGAGCG
CAAAAAACCCCGCCCCTGACAGGGCGGGGTTTTTTCGCGATCGGAGGATCCTTTTTTGTAGAA
ATGTCTTGGTGTCCTCGTCCAATCAGGTAGCCATCTCTGAAATATCTGGCTCCGTTGCAACTCC
GAACGACCTGCTGGCAACGTAAAATTCTCCGGGGTAAAACTTAAATGTGGAGTAATGGAACC
AGAAACGTCTCTTCCCTTCTCTCTCCTTCCACCGCCCGTTACCGTCCCTAGGAAATTTTACTCT
GCTGGAGAGCTTCTTCTACGGCCCCCTTGCAGCAATGCTCTTCCCAGCATTACGTTGCGGGTA
AAACGGAGGTCGTGTACCCGACCTAGCAGCCCAGGGATGGAAAAGTCCCGGCCGTCGCTGG
CAATAATAGCGGGCGGACGCATGTCATGAGATTATTGGAAACCACCAGAATCGAATATAAAA
GGCGAACACCTTTCCCAATTTTGGTTTCTCCTGACCCAAAGACTTTAAATTTAATTTATTTGTCC

CTATTTCAATCAATTGAACAACTATCAAAACACCATGGTGAGCAAGGGCGAGGAGGATAACA
TGGCCATCATCAAGGAGTTCATGCGCTTCAAGGTGCACATGGAGGGCTCCGTGAACGGCCAC
GAGTTCGAGATCGAGGGCGAGGGCGAGGGCCGCCCCTACGAGGGCACCCAGACCGCCAAG
CTGAAGGTGACCAAGGGTGGCCCCCTGCCCTTCGCCTGGGACATCCTGTCCCCTCAGTTCAT
GTACGGCTCCAAGGCCTACGTGAAGCACCCCGCCGACATCCCCGACTACTTGAAGCTGTCCT
TCCCCGAGGGCTTCAAGTGGGAGCGCGTGATGAACTTCGAGGACGGCGGCGTGGTGACCGT
GACCCAGGACTCCTCCCTGCAGGACGGCGAGTTCATCTACAAGGTGAAGCTGCGCGGCACC
AACTTCCCCTCCGACGGCCCCGTAATGCAGAAAAAGACCATGGGCTGGGAGGCCTCCTCCG
AGCGGATGTACCCCGAGGACGGCGCCCTGAAGGGCGAGATCAAGCAGAGGCTGAAGCTGA
AGGACGGCGGCCACTACGACGCTGAGGTCAAGACCACCTACAAGGCCAAGAAGCCCGTGCA
GCTGCCCGGCGCCTACAACGTCAACATCAAGTTGGACATCACCTCCCACAACGAGGACTACA
CCATCGTGGAACAGTACGAACGCGCCGAGGGCCGCCACTCCACCGGCGGCATGGACGAGCT
GTACAAGTAAGCTTCTGTTTAATTCTGTAGGATAAACGAAATAACTGTATCTTTATAATCATCCT
TAAGATAGCCCTGGACTCGTCAGACATTAGTTCGCGCACATTATTTTCCTCCACTGTCTGCAAA
CAGCCACACGGCATTTCCCTCTCTAGTTTCATTTATCAAAGTAGCACAACACGCTCCGGCATG
GACATGGAGGCCCAGAATACCCTCCTTGACAGTCTTGACGTGCGCAGCTCAGGGGCATGATG
TGACTGTCGCCCGTACATTTAGCCCATACATCCCCATGTATAATCATTTGCATCCATACATTTT
GATGGCCGCACGGCGCGAAGCAAAAATTACGGCTCCTCGCTGCAGACCTGCGAGCAGGGAA
ACGCTCCCCTCACAGACGCGTTGAATTGTCCCCACGCCGCGCCCCTGTAGAGAAATATAAAA
GGTTAGGATTTGCCACTGAGGTTCTTCTTTCATATACTTCCTTTTAAAATCTTGCTAGGATACAG
TTCTCACATCACATCCGAACATAAACAACCATGGGTACCACTCTTGACGACACGGCTTACCGG
TACCGCACCAGTGTCCCGGGGGACGCCGAGGCCATCGAGGCACTGGATGGGTCCTTCACCA
CCGACACCGTTTTTCGCGTCACCGCCACCGGGGACGGCTTCACCCTGCGGGAGGTGCCGGT
GGACCCGCCCCTGACCAAGGTGTTCCCCGACGACGAATCGGACGACGAATCGGACGACGGG
GAGGACGGCGACCCGGACTCCCGGACGTTCGTCGCGTACGGGGACGACGGCGACCTGGCG
GGCTTCGTGGTCGTCTCGTACTCCGGCTGGAACCGCCGGCTGACCGTCGAGGACATCGAGGT
CGCCCCGGAGCACCGGGGGCACGGGGTCGGGCGCGCGTTGATGGGGCTCGCGACGGAGTT
CGCCCGCGAGCGGGGCGCCGGGCACCTCTGGCTGGAGGTCACCAACGTCAACGCACCGGC
GATCCACGCGTACCGGCGGATGGGGTTCACCCTCTGCGGCCTGGACACCGCCCTGTACGAC
GGCACCGCCTCGGACGGCGAGCAGGCGCTCTACATGAGCATGCCCTGCCCCTAATCAGTACT
GACAATAAAAAGATTCTTGTTTTCAAGAACTTGTCATTTGTATAGTTTTTTTATATTGTAGTTGTT
CTATTTTAATCAAATGTTAGCGTGATTTATATTTTTTTTCGCCTCGACATCATCTGCCCAGATGC
GAAGTTAAGTGCGCAGAAAGTAATATCATGCGTCAATCGTATGTGAATGCTGGTCGCTATACT
GGCTTCAGCAACGTTGTCACTGAAGTTGGCATCAGTATCTACAAACCCTACATACCTCTTGAA
GGTCCAGAAGGTAAGCATTTGCCCGCTCCATTCCTACCCGACATATCCACTCTACATTTTATA
GTTCAAGGCACCGAAAAAGTTCGAAACAAGAAGTTTGTTCCTGATAACAAGGATTTCTTTATT
GGTGGGACTTCATTTACTGTCTCCAAGAAGGACATTTCTGCTGTCATAACCGAGATAGTCTCA
CAGTTCGAGTCTACTGATGACCAAAAGTCAGAGAAATTTACCATGAAATCTCCTCCTCCGGTT
GCTTTAGTTGGGCATAATCTGATAGGAGGCCTCAAGACACTGAAGAATGCGGGTATCACCATT
CCCATACTTCCAATCATAGGATCCGGCGCGCCGATACTCGAGAATTATGGCTTAATCAAGTGA
ATACATCAAAGTCAAACTTAAAATACATTCTTCGCAGGCTTGGCTTGCCACATAGTTTTCTTCA
CAATGCTGCAAATGACGCTTATTATACCCTTTTGGCATCGTTGAAGCTTGCACAAACGAACGT
CCCACTTAATCTTCTGTACTCTGAAGAGGAGTGGGAAATACCAAGAAAAACATCAAACTCGAA
TGATTTTCCCGAACCCCTACCACAAGATATTCATCAGCTGCGAGATAGGCTGATCAGGAGCAA
GCTCGTACGAGAAGAAACAAAATGACAAAAAAAATCCTATACTATATAGGTTACAAATAAAAA
AGTATCAAAAATGAAGCCTGCATCTCTCAGGCAAATGGCATTCTGACATCCTCT-3´

5.1.2. BB3aK_FS4_pGAP_eGFP_RPS3TT



5´-CGCTGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCG
TTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTG
GCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCT
CTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGG
CGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGG
GCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTG
AGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGC
AGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACAC
TAGAAGAACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGG
TAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCA
GATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTGGAGCG
CAAAAAACCCCGCCCCTGACAGGGCGGGGTTTTTTCGCGATCGGAGGATCCTTTTTTGTAGAA
ATGTCTTGGTGTCCTCGTCCAATCAGGTAGCCATCTCTGAAATATCTGGCTCCGTTGCAACTCC
GAACGACCTGCTGGCAACGTAAAATTCTCCGGGGTAAAACTTAAATGTGGAGTAATGGAACC
AGAAACGTCTCTTCCCTTCTCTCTCCTTCCACCGCCCGTTACCGTCCCTAGGAAATTTTACTCT
GCTGGAGAGCTTCTTCTACGGCCCCCTTGCAGCAATGCTCTTCCCAGCATTACGTTGCGGGTA
AAACGGAGGTCGTGTACCCGACCTAGCAGCCCAGGGATGGAAAAGTCCCGGCCGTCGCTGG
CAATAATAGCGGGCGGACGCATGTCATGAGATTATTGGAAACCACCAGAATCGAATATAAAA
GGCGAACACCTTTCCCAATTTTGGTTTCTCCTGACCCAAAGACTTTAAATTTAATTTATTTGTCC
CTATTTCAATCAATTGAACAACTATCAAAACACCATGGTGAGCAAGGGCGAGGAGCTGTTCAC
CGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTG
TCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCA
CCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGTGCAGTG
CTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAG
GCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGACCCGCGCCGA
GGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAG

GAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACAGCCACAACGTCTATAT
CATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCGCCACAACATCGAG
GACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCATCGGCGACGGCCCCG
TGCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGA
GAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGGGATCACTCTCGGCATGG
ACGAGCTGTACAAGTAAGCTTCTGTTTAATTCTGTAGGATAAACGAAATAACTGTATCTTTATA
ATCATCCTTAAGATAGCCCTGGACTCGTCAGACATTAGTTCGCGCACATTATTTTCCTCCACTG
TCTGCAAACAGCCACACGGCATTTCCCTCTCTAGTTTCATTTATCAAAGTAGCACAACACGCT
AGCTCATGGACATGGAGGCCCAGAATACCCTCCTTGACAGTCTTGACGTGCGCAGCTCAGGG
GCATGATGTGACTGTCGCCCGTACATTTAGCCCATACATCCCCATGTATAATCATTTGCATCCA
TACATTTTGATGGCCGCACGGCGCGAAGCAAAAATTACGGCTCCTCGCTGCAGACGTGCGAG
CAGGGAAACGCTCCCCTCACAGACGCGTTGAATTGTCCCCACGCCGCGCCCCTGTAGAGAAA
TATAAAAGGTTAGGATTTGCCACTGAGGTTCTTCTTTCATATACTTCCTTTTAAAATCTTGCTAG
GATACAGTTCTCACATCACATCCGAACATAAACAACCATGGGTAAGGAAAAGACTCACGTTTC
GAGGCCGCGATTAAATTCCAACATGGATGCTGATTTATATGGGTATAAATGGGCTCGCGATAA
TGTCGGGCAATCAGGTGCGACAATCTATCGATTGTATGGGAAGCCCGATGCGCCAGAGTTGT
TTCTGAAACATGGCAAAGGTAGCGTTGCCAATGATGTTACAGATGAGATGGTCAGACTAAACT
GGCTGACGGAATTTATGCCTCTTCCGACCATCAAGCATTTTATCCGTACTCCTGATGATGCATG
GTTACTCACCACTGCGATCCCCGGCAAAACAGCATTCCAGGTATTAGAAGAATATCCTGATTC
AGGTGAAAATATTGTTGATGCGCTGGCAGTGTTCCTGCGCCGGTTGCATTCGATTCCTGTTTGT
AATTGTCCTTTTAACAGCGATCGCGTATTTCGTCTCGCTCAGGCGCAATCACGAATGAATAAC
GGTTTGGTTGATGCGAGTGATTTTGATGACGAGCGTAATGGCTGGCCTGTTGAACAAGTCTGG
AAAGAAATGCATAAGCTTTTGCCATTCTCACCGGATTCAGTCGTCACTCATGGTGATTTCTCAC
TTGATAACCTTATTTTTGACGAGGGGAAATTAATAGGTTGTATTGATGTTGGACGAGTCGGAAT
CGCAGACCGATACCAGGATCTTGCCATCCTATGGAACTGCCTCGGTGAGTTTTCTCCTTCATT
ACAGAAACGGCTTTTTCAAAAATATGGTATTGATAATCCTGATATGAATAAATTGCAGTTTCAT
TTGATGCTCGATGAGTTTTTCTAATCAGTACTGACAATAAAAAGATTCTTGTTTTCAAGAACTTG
TCATTTGTATAGTTTTTTTATATTGTAGTTGTTCTATTTTAATCAAATGTTAGCGTGATTTATATTT
TTTTTCGCCTCGACATCATCTGCCCAGATGCGAAGTTAAGTGCGCAGAAAGTAATATCATGCG
TCAATCGTATGTGAATGCTGGTCGCTATACTGGCTTCAGCAACGTTGTCACTGAAGTTGGCAT
CAGTATCTACAAACCCTACATACCTCTTGAAGGTCCAGAAGGTAAGCATTTGCCCGCTCCATT
CCTACCCGACATATCCACTCTACATTTTATAGTTCAAGGCACCGAAAAAGTTCGAAACAAGAA
GTTTGTTCCTGATAACAAGGATTTCTTTATTGGTGGGACTTCATTTACTGTCTCCAAGAAGGAC
ATTTCTGCTGTCATAACCGAGATAGTCTCACAGTTCGAGTCTACTGATGACCAAAAGTCAGAG
AAATTTACCATGAAATCTCCTCCTCCGGTTGCTTTAGTTGGGCATAATCTGATAGGAGGCCTCA
AGACACTGAAGAATGCGGGTATCACCATTCCCATACTTCCAATCATAGGATCCGGCGCGCCG
ATACTCGAGAATTATGGCTTAATCAAGTGAATACATCAAAGTCAAACTTAAAATACATTCTTCG
CAGGCTTGGCTTGCCACATAGTTTTCTTCACAATGCTGCAAATGACGCTTATTATACCCTTTTG
GCATCGTTGAAGCTTGCACAAACGAACGTCCCACTTAATCTTCTGTACTCTGAAGAGGAGTGG
GAAATACCAAGAAAAACATCAAACTCGAATGATTTTCCCGAACCCCTACCACAAGATATTCAT
CAGCTGCGAGATAGGCTGATCAGGAGCAAGCTCGTACGAGAAGAAACAAAATGACAAAAAA
AATCCTATACTATATAGGTTACAAATAAAAAAGTATCAAAAATGAAGCCTGCATCTCTCAGGC
AAATGGCATTCTGACATCCTCT-3´

5.2. Primers

5.2.1. 468_AOX1tt_locus_fwd

5´-GGGCTTGAAGGTTGGTGA-3´

5.2.2. 469_AOX1tt_locus_rev

5´-CGAGCTGATTTGGGGTTT-3´

5.2.3. 470_ AOX1tt_locus_fwd2

5´-ATGGGGTGGTGTTTTGGA-3´

5.2.4. 488_ AOX1tt_locus_fwd3

5´-CTGAAGAAACAGTAGTGGTCTTGGAAGTTT-3´

5.2.5. 498_ AOX1tt_locus_rev2

5´-CTCTAAAACAAGATAAGTGCGTCTCA-3´

5.2.6. 490_ AOX1tt_locus_fwd4

5´-AAACTTCCAAGACCACTACTGTTTCTTCAG-3´

5.3. Spotting assay results

On the following pages, summarized results of the spotting assays are shown. The rating was set in relation to the reference strain *K. phaffii* CBS 2612 on the respective plate, whereby the same growth was rated with a 0. Better growth than the reference was rated in ascending order from 1 to 2. Growth worse than the reference was rated in descending order from -1 to -2.

Table A1: Rating of the spotting assays in relation to the growth of K. phaffii CBS 2612

	rating
best growth	2
good growth	1
same growth	0
bad growth	-1
worst growth (nearly no growth)	-2

5.3.1. After 2 days on different carbon sources

after 2 days											
		30 °C					37 °C				
species name	strain code	Glucose	Xylose	Glycerol	MeOH	EtOH	Glucose	Xylose	Glycerol	MeOH	EtOH
K. phaffii	CBS 2612	0	0	0	0	0	0	0	0	0	0
K. phaffii	CBS 7435	0	1	0	0	0	0	0	0	0	0
K. phaffii	85- 263.1	0	1	0	0	0	1	1	1	1	0
K. phaffii	03- 318t1	0	1	0	0	0	1	0	1	0	0
K. phaffii	03- 328y3	0	1	0	0	0	1	0	1	0	0
K. phaffii	81- 86	0	0	0	0	0	1	1	1	1	1
K. phaffii	81- 18	0	0	0	0	0	1	1	1	1	1
K. phaffii	82- 16	0	0	0	0	0	1	0	1	1	1
K. phaffii	85- 348.1	0	0	0	0	0	0	1	0	0	0
K. phaffii	85- 926.1	0	0	0	0	0	0	0	0	-2	-2
K. phaffii	91- 119.3	0	0	0	0	0	1	1	1	1	1
K. phaffii	91- 132.2	0	0	0	0	0	1	0	1	1	0
Komagataella	HA71	0	0	0	-1	-1	-2	0	-1	-2	-2
K. pastoris	DSMZ 70382 (CBS 704)	0	2	0	-1	0	-2	0	-1	-2	-2
K. pastoris	DSMZ 70877	0	-1	0	0	-1	0	0	0	0	0
K. pastoris	CBS 9178	0	0	0	0	0	0	0	0	0	0
K. pastoris	CBS 9173	0	0	0	-1	0	0	0	0	-1	0
K. pastoris	CBS 9185	0	0	0	0	0	0	0	0	0	0
K. ulmi	CBS 12361	0	-1	0	0	0	-1	0	-1	-2	-2
K. ulmi	91- 206.2	0	0	0	0	0	-2	0	-2	-2	-2
K. ulmi	03- 338t1	0	0	-1	-1	0	-2	0	-1	-2	-1
K. pseudopastoris	CBS 9187	0	1	0	0	0	-1	0	-1	-2	-2
K. pseudopastoris	CBS 9189	0	1	0	0	0	0	0	0	-1	-2
K. kurtzmanii	CBS 12817	0	-2	-1	-1	-1	-2	-2	-2	-2	-2
K. populi	CBS 12362	0	2	0	0	1	1	2	1	1	1

5.3.2. After 2 days with stressors

		after 2 days									
		30 °C					37 °C				
species name	strain code	1 M NaCl	pH 9	pH 4	0.5 mM H_2O_2	1 mM H_2O_2	1 M NaCl	pH 9	pH 4	0.5 mM H_2O_2	1 mM H_2O_2
K. phaffii	CBS 2612	0	0	0	0	0	0	0	0	0	0
K. phaffii	CBS 7435	0	0	0	0	-1	1	0	0	-1	0
K. phaffii	85- 263.1	0	0	0	0	1	0	0	0	0	0
K. phaffii	03- 318t1	0	0	0	0	1	0	0	-1	-1	0
K. phaffii	03- 328y3	0	0	0	1	1	0	0	0	1	0
K. phaffii	81- 86	0	0	0	0	1	0	0	1	1	2
K. phaffii	81- 18	0	0	0	0	0	0	0	1	1	1
K. phaffii	82- 16	0	0	0	-1	0	0	0	-1	-1	0
K. phaffii	85- 348.1	0	0	0	0	0	0	0	1	1	0
K. phaffii	85- 926.1	0	0	0	0	0	0	0	1	1	1
K. phaffii	91- 119.3	0	0	0	0	0	0	0	-1	0	0
K. phaffii	91- 132.2	0	0	0	0	0	0	0	0	0	0
Komagataella	HA71	-1	-1	0	-1	-1	0	0	-1	-2	0
K. pastoris	DSMZ 70382 (CBS 704)	0	-1	0	0	1	1	0	-1	-2	0
K. pastoris	DSMZ 70877	0	0	0	0	0	1	0	0	-1	0
K. pastoris	CBS 9178	0	0	0	0	0	1	0	0	-1	0
K. pastoris	CBS 9173	-1	-2	0	-1	0	0	-1	0	-1	0
K. pastoris	CBS 9185	0	-1	0	0	1	1	0	0	-1	0
K. ulmi	CBS 12361	0	-1	0	-1	-2	0	0	-1	-2	0
K. ulmi	91- 206.2	0	0	0	0	1	0	0	1	1	0
K. ulmi	03- 338t1	0	0	0	0	1	0	0	1	1	1
K. pseudopastoris	CBS 9187	0	0	0	-2	-2	0	-1	-1	-2	0
K. pseudopastoris	CBS 9189	0	2	0	-1	-2	1	0	-1	-2	0
K. kurtzmanii	CBS 12817	0	-2	-1	-2	-2	0	-2	-2	-2	0
K. populi	CBS 12362	0	-1	0	0	-1	1	0	0	-2	0

5.3.3. After 7 days with stressors

		after 7 days									
		30 °C					37 °C				
species name	strain code	1 M NaCl	pH 9	pH 4	0.5 mM H_2O_2	1 mM H_2O_2	1 M NaCl	pH 9	pH 4	0.5 mM H_2O_2	1 mM H_2O_2
K. phaffii	CBS 2612	0	0	0	0	0	0	0	0	0	0
K. phaffii	CBS 7435	0	0	0	0	0	2	1	0	0	0
K. phaffii	85- 263.1	0	-1	0	0	0	2	-2	0	1	1
K. phaffii	03- 318t1	0	-1	0	0	0	1	-1	0	0	2
K. phaffii	03- 328y3	-1	-1	0	0	0	1	-1	0	0	2
K. phaffii	81- 86	0	-1	0	0	0	1	-2	0	0	1
K. phaffii	81- 18	0	0	0	0	0	1	-1	0	0	1
K. phaffii	82- 16	0	-1	0	0	0	1	-1	0	0	1
K. phaffii	85- 348.1	0	-1	0	0	0	0	-1	0	0	1
K. phaffii	85- 926.1	0	-1	0	0	0	0	-2	0	0	0
K. phaffii	91- 119.3	0	-1	0	0	0	1	-1	0	0	1
K. phaffii	91- 132.2	-1	-1	0	0	0	1	-2	0	0	1
Komagataella	HA71	-1	0	0	0	0	-2	-2	-2	-1	-1
K. pastoris	DSMZ 70382 (CBS 704)	0	2	0	0	0	-1	-2	-2	-2	-2
K. pastoris	DSMZ 70877	0	1	0	0	0	2	-1	0	0	0
K. pastoris	CBS 9178	0	1	0	0	0	2	0	0	0	0
K. pastoris	CBS 9173	0	0	0	0	0	2	-2	0	0	0
K. pastoris	CBS 9185	0	1	0	0	0	2	-1	0	0	-1
K. ulmi	CBS 12361	0	1	0	0	0	0	0	0	-1	-2
K. ulmi	91- 206.2	0	-1	0	0	0	0	-2	0	-1	-2
K. ulmi	03- 338t1	0	-2	0	0	0	0	-2	0	-1	-2
K. pseudopastoris	CBS 9187	-1	2	0	0	0	-2	0	-1	-1	-2
K. pseudopastoris	CBS 9189	0	2	0	0	0	-1	2	-2	0	-1
K. kurtzmanii	CBS 12817	0	2	0	0	0	-2	-2	-2	-2	-2
K. populi	CBS 12362	-1	1	0	0	0	1	1	0	0	-2

REFERENCES

1. Bengtsson, Oskar; Hahn-Hägerdal, Bärbel; Gorwa-Grauslund, Marie F. (2009): Xylose reductase from Pichia stipitis with altered coenzyme preference improves ethanolic xylose fermentation by recombinant Saccharomyces cerevisiae. In Biotechnol Biofuels 2 (1), p. 9. DOI: 10.1186/1754-6834-2-9.
2. Biolabs, New England (2020a): NEB Tm Calculator. Available online at https://tmcalculator.neb.com/#!/main, updated on 7/25/2020, checked on 12/16/2020.
3. Biolabs, New England (2020b): 1 kb Plus DNA Ladder | NEB. Available online at https://international.neb.com/products/n3200-1-kb-plus-dna-ladder#Product%20Information, updated on 12/16/2020, checked on 12/16/2020.
4. Boekhout, Teun; Kurtzman, Cletus P. (1996): Principles and Methods Used in Yeast Classification, and an Overview of Currently Accepted Yeast Genera. In Klaus Wolf (Ed.): Nonconventional Yeasts in Biotechnology. A Handbook. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 1–81.
5. Braun-Galleani, Stephanie; Dias, Julie A.; Coughlan, Aisling Y.; Ryan, Adam P.; Byrne, Kevin P.; Wolfe, Kenneth H. (2019): Genomic diversity and meiotic recombination among isolates of the biotech yeast Komagataella phaffii (Pichia pastoris). In Microb Cell Fact 18 (1), p. 211. DOI: 10.1186/s12934-019-1260-4.
6. Briza, P.; Ellinger, A.; Winkler, G.; Breitenbach, M. (1988): Chemical composition of the yeast ascospore wall. The second outer layer consists of chitosan. In The Journal of biological chemistry 263 (23), pp. 11569–11574.
7. Cao, Liang; Alani, Eric; Kleckner, Nancy (1990): A pathway for generation and processing of double-strand breaks during meiotic recombination in S. cerevisiae. In Cell 61 (6), pp. 1089–1101. DOI: 10.1016/0092-8674(90)90072-M.
8. Cregg, J. M.; Barringer, K. J.; Hessler, A. Y.; Madden, K. R. (1985): Pichia pastoris as a host system for transformations. In Molecular and Cellular Biology 5 (12), pp. 3376–3385. DOI: 10.1128/mcb.5.12.3376.
9. Cregg, James M.; Vedvick, Thomas S.; Raschke, William C. (1993): Recent Advances in the Expression of Foreign Genes in Pichia pastoris. In Nat Biotechnol 11 (8), pp. 905–910. DOI: 10.1038/nbt0893-905.
10. Duret, Laurent; Galtier, Nicolas (2009): Biased gene conversion and the evolution of mammalian genomic landscapes. In Annual review of genomics and human genetics 10, pp. 285–311. DOI: 10.1146/annurev-genom-082908-150001.
11. Gasch, A. P.; Spellman, P. T.; Kao, C. M.; Carmel-Harel, O.; Eisen, M. B.; Storz, G. et al. (2000): Genomic expression programs in the response of yeast cells to environmental changes. In Molecular biology of the cell 11 (12), pp. 4241–4257. DOI: 10.1091/mbc.11.12.4241.

12. Greig, D. (2009): Reproductive isolation in Saccharomyces. In Heredity 102 (1), pp. 39–44. DOI: 10.1038/hdy.2008.73.

13. Guilliermond A. (1920): Zygosaccharomyces pastori, nouvelle espèce de levures copulation hétérogamique. (36), pp. 203–211.

14. Hanson, Sara J.; Byrne, Kevin P.; Wolfe, Kenneth H. (2014): Mating-type switching by chromosomal inversion in methylotrophic yeasts suggests an origin for the three-locus Saccharomyces cerevisiae system. In Proceedings of the National Academy of Sciences of the United States of America 111 (45), E4851-8. DOI: 10.1073/pnas.1416014111.

15. Hanson, Sara J.; Byrne, Kevin P.; Wolfe, Kenneth H. (2017): Flip/flop mating-type switching in the methylotrophic yeast Ogataea polymorpha is regulated by an Efg1-Rme1-Ste12 pathway. In PLoS genetics 13 (11), e1007092. DOI: 10.1371/journal.pgen.1007092.

16. Heistinger, Lina; Gasser, Brigitte; Mattanovich, Diethard (2018): Creation of Stable Heterothallic Strains of Komagataella phaffii Enables Dissection of Mating Gene Regulation. In Molecular and Cellular Biology 38 (2). DOI: 10.1128/MCB.00398-17.

17. Kurtzman, C. P.; Fell, Jack W.; Boekhout, T. (2010): The yeasts. A taxonomic study, vol. 1 - 3. 5th ed. Amsterdam, Oxford: Elsevier. Available online at https://books.google.at/books?id=yfg79rlIFIkC.

18. Li, Pengfei; Sun, Hongbing; Chen, Zao; Li, Yin; Zhu, Taicheng (2015): Construction of efficient xylose utilizing Pichia pastoris for industrial enzyme production. In Microb Cell Fact 14, p. 22. DOI: 10.1186/s12934-015-0206-8.

19. Merlini, Laura; Dudin, Omaya; Martin, Sophie G. (2013): Mate and fuse: how yeast cells do it. In Open Biology 3 (3), p. 130008. DOI: 10.1098/rsob.130008.

20. Mrak, E. M.; MILLER, M. W.; PHAFF, H. J. (1978): The Life of Yeasts. 2. Aufl., 2nd ed., rev. a. enl. Reprint 2014. s.l.: Harvard University Press.

21. Naumov, Gennadi I.; Naumova, Elena S.; Boundy-Mills, Kyria L. (2018): Description of Komagataella mondaviorum sp. nov., a new sibling species of Komagataella (Pichia) pastoris. In Antonie van Leeuwenhoek 111 (7), pp. 1197–1207. DOI: 10.1007/s10482-018-1028-6.

22. Neiman, A. M. (1998): Prospore membrane formation defines a developmentally regulated branch of the secretory pathway in yeast. In The Journal of cell biology 140 (1), pp. 29–37. DOI: 10.1083/jcb.140.1.29.

23. Neiman, Aaron M. (2005): Ascospore formation in the yeast Saccharomyces cerevisiae. In Microbiology and molecular biology reviews : MMBR 69 (4), pp. 565–584. DOI: 10.1128/MMBR.69.4.565-584.2005.

24. Phaff, H. J.; Miller, M. W.; Shifrine, M. (1956): The taxonomy of yeasts isolated from Drosophila in the Yosemite region of California. In Antonie van Leeuwenhoek 22 (2), pp. 145–161. DOI: 10.1007/BF02538322.

25. Prielhofer, Roland; Barrero, Juan J.; Steuer, Stefanie; Gassler, Thomas; Zahrl, Richard; Baumann, Kristin et al. (2017): GoldenPiCS: a Golden Gate-derived modular cloning system for applied synthetic biology in the yeast Pichia pastoris. In BMC systems biology 11 (1), p. 123. DOI: 10.1186/s12918-017-0492-3.

26. Sibirny, Andriy (2019): Non-conventional Yeasts: from Basic Research to Application. 1st ed. 2019: Springer International Publishing. Available online at https://books.google.at/books?id=l-eoDwAAQBAJ.

27. Steensels, Jan; Snoek, Tim; Meersman, Esther; Picca Nicolino, Martina; Voordeckers, Karin; Verstrepen, Kevin J. (2014): Improving industrial yeast strains: exploiting natural and artificial diversity. In Fems Microbiology Reviews 38 (5), pp. 947–995. DOI: 10.1111/1574-6976.12073.

28. Swinnen, Steve; Thevelein, Johan M.; Nevoigt, Elke (2012): Genetic mapping of quantitative phenotypic traits in Saccharomyces cerevisiae. In FEMS yeast research 12 (2), pp. 215–227. DOI: 10.1111/j.1567-1364.2011.00777.x.

29. taxonomy | Definition, Examples, Levels, & Classification (2020). Available online at https://www.britannica.com/science/taxonomy, updated on 6/3/2020, checked on 6/3/2020.

30. Tsai, Isheng J.; Bensasson, Douda; Burt, Austin; Koufopanou, Vassiliki (2008): Population genomics of the wild yeast Saccharomyces paradoxus: Quantifying the life cycle. In Proceedings of the National Academy of Sciences of the United States of America 105 (12), pp. 4957–4962. DOI: 10.1073/pnas.0707314105.

31. Yamada, Y.; Matsuda, M.; Maeda, K.; Mikata, K. (1995): The phylogenetic relationships of methanol-assimilating yeasts based on the partial sequences of 18S and 26S ribosomal RNAs: the proposal of Komagataella gen. nov. (Saccharomycetaceae). In Bioscience, biotechnology, and biochemistry 59 (3), pp. 439–444. DOI: 10.1271/bbb.59.439.