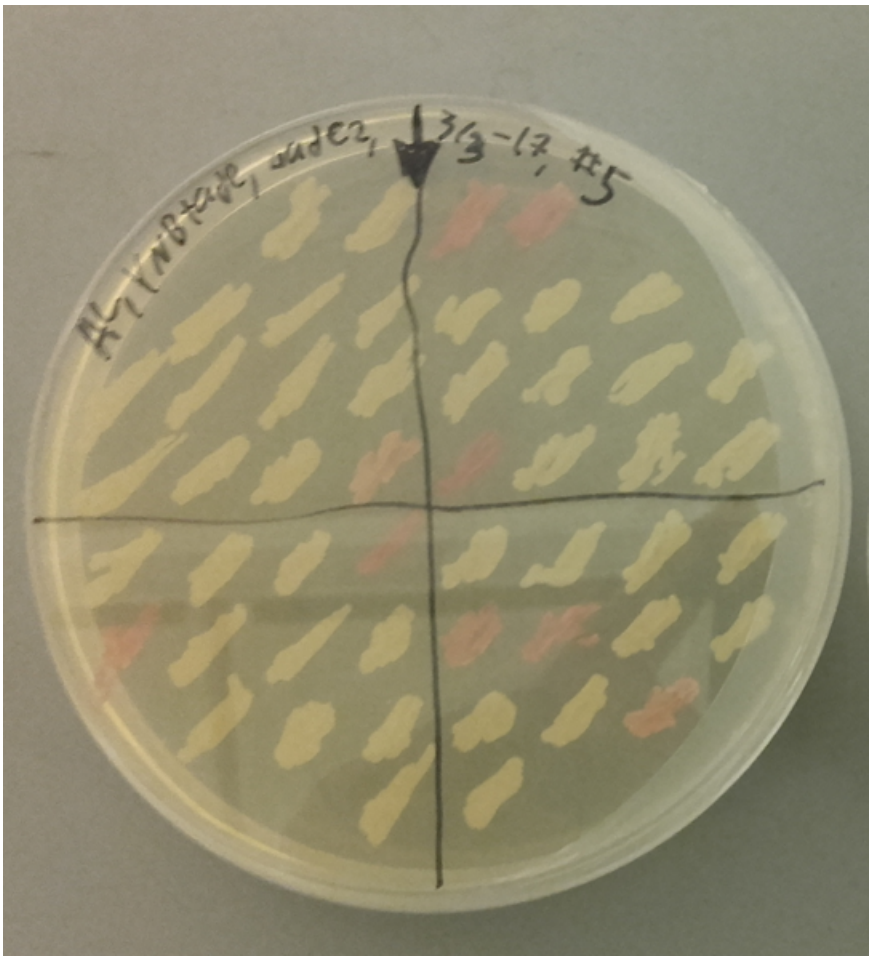
CHALMERS
UNIVERSITY OF TECHNOLOGY



Development & application of a molecular toolbox for the unconventional yeast *Candida intermedia*

Master's thesis in Biotechnology

ADAM LARSSON

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

Master's thesis 2017:June

Development & application of a molecular toolbox for the unconventional yeast *Candida intermedia*

ADAM LARSSON

CHALMERS
UNIVERSITY OF TECHNOLOGY

Department of Biology and Biological Engineering
Division of Industrial Biotechnology
Chalmers University of Technology
Gothenburg, Sweden 2017

Development & application of a molecular toolbox for the unconventional yeast *Candida intermedia*

ADAM LARSSON

Supervisor: Cecilia Geijer, Industrial Biotechnology
Supervisor: Fábio Faria Oliveira, Industrial Biotechnology
Examiner: Lisbeth Olsson, Industrial Biotechnology

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

Cover: Screening of strains successfully deleted of the *ADE2* gene.

Typeset in LaTeX
Gothenburg, Sweden 2017

Abstract

Over the last years there have been extensive attempts at replacing fossil fuels with the more sustainable option bioethanol. Traditionally bioethanol is produced using *Saccharomyces cerevisiae* to ferment glucose from crops such as sugar cane, but this is unsustainable as it doesn't utilize the whole biomass and the resource generation competes with food production. In search of more sustainable substrates the scientific community has turned to more unconventional microorganisms as *S. cerevisiae* is less adept at using sugars other than glucose.

In this project an interesting *Candida intermedia* strain, named C11, was studied. The strain has exhibited interesting traits for utilization of lactose/galactose and xylose. Three aldose reductase genes, *XYL1*, *XYL1_2* & *XYL1_3*, have been identified to be of especial interest for xylose metabolism. *XYL1_2* also seems to be of interest for lactose/galactose metabolism as the gene is located in a novel galactose gene cluster. The aim of this project was thus to study the physiological roles of the three aldose reductases Xyl1p, Xyl1_2p & Xyl1_3p.

The expression levels of the three aldose reductase genes were measured in cells grown in glucose, lactose & xylose to find hints as to the role of the enzymes. *XYL1* was found to be upregulated during growth on xylose, whereas *XYL1_2* was upregulated in lactose containing growth medium. This suggests Xyl1p's role to be that of a xylose reductase used in xylose metabolism. It also suggests that Xyl1_2p is important for lactose/galactose utilization. To further study the physiological role of the aldose reductases the goal was to delete them from the genome to obtain a phenotype indicating the role of the enzymes.

C. intermedia strongly favors non-homologous end joining (NHEJ) over homologous recombination (HR) making targeted gene deletions near-impossible (<1 % of transformed cells). For this reason genetic tools were developed, at great success, for *C. intermedia* C11 to facilitate gene deletions with a high gene target efficiency. The best obtained deletion protocol was based on an unconventional method of using a split marker deletion cassette. The resulting protocol is quick and provides a gene targeting efficiency of up to 67 %. The deletion protocol was applied in the deletion of aldose reductase gene *XYL1_2*. The deletion strain did however not produce a visible phenotype when grown on galactose.

During the project suggestions as to the metabolic role of Xyl1p & Xyl1_2p were found. A very efficient gene deletion protocol was developed with a gene targeting efficiency of 67 % and it was applied to delete *XYL1_2*. However, more work is needed to fully evaluate the three aldose reductases and the novel galactose gene cluster, for example by producing more deletion strains.

Keywords: yeast, homologous recombination, non-homologous end joining, split-marker, *Candida intermedia*, xylose, galactose.

Acknowledgements

First and foremost I must show my most humbly appreciation of my supervisors Cecilia Geijer & Fábio Faria Oliveira. My thesis would not be the same if it wasn't for their hard work and for them always being available for discussions and troubleshooting while also contributing with new ideas.

I must also thank all of the Industrial Biotechnology research group, and especially its head Lisbeth Olsson, at Chalmers for giving me the opportunity to perform this project in such an experienced group.

This project was made possible by funding from the Swedish energy agency (Energimyndigheten).

Adam Larsson, Gothenburg, 12th of June 2017

Contents

List of Figures

List of Tables

1

Introduction

Over the last decades there has been an ever increasing interest in xylose fermentation which originates from the production of bioethanol as an alternative to fossil fuels. By replacing fossil fuels in the transport sector with bioethanol it would enable a lower environmental impact [1]. The usage of lignocellulose in the process of producing second-generation bioethanol would make this fuel even more sustainable, as the current production of first-generation bioethanol is based on fermenting the sugars in crops such as sugar cane, which doesn't utilize the whole plant biomass [2]. Another concern with the current production of bioethanol is that many of the crops currently used are otherwise used by the food industry, meaning that there is an increase in food prices leading to problems with the social sustainability of bioethanol production [3]. But if lignocellulose was used it would mean that there is a wider selection of plants available that don't compete with food production. However, for production of bioethanol from lignocellulosic materials to be both sustainable and economically competitive it is required to have a strain that can efficiently ferment both glucose and xylose as they together make up the majority of the sugar content of many lignocellulosic materials, for example when obtained from annual crops or hardwood. The wild type (WT) of *Saccharomyces cerevisiae* is excellent at fermenting glucose and has a long history of industrial use as an ethanol producing organism, but it is unable to grow on and ferment xylose [4]. To overcome the inability of the WT to utilize xylose genetically engineered strains of *S. cerevisiae* have been developed to possess this capacity [5]. However, many of the engineered *S. cerevisiae* strains have exhibited low ethanol yields due to by-product formation and poor xylose consumption rates [6, 7].

One of the main concerns and causes of problems in many xylose fermentative strains of *S. cerevisiae* is the fact that there is a co-factor imbalance in the fungal xylose metabolic pathway [8]. In the xylose pathway xylose is converted to xylitol by a xylose reductase (XR), a xylitol dehydrogenase (XDH) then converts xylitol into xylulose [9] that can further be led into the pentose phosphate pathway and glycolysis [10]. XR's commonly use NADPH in order to convert xylose, while the XDH's use NAD^+ to further convert xylitol, this gives rise to said co-factor imbalance and causes cells to accumulate xylitol [8].

In 2014 an unknown yeast was isolated from an evolutionary engineering experiment. The aim of the experiment had been to obtain a *S. cerevisiae* strain that was capable

of efficiently fermenting xylose. At first the evolution seemed to be a success, but upon further inspection it was discovered that the evolved culture actually contained a contamination. What was remarkable about the contamination was that it was able to grow in a competitive environment with a *S. cerevisiae* strain engineered for xylose fermentation in an industrial setting.

The microorganisms that had contaminated the evolutionary engineering experiment were isolated and found to be of the unconventional yeast species *Candida intermedia*. Several strains of the isolated yeast were tested and the best performing one, named C11, was selected to proceed with for the following experiments. It has long been known that many yeasts of the *Candida* genus are adept xylose fermenters and in this aspect *C. intermedia* is no different [11]. But the strain isolated from the evolutionary engineering that was later to be used in this project displayed even more interesting characteristics. When the genome was sequenced three separate aldose reductases, *XYL1*, *XYL1_2* & *XYL1_3*, were found that are potentially used in the xylose metabolic pathway [12, 13]. Aldose reductases are enzymes that reduce aldoses, a group of monosaccharides. Examples of aldose reductases are xylose reductases which reduce the aldose xylose into xylitol by adding a hydrogen to its oxygen group [14]. Much of the initial interest in the C11 strain was generated with hopes of finding a XR among the aldose reductases that could reduce xylitol accumulation by having a co-factor preference for NADH instead of NADPH. To the best of our knowledge, there are no other species of yeast with three so closely related aldose reductases [13], thus hinting at the possibility of them having different roles but also at the potential of one of them displaying favourable traits. As C11's aldose reductases were further investigated it was discovered that one of them, namely *XYL1_2*, has a mutation in the amino acid sequency in a highly conserved region that has previously been linked to a change in co-factor preference from NADPH to NADH [13, 15]. Aldose reductase *XYL1_2* was thus identified as a high-value target with potential industrial use as it could potentially be cloned into *S. cerevisiae* to solve the xylitol accumulation problem.

One point of interest worth noting with one of the aldose reductases, *XYL1_2*, is that it is located inside of a gene cluster. The gene cluster contains, besides the aldose reductase, several genes associated with galactose utilization [12]. The contents of the gene cluster thus hints at aldose reductase *XYL1_2* being important in the metabolism of galactose, which means that it is also indirectly important for the metabolism of lactose as the biochemical conversion of one lactose yields a glucose and a galactose. The genes in the cluster are however putative and need to be confirmed with for example gene expression data. An interesting point here is that C11 also contains, except for the galactose gene cluster where *XYL1_2* is found, the conserved galactose gene cluster that can be found across other galactose utilizing yeasts [16]. So C11 contains two separate galactose gene clusters which, to the best of our knowledge, has not been observed in other yeast species. This means that the isolated strain C11 displays potential for being useful in industry for both its capacity to utilize xylose as well as lactose.

The industrial interest in microbial growth on lactose stems from the fact that

the dairy industry produces large quantities of lactose-containing waste. This is especially relevant in cheese production which produces huge amounts of cheese whey with high contents of lactose and proteins [17, 18]. Having strains with efficient growth or fermentation on lactose means that the waste can utilized as a useful resource in the production of biomass, bioethanol or other valuable chemicals. As such any microorganisms that display capacity for metabolizing lactose and galactose are potentially industrially relevant microorganisms.

As here laid out, the C11 strain of *C. intermedia* displays interesting traits that can be industrially relevant both for the utilization of xylose and lactose making it a prime target for research. However, with *C. intermedia* being an unconventional organism it lacks many of the genetic tools readily available in other more commonly used organisms. There has been previous work by our research group on C11 where protocols for overexpression of genes have been developed [19], but there still isn't a way to perform efficient deletions of genes in the genome. Observing the phenotype after deletion of a gene of interest is one of the most commonly adopted methods of studying its role in the cell.

Gene deletions are performed by inserting a DNA fragment into the cell via transformation upon which one relies on the existing DNA repair mechanisms of the cell to insert the DNA into the genome. There are two different methods of DNA repair that a microorganism can employ, homologous recombination (HR) and non-homologous end joining (NHEJ). HR can be utilized to insert DNA into a specific targeted location based on homologous regions. However, the mechanisms of NHEJ cause inserted DNA to be integrated at the location of any DNA break meaning that insertion is basically random. This makes NHEJ preferring microorganisms much more difficult to study using targeted genome editing techniques, and it is often required to screen a large amount of transformants to identify a strain with the desired mutation [20].

Although most cells can use both HR and NHEJ they tend to favor using one over the other. *C. intermedia*'s favoured DNA repair pathway is NHEJ [19] meaning that HR frequency is low and targeted genome editing and deletions are difficult to achieve.

1.1 Aim

The overall aim of the research project is to determine the role of the three aldose reductases, *XYL1*, *XYL1_2* & *XYL1_3*, in *C. intermedia*'s utilization of xylose and lactose. For this purpose a molecular toolbox for efficient targeted genome editing has to be developed.

The project aims to answer the following research questions:

1. How can efficient targeted genome editing be achieved in *C. intermedia*? Can the frequency of homologous recombination be increased or can background with incorrect inserts be minimized?

2. What is the physiological role of the three aldose reductases *XYL1*, *XYL1_2* & *XYL1_3*, and how do they contribute to metabolism of xylose or lactose? What are the expression patterns of the aldose reductases when grown on different carbon sources, such as xylose and lactose?

2

Theory

2.1 Xylose metabolism

Xylose is metabolized into D-xylulose-5-phosphate via three separate enzymatic reactions. The first step is performed by a xylose reductase (XR) converting D-xylose into xylitol at the cost of a NADPH. Xylitol then acts as substrate for the second reaction which converts it into D-xylulose by a xylitol dehydrogenase (XDH) while producing a NADH in the process. The final step of the xylose metabolic pathway is performed by a xylulose kinase enzyme and the D-xylulose is thus converted into D-xylulose-5-phosphate at the cost of an ATP [9]. D-xylulose-5-phosphate can then be further converted by a phosphoketolase into D-glyceraldehyde-3-P (GAP) for use in the pentose phosphate pathway and the glycolysis [10]. The xylose pathway described here is displayed in Figure 2.1.

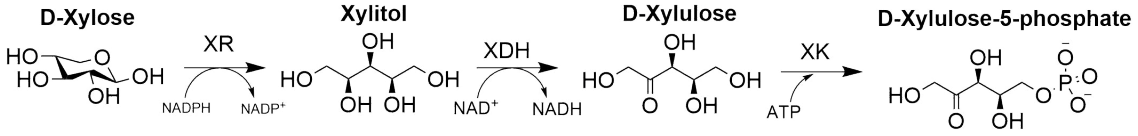


Figure 2.1: The xylose pathway for conversion of D-xylose into D-xylulose-5-P. Conversion is performed by the enzymes xylose reductase (XR), xylitol dehydrogenase (XDH) & xylulosekinase (XK) [21]. The co-factor usage in the first and second enzymatic conversion-steps differ as the first preferentially uses NADPH/$NADP^+$ and the second uses NADH/NAD^+. This causes an imbalance and an accumulation of xylitol [8].

It is worth noting that the first two steps of the xylose pathway have a different co-factor usage. The enzyme in the first step has a preference for using NADPH while the second step produces a NADH. This means an imbalance is formed that leads to the cell having to divert resources into producing NADPH and consuming NADH in order to maintain the flow through the xylose pathway [8]. There are reports of some yeast, e.g. *Pichia stipitis*, having dual co-factor usage on their xylose reductases potentially allowing them to avoid the co-factor imbalance problem by using both NADH and NADPH. But the affinity for NADPH still seems to be higher

meaning that even in these yeast there is still a problem with xylitol accumulation [21].

2.2 Lactose & galactose metabolism

Organisms that can metabolize lactose do so by first enzymatically degrading lactose using e.g. a beta-galactosidase. The lactose is converted into one glucose and one galactose [22]. The produced glucose is used in the glycolysis and the galactose is shuttled into a galactose pathway.

The pathway for converting galactose into glucose, and thus facilitating growth on galactose, commonly found in yeast is the Leloir pathway. In the Leloir pathway the conversion of galactose into glucose happens via three enzymatic steps. Firstly, galactokinase phosphorylates galactose to galactose-1-P at the cost of an ATP, following that the galactose-1-P is converted into glucose-1-P by galactose-1-P uridylyltransferase and the co-factor UDP-glucose. In the process of converting galactose-1-P the UDP-glucose is converted into UDP-galactose. The regeneration of UDP-glucose is facilitated by the third and final enzyme UDP-galactose-4-epimerase. Although not directly part of the patway there is also a fourth enzyme, galactose mutarotase, of importance that supports the conversion of β-D-galactose into the α-anomer which acts as substrate for the Leloir pathway. As it takes 1 ATP to convert galactose to glucose-1-P both galactose and glucose are equal energy sources with respect to their ATP yield [23].

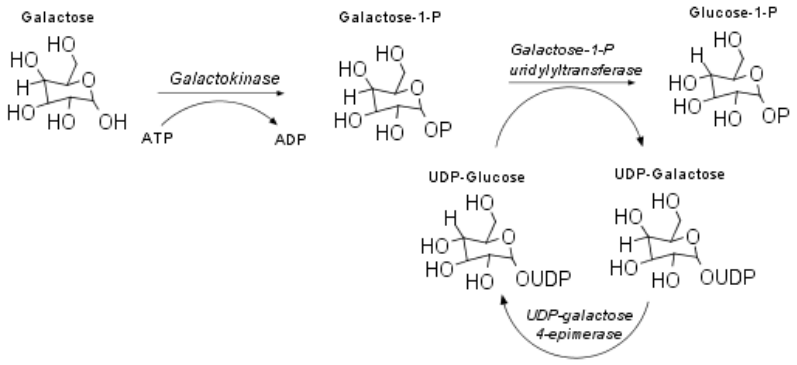


Figure 2.2: The Leloir pathway for conversion of galactose into glucose [23]. Galactose is converted into glucose-1-P by two enzymes, namely galactokinase followed by galactose-1-P uridylyltransferase. Galactokinase also uses an ATP while galactose-1-P uridylyltransferase consumes an UDP-glucose. The UDP-glucose is restored by the enzyme UDP-galactose 4-epimerase.

For a long time it was believed that the Leloir pathway was the only enzymatic system that had evolved for the specific purpose of metabolizing galactose [23]. Although it is still true for many species of microorganisms that the Leloir pathway is the only pathway for metabolizing galactose there have been a few exceptions

found in bacteria [24, 25] and filamentous fungi [26, 27]. Such alternative pathways have however, to our knowledge, never been previously identified in yeast.

2.3 Second generation bioethanol & *Saccharomyces cerevisiae*

Among the potential options for finding a viable alternative to fossil fuels, the production of bioethanol is one of the most heavily studied. The first generation of bioethanol was designed for fermentation of sugars from crops such as sugar cane. There are however several problems with this approach. It doesn't utilize the whole plant biomass and it also competes with food production for land use. These points combined cause first generation bioethanol to be less sustainable environmentally, economically and socially. Second generation bioethanol attempts to respond to some of these problems existing in the first generation. This is to be done by using lignocellulosic material extracted from plant biomass. This would allow for a wider selection of plants to be used in the process as well as allowing for a larger part of the biomass to be utilized [3]. However, as second generation bioethanol is produced from lignocellulosic materials it is required to have a strain that can efficiently ferment both glucose and xylose which together make up the majority of the sugar content of many lignocellulosic materials.

When it comes to the production of bioethanol the most used organism is *Saccharomyces cerevisiae*, especially in the case of first-generation bioethanol. *S. cerevisiae* has also been used for thousands of years in the production of both food and beverages, which in turn has led to it being the most heavily studied yeast species. However, the WT of *S. cerevisiae* cannot utilize xylose for growth or fermentation, but it can naturally take up xylose through the use of non-specific hexose transporters (HXT, Km approx. 100 mM) [28]. Thanks to the existing transporters it is relatively easy to engineer *S. cerevisiae* to utilize xylose and there are many examples of engineered strains of *S. cerevisiae* that are able to both grow and ferment on xylose. An example of how this was achieved is for example through the cloning of xylose dehydrogenase and xylose reductase into *S. cerevisiae* from the xylose fermenting yeast *Pichia stipitis* [29]. But throughout the engineered strains of *S. cerevisiae* there are reported cases of problems such as low ethanol yields, poor xylose consumption rates and by-product formations [6, 7]. A study from 2007 compared multiple strains of engineered *S. cerevisiae* which displayed xylose consumption rates varying from 0.05 to 0.67 (g xylose/g biomass h) [7]. In perspective, glucose consumption rates for *S. cerevisiae* have long been known to be over 2 (g glucose/g biomass h) [30].

Attempts have been made to improve *S. cerevisae's* ability to grow on xylose through cloning of glucose/xylose transporters Gxf1p & Gxs1p from *C. intermedia* into *S. cerevisiae*. These attempts have showed little or no success in improving the growth [31]. This might indicate that the transporter is not as effective when expressed

in *S. cerevisiae* as when expressed in *C. intermedia*, or that there are other factors limiting the xylose consumption rate such as the xylose reductases in the xylose pathway.

2.4 *Candida intermedia* & xylose

With the rise of second-generation bioethanol the research community and the industry have started looking at other non-conventional organisms to either fill the role of *S. cerevisiae* or to learn from in order to improve the engineered *S. cerevisiae* strains. There have been extensive studies comparing the capacity of different organisms to grow and ferment on xylose [16, 32, 33]. But despite the large amount of study in the field there is still a lack of solutions to problems associated with xylose fermentation, such as by-product formation & low ethanol yields.

Many yeasts of the *Candida* genus exhibit strong growth on xylose [11], and it is thus no suprise that *C. intermedia* also exhibit these traits and is ample at both transporting and metabolizing xylose. *C. intermedia* transports xylose via two separate transport systems. The first of these being high-capacity and low-affinity facilitated diffusion (Gxf1p, Km approx. 50 mM for xylose), and the other one being a high-affinity glucose/xylose-proton symporter (Gxs1p, Km approx. 0.2 mM for xylose & approx. 0.1-0.5 mM for glucose). *GXS1* expression is increased 8-fold in presence of low xylose and is absent in high-glucose [31].

The C11 strain of *C. intermedia* used in this project has three putative xylose (aldose) reductase genes, namely *XYL1*, *XYL1_2* & *XYL1_3*. These were initially identified due to their similarity to the xylose reductase found in *Scheffersomyces stipitis*. To our knowledge, there are no other species of yeast with three so closely related aldose reductases [13]. This could mean that the three putative aldose reductases have different roles in metabolism, but it could also hint at one of them having potential for favourable traits for xylose utilization.

2.5 *Candida intermedia* & galactose

C. intermedia can naturally utilize galactose as its carbon source [13]. This is not unlike most yeast species, and the galactose metabolism that can be found throughout most galactose utilizing yeasts is performed through the Leloir pathway [34]. Across the yeasts that possess the Leloir pathway a conserved gene cluster can be found, and this conserved gene cluster can also be observed in *C. intermedia* indicating that it too utilizes the Leloir pathway [13].

Except for the conserved Leloir gene cluster *C. intermedia* also contains a second gene cluster that seems to be connected to galactose metabolism as it contains for

example a *GAL10* gene [13], which is also found in the conserved gene cluster [34]. This second gene cluster also contains the aldose reductase *XYL1_2* [13], a fact made interesting as the Leloir pathway for galactose metabolism does not contain any aldose reductases [23]. This secondary galactose gene cluster does not seem to appear in other yeasts [16].

Apart from the Leloir pathway there is also an alternative galactose pathway that was recently discovered in some fungal species [35]. There are several connections that can be drawn between this fungi galactose pathway and the secondary galactose gene cluster found in *C. intermedia* such as the fact that the fungi pathway does contain an aldose reductase [35, 13]. This raises the question whether *C. intermedia* possess two separate pathways for galactose metabolism.

2.6 Homologous recombination

Homologous recombination (HR) is a mechanism for DNA repair. It functions on various problems that can arise in DNA, such as double stranded breaks and inter-strand crosslinking [36].

During HR a loose end of a strand, for example formed due to a double strand break, invades another DNA molecule and binds to a complementary DNA strand. DNA synthesis takes place elongating the bound DNA strand. From this point two different paths of action can take place, either synthesis-dependent strand annealing (SDSA) or double-strand break repair (DSBR) [36]. For a visualization of the described HR mechanism or of the possible resulting sequences see Figure 2.3.

During SDSA the invading strand dissociates from the other strand and returns to bind with the initial complementary strand. A DNA polymerase can then resynthesize the missing bases successfully repairing the double-strand break. During DSBR while the invading strand binds to the invaded DNA molecule the complementary strand that is pushed away from the invaded molecule also binds to the complementary strand of the invading strand. DNA polymerases can that way elongate both strands of the double-stranded break. The two locations where the strands cross over to the other molecule form two crossings. The DSBR is resolved by the crossings being cut apart, but depending on the angle of the cut the content of the resulting sequences will differ [36].

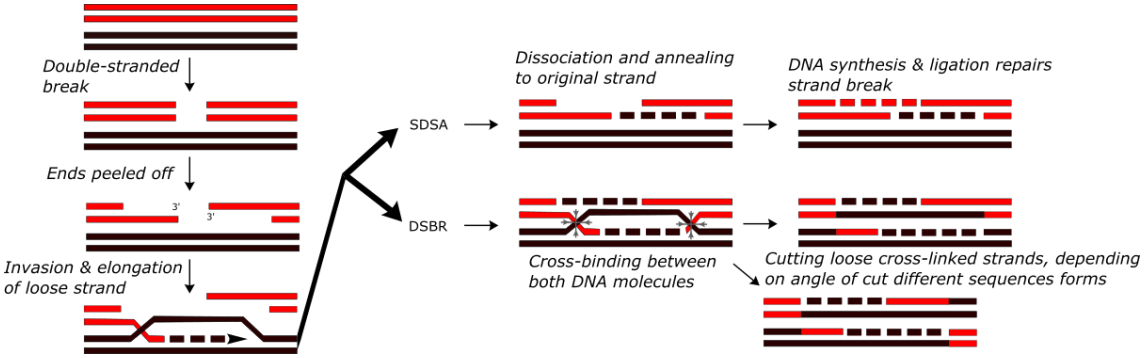


Figure 2.3: The mechanism of DNA repair through homologous recombination. After a double stranded break one of the strands invade a homologous region on another DNA molecule and is elongated to facilitate repair. The repair pathway then splits in two scenarios, synthesis-dependent strand annealing (SDSA) & double-strand break repair (DSBR) which yield different resulting sequences [36].

An important trait of the HR DNA repair pathway is that it acts based on the homologous regions for the strand invasion. This allows it to be used for genome engineering to integrate and replace desired sequences of DNA. As such it can be a very powerful tool for targeted genome editing [37].

2.7 Non-homologous end-joining

A second major pathway for DNA repair that can be employed by cells is non-homologous end joining (NHEJ), Figure 2.4. Although many organisms can employ both HR and NHEJ they tend to favor one over the other [38]. NHEJ works by acting on blunt ends. In the case of a double stranded break the NHEJ repair initiates by a dimer protein formed from the Ku70 and Ku80 proteins bind to the ends of the two fragment. The two dimers bound to the two fragments bring the two pieces together. Ligases, nucleases and polymerases bind to the Ku-dimer and thus elongate and ligate the two strands together. For the NHEJ repair pathway to funcion the *KU70* and *KU80* genes are essential as the Ku-dimer will not form without one of the two proteins [39].

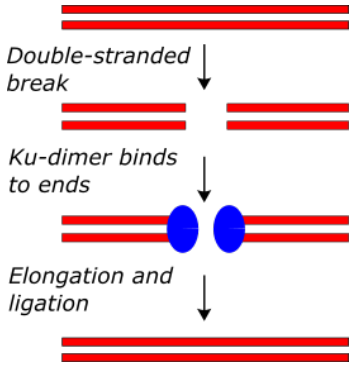


Figure 2.4: The mechanism of DNA repair through non-homologous end joining. After a double stranded break a Ku-dimer protein binds to the ends of the DNA fragments. The two fragments are brought together and repaired through elongation and ligation by DNA polymerases and ligases [39].

The fact that NHEJ acts on blunt ends leads to the fact that breaks can be repaired not only by the use of homologous DNA strands but by any DNA sequence with a blunt end. This in turn leads to DNA pieces inserted into cells during transformations can be randomly inserted at the position of any break in the DNA, meaning that performing gene deletions in organisms favoring NHEJ is much more difficult than in those favoring HR [38].

2.8 Improving gene target efficiency

With an efficient way of performing targeted genome editing in C11 it would be possible to perform gene deletions of various genes of interest, such as the aldose reductases. Thus assisting in the characterization of these genes. In order to improve gene target efficiency the ratio of homologous combination has to be improved.

Increasing the length of flanking homologous regions is one way to promote integration through the use of HR. This positive effect on gene target efficiency has been reported across several organisms [38, 40].

Another effective way of improving the cells capacity of using HR is by inhibiting the NHEJ pathway. As the *KU70* and *KU80* genes are essential for NHEJ one of the more common ways of inhibiting NHEJ is to delete one of those genes [41, 42, 43]. Despite the effectiveness of this method it comes with a few drawbacks. Several negative side effects have been noted in some strains deficient in either *KU70* or *KU80* such as changed growth rates, worse transformation rates, phenotypic instability or an increased sensitivity to UV light [20]. It can also be very difficult to obtain a *KU70* deletion in unconventional organisms favoring NHEJ that are lacking in satisfactory tools for gene deletions.

During different phases of the cell cycle yeasts have a different ratio between performed NHEJ & performed HR. The S phase, see Figure 2.5, is of particular interest since during it much more HR takes place compared to NHEJ. By using a chemical, in this case hydroxyurea, the cells progress through the cell cycle can be halted reversibly for a short period of time [44]. The mechanism by which hydroxyurea arrests the cell cycle is as a ribonucleotide reductase inhibitor [45]. Inhibition of the ribonucleotide reductases decreases the concentration of dNTPs available for synthesis of DNA which triggers a signaling cascade halting the cells progress forcing it to not pass between the S and G2 cell cycle phase [46]. As the effect is reversible it can be utilized for increasing the gene targeting efficiency during for example a transformation protocol [44].

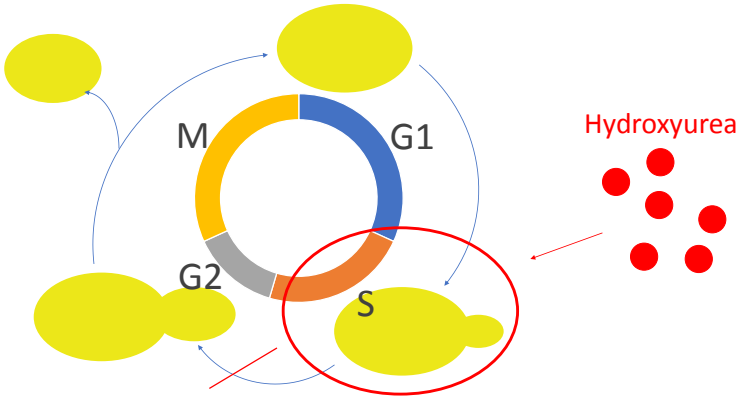


Figure 2.5: Illustration of the cell cycle of budding yeasts such as *C. intermedia* & *S. cerevisiae*. The S-phase is of especial interest as there is an increased ratio of HR taking place during it. When added to a culture, hydroxyurea arrests the cell cycle in the S-phase [44].

3

Methods

3.1 Strains & media

3.1.1 Strains

The parental WT used across all experiments with *C. intermedia* was the in-house isolated C11 strain. Several deletion strains were produced from C11 including deletions of *ADE2*, *XYL1_2* as well as the novel galactose gene cluster where *XYL1_2* is situated.

Escheria coli strain *DH5-α* was used for amplification of plasmids.

3.1.2 Media

C. intermedia strains were routinely grown and maintained using YPD medium containing 1 % yeast extract, 2 % peptone and 2 % glucose.

For screening of *ADE2* deletions, cells were grown on YNB medium containing yeast nitrogen base without amino acids. The YNB media was supplemented with 0.015 % of adenine and 0.5 % of $(NH_4)_2SO_4$.

E. coli was grown on either LB or SOC medium. LB was made to contain 1 % of tryptone, 1 % of NaCl and 0.5 % of yeast extract. SOC was made to contain 2 % of tryptone, 0.5 % of yeast extract, 10 mM of NaCl, 2.5 mM of KCl, 10 mM of $MgCl_2$ and 20 mM of glucose.

For solid versions of any of the above media 2 % of agar was also added. Following transformations appropriate antibiotics were also added to the medium for selection.

3.2 qPCR

A qPCR study was performed in order to examine the gene expression levels of C11's xylose reductases in different conditions. The resulting information could be used to identify high priority targets for gene deletion later in the project.

All qPCR runs were performed on an Agilent technologies Stratagene Mx3005P qPCR system. In each experiment the DNA was stained using a Agilent technologies SYBR Green Brilliant III Ultra-Fast QPCR Master Mix.

3.2.1 Design of qPCR experiments

Three genes of interest were chosen to be examined, specifically three aldose reductase genes from C11; *XYL1*, *XYL1_2* & *XYL1_3*. For quantifying the gene expression the relative quantification method outlined by Livak et. al [47] was utilized. The conditions chosen for the genes to be studied in were 20 % xylose, 20 % glucose and 2 % lactose.

In order to compare gene expression across multiple conditions the expression of the gene of interest has to be normalized against the expression of a reference gene which has a stable expression level across multiple conditions. Reference genes are usually chosen from genes that are part of fundamental cell functions, for example genes that are required for the production of organelles or structural proteins (such as actin). As *C. intermedia* does not have any previously established reference genes for relative quantification two reference genes were selected from the literature that have been used for the similar organism *Pichia stipitis* [48]. The reference genes were thus selected to be *ACT1* & *RPL13*.

Primers for each of the genes of interest as well as the reference genes were designed according to Table 3.1.

Table 3.1: Primers used in the qPCR analysis.

Gene	Direction	Sequence
ACT1	Fwd	CGTCCCAGCATTCTACGTCA
ACT1	Rev	GGTTTCTCTTTGCCACACGG
RPL13	Fwd	CTAAGGTTGCCCCAAAGCCA
RPL13	Rev	ACTTAGCAGACAAGCCAGCA
XYL1	Fwd	ACGAGGATGTCCCATTGCTTAG
XYL1	Rev	CTTGACCAATCTTGGCTGCTG
XYL1_2	Fwd	CCAACGAAAAGGAGGTCGGT
XYL1_2	Rev	GCAGCTTCGACATTGTCTGG
XYL1_3	Fwd	GGTGACGGTGACAAATTCCAG
XYL1_3	Rev	CGGCAGGAGGAATAGTAGCC

3.2.2 Validation of reference genes

A qPCR was run on the reference genes in five conditions (20 % glucose, 20 % xylose, 3 % glucose, 2 % xylose + 1 % glucose & 2 % lactose). Each condition had samples prepared for triple biologic replicates and double technical replicates. Averages of technical replicates were calculated. The lowest Ct value obtained was noted and all the measured Ct values were normalized according to Equation 3.1 where X is the value to be normalized and *min* is the lowest obtained Ct in the data set.

$$2^{X-min} \tag{3.1}$$

After normalization values were run through the GeNORM algorithm. GeNORM is used to automatically calculate the gene-stability measure M and ranks the reference genes based on it [48].

3.2.3 qPCR primer efficiency

Each primer pair was run in the qPCR using a dilution series with six separate concentrations of cDNA. Each concentration was run with two technical replicates. An average Ct was calculated across the technical replicates. For each primer pair a graph was plotted where the logarithm of the dilutions was used on the x-axis and the average Ct was used on the y-axis. The resulting slope of the graph was used in Equation 3.2.

$$10^{(\frac{-1}{\text{slope}})} - 1 = \text{qPCR primer efficiency} \tag{3.2}$$

The obtained value from Equation 3.2 gave the qPCR primer efficiency.

3.2.4 qPCR lab execution

Samples were loaded onto 96-well plates. For each condition three biological replicates were used and two technical replicates was used for each sample. The entire sample set up was then repeated once more but without the addition of cDNA, this was to be used as negative controls. This experimental design was repeated for each primer pair that wanted to be examined.

To each 96-well qPCR plate two samples were also loaded that were meant to be used as "interplate samples" which are identical across several plates to be able to be used as points of reference for normalization between plates.

3.2.5 qPCR data analysis

An interplate normalization number was calculated by using equation 3.3 for both interplate samples and then taking the average of both answers. Each measured Ct on the plate was then multiplied with the interplate normalization number.

$$\frac{\text{(Measured Ct on reference plate)}}{\text{(Measured Ct on plate of interest)}} = \text{Interplate normalization number} \quad (3.3)$$

An average Ct was taken between all replicates of the same condition. The average Ct for each condition was then used to calculate a $\Delta\Delta$Ct for each condition according to the method described by livak et. al [47]. 20 % glucose was used as the reference condition.

3.3 Optimization of cell cycle synchronization using hydroxyurea.

Based on experiments previously performed in the project a range of hydroxyurea concentrations as well as incubation times were identified as interesting for achieving a consistent cell cycle synchronization. Experimental design was prepared according to Table 3.2.

Table 3.2: Experimental design of the cell cycle synchronization optimization.

Culture	Hydroxyurea concentration	Replicates
#1	0 mM	2
#2	100 mM	2
#3	150 mM	2
#4	200 mM	2

Liquid cultures of C11 were prepared by inoculating 20 ml of YPD with fresh C11 from YPD plates. Cultures were left overnight in 30°C with 200 rpm of shaking. The obtained cultures were diluted to OD_{600} = 0.2 and left to grow for 3h in 30°C with 200 rpm of shaking. Cells were harvested and resuspended in 16 ml of YPD. To each culture 4 ml of sterile MQ-water and/or 1 M hydroxyurea in MQ-water were added to obtain desired concentration of hydroxyurea. Cultures were incubated at 30°C with 200 rpm of shaking. Samples were collected at intervals from the cultures, each sample had its OD_{600} measured and images of the cells were taken using a microscope and a smart phone camera.

3.4 Plasmid extraction

The method for extraction of plasmids was based on a study by Pronobis et. al. where addition of ethanol showed to greatly increase the yields from standard Miniprep kits [49]. A GeneJET Plasmid Miniprep Kit from Thermo Scientific was used, the wash-, resuspension-, neutralization- & lysis solution as well as the elution buffers used were all obtained from the Miniprep kit.

A culture was prepared by inoculating 4 ml of LB media using a frozen glycerol stock of *E. coli* with the desired plasmid. Culture was left to grow overnight in 37°C with 200 rpm of shaking. The following day the culture was divided into 2ml aliquots and cells were harvested and resuspended in 250 µl resuspension solution. 250 µl of lysis solution was added and the sample was mixed via inversion. 350 µl of neutralization solution was added and the sample was again mixed via inversion. The sample was centrifuged for 5 minutes at 12000 rpm. Supernatant was transfered to a microcentrifuge tube and equal volumes of 99 % ethanol was added to it. Entire sample was loaded onto a ThermoFisher Scientific GeneJET Spin Column and centrifuged for 1 minute at 12000 rpm. Column was washed twice with 500 µl wash solution. After washing the column was moved to a microcentrifuge tube and 20-50 µl elution buffer was added to the membrane before being centrifuged down for 1 minute at 12000 rpm.

3.5 Preparation of short ssDNA fragments

Salmon sperm ssDNA was mixed with AluI (1 U/µl reaction mix) restriction enzyme and Tango buffer (10X buffer diluted 10 times). The reaction mix was incubated at 37°C for 2,25 h. The complete mixture was run on gel electrophoresis with GelGreen DNA staining. Bands shorter than 1000 bp in length were cut out from gel and purified using GeneJET Gel Extraction Kit from ThermoFisher Scientific.

3.6 Preparation of linear deletion cassettes

Plasmids containing the homologous regions and CLONAT resistance were linearized by the use of restriction enzymes. Plasmids were only cut using a single restriction enzyme cutting at a single restriction site giving rise to long linear fragments. Homologous regions were designed to each be ∼1000 bp long.

3.7 Preparation of split marker deletion cassette for ADE2

The plasmid for deletion of the gene of interest in C11 was used as a template for two separate PCR reactions. Each product contained roughly half of a CloNAT resistance gene and either of the two flanking homologous regions. The fragments were verified to be correct using gel electrophoresis. Homologous regions were designed to each be ∼1000 bp long. Visualization of split marker design can be seen in Figure 3.1.

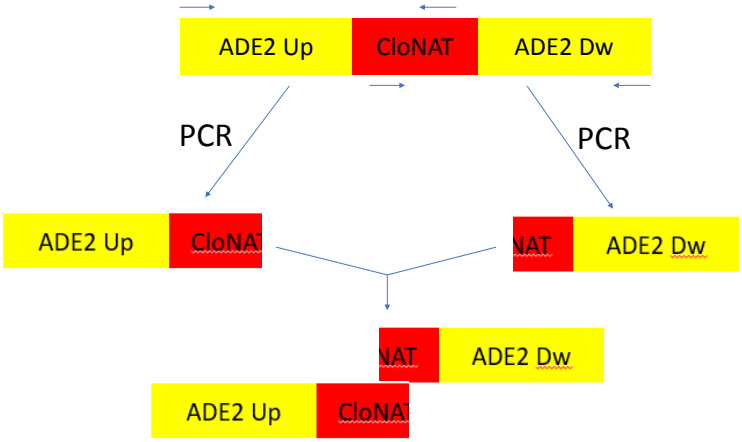


Figure 3.1: Visualization of split marker cassette design. The deletion cassette is split into two separate fragments using PCR. The two fragments have an overlapping region meaning that they can align and be repaired by HR in the cells.

3.8 Preparation of split marker deletion cassettes using fusion PCR

The three fragments of interest, upstream homology, downstream homology & CLONAT resistance, were separately amplified using PCR. Between the fragments that were to be fused an artificial overhang was added during the PCR that has been shown to be effective for the use in assembly of constructs using fusion PCR reactions [50]. The artificial overhangs used were "GGGGGCCCCCGGGGG" and "GGGCCCGGGCCCGGG" as well as their reverse complements. The fusion of constructs were designed to make sure that each of the two artificial overhangs were only used upstream or downstream of the CLONAT resistance, thus enabling a system that was easily modifiable.

The fusion of fragments was done using a fusion PCR reaction in which one of the homology fragments were run with the CLONAT resistance. Primers were chosen as to enable amplification of the entire homology region but only roughly half of the CLONAT resistance, thus giving rise to a split marker deletion cassette.

3.9 Transformation of *C. intermedia* C11 using electroporation

Two separate cultures were prepared by inoculating 100 ml of YPD with fresh C11 cells from YPD plates. Cultures were left to grow overnight at 30°C with 200 rpm of shaking. After overnight incubation cultures were diluted using YPD to OD_{600}=0.2 with a volume of 100 ml. Cells were left to grow for 3-4 h to OD_{600}=1. Cells were harvested and washed once with ice cold sterile MQ-water and washed once with ice cold 1 M sorbitol. After washing with sorbitol cells were resuspended in 40 µl of 1 M sorbitol per desired transformation. The resuspended cells were divided into 40 µl aliquotes and transfered to microcentrifuge tubes. To each aliquot 8 µl of insert-DNA (with an optimal amount of 2 µg) was added. Each aliquot also had 2 µl of short ssDNA fragments added to it. The transformation mix was transfered to a chilled electroporation cuvette and was left to incubate on ice for 15-30 min. Cuvettes were wiped clean and electroporated with a BioRad MicroPulser at 3 kV for ∼5 ms. Cells were gently transfered to 4 ml of 0.5 M sorbitol in YPD and were then left to incubate overnight at 30°C with 200 rpm of shaking. Overnight cultures were spread on YPD plates with appropriate antibiotics for selection and left to grow in 30°C for 48 h.

3.10 Transformation of *C. intermedia* C11 using cell cycle synchronization & electroporation

Two seperate cultures were prepared by inoculating 100 ml of YPD with fresh C11 cells from YPD plates. Cultures were left to grow overnight at 30°C with 200 rpm of shaking. After overnight incubation cultures were diluted using YPD to OD_{600}=0.2 with a volume of 80 ml. Cells were left to grow for 2 h at 30°C with 200 rpm of shaking after which 20 ml of 1 M hydroxyurea was added to the culture with the highest OD_{600} while 20 ml of sterile MQ-water was added to the other culture as a control. Cells were incubated at 30°C for 2 h with 200 rpm of shaking, cells were then harvested after which the protocol followed the steps from the washing of cells and onward in 3.9.

3.11 Deletion of ADE2 in C11

Transformation protocol was used according to 3.10. After the washing steps the cells were resuspended in 240 µl of 1 M sorbitol to have enough cell solution for 5 separate transformation aliqouts from each cell culture. Each transformation aliquot contained 40 µl of cells and then 10 µl of DNA was added. DNA was added to the transformation aliquots according to the experimental design presented in Table 3.3. Selection was done on YPD plates with CloNAT (50 µg/ml) which were left to grow for 48 h in 30°C.

Table 3.3: Experimental design of *ADE2* deletion in C11.

Transformation	Hydroxyurea	Added DNA	Replicates
#1 (negative control)	No	None	1
#2 (negative control)	Yes	None	1
#3	No	Linear *ADE2* deletion cassette	2
#4	No	Split *ADE2* deletion cassette	2
#5	Yes	Linear *ADE2* deletion cassette	2
#6	Yes	Split *ADE2* deletion cassette	2

The colonies that appeared on the YPD + CloNAT plates were restreaked onto YNB + adenine (0.015 %) plates. This was in order to screen for colonies with correct integration of the deletion cassette for deletion of the *ADE2* gene. Successful deletions turns the colonies red, while incorrect inserts cause the colonies to remain white, Figure 3.2. The amount of red and white colonies was counted and a target efficiency ratio was calculated for each condition.

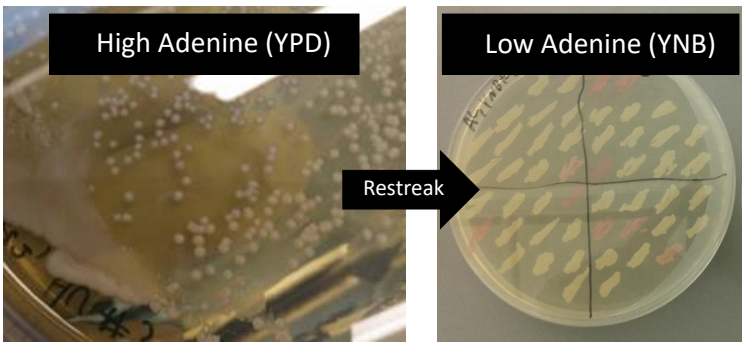


Figure 3.2: Cells deleted of *ADE2* restreaked from YPD to low adenine media. By picking and restreaking colonies from YPD onto a low-adenine solid media like YNB with only small amounts (∼0.015 %) of adenine added some of the white colonies turn red revealing a correctly targeted gene deletion of *ADE2*.

3.12 Deletion of galactose gene cluster and aldose reductase *XYL1_2* in C11

Transformation protocol was used according to 3.9. After washing the cells were resuspended in 300 µl of 1 M sorbitol to have enough cell solution for 6 seperate transformation aliqouts from each cell culture. Each transformation aliquot contained 40 µl of cells and then 10 µl of DNA was added. DNA was added to the transformation aliquots according to the experimental design presented in Table 3.4. Selection was done on YPD plates with CloNAT (50 µg/ml) which were left to grow for 48 h in 30°C.

Table 3.4: Experimental design of galactose gene cluster deletion in C11.

Transformation	**Added DNA**	**Replicates**
#1 (negative control)	None	1
#2 (positive control)	Linear *ADE2* deletion cassette	1
#3	Split *XYL1_2* deletion cassette	2
#4	Split Gal cluster deletion cassette	2

The colonies that appeared for the deletion on the YPD + CloNAT plates were resuspended in YPD and grown overnight on a sterile 96-well plate. Overnight cultures had their genomic DNA extracted according to a protocol outlined by Drumonde-Neves et al. [51]. The genome extract was then used as template in a PCR reaction to verify the correct deletion. Primers used in verification were designed as to have one primer binding to the genome upstream of the deletion and one primer binding inside of the deletion cassette.

3.13 Bioscreen growth trial of deletion strains

To examine the phenotype of obtained deletion strains a growth trial was performed using a Bioscreen C MBR. A WT strain of *C. intermedia* was also included as comparison. Three technical replicates was included for each used strain. Cells were grown for 66 hours in 30°C while optic measurements of cell density were made every 10 minutes.

Raw data obtained from the bioscreen was converted to OD_{600}. The logarithm of the OD_{600} values was used to compile growth curves.

4

Results

The results of the project have here been divided into four sections. Each section relates to a distinct part of the project meant to work towards fulfilling the aim of the project. The four sections are as follows: (1) Examining the physiological role of the three aldose reductases *XYL1*, *XYL1_2* & *XYL1_3* using qPCR measurements. (2) Developing a gene deletion protocol with high target efficiency. (3) Application of the developed gene deletion protocol to delete genes of interest. (4) Studying the phenotype of the generated deletion strains using a bioscreen.

4.1 qPCR

C. intermedia possess no less than three closely related aldose reductases, *XYL1*, *XYL1_2* & *XYL1_3*. A qPCR experiment was set up with the purpose of examining expression levels of the three genes during growth in three different carbon sources: xylose, lactose & glucose. Xylose and lactose were chosen as conditions due to the hypothesis that the aldose reductases would be important for the utilization of both. Glucose was chosen as comparison since it was thought that the aldose reductases would be repressed in that condition. Two reference genes, *ACT1* & *RPL13*, were also chosen to have their expression measured to use as points of reference for normalization of gene expression. As no established qPCR reference genes could be found in literature for *C. intermedia ACT1* and *RPL13* were chosen since they are have been used as reference genes in similar organisms.

4.1.1 qPCR primer efficiency

For successful quantification of gene expression, qPCR oligonuecleotide primers are needed that efficiently can amplify the cDNA from genes of interest. A primer efficiency experiment was run to confirm the efficiency of the primers. These experiments were executed by using a serial dilution of cDNA in a qPCR run with the primers of interest to verify that the decrease in amplified expression is proportional to the dilution of cDNA. This way it shows that the primers efficiently binds to and amplifies all of the present cDNA.

All of the primer pairs designed for the qPCR experiments showed primer efficiencies between 90 and 110 % which is the desired range, Table 4.1. The lowest obtained efficiency was 92 % for the primers used for aldose reductase XYL1_3 and the highest obtained efficiency was for the *RPL13* primers with 109 %.

Table 4.1: Efficiency of qPCR primer pairs for the two reference genes *ACT1* & *RPL13* as well as the three aldose reductase genes *XYL1*, *XYL1_2* & *XYL1_3*. An optimal efficiency would be 100 %, but values in the rough range between 90 % and 110 % are considered satisfactory.

Gene	Primer efficiency
ACT1	95 %
RPL13	109 %
XYL1	106 %
XYL1_2	94 %
XYL1_3	92 %

4.1.2 qPCR reference gene verification

In order to use the reference genes for normalization of gene expression, the stability of their gene expression has to be confirmed across several conditions. If their expression is consistent over many conditions it means that you can use their expression for comparison to infer the changes in expression of other genes across sets of conditions.

Expression data was collected for the two reference genes, *ACT1* & *RPL13*, at five different conditions: 20 % glucose, 20 % xylose, 3 % glucose, 2 % xylose + 1 % glucose & 2 % lactose. Data was analyzed using an Excel plugin based on the GENORM algorithm which provides a value M which indicates the stability of the gene [48]. In this analysis the lower the value of M the more stable the gene's expression is. According to the analysis M for both genes had a value of 0.782, which falls below the threshold of 1.5 which the plugin uses to indicate unstable genes unfit for use as reference genes.

4.1.3 qPCR data analysis

For the actual measurements of expression of the aldose reductases, *ACT1* proved to give somewhat more consistent and stable expression levels when it was used as reference genes. Only one reference gene can be used for normalization and as such *ACT1* was chosen due to the stated reasons. Negative controls in which no DNA was present gave expected results for measurements of *XYL1* & *XYL1_2*. However there were some unexpected measurements for the negative controls for *XYL1_3* with some readings of higher expression of *XYL1_3* in the negative control than in the actual samples.

Expression for the three aldose reductase genes *XYL1*, *XYL1_2* & *XYL1_3* was normalized against *ACT1* as the reference gene and the condition 20 % glucose. The expression is normalized against glucose as it was believed that the expression of the aldose reductases should be repressed in the presence of glucose. The readings are displayed in Figure 4.1. The data clearly showed that the aldose reductase gene *XYL1* had an increased expression while in the presence of xylose when compared to when the cells were grown on lactose or glucose. Aldose reductase gene *XYL1_2* is expressed strongly in the presence of lactose while the expression in the presence of either glucose or xylose is much lower. *XYL1_3* showed a lower expression in glucose while conditions with lactose or xylose showed a higher expression.

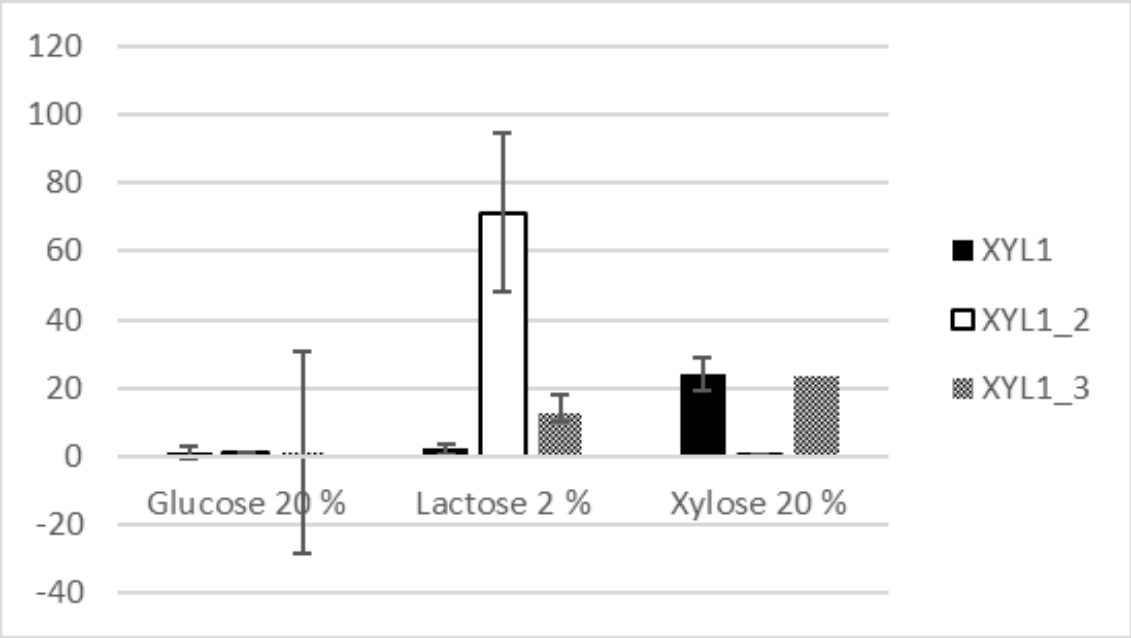


Figure 4.1: Relative gene expression of aldose reductase genes *XYL1*, *XYL1_2* & *XYL1_3* in the conditions 20 % glucose, 2 % lactose & 20 % xylose. Expression is normalized against *ACT1* expression and the 20 % glucose condition. Standard deviation is displayed for all conditions and genes except for 20 % xylose for *AR_3a* as only a single replicate was successfully measured.

The fact that negative controls for *XYL1_3* in some conditions gave higher readings of the expression levels than what was found in the actual samples indicates that one should be suspicious of the relative quantification of *XYL1_3* showed in the data and in Figure 4.1. A reasonable explanation of the negative control readings could be that *XYL1_3* is not significantly expressed in any of the conditions. If this is the case then any perceived patterns of expression of *XYL1_3* found between the conditions would simply be an effect of random deviation either due to technical circumstances or caused by impurities in the samples. This would also explain high error observed in 20 % glucose for *XYL1_3*. Another possible explanation is that the negative control samples were contaminated with cDNA or RNA thus giving off readings of gene expression.

XYL1_2 is clearly more expressed in the lactose 2 % condition than in the xylose and glucose conditions. This points towards *XYL1_2* being important for growth on lactose, potentially through the lactose/galactose pathway. This is further supported by the position of *XYL1_2* inside what is believed to be a gene cluster for growth on galactose.

According to the qPCR data the expression of *XYL1* is upregulated in the presence of xylose, but not glucose or lactose. This indicates that the role of aldose reducase Xyl1p is that of a xylose reductase.

4.2 Development of a gene deletion protocol with high target efficiency

To further study the role of the aldose reductases of *C. intermedia* C11 the aim was to delete their genes from the genome and follow up by observing the acquired phenotype. In order to perform these deletions a protocol which allows for high gene target efficiency must be developed as NHEJ is the heavily favored DNA repair pathway in *C. intermedia*.

4.2.1 Optimization of cell cycle synchronization using hydroxyurea

A study was performed to find optimal conditions for sychronization to the S-phase of the cell cycle in a culture of *C. intermedia* using hydroxyurea. The cell cycle synchronization was then to be used in cell transformations in order to increase homologous recombination frequency, and thus gene targeting efficiency, in the cells. Different concentrations (0-200 mM) of hydroxyurea was added to *C. intermedia* cultures after which OD_{600} was measured and microscopy pictures were taken at regular intervals over 3.5 hours.

The OD_{600} measurements from all the samples were collected and assembled into a growth curve for each individual culture. When inspecting the growth curves for the individual cultures one of the cultures with a hydroxyurea concentration of 150 mM were found to not follow the same pattern as the rest of the cultures, and it was thus considered to be an outlier. An average was calculated per hydroxyurea concentration and a new growth curve was assembled for each hydroxyurea concentration, Figure 4.2. The control culture, with no added hydroxyurea, exhibited the expected exponential growth. In the cases with hydroxyurea concentrations between 100 mM and 200 mM a very slow linear growth was displayed. The cultures grown with hydroxyurea concentrations of 150 mM or above also exhibited a pattern of oscillation in the OD. The microscopy images showed increasing homogeneity in

the distribution of cell cycle phases with an increase in hydroxyurea concentration, Figure 4.3.

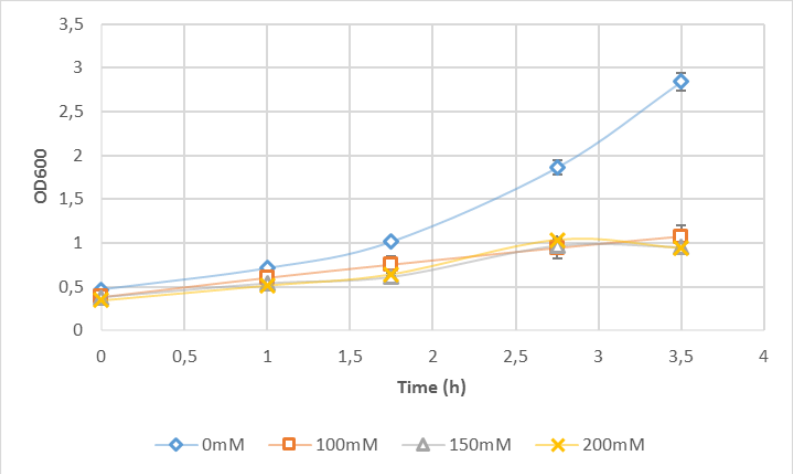


Figure 4.2: Growth curve of C11 while in the presence of varying concentrations of hydroxyurea. Without hydroxyurea the cells grew exponentially, which was expected. All cultures that had hydroxyurea added to them grew linearly at a slower rate while the cultures with 150 and 200 mM of hydroxyurea also oscillated in their OD_{600}.

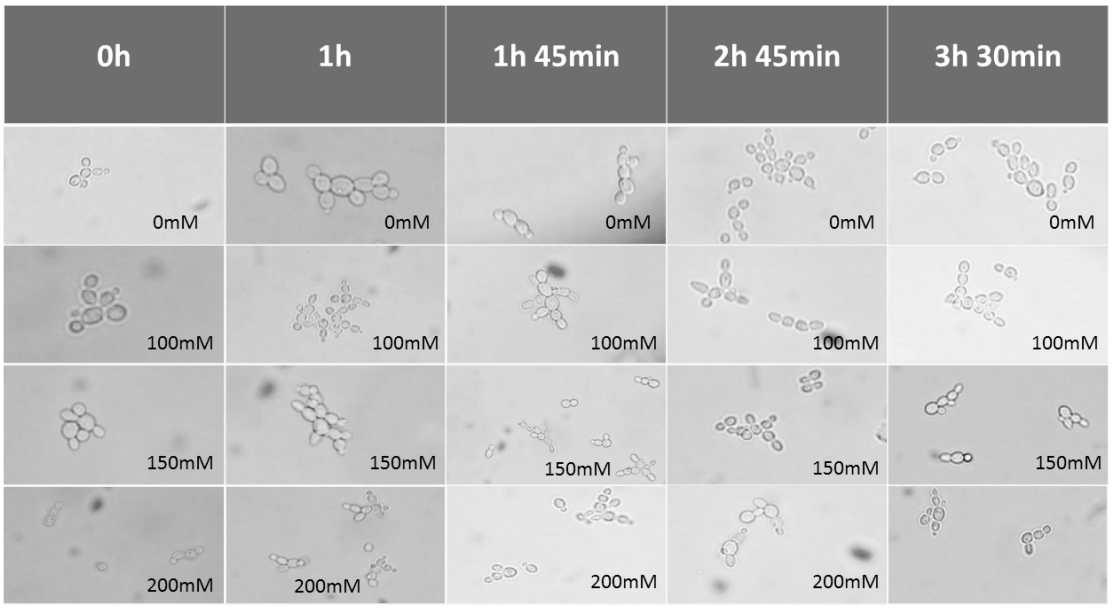


Figure 4.3: Micrographs of *C. intermedia* cultures incubated with different concentrations of hydroxyurea (100-200 mM) over time. The images indicate an increase of cells in the S-phase with increasing concentrations of hydroxyurea. A S-phase cell can be recognized by a large cell with an attached smaller daughter-cell that yet has not deattached itself. Images were taken through a microscope using a smart phone camera. For each image a small section was selected that was considered to be representative of the sample.

In the growth curves, Figure 4.2, it is clearly visible that all the cultures with added hydroxyurea grow much slower than the ones without it. This can be seen as a first indication that the cell cycle arrest is successful as the cells will slow down their growth as more and more gets synchronized in the S-phase of the cell cycle. It is interesting to note the oscillating shape of the growth for the 150 and 200 mM cultures. This pattern occurred in all cultures above 100 mM of Hydroxyurea and has also been observed in earlier cell cycle synchronization experiments in the project. It is however unclear what causes this and whether it is a sign of a successful cell cycle synchronization. Considering that OD measures the optical density of the culture and not actual cell concentration it is possible that the oscillating OD_{600} is an effect of for example the cells clumping together as an initial response to the hydroxyurea.

Despite the growth curves looking very similar for the different hydroxyurea concentrations a wider difference was observed when inspecting the samples in a microscope. Although it was clear that there were cells synchronized to the S-phase at all three examined hydroxyurea concentrations, especially after ∼2 hours of incubation, the ratios of synchronized cells seemed to differ somewhat. In the 100 mM concentration there were clearly visible cells of different cell cycle phases throughout the experiment indicating that this concentration might not be sufficient. The 150 & 200 mM cultures were quite similar with most cells appearing to be arrested in the S-phase. However, over the course of looking through multiple samples an overall sense was that the cells were slightly more synchronized in the 200 mM concentration, especially after long incubation times (3 h or more) where the effect sometimes starts disappearing. To investigate this further an extensive study with counting of individual cells would be required to find proper ratios of different cell cycle phases for the different concentrations. But for the uses in this project the 200 mM with an incubation time of 2 hours was deemed to be the most promising condition and work proceeded with this for the cell cycle synchronization transformations.

4.2.2 Deletion of *ADE2* in C11 using cell cycle synchronization

Using the found optimal conditions for cell cycle synchronization using hydroxyurea an experiment was planned to transform cell cycle synchronized cells. The hope was here that the synchronized cells would have an improved gene targeting efficiency which would be beneficial for a deletion protocol. The *ADE2* gene was chosen as deletion target for the development of the deletion protocol. When *ADE2* is knocked out cells grown while starved for adenine will display a red pigment due to the accumulation of the intermediary p-ribosyl amino imidazole in the adenine pathway [52], Figure 4.4. This makes the *ADE2* deletion ideal for quick and easy large scale screening.

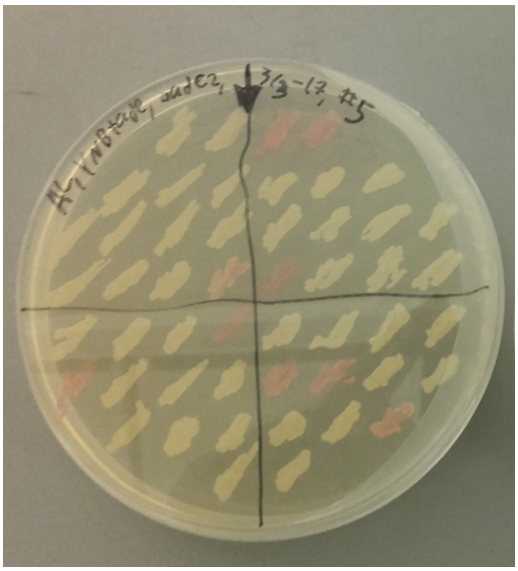


Figure 4.4: *C. intermedia* colonies streaked on a YNB plate with low adenine content. When grown on media with a low concentration of adenine colonies that have a deleted *ADE2* gene turn red.

Deletion cassettes were transformed with electroporation and both cultures that had and hadn't been cell cycle synchronized with hydroxyurea were used. Cells were transformed with 8 µl of DNA with a total weight of roughly 2 µg, 2 µl (roughly 200 ng) of short (<1000 bp) ssDNA fragments were also added. For the deletion cassette a linear cassette was used that included a selection marker flanked on both sides by regions homologous to the targeted region in the genome. Selection marker used was nourseothricin (CLONAT) resistance.

The negative control (no DNA added during transformation) did not produce any colonies when plated on plates with CLONAT indicating that the selection worked as intended. Linear cassette transformations provided several hundred or even above a thousand colonies on YPD+CLONAT plates, this occurred both with and without cell cycle synchronization. The colonies were restreaked on YNB + adenine (0.015 %), where adenine was limiting, and were counted based on their color (red or white). Red streaks were considered to be successfully deleted of the *ADE2* gene while white colonies were considered to have a random insertion and thus not a deletion of the *ADE2* gene. For both conditions (with or without cell cycle synchronization) more than 100 colonies were screened. The gene deletion target efficiency of *ADE2* for the different conditions was calculated and is presented in Table 4.2. In the culture that was not cell cycle synchronized at the time of transformation there was no successful deletions. Several colonies that had a correct gene deletion was however found for the cell cycle synchronized cells, and the achieved gene targeting effieciency was ∼3.7 %.

Table 4.2: Gene targeting efficiency for deletion of *ADE2* in C11 using cell cycle synchronization.

DNA insert	Cell cycle synchronized	Screened colonies	Correct deletions
Linear cassette	No	124	0 %
Linear cassette	Yes	164	3.7 %

Although 3.7 % is not a great gene targeting efficiency it is definitely an improvement upon the initial case of having no successful deletions. With 3.7 % it is possible to obtain deletions even with other genes except for *ADE2*, but it requires screening of a large amount of colonies which can be very labor intensive with deletion-phenotypes that aren't visually detectable. It would also be possible to use the cell cycle synchronization to delete *KU70* to enable easier future deletions.

4.2.3 Deletion of *ADE2* in C11 using a split marker cassette

It was desired to further improve upon the deletion protocol developed using cell cycle synchronization, either through improving gene targeting efficiency or by removing background colonies that did not have the successful gene deletion. It was hypothesized that at any time there is a subset of cells effectively performing HR, and if it was possible to specifically target only these during transformation a lot of background colonies could be eliminated.

A new approach was designed based on a concept known as split marker deletion cassettes. These deletion cassettes are split in two separate fragments in order to make sure that each fragment of the deletion cassette only contains half of the selection marker, CLONAT resistance. The fragments are designed with a homologous region overlap so that cells using HR can repair the cassette and utilize the resistance. Since it is required of cells to use HR to repair the cassette the idea is that the chances of them also using HR to successfully delete the gene of interest should be higher. On the other hand, cells not using HR cannot repair the cassette and will be eliminated when grown on selection medium, thus removing many background colonies.

Cells were transformed with electroporation and both cultures that had and hadn't been cell cycle synchronized with hydroxyurea were used. For deletion cassette the split marker cassette was obtained by two separate PCR reactions amplifying the two separate fragments from a plasmid containing the cassette. CLONAT was used as the selection marker.

The negative controls (no DNA added during transformation) did not produce any colonies when streaked on plates with CLONAT indicating that the selection worked as intended. When the split cassette was used instead there were only a few, 10-40, colonies found on each selection plate which is much fewer transformants detected

than for the case of using a linear casstte. Cell synchronization increased the amount of obtained colonies on CLONAT from the split cassette transformation, 10-70 per plate, but it was still not close to the linear cassette amounts. The colonies were restreaked on YNB + adenine (0.015 %), where adenine was limiting, and were counted based on their color (red or white). Red colonies were considered to be successfully deleted of the *ADE2* gene while white colonies were considered to have a random insertion and thus not a deletion of the *ADE2* gene. All obtained colonies were screened and the gene deletion targeting efficiency was calculated and is presented in Table 4.3.

Gene targeting efficiency was very similar for both the culture that had been cell cycle synchronized and the one that hadn't. The one that hadn't been cell cycle synchronized had an efficiency of ∼67 % and produced 43 colonies across two replicates. The cell cycle synchronized culture had a gene targeting efficiency of ∼64 % and produced 119 colonies across two replicates.

Table 4.3: Gene targeting efficiency for deletion of *ADE2* in C11 using a split marker deletion cassette. The results from when a linear cassette was used is also included for comparison.

DNA insert	**Cell cycle synchronized**	**Screened colonies**	**Correct deletions**
Linear cassette	No	124	0 %
Linear cassette	Yes	164	3.7 %
Split cassette	No	43	67 %
Split cassette	Yes	119	64 %

The gene targeting efficiency is substantially increased with the split marker cassette when compared with the linear cassette. Across all conditions the lowest target efficiency achieved was transformation with the linear cassette without cell cycle synchronization which didn't provide any successful transformants. Highest target efficiency was achieved with transformations with the split deletion cassette which provided efficiencies above 60 %. Cell cycle synchronized cells transformed with the linear cassette achieved a targeting efficiency of 3.7 %.

The transformations with a linear cassette only gave rise to successful deletions of *ADE2* when the cells were synchronized to the S-phase of the cell cycle, even then the targeting efficiency was very low. The split cassette approach provided a much higher success rate but at the cost of having fewer colonies appear per transformation. In the case of *ADE2* the lower number of obtained transformants is not a problem as even with only 20-50 transformants it was always possible to find a correct deletion due to the high target efficiency. Based on this the optimal approach was decided to be to use the split cassette without hydroxyurea cell cycle synchronization as this provides a satisfactory success rate and provides enough transformants, also by avoiding the use of hydroxyurea the transformation protocol is shorter and it avoids the use of a toxic chemical. If a deletion other than *ADE2* proves to be

more difficult either due to it having fewer transformants obtained or a lower gene deletion targeting efficiency it is worth considering to use the hydroxyurea cell cycle synchronization together with the split cassette. This can increase the number of obtained transformants to ensure that a correct deletion can be obtained.

4.3 Application of the gene deletion protocol

With the obtained protocol for deletion of genes using a split marker cassette it was possible to further study the aldose reductases Xyl1p, Xyl1_2p & Xyl1_3p in greater detail using gene deletions. Following a deletion hopefully a phenotype would become clear that can give greater insight into the role of the gene.

Due to time restrictions in the project it wasn't possible to delete all three aldose reductase genes. But based on the information seen so far aldose reductase gene *XYL1_2* with its interesting location in what appeared to be a galactose gene cluster seemed to be the most novel and interesting target. The deletion experiments were focused on deleting both the *XYL1_2* gene and the gene cluster in which it is situated.

4.3.1 Design & assembly of deletion cassettes

Deletion cassettes were designed for both the *XYL1_2* gene and the gene cluster where it's situated (thought to be a galactose gene cluster). In each construct a CLONAT resistance was included as a selection marker. Assembly of deletion plasmids was done using a NEB DNA assembly cloning kit, but no succesful assemblies were ever obtained this way. Had this approach worked the work would have continued in a similar manner as for the deletion of *ADE2* where the plasmid was used as a template for PCR reactions to construct the split marker deletion cassettes.

Instead of using the assembly kit a new approach was designed where a single homologous region was combined with the CLONAT resistance using fusion PCR. The primers used in fusion PCR was designed such as to only include half the CLONAT resistance thus resulting in the final fragment being one of the two fragments required for the split marker deletion cassette approach. This was done using both up- and downstream homology to obtain both split marker fragments. This approach of constructing the deletion cassette showed much greater success leading to constructs for both *XYL1_2* and the gene cluster being successfully assembled.

4.3.2 Deletion of aldose reductase gene *XYL1_2* & galactose gene cluster in C11

Cell transformation was done using electroporation and the assembled split marker deletion cassette as outlined in the *ADE2* deletion experiment. Deletion cassettes used in the transformation were the ones both for aldose reductase gene *XYL1_2a* and the galactose gene cluster, although both cassettes were added to separate transformation aliquots. Transformed cells were plated on YPD + CLONAT plates for selection. Colonies appeared indicating that cells were successfully transformed with the repaired CLONAT resistance, this was true for both the *XYL1_2* & the gene cluster deletion. About 30 colonies appeared in total for each transformation over two replicates. Negative control (no dna insert) and positive control (linear *ADE2* deletion cassette) both looked as expected further showing that the transformation was a success.

Colonies were picked from the selection plates and the genome was extracted using a quick genomic extraction protocol [51] allowing many colonies to be screened simultaneously. To verify the deletion a PCR was run with the genomic extract from the colonies as a template. The primers used were designed to have one binding to the integrated CLONAT resistance and the other to bind outside of the homologous regions, thus only providing an amplification of the correct fragment size (2111 bp for *XYL1_2* & 1798 bp for galactose cluster) if the deletion cassette was integrated at the correct location.

10 colonies were tested for verification of the deletion for both the *XYL1_2* and the galactose cluster deletions. When the amplified PCR products were run on gel the *XYL1_2* deletion displayed that all ten tested colonies had been successfully deleted, Figure 4.5. Sadly, for the galactose cluster deletion it was not possible to verify any deletions and due to time restrictions further attempts were cancelled.

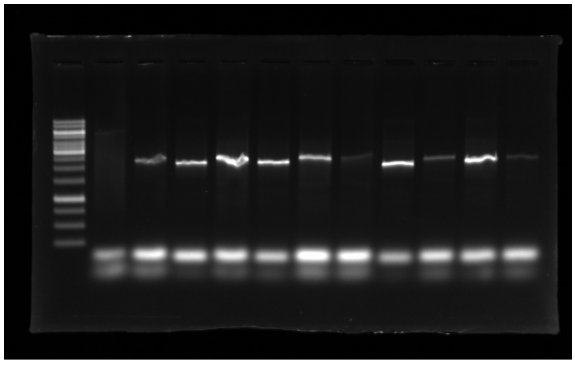


Figure 4.5: Gel electrophoresis for the PCR amplification meant to verify the *XYL1_2* deletion. First lane from the left is a DNA ladder where the three distinct bands starting from the bottom are of sizes 1000, 3000 & 6000 bp. Second lane is a negative control with WT genomic DNA. The following ten lanes are genomic extracts from the transformants, each displaying a band of the expected length close to the 2000 bp ladder band.

It was known after the deletion of *ADE2* that the split marker deletion approach to delete genes was effective, but having all ten tested colonies be correct was even better than expected. It is of course difficult to argue for any statistical significance when only 10 colonies were tested, but this does seem to hint that the gene targeting efficiency might be as good or even better for some genes than it was for *ADE2*.

The lack of verified deletions of the galactose cluster is unfortunate and it is not entirely clear as to whether it was only difficult to verify deletions or if there weren't any deletions. Another explanation could also be that the genomic extraction failed for the tested samples. If the deletions were indeed unsuccessful a possible reason as to why could be the fact that the galactose gene cluster is a much larger region of the genome when compared to single genes such as *XYL1_2* or *ADE2*, and the deletion cassette being shorter than the targeted region might simply not be able to successfully bind to both homologous regions simultaneously. If this was the case it would explain why a similar amount of colonies were obtained on selection plates as roughly an equal amount of cells should still be able to repair the cassette but they might integrate it based on only one of the homologous regions or integrate it into a random location in the genome.

4.4 Bioscreen growth trial of deletion strains

With the successful obtainment of a *XYL1_2* deleted strain work continued with examining a potential phenotype. For this a bioscreen was used as it allows for simultaneous screening of multiple strains. Three colonies that had been verified as having their *XYL2* gene deleted were selected to be examined in the bioscreen. For comparison a WT of *C. intermedia* C11 was also included in the bioscreen experiment. All 4 strains were grown in the bioscreen in two different conditions, namely in minimal medium with 2 % of glucose or minimal medium with 2 % of galactose.

Cells were grown in the bioscreen for 66 hours at 30°C. The collected data was converted to OD_{600} and compiled in graphs to compare the specific growth rate of the four strains, see Figure 4.6 & Figure 4.7.

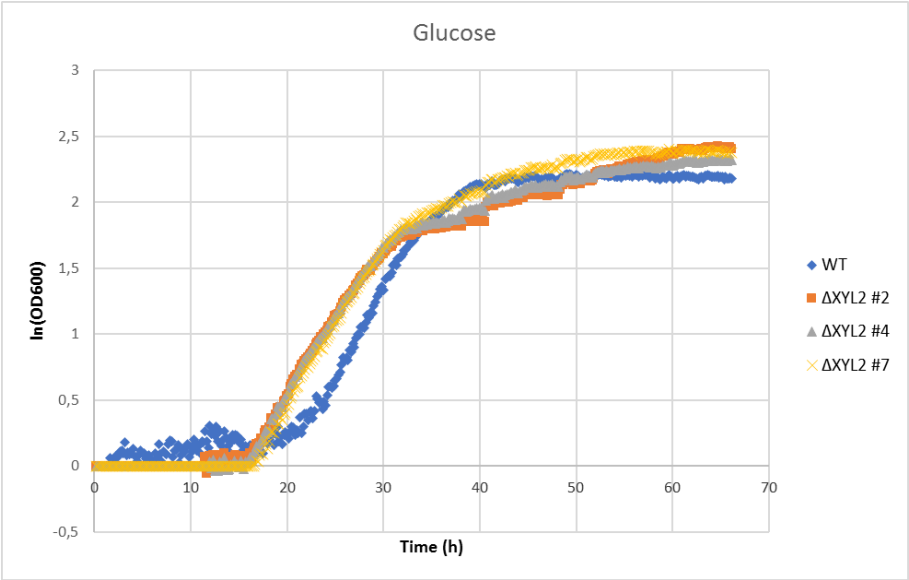


Figure 4.6: Growth curves of the WT of C11 and three strains where *XYL1_2* has been deleted while grown on glucose. All four strains exhibit very similar growth indicating that the *XYL1_2* deleted strains still grow well on glucose.

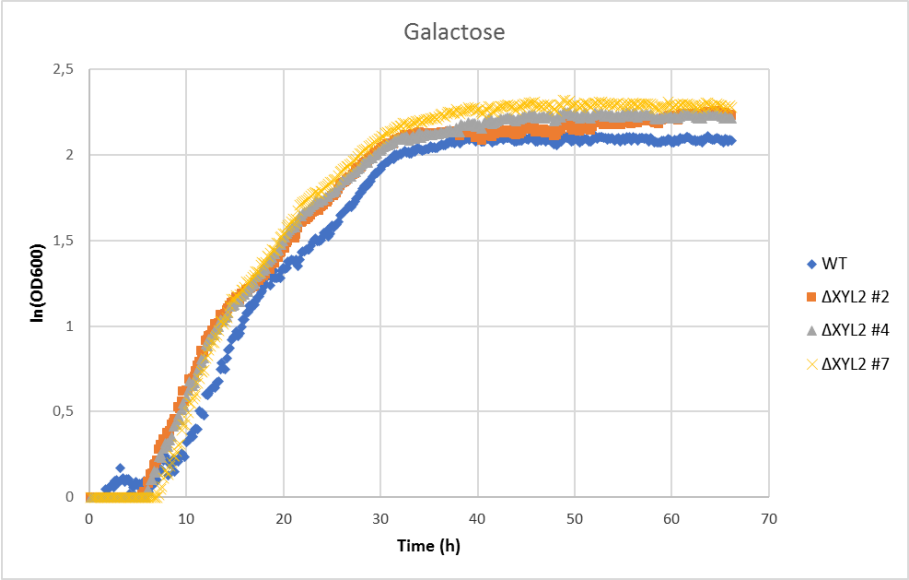


Figure 4.7: Growth curves of the WT of C11 and three strains where *XYL1_2* has been deleted while grown on galactose. All four strains exhibit very similar growth indicating that the deletion of *XYL1_2* does not give rise to a clear phenotype by itself in the tested condition.

All four strains examined gave rise to very similar growth curves on both glucose and galactose. All cultures undergo three distinct growth phases, the first one being a lag phase where cells are acclimatizing to the media which is then follow by the exponential growth phase during which most of the increase in cell density takes place. The third and final growth phase is the stationary phase where growth slows down to a halt. In the growth curves it is clear that the growth phases are somewhat different when the cells are grown on glucose compared to when grown on galactose, most notable is the shorter lag phase when grown on galactose.

The results from the glucose growth trial indicates that the deletion did not have a negative effect on normal growth on glucose. The results from the galactose growth trial on the other hand indicates that there were no visible effects on growth rate while growing on galactose. The results from the galactose growth trial thus shows that the *XYL1_2* gene is not essential for utilization of galactose. This could hint to the fact that the more traditional Leloir pathway for metabolism of galactose is the primary pathway for galactose utilization. Another possibility is that there are other enzymes that can fill the roll of the Xyl1_2p aldose reductase in galactose metabolism. To completely grasp the role of *XYL1_2* and the relevance of the two perceived galactose pathways more deletion strains have to be produced. It would be very interesting to repeat the galactose growth trial with strains deleted in either or both of the two putative galactose pathways.

5

Discussion

5.1 Split marker deletion cassette

The method of transforming cells using a split selection marker was first developed by Fairhead et al. in 1996 [53]. Since then there have been a few examples in literature of split markers being used to either allow for rapid one-step cloning or to increase gene targeting efficiency [54, 55]. Although the technique was initially developed for the yeast *Saccharomyces cerevisiae* the most prominent use seems to have been in filamentous fungi, especially when it comes to the application of increasing gene targeting efficiency. There are cases of the split marker technique having been used in yeast [56, 57], but overall there seems to be a lacking awareness of the technique among yeast researchers which is striking considering the amount of reserach that has been done on increasing homologous recombination in various yeast species [44, 58, 38]. In this thesis it has been shown that the split marker approach can be a powerful tool to increase gene targeting efficiency in yeast. This can hopefully assist in raising awareness for the split marker technique for future research as studies delve further into unconventional yeasts with industrial potential.

In the literature there are studies reporting an increase in gene targeting efficiency when using the split marker with efficiencies ranging from a few percent up to 63 % [54, 59]. This is somewhat surprising considering the highest target efficiency obtained during optimization in this project was ∼67 %. It might be the case that that number would have gone up or down somewhat if there was time to perform more replicates of the transformation. It is also possible that *C. intermedia* is an especially ripe target for this technique, for example by having a phase in the cell cycle with an abnormally high ratio of homologous recombination, which would mean that this method might not be as useful for other yeast species as one might hope.

One of the most influential techniques for genomic modification in recent years is CRISPR-Cas which allows for easy and quick construction of strains with single or multiple deletions [60, 61]. However, one of the lesser talked about aspects of CRISPR-Cas is the fact that it for certain organisms require a lot of preparation, especially in the case of unconventional organisms with fewer developed genetic

tools. One of the main problems one has to work around in order to make CRISPR-Cas function is that the cells need to have an effective homologous recombination repair system. If the main mode of DNA gap repair is NHEJ it can easily lead to the cuts introduced by Cas9 simply being repaired by random insertion or deletion of a few base pairs thus making gene deletions difficult. Today the most common fix to this problem is to use a strain where *KU70* or *KU80* has been knocked out. But a *KU70* or *KU80* deletion can be tricky to achieve in many species with extremely low ratios of HR compared to NHEJ, and it can give rise to unwanted side-effects such as poor specific growth rates and an increased sensitivity to mutations [20]. The split marker approach here seems to be a very attractive option to either be used to assist in the deletion of *KU70* or even to directly use alongside CRISPR-Cas to quickly proceed to genome editing with a minimal amount of preparatory work needed. Perhaps a more widespread awareness of the split marker approach in conjunction with CRISPR-Cas can lead to limitless potential of genome editing in even the most novel organisms.

5.2 Aldose reductases

For many yeast species aldose reductases have traditionally been linked to the utilization of xylose [62, 63, 64]. Aldose reductases have however also been shown to have activity towards a wide range of substrates in yeast [62, 63, 64], this can make it difficult to pinpoint a single clear physiological role of an aldose reductase as it might be important to many functions. It might thus not be possible to solely based on this project provide a definitive answer as to what the physiological role of the aldose reductases is in *C. intermedia* C11, several hints have however been collected that allows us to establish a rough outline. *XYL1_2* seems to be intrinsically connected to utilization of galactose. This was suspected early due to the gene's location in the genome which is inside of what seems to be a galactose gene cluster. The qPCR data further supports this idea as *XYL1_2* is clearly more expressed in the presence of lactose which traditionally is converted to galactose & glucose.

XYL1 was the only of the aldose reductase genes that had a clear increase in expression while in the presence of xylose. Although *XYL1_3* also had an increase in expression the data was questionable due to the problematic negative controls. Considering the fact that C11 can effectively grow on xylose it makes complete sense to attribute at least one of the three aldose reductases as a xylose reductase to be utilized in the xylose pathway. Although more data would be preferable it isn't too long a stretch to say that *XYL1* very likely fits that role as a xylose reductase due to its expression profile.

The more ambigious of the aldose reductases is *XYL1_3*. The problematic negative controls causes it to be difficult to draw any clear conclusions as to its role. The thing that however seems reasonable is that it is in fact not significantly expressed in any of conditions examined during the qPCR experiments. If *XYL1_3* is not expressed

in any of the conditions it could potentially indicate the role of the gene to be that of a backup for one of the other aldose reductases to increase robustness, for example in case the other enzyme being inhibited or put under extreme stress. Another explanation could be that *XYL1_3* is a gene duplication that lacks a defined role. A final perspective on the matter could be that *XYL1_3* is active in a condition that has yet to be studied, for example due to it having a different substrate-affinity.

Sadly the time limit of the project makes itself profoundly noticeable when it comes to the study of the role of the aldose reductases. Clear and definitive conclusions would require deletions of all three genes and a study of the phenotype of the resulting strains. It would also be very interesting to see the qPCR experiments be expanded with more conditions; some of interest would be to examine galactose to see the expression of *XYL1_2* as well a wider selection of carbon sources such as arabinose to see if a role of *XYL1_3* could be identified from it.

5.3 Galactose utilization

If *XYL1_2* is important for the utilization of galactose it raises the question of what its role in that metabolism is. In the pathway normally found for the metabolism of galactose in yeast, the Leloir pathway, there isn't any aldose reductases that partakes in conversion of galactose, see Figure 2.2.

In our research group there has previously been made the connection that aldose reductase *XYL1_2* might actually partakes in a separate galactose pathway previously unseen in yeast [13]. In filamentous fungi an alternative pathway for the metabolism of galactose has been identified [35]. There are several interesting parallels between this pathway in filamentous fungi and the observations regarding *XYL1_2* in C11 [13, 35, 65] the first of which is that the first step of the pathway is performed by an aldose reductase. Another point supporting the hypothesis is that in the second galactose gene cluster of C11 where *XYL1_2* is situated a *GAL10* gene is found that only has a single domain. In yeast *GAL10* traditionally has two domains that together make up the protein when transcribed, this is also true for the *GAL10* found in the more traditional gene cluster that does not contain *XYL1_2* in C11. However, in filamentous fungi *GAL10* is only made up from a single domain, same as in the novel gene cluster in C11.

The potential discovery of a galactose pathway previously unseen in yeast is indeed very exciting, but based on previous work on C11 [13] as well as this project the data is as of yet indecisive on whether this is the case. The deletion of aldose reductase gene *XYL1_2* didn't provide any visible phenotype, more work is thus required to study galactose utilization in C11 for example by producing knock-out strains lacking either or both of the potential galactose pathways. This is work that hopefully can be performed easily thanks to the progress made during this project with the development of new deletion protocols with high gene targeting efficiency.

6

Conclusion

Deletion protocols were successfully developed that have a very high gene targeting efficiency. The initial attempts at using a linear deletion cassette with homologous regions flanking the insert provided no successful deletions, but this was improved up to a gene target efficiency of 3.7 % using cell cycle synchronization to the S-phase. This was even further improved upon using the novel approach of utilizing a split marker deletion cassette which pushed the gene targeting efficiency to a staggering 67 %.

The developed deletion protocol was applied for the deletion of aldose reductase *XYL1_2* where each out of ten transformant colonies tested for deletion verified the correct insertion. However, the second application of the deletion protocol which aimed at deleting the galactose gene cluster where *XYL1_2* is situated couldn't be verified as having a successful deletion. When examined further the strain with a deletion of *XYL1_2* didn't provide a visible phenotype.

Probable physiological roles of aldose reductases Xyl1p & Xyl1_2p were found using the qPCR data acquired in the project as well as data obtained prior by the research group [13]. *XYL1* seems to be connected to xylose utilization, most likely as a xylose reductase. The data indicates *XYL1_2* as being a part of galactose utilization, possibly as part of a galactose pathway traditionally found in filamentous fungi that previously hasn't been observed in yeast. More work is however desirable to verify the conclusions drawn regarding aldose reductases *XYL1* & *XYL1_2*.

The role of *XYL1_3* is as of yet unclear. But it appears to not be significantly expressed in the presence of xylose or lactose, the two main conditions of interest throughout this project. To define the role of *XYL1_3* more work with different deletion strains and a broader range of conditions is necessary.

7

Bibliography

[1] L. D. Gomez, C. G. Steele-King, and S. J. McQueen-Mason, "Sustainable liquid biofuels from biomass: the writing's on the walls," *New Phytologist*, vol. 178, no. 3, pp. 473–485, 2008.

[2] A. Y. Krylova, E. Kozyukov, and A. Lapidus, "Ethanol and diesel fuel from plant raw materials: a review," *Solid Fuel Chemistry*, vol. 42, no. 6, pp. 358–364, 2008.

[3] S. N. Naik, V. V. Goud, P. K. Rout, and A. K. Dalai, "Production of first and second generation biofuels: a comprehensive review," *Renewable and Sustainable Energy Reviews*, vol. 14, no. 2, pp. 578–597, 2010.

[4] A. J. Van Maris, A. A. Winkler, M. Kuyper, W. T. De Laat, J. P. Van Dijken, and J. T. Pronk, "Development of efficient xylose fermentation in saccharomyces cerevisiae: xylose isomerase as a key component," in *Biofuels*, pp. 179–204, Springer, 2007.

[5] P. Kötter and M. Ciriacy, "Xylose fermentation by saccharomyces cerevisiae," *Applied microbiology and biotechnology*, vol. 38, no. 6, pp. 776–783, 1993.

[6] T. Jeffries and Y.-S. Jin, "Metabolic engineering for improved fermentation of pentoses by yeasts," *Applied Microbiology and Biotechnology*, vol. 63, no. 5, pp. 495–509, 2004.

[7] B. Hahn-Hägerdal, K. Karhumaa, M. Jeppsson, and M. F. Gorwa-Grauslund, "Metabolic engineering for pentose utilization in saccharomyces cerevisiae," in *Biofuels*, pp. 147–177, Springer, 2007.

[8] J.-P. Pitkänen, A. Aristidou, L. Salusjärvi, L. Ruohonen, and M. Penttilä, "Metabolic flux analysis of xylose metabolism in recombinant saccharomyces cerevisiae using continuous culture," *Metabolic engineering*, vol. 5, no. 1, pp. 16–31, 2003.

[9] L. Veiga, M. Bacila, and B. Horecker, "Pentose metabolism in candida albicans. i. the reduction of d-xylose and l-arabinose," *Biochemical and Biophysical Research Communications*, vol. 2, no. 6, pp. 440–444, 1960.

[10] E. Heath, J. Hurwitz, B. Horecker, and A. Ginsburg, "Pentose fermentation by lactobacillus plantarum i. the cleavage of xylulose 5-phosphate by phosphoketolase," *Journal of Biological Chemistry*, vol. 231, no. 2, pp. 1009–1029, 1958.

[11] A. Toivola, D. Yarrow, E. Van Den Bosch, J. P. Van Dijken, and W. A. Scheffers, "Alcoholic fermentation of d-xylose by yeasts," *Applied and Environmental Microbiology*, vol. 47, no. 6, pp. 1221–1223, 1984.

[12] A. D. Moreno, C. Tellgren-Roth, L. Soler, J. Dainat, L. Olsson, and C. Geijer, "Complete genome sequences of the xylose-fermenting candida intermedia strains cbs 141442 and pycc 4715," *Genome Announcements*, vol. 5, no. 14, pp. e00138–17, 2017.

[13] C. Geijer, "Unpublished results," 2017.

[14] H. Lee, "Review: The structure and function of yeast xylose (aldose) reductases," *Yeast*, vol. 14, no. 11, pp. 977–984, 1998.

[15] S. Watanabe, A. A. Saleh, S. P. Pack, N. Annaluru, T. Kodaki, and K. Makino, "Ethanol production from xylose by recombinant saccharomyces cerevisiae expressing protein-engineered nadh-preferring xylose reductase from pichia stipitis," *Microbiology*, vol. 153, no. 9, pp. 3044–3054, 2007.

[16] R. Riley, S. Haridas, K. H. Wolfe, M. R. Lopes, C. T. Hittinger, M. Göker, A. A. Salamov, J. H. Wisecaver, T. M. Long, C. H. Calvey, *et al.*, "Comparative genomics of biotechnologically important yeasts," *Proceedings of the National Academy of Sciences*, p. 201603941, 2016.

[17] S. Ostojić, M. Pavlović, M. Živić, Z. Filipović, S. Gorjanović, S. Hranisavljević, and M. Dojčinović, "Processing of whey from dairy industry waste," *Environmental Chemistry Letters*, vol. 3, no. 1, pp. 29–32, 2005.

[18] M. G. Siso, "The biotechnological utilization of cheese whey: a review," *Bioresource Technology*, vol. 57, no. 1, pp. 1–11, 1996.

[19] A. Plovie, "Creating a molecular toolbox in *Candida intermedia* to optimize xylose fermentation," *MSc Thesis, University of Gothenburg*, 2015.

[20] J. M. Wagner and H. S. Alper, "Synthetic biology and molecular genetics in non-conventional yeasts: current tools and future advances," *Fungal Genetics and Biology*, vol. 89, pp. 126–136, 2016.

[21] F. K. Agbogbo and G. Coward-Kelly, "Cellulosic ethanol production using the naturally occurring xylose-fermenting yeast, pichia stipitis," *Biotechnology letters*, vol. 30, no. 9, pp. 1515–1524, 2008.

[22] J. Blakely and S. MacKenzie, "Purification and properties of a β-hexosidase from sporobolomyces singularis," *Canadian journal of biochemistry*, vol. 47,

no. 11, pp. 1021–1025, 1969.

[23] P. A. Frey, "The leloir pathway: a mechanistic imperative for three enzymes to change the stereochemical configuration of a single carbon in galactose.," *The FASEB Journal*, vol. 10, no. 4, pp. 461–470, 1996.

[24] B. M. Chassy and J. Thompson, "Regulation and characterization of the galactose-phosphoenolpyruvate-dependent phosphotransferase system in lactobacillus casei.," *Journal of bacteriology*, vol. 154, no. 3, pp. 1204–1214, 1983.

[25] K. Bettenbrock and C.-A. Alpert, "The gal genes for the leloir pathway oflactobacillus casei 64h," *Applied and environmental microbiology*, vol. 64, no. 6, pp. 2013–2019, 1998.

[26] A. M. Elshafei and O. M. Abdel-Fatah, "Evidence for a non-phosphorylated route of galactose breakdown in cell-free extracts of aspergillus niger," *Enzyme and microbial technology*, vol. 29, no. 1, pp. 76–83, 2001.

[27] E. Fekete, L. Karaffa, E. Sándor, I. Bányai, B. Seiboth, G. Gyémánt, A. Sepsi, A. Szentirmai, and C. P. Kubicek, "The alternative d-galactose degrading pathway of aspergillus nidulans proceeds via l-sorbose," *Archives of microbiology*, vol. 181, no. 1, pp. 35–44, 2004.

[28] A. Kotyk, "Mobility of the free and of the loaded monosaccharide carrier in saccharomyces cerevisiae," *Biochimica et Biophysica Acta (BBA)-Biomembranes*, vol. 135, no. 1, pp. 112–119, 1967.

[29] B. Hahn-Hägerdal, J. Hallborn, H. Jeppsson, L. Olsson, K. Skoog, and M. Walfridsson, "Pentose fermentation to alcohol," *Biotechnology in Agriculture*, pp. 231–231, 1993.

[30] P. Peringer, H. Blachere, G. Corrieu, and A. Lane, "A generalized mathematical model for the growth kinetics of saccharomyces cerevisiae with experimental determination of parameters," *Biotechnology and bioengineering*, vol. 16, no. 4, pp. 431–454, 1974.

[31] M. J. Leandro, P. Gonçalves, and I. Spencer-Martins, "Two glucose/xylose transporter genes from the yeast candida intermedia: first molecular characterization of a yeast xylose–h+ symporter," *Biochemical Journal*, vol. 395, no. 3, pp. 543–549, 2006.

[32] D. Harcus, D. Dignard, G. Lépine, C. Askew, M. Raymond, M. Whiteway, and C. Wu, "Comparative xylose metabolism among the ascomycetes c. albicans, s. stipitis and s. cerevisiae," *PloS one*, vol. 8, no. 11, p. e80733, 2013.

[33] D. J. Wohlbach, A. Kuo, T. K. Sato, K. M. Potts, A. A. Salamov, K. M. LaButti, H. Sun, A. Clum, J. L. Pangilinan, E. A. Lindquist, *et al.*, "Comparative genomics of xylose-fermenting fungi for enhanced biofuel production,"

Proceedings of the National Academy of Sciences, vol. 108, no. 32, pp. 13212–13217, 2011.

[34] M. Johnston, "A model fungal gene regulatory mechanism: the gal genes of saccharomyces cerevisiae.," *Microbiological reviews*, vol. 51, no. 4, p. 458, 1987.

[35] B. Seiboth, B. S. Pakdaman, L. Hartl, and C. P. Kubicek, "Lactose metabolism in filamentous fungi: how to deal with an unknown substrate," *Fungal Biology Reviews*, vol. 21, no. 1, pp. 42–48, 2007.

[36] J. San Filippo, P. Sung, and H. Klein, "Mechanism of eukaryotic homologous recombination," *Annu. Rev. Biochem.*, vol. 77, pp. 229–257, 2008.

[37] A. De Lozanne and J. A. Spudich, "Disruption of the dictyostelium myosin heavy chain gene by homologous recombination," *Science*, vol. 236, no. 4805, pp. 1086–1091, 1987.

[38] R. Kooistra, P. J. Hooykaas, and H. Y. Steensma, "Efficient gene targeting in kluyveromyces lactis," *Yeast*, vol. 21, no. 9, pp. 781–792, 2004.

[39] M. R. Lieber, "The mechanism of double-strand dna break repair by the non-homologous dna end-joining pathway," *Annual review of biochemistry*, vol. 79, pp. 181–211, 2010.

[40] M. D. Krawchuk and W. P. Wahls, "High-efficiency gene targeting in schizosaccharomyces pombe using a modular, pcr-based approach with long tracts of flanking homology," *Yeast (Chichester, England)*, vol. 15, no. 13, p. 1419, 1999.

[41] L. Näätsaari, B. Mistlberger, C. Ruth, T. Hajek, F. S. Hartner, and A. Glieder, "Deletion of the pichia pastoris ku70 homologue facilitates platform strain generation for gene expression and synthetic biology," *PloS one*, vol. 7, no. 6, p. e39720, 2012.

[42] C. M. J. Koh, Y. Liu, M. Du, L. Ji, *et al.*, "Molecular characterization of ku70 and ku80 homologues and exploitation of a ku70-deficient mutant for improving gene deletion frequency in rhodosporidium toruloides," *BMC microbiology*, vol. 14, no. 1, p. 50, 2014.

[43] J. Verbeke, A. Beopoulos, and J.-M. Nicaud, "Efficient homologous recombination with short length flanking fragments in ku70 deficient yarrowia lipolytica strains," *Biotechnology letters*, vol. 35, no. 4, pp. 571–576, 2013.

[44] V. Tsakraklides, E. Brevnova, G. Stephanopoulos, and A. J. Shaw, "Improved gene targeting through cell cycle synchronization," *PloS one*, vol. 10, no. 7, 2015.

[45] H. S. Rosenkranz and J. A. Levy, "Hydroxyurea: a specific inhibitor of deoxyribonucleic acid synthesis," *Biochimica et Biophysica Acta (BBA)-Nucleic Acids*

and Protein Synthesis, vol. 95, no. 1, pp. 181–183, 1965.

[46] A. Galli and R. H. Schiestl, "Hydroxyurea induces recombination in dividing but not in g1 or g2 cell cycle arrested yeast cells," *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis*, vol. 354, no. 1, pp. 69–75, 1996.

[47] K. J. Livak and T. D. Schmittgen, "Analysis of relative gene expression data using real-time quantitative pcr and the 2- ddct method," *methods*, vol. 25, no. 4, pp. 402–408, 2001.

[48] J. Vandesompele, K. De Preter, F. Pattyn, B. Poppe, N. Van Roy, A. De Paepe, and F. Speleman, "Accurate normalization of real-time quantitative rt-pcr data by geometric averaging of multiple internal control genes," *Genome biology*, vol. 3, no. 7, pp. research0034–1, 2002.

[49] M. I. Pronobis, N. Deuitch, and M. Peifer, "The miraprep: a protocol that uses a miniprep kit and provides maxiprep yields," *PLoS One*, vol. 11, no. 8, 2016.

[50] K. Cha-aim, H. Hoshida, T. Fukunaga, and R. Akada, "Fusion pcr via novel overlap sequences," *Gene Synthesis: Methods and Protocols*, pp. 97–110, 2012.

[51] J. Drumonde-Neves, E. Vieira, M. Lima, I. Araujo, M. Casal, and D. Schuller, "An easy, quick and cheap high-throughput method for yeast dna extraction from microwell plates," *Journal of microbiological methods*, vol. 93, no. 3, pp. 206–208, 2013.

[52] A. Gedvilaite and K. Sasnauskas, "Control of the expression of the ade2 gene of the yeast saccharomyces cerevisiae," *Current genetics*, vol. 25, no. 6, pp. 475–479, 1994.

[53] C. Fairhead, B. Llorente, F. Denis, M. Soler, and B. Dujon, "New vectors for combinatorial deletions in yeast chromosomes and for gap-repair cloning using'split-marker'recombination," *Yeast*, vol. 12, no. 14, pp. 1439–1457, 1996.

[54] J. Fu, E. Hettler, and B. Wickes, "Split marker transformation increases homologous integration frequency in cryptococcus neoformans," *Fungal Genetics and Biology*, vol. 43, no. 3, pp. 200–212, 2006.

[55] N. L. Catlett, B.-N. Lee, O. Yoder, and B. G. Turgeon, "Split-marker recombination for efficient targeted deletion of fungal genes," *Fungal Genetics Reports*, vol. 50, no. 1, pp. 9–11, 2003.

[56] Q.-R. Xu, L. Yan, Q.-Z. Lv, M. Zhou, X. Sui, Y.-B. Cao, and Y.-Y. Jiang, "Molecular genetic techniques for gene manipulation in candida albicans," *Virulence*, vol. 5, no. 4, pp. 507–520, 2014.

[57] L. T. Colombo, J. C. C. Rosa, C. R. Bragança, R. P. Ignacchiti, M. C. Alvim, W. B. Silveira, M. V. de Queiroz, D. M. Bazzolli, and F. M. Passos, "Con-

struction of a kluyveromyces lactis ku80- host strain for recombinant protein production: Extracellular secretion of pectin lyase and a streptavidin–pectin lyase chimera," *Molecular biotechnology*, vol. 56, no. 4, pp. 319–328, 2014.

[58] A. Kretzschmar, C. Otto, M. Holz, S. Werner, L. Hübner, and G. Barth, "Increased homologous integration frequency in yarrowia lipolytica strains defective in non-homologous end-joining," *Current genetics*, vol. 59, no. 1-2, pp. 63–72, 2013.

[59] B.-J. You, M.-H. Lee, and K.-R. Chung, "Gene-specific disruption in the filamentous fungus cercospora nicotianae using a split-marker approach," *Archives of microbiology*, vol. 191, no. 7, pp. 615–622, 2009.

[60] E. S. Lander, "The heroes of crispr," *Cell*, vol. 164, no. 1, pp. 18–28, 2016.

[61] E. Charpentier and J. A. Doudna, "Biotechnology: Rewriting a genome," *Nature*, vol. 495, no. 7439, pp. 50–51, 2013.

[62] W. NEUHAUSER, D. HALTRICH, K. D. KULBE, and B. NIDETZKY, "Nad (p) h-dependent aldose reductase from the xylose-assimilating yeast candida tenuis: isolation, characterization and biochemical properties of the enzyme," *Biochemical Journal*, vol. 326, no. 3, pp. 683–692, 1997.

[63] H. Lee, "The structure and function of yeast xylose (aldose) reductases," *Yeast*, vol. 14, no. 11, pp. 977–984, 1998.

[64] C. Verduyn, R. Van Kleef, J. Frank, H. Schreuder, J. Van Dijken, and W. Scheffers, "Properties of the nad (p) h-dependent xylose reductase from the xylose-fermenting yeast pichia stipitis," *Biochemical journal*, vol. 226, no. 3, pp. 669–677, 1985.

[65] J. C. Slot and A. Rokas, "Multiple gal pathway gene clusters evolved independently and by different mechanisms in fungi," *Proceedings of the National Academy of Sciences*, vol. 107, no. 22, pp. 10136–10141, 2010.