The Dissertation Committee for Sun-Mi Lee Certifies that this is the approved version of the following dissertation :

Combinatorial Engineering of *Saccharomyces cerevisiae* for efficient pentose catabolism

Committee:

Hal S. Alper, Supervisor

Lydia Contreras

George Georgiou

Jennifer Maynard

Marvin Whiteley

Combinatorial Engineering of *Saccharomyces cerevisiae* for efficient pentose catabolism

by

Sun-Mi Lee, B.S. Agriculture; M.S. Environ. Sci. Ecol. E.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

August 2014

Dedication

To my husband and son, Jehyeong and Junu Lee

We live under the same roof!

Acknowledgements

I would like to thank my advisor, Dr. Hal Alper, for all his support throughout this journey. Without his guidance, I would not have accomplished this work. His enthusiasm, kindness, and in-depth knowledge of research have inspired me, and he has been a great role model. He was the reason I came to UT for my PhD study, and it was the right choice.

I would also like to thank my thesis committee members. Special thanks to Dr. Jennifer Maynard for her supports and advices in both my research and life. I am also grateful to Dr. Lydia Contreras for her warmth and kindness. Her support at AIChE conference in San Francisco was really helpful. I especially acknowledge Dr. George Georgiou for his advice and time during his busy schedule. I sincerely thank Dr. Marvin Whiteley for his thought-provoking suggestions and advices.

I would like to thank Taylor Jellison for her help in experiments. We worked hard and had so much fun in the lab. I would also like to thank Dr. Cheulhee Jung and his wife-to-be Boyoung Park for their help in enzyme activity assay and GC analysis.

I would also like acknowledge Korea Institute of Science and Technology (KIST) for letting me to pursue my PhD study. I would like to thank those from KIST who visited and encouraged me while I am in Austin.

I would also like to thank all of the members of the Alper Group. Thanks: to Dr. Eric Young for his help in research and writing, to Dr. Jie Sun for her friendship, help in finishing up my works, and delicious dishes, to Dr. Nathan Crook for friendship and sharing lots of fun, to Dr. Kate Curran for friendship and delicious baked goods, to Leqian Liu for scientific/non-scientific discussion and friendship, Joseph Cheng for

friendship and organizing/hosting dinner parties, Andrew Hill for help in bioreactor running, Joe Abatemarco for Kimchi, Haibo Li for experiments and the book, John Leavitt for night experiment companion, Heidi Redden for sequencing samples, also Dr. Amanda Lanza, Dr. John Blazeck, Rebecca Knight, Kelly Markham, and Nick Morse.

I would like to thank my friends outside of the Alper Lab who have made my time in Austin very enjoyable: Lynn Li, Beatrice Mabrey, Wenzong Li, Peach Kasemset, and Jongsuk Kim. Let's keep in touch!

Most importantly, I want to thank my family for their love and support. I want to thank my parents, Hee-Young Lee and Sung-Ja Kang. I cannot express how much I appreciate your love and supports. I also want to thank my sister and brother, Eunkyung and Jungmin Lee, for their love and being such good friends. I appreciate my in-laws, Kyu-Chul Chung and Sung-Nam Jung, for their love and understanding. I want to thank my son, Junu Lee Jung, for being my happiness. Finally, my most sincere thanks go to my husband, Je Hyeong Jung. I love you. Thank you for being my rock.

Combinatorial Engineering of *Saccharomyces cerevisiae* for efficient pentose catabolism

Sun-Mi Lee, Ph. D.

The University of Texas at Austin, 2014

Supervisor: Hal S. Alper

The efficient fermentation of lignocellulosic biomass would enable more economically and environmentally friendly production of biofuels and biochemicals. Yet, *Saccharomyces cerevisiae*, a platform organism for biofuels and biochemicals production, is unable to convert all of the sugars in lignocellulosic biomass into biofuels and biochemicals mainly due to the lack of a pentose catabolic pathway. Though the advance of genetic engineering enabled *S. cerevisiae* to utilize pentose sugars, the efficiency of pentose sugar catabolism in *S. cerevisiae* is still limited. Here, the goal of this research was to confer efficient pentose sugar catabolism to *S. cerevisiae* by combinatorial and evolutionary engineering. To this end, pentose catabolic pathways were 1) constructed by heterologous expression of pentose catabolic genes, 2) optimized through rational engineering, and 3) further improved through evolutionary engineering.

Through these efforts, we reported the highest ethanol yield (0.45 g ethanol / g xylose) and the second highest xylose consumption and ethanol production rates (0.98 g xylose g cell^{-1} h^{-1} and 0.44 g ethanol g cell^{-1} h^{-1}, respectively) in xylose fermentation

reported to date. The high performance in xylose fermentation was achieved based on the mutant xylose isomerase (*xylA3*), which showed 77% increased enzyme activity, engineered through directed evolution. In addition, we have established the first cells capable of growing on arabinose in mimimal medium and demonstrated ethanol production from xylan in minimal medium. The arabinose and xylan catabolic pathways were constructed in *S. cerevisiae* by expressing novel pentose catabolic genes from a strain with remarkable pentose catabolic potential that we isolated and named *Ustilago bevomyces*. In doing so, a complete workflow of bioprospecting to pathway engineering and evolution was detailed as an effective way to transfer a desired phenotype from a non-model organism to a model organism.

This study substantially improved the prospect of biofuels and biochemicals production from lignocellulosic biomass by developing efficient pentose utilizing strains, finding new pentose catabolic genes, and suggesting alternative pentose catabolic pathway. Furthermore, the general tools for metabolic engineering demonstrated in this study would also advance microbial strain engineering.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction and Background

1.1 Biofuels and Biochemicals Production from Lignocellulosic Biomass

Bioconversion of lignocellulosic biomass offers an alternative means to produce fuels and chemicals which are now produced during petroleum refining processes. Moreover, this bioconversion is regarded as an environmentally friendly and sustainable approach [1, 2]. A major challenge in bioconversion of lignocellulosic biomass, especially for yeast, comes from inefficient conversion of the pentose sugars that comprise a substantial portion of lignocellulosic biomass. This prevents economical biofuels and biochemicals production.

1.1.1 Lignocellulosic biomass as a sustainable feedstock

Bioethanol fermentation from starch and sugar crops is currently the industry standard for the conversion process of biomass to biofuel. With well-developed fermentation technologies, industrial bioethanol production is over 15 billion gallons per year by baker's yeast *Saccharomyces cerevisiae* [3]. However, there is a paradigm shift in the feedstock for the biofuel production, from starch and sugar crops to lignocellulosic biomass, due to the food vs fuels and land usages issues [4].

Lignocellulosic biomass, such as agricultural residues, forestry wastes, waste paper, and energy crops, is regarded as economical and sustainable feedstock for biofuel production since it has a plentiful source of supply, competes less with food and feed crops, and requires less fossil fuel inputs [5]. While glucose is the most common sugar in

lignocellulosic biomass (32 ~ 50 %), large quantities of other hexose sugars such as galactose and mannose (2 ~ 14%) and pentose sugars such as xylose and L-arabinose (11~ 25%) exist in varying fractions depending on the sources of lignocellulosic biomass [4, 6-8]. Additionally, lignocellulosic biomass derivatives contain other molecules besides monomeric sugars generated from prior processes of pretreatment and hydrolysis for fermentation. Phenolic lignin derivatives are present in large fractions (10-20%), partial depolymerization during hydrolysis produces small carbohydrate polymers such as cellobiose, and harsh reaction conditions can degrade sugars to furfurals. Traditional metabolic engineering has relied upon the conversion of readily consumed carbon sources such as glucose. However, the composition of lignocellulosic biomass demands the development of novel catabolic pathways to consume non-native, or exogenous, carbon sources.

1.1.2 Engineering *S. cerevisiae* for utilizing pentose sugars in lignocellulosic biomass

S. cerevisiae readily ferments glucose to carbon dioxide and ethanol, so it has been used as an organism of choice for industrial production of commodities such as bread, wine and beer. This fermentative capacity also makes *S. cerevisiae* an attractive host for starch and sugar based ethanol production since these feedstocks are primarily glucose. With these feedstocks, *S. cerevisiae* produce ethanol with high yield and titers (over 90 % of the theoretical yield and up to 160 g/L of titiers) in industrial scale processes [3]. However, ethanol production from lignocellulosic biomass remains

suboptimal due to incomplete conversion of all available sugars in lignocellulosic biomass to ethanol.

Pentose sugars such as xylose and arabinose constitute significant portion of lignocellulosic biomass. In particular, xylose is the second most abundant sugar in lignocellulosic biomass constituting up to 22 % of dry mass [4]. However, native *S. cerevisiae* are unable to utilize pentose sugars due to the lack of pentose catabolic pathways. Therefore, *S. cerevisiae* has been engineered to facilitate pentose sugar utilization for efficient biofuel production from lignocellulosic biomass. For the last few decades, two types of pentose pathways have been constructed in yeast: the oxidoreductase pathway and the isomerase pathway (**Figure 1.1**). Both xylose and arabinose can be metabolized through each of these pathways, although arabinose assimilation involves additional steps in both cases [9]. All four possible pathway variants have been previously constructed [9-12], and all feed into native *S. cerevisiae* metabolism via D-xylulose or D-xylulose-5-phosphate. Once converted to xylulose-5-phosphate, these sugars are further metabolized through the native pentose phosphate pathway (PPP).

The pentose oxidoreductase pathways are conserved between certain species of native fungi, and employ common enzymes and redox cofactors to catalyze substrate conversion. The xylose oxidoreductase pathway was the first heterologous pentose pathway constructed in *S. cerevisiae* [10]. In this pathway, xylose is reduced to xylitol by an aldose reductase (AR), and then xylitol is oxidized to xylulose by xylitol dehydrogenase (XDH). The AR most commonly used is encoded by *Scheffersomyces stipitis* (formerly

Pichia stipitis) *XYL1* (xylose reductase, XR) which prefers the cofactor NADPH over NADH. The XDH is encoded by *S. stipitis XYL2*, which is NAD^+ dependent [13].

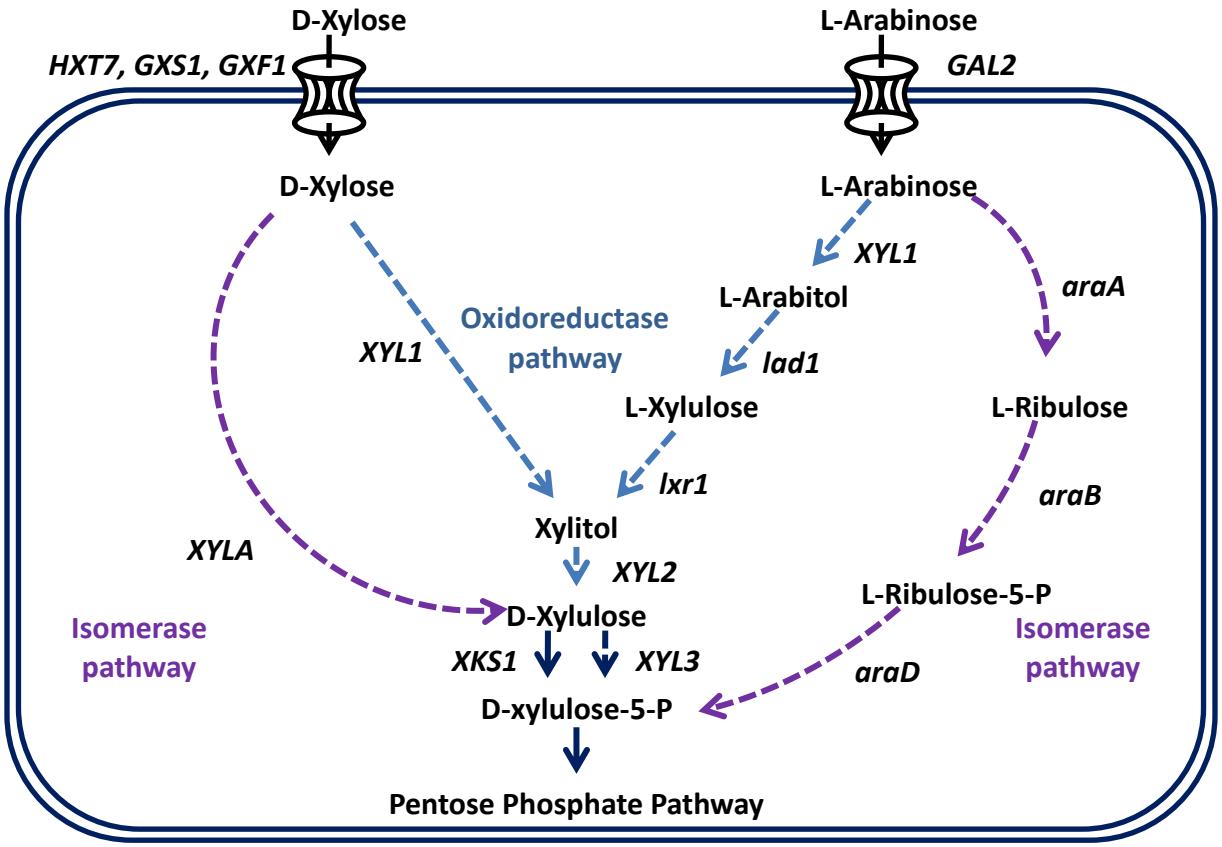


Figure 1.1: Successful exogenous transport and metabolic pathways introduced *in S. cerevisiae.* Reported improvements of pentose utilization in yeast (as described in the text) are depicted in this schematic. Left: Bacterial xylose isomerase pathway. Middle left: Fungal xylose oxidoreductase pathway. Middle Right: Fungal arabinose oxidoreductase pathway. Right: Bacterial arabinose isomerase pathway. The genes used to accomplish the enzymatic step are italicized. Heterologous steps are indicated by dashed lines.

The L-arabinose oxidoreductase pathway was also the first variant constructed in *S. cerevisiae* for arabinose conversion [14]. The two enzymes from the xylose oxidoreductase pathway AR and XDH serve as catalysts of the first and last reactions, respectively. The remaining two steps have recently been constructed using two genes from the fungus *Trichoderma reesei* (scientific name of *Hypocrea jecorina*). The first gene, *LAD1*, encodes arabinitol 4-dehydrogenase (ADH) which reduces arabitol to L-xylulose using NADPH as a cofactor[15]. The second is an L-xylulose reductase (XR) encoded by *LXR1* , which reduces L-xylulose to D- xylulose using NAD^+ as a cofactor [14]. The D-xylulose from both the xylose and arabinose variants is then utilized by native metabolic steps, culminating in the production of biomass, carbon dioxide and ethanol.

The main challenges in xylose and arabinose oxidoreductase pathways are cofactor imbalances limiting theoretical and actual pentose conversion by *S. cerevisiae* [2, 16]. Recent studies have revealed AR and XDH genes from other organisms such as *Rhodototurula mucilaginosa* may have more favorable cofactor usage relieving cofactor imbalance [17]. Additional homologs have also been reported [18-23]. The pathways constructed from these homologous genes showed the potential to function better than the pathway constructed from *S. stipitis* AR and XDH. Additionally, discovery of more efficient LAD and LXR homologs may enable faster rates of L-arabinose consumption. Significant work has gone into discovering more LAD and LXR [24-30]; however, this pathway remains very slow and inefficient. Furthermore, there has been some disagreement over whether some proteins are in fact LXR [30]. This inability to replicate the arabinose growth rate of native organisms by constructing the arabinose

oxidoreductase pathway could indicate that the enzymes are not functioning well in *S. cerevisiae*, or that the pathway has not been fully elucidated.

Alternatively, isomerase pathways may be constructed to facilitate the consumption of xylose and arabinose. In contrast to the pentose oxidoreductase pathway, the isomerase pathway variants require no cofactors. The pathway is native to bacterial species and to rare yeasts. The heterologous xylose isomerase pathway minimally consists of one enzyme, xylose isomerase (XI), which directly converts xylose to xylulose. Because most XIs are native to bacteria, difficulties for heterologous expression in yeast exist [31], yet work has demonstrated functional bacterial XI pathways [11, 32] and yielded functional heterologous XIs isolated from rare fungi [33, 34]. Recently, pathway function and ethanol yields were shown to be improved via the directed evolution of xylose isomerase [35]. This represents a step forward in xylose pathway engineering, especially for the more attractive isomerase-based pathway. Furthermore, as with the oxidoreductase pathway, the complementation of a xylulokinase can further improve yields and assimilation rates. This pathway has improved ethanol conversion yields over the oxidoreductase pathway; however, strains show lower growth and sugar uptake rates. Nevertheless, this pathway is attractive because of its lack of cofactor imbalance.

Whereas the xylose isomerase pathway involves one step catalyzed by xylose isomerase, the arabinose isomerase pathway consists of three steps requiring three enzymes of arabinose isomerase, ribulosekinase, and ribulose-5-P-4-epimerase encoded by *araA, araB and araD*, respectively. Two variations of this pathway have been constructed in yeast using distinct sets of heterologous genes from *Bacillus subtilis*

(*araA*) and *Escherichia coli* (*araB* and *araD*) [12] and from *Lactobacillus plantarum* (*araA*, *araB* and *araD*) [36]. However, unlike the other pathways described above, evolutionary engineering was needed in both cases to isolate a yeast strain with an active arabinose isomerase pathway. Thus, only after mutations is a functional arabinose isomerase pathway in yeast possible. As a result, more work is necessary to describe a fully stand-alone arabinose catabolic pathway.

Whereas the enzymes discussed above alone are sufficient to enable xylose or arabinose metabolism in yeast, pentose assimilation remains suboptimal. This has led researchers to investigate downstream enzymatic steps in the PPP. The xylulokinase gene from *S. stipitis* (XYL3) is often complemented to further reduce xylitol production [37-42]. Transaldolase (TAL) and transketolase (TKL) alterations have also been shown to improve pathway function [43-45]. Another step to improve pentose assimilation is to optimize a step not typically included in metabolic pathways – transport across the cellular membrane. *S. cerevisiae* possesses many monosaccharide transporters, but these are almost exclusively hexose transporters [46-50]. Due to the broad substrate specificity of many of these transporters, xylose and arabinose leak in through several hexose transporters [51]. This is not optimal since these transporters are not adapted to efficiently uptake pentoses, and glucose can competitively inhibit pentose uptake limiting pentose assimilation. Heterologous expression of sugar transporters has produced some encouraging results [17, 52-60]. However, the most potential likely lies in the directed evolution of transporters for pentose efficiency and specificity [61, 62], and there is still much improvement to be made for pentose uptake to rival that of glucose uptake.

Even with an efficient xylose catabolic pathway, *S. cerevisiae* is not able to utilize xylose in lignocellulosic biomass unless xylan is hydrolyzed to xylose by hemicellulases. This requires an enzyme hydrolysis step prior to fermentation in a biofuel production process. Therefore, engineering *S. cerevisiae* for xylan utilization, which allows consolidated bioprospecting processes, has also been attracted interest of researchers. Depolymerization of xylan requires a variety of hemicellulases. A minimum combination of hemicellulases for enzymatic hydrolysis includes xylanases, xylosidases, and arabinofuranosidases, which degrade a main chain of xylan polymer, xylose disaccharides, and a side chain of xylan, respectively [63]. The most commonly used genes for the expression of these enzymes in *S. cerevisiae* involve *XynII* (xylanase) from *Trichoderma reesei* [64], *XlnD* (xylosidase) from *Aspergillus niger* [65], and *AbfB* (arabinofuranosidase) from *A. niger* [64]. Yet, xylan is still minimally utilized [67, 68] and hardly converted into ethanol [63] even with engineered *S. cerevisiae* using surface display technique. Most studies have reported only enzyme activity of hemicellulases when expressed in engineered *S. cerevisiae* or have demonstrated the conversion of xylan to reducing sugars [66, 67].

1.1.3 Engineering *S. cerevisiae* for advanced biofuels and biochemical production

Developing *S. cerevisiae* with efficient pentose catabolism would enable complete utilization of all available carbon sources in lignocellulosic biomass. Once developed and combined with downstream engineering, this strain could be used to

produce not only bioethanol but also other biofuels and biochemicals from lignocellulosic biomass.

Bioethanol is currently the most common but not the most attractive biofuel. The low energy content, which is 70% of the energy content of gasoline, is a major drawback of bioethanol. The hygroscopicity, which is the tendency to attract water molecules, also causes problems with storage and transportation of bioethanol with widely distributed conventional infrastructure. In addition, the high miscibility and formation of an azeotrope with water makes bioethanol difficult to separate in distillation processes. As a result, there has been increasing interest in producing advanced biofuels that are more similar to petroleum-derived compounds with high energy content and drop-in capability for current infrastructures.

Despite the advantages of using advanced biofuels, production in microorganisms has been difficult. Unlike bioethanol, a native byproduct of fermentation for some yeasts including *S. cerevisiae*, most advanced biofuels are not naturally produced in microorganisms - especially at the industrial scale. This is not surprising considering the formation of petroleum over eons, under high temperature and high pressure. In addition, microorganisms have not been forced by natural selection to produce large amounts of petroleum-like products in nature, so the possession of a complete metabolic pathway for advanced biofuels is rare or does not exist. Though some organisms possess a native pathway for certain advanced biofuels such as short-chain alcohols, these natural biofuels still have less attractive features as transportation fuels than other advanced biofuels that

are not produced natively in microorganisms. Therefore, metabolic engineering for the advanced biofuels production commonly requires novel pathway construction.

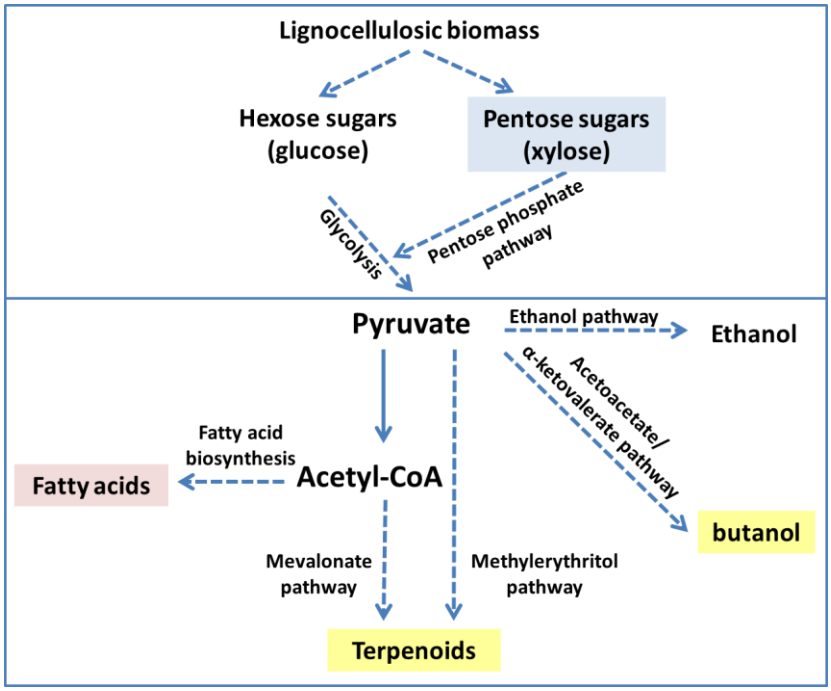


Figure 1.2: Advanced biofuels production in engineered *S. cerevisiae*. Pentose sugar utilization (blue), butanol, and terpenoids production (yellow) have been demonstrated, but require further improvement for industrial scale applications. Fatty acids production (red) has not been successfully demonstrated in *S. cerevisiae*.

To construct a novel pathway for advanced biofuels production, metabolic engineers often use bioethanol production pathways as benchmarks. Through metabolic engineering, advanced biofuels production is possible in *S. cerevisiae* (**Figure 1.2**). The inherent ability of ethanol production and high tolerance to alcohol makes *S. cerevisiae* a suitable host for alcohol-based biofuels. Butanol production in *S. cerevisiae* has been

demonstrated by re-routing acetyl-coA flux away from ethanol and toward butanol. Butanol is a short chain alcohol with higher energy content and lower hygroscopicity than ethanol. By expressing genes from butanol producing *Clostridium* bacteria, *S. cerevisiae* was able to produce 2.5 mg/L of butanol [68]. Later, re-routing the valine pathway for butanol production improved titers to 143 mg/L [69], and re-locating valine synthesis enzymes further increased butanol titers in *S. cerevisiae* to 630 mg/L [70]. While this is one of the leading achievements for advanced biofuels production, titer remains a limitation. Therefore, further metabolic engineering is required for pathway and metabolic network optimization.

Despite the possibility of producing these molecules, the vast majority of alternative biofuels in yeast are produced at inferior yields, titers, and rates when compared with bioethanol production. At present, these values are not high enough to meet the demand for liquid fuels in the transportation sector or even to be comparable to those from metabolically engineered *E. coli*. For example, an engineered *E. coli* strain expressing heterologous genes from *Clostridium acetobutylicum*, the most well-known butanol producing bacteria, produced 30 g/L of butanol [71]. The butanol titer of engineered *E. coli* is 50 fold higher than that of engineered *S. cerevisiae* (0.63 g/L) [70]. However, the robustness of yeast makes them promising hosts for advanced biofuels especially when we consider using lignocellulosic biomass as feedstocks. Yeast strains are commonly more tolerant to products and inhibitors from lignocellulosic biomass hydrolysates [3]. Therefore, once the yields, titers, and production rates are improved, advanced biofuel production will be more feasible with yeasts as a producing host. Here, metabolic

engineering could play a role to make yeasts more promising host in producing advanced biofuels.

1.2 METABOLIC ENGINEERING TOOLS FOR A NEW PATHWAY

Microorganisms have a long history of being used as a 'cell factory' to produce numerous compounds humans desire such as ethanol, and now the range of products and feedstocks can be expanded with the help of remarkable advances in metabolic engineering [72]. The transition of a 'cell factory' to an 'advanced cell factory' adds a value of products and lowers the cost of feed stock[73]. Introduction of heterologous pathways into platform organisms such as *S. cerevisiae* confer the ability to expand possible products and feedstock, but the introduced pathways need to be well-incorporated within a complex cellular network of the host organism to maximize yield and productivity of a desired product [74].

1.2.1 New enzymes for a new metabolic pathway

The abilities to utilize or to produce biofuels are often conferred by new metabolic pathways that are commonly based on the expression of foreign enzymes adapted from organisms with desired features. The success in the functional expression of foreign enzymes in yeasts can be achieved by choosing promising candidate genes that encode enzymes [17]. Even with a proper choice of the enzyme, however, alternative or modified enzymes are required for high efficiency. Yet, this has been hardly investigated

to date. For example, researchers have only begun looking into alternative xylose reductases, a key enzyme in the initial step of xylose catabolism [22].

The search for the new enzymes can be done by using computational tools or experimental bioprospecting approach. With an ever increasing sequence database, enzymes homologous to the previously known enzymes could be found through BLAST search. These newly found enzymes could be isolated from the source organism or they could be synthesized without the need to culture the source organism. Recently, a functional survey of sugar transporters based on BLAST search offered efficient xylose transporter candidates[17]. If *a priori* predictive capability is not available, bioprospecting can be used to discover required components to construct the best pathway [13]. For example, bioprospecting of the xylose fermenting fungi of *S. stipitis* found genes necessary to improve xylose utilization [45]. In the metabolic engineering field, bioprospecting is commonly conducted by building a cDNA library of the organism with desired phenotype. Though this approach does not require *a priori* knowledge, it is a labor and time intensive method and often restrictive for isolating single genes rather than pathways for the case of eukaryotic host DNA. Therefore, a new approach of bioprospecting would be required to facilitate engineering model organisms. With the decreasing cost and rapid pace of whole genome sequencing, an increasing number of whole genome sequences is available. In the biomedical field, genome-based bioprospecting is already used for finding new therapeutics [75]. Therefore, bioprospecting based on whole genome sequencing could offer a new direction which could be used for finding novel genes and pathways to engineer a model organism.

1.2.2 Successful introduction of heterologous pathways

Heterologous pathways are usually adapted from native organisms with a desired phenotype. DNA synthesis technology and inverse metabolic engineering enables the introduction of any known and unknown heterologous genes into host organisms. The success of the heterologous gene expression, however, depends on the compatibility of heterologous genes as an agreeable part in a complex host cellular network [72, 74, 76]. The enzymes showing high activity and stability in the native organisms often fail to be expressed appropriately in a heterologous expression host. The unsuccessful expression of heterologous xylose isomerases from *E. coli* [77], *Bacillus subtilis* [78], and *Thermus thermophiles* [11] into *S. cerevisiae* are good examples of incompatible heterologous gene expression in a nonnative host. Thus, tailoring genes is often necessary to ensure desired performance of heterologous enzymes, leading to the success of the heterologous pathway introduction. Direct modification of foreign genes by codon-optimization and directed evolution improve the fitness of the foreign enzymes in a host cell system leading to an efficient enzymatic step and high metabolic flux toward desired product, thus finally improving titers and production rates. Of gene tailoring methods, directed evolution offers a powerful tool for tailoring proteins that require refined functions. The iterative application of directed evolution allows for proteins with significantly improved function to be easily obtained in a short period of time [79]. By harnessing natural selection process in an intensive manner, therefore, the final version of enzyme can have improved

properties in enzyme activity, stability, and compatibility in a cellular network system of a host organism [79].

1.2.3 Optimization of metabolic pathways

Once introduced successfully, heterologous genes should co-operate with other endogenous genes within a cellular network system of a host strain. Specifically, the changes in the metabolic flux disrupted by newly introduced heterologous pathway need to be balanced through optimizing expression of endogenous genes both directly and indirectly related to the heterologous pathway [76]. In traditional metabolic pathway engineering, optimization of metabolic flux has been focused on the genes involved in the target pathway, and this approach for xylose catabolic pathway has shown to improve cell growth, xylose consumption and ethanol fermentation rates [40, 45, 80-83]. Of these genes, xylulokinase, the enzyme involved in the last step of the heterologous xylose pathway, is regarded as the primary target for pathway optimization to improve xylose catabolism given that the native xylulokinase encoded by *XKS1* is insufficient to support xylose metabolism in *S. cerevisiae*. Other than the xylose pathway, engineering of the genes encoding the enzymes involved in the pentose phosphate pathway and sugar transporters have also improved the xylose catabolism in *S. cerevisiae* suggesting the possibility that additional genes can be engineered for imporved xylose catabolism [84]. Moreover, genes that are not directly related to the target pathway, such as *PHO13* (ρ-nitrophenyl phosphate), also improved xylose fermentation [85].

In a complex cellular network system, however, searching for the right targets for pathway optimization is often hard to achieve, and sometimes the genes may not be present in the host strain which limits the possible options for the cellular network optimization [2, 86]. In these instances, bioprospecting – searching for novel enzymes from the strains with desired phenotypes in nature, provides an excellent solution [1, 86]. For example, transaldolase in the pentose phosphate pathway, which improved xylose utilization in *S. cerevisiae* when overexpressed, has been obtained from a bioprospecting approach by using cDNA library of *S. stipitis* [45]. Bioprospecting can also be used for finding novel genes involved not only in metabolic pathways but also in regulatory systems. Evolutionary engineering has also shown that cells often choose to modify their cellular network by modifying gene control systems as a response to the changing environment, not by gene modification itself [87]. Recent reports on the relationships of transcription factors with sugar utilization efficiency [87], ethanol production [88], and tolerances [89] support the need to broaden the focus of metabolic engineering approach from structural genes to regulatory genes.

1.2.4 Whole cell level optimization

The final goal of metabolic engineering is to develop a strain that produces the desired product with high yield and productivity. However, there is a limit in the rational engineering due to high complexity of cellular metabolic network. Therefore, combinatorial engineering is often applied to develop a strain with a desired

phenotype(s). In combinatorial engineering approach, rationally engineered strains are subjected to evolutionary engineering.

Evolutionary engineering is one of the effective ways to improve strains for a desired phenotype, and has been routinely applied to rationally engineered strains to boost their performance. For the case of biofuel production, evolutionary engineering has led to strains with improved tolerances to products and utilization of substrates [87, 90-92]. However, the adaptive process is rather intuitively conducted and requires a long period of time. Therefore, a generalized and effective method for adaptive evolutionary engineering should be developed for more successful strain engineering for industrial applications. Recently, adaptive engineering combined with or without metabolic engineering has been successfully applied to improve biochemical productivity and to broaden fermentable substrate ranges [91, 93, 94]. In addition, the efforts to understand 'irrational' adaptive engineering in a 'rational' way, such as whole genome sequencing, transcriptome or metabolome analysis, help to unravel underlying mechanisms happened during adaptive engineering processes [87]. Yet, the general application of the improved strain is limited since the adapted stain is highly specific for the target purpose. Therefore, the power of evolutionary engineering will be significantly improved if general approaches for efficient evolutionary engineering are determined.

Chapter 2: Improving the heterologous xylose isomerase pathway

2.1 Chapter Summary

The heterologous expression of a highly functional xylose isomerase pathway in *Saccharomyces cerevisiae* would have significant advantages for ethanol yield since this pathway bypasses cofactor requirements found in the traditionally-used oxidoreductase pathways. However, nearly all reported xylose isomerase-based pathways in *S. cerevisiae* suffer from poor ethanol productivities, low xylose consumption rates, and poor cell growth when compared with an oxidoreductase pathway and additionally often require adaptive strain evolution. Here, we report on the directed evolution of the *Piromyces sp.* xylose isomerase (*xylA*) for use in yeast. After three rounds of mutagenesis and growth-based screening, we isolated a variant containing six mutations (E15D, E114G, E129D, T142S, A177T, and V433I) that exhibited a 77% increase in enzymatic activity. When expressed in a minimally engineered yeast host containing a *gre3* knockout, and *tal1* and *XKS1* overexpression, the strain expressing this mutant enzyme improved aerobic growth rate by 61 fold and both ethanol production and xylose consumption rates by nearly 8 folds. Moreover, this mutant enzyme enabled ethanol production by these yeasts in an oxygen-limited fermentation condition, unlike the wild-type enzyme. Under micro-aerobic conditions, ethanol production rates of the strain expressing the mutant xylose isomerase were considerably higher than previously reported values for yeast harboring a xylose isomerase pathway, and also comparable to the strains harboring an

oxidoreductase pathway. Consequently, this study shows the potential to evolve a xylose isomerase pathway for more efficient xylose utilization.

2.2 INTRODUCTION

Efficient utilization of all available carbon in lignocellulosic biomass is one of the major challenges preventing economically viable biofuels production [1, 2]. Commonly used organisms for biofuel production such as the yeast *Saccharomyces cerevisiae* are unable to natively utilize the pentose sugars, which comprises a substantial portion of lignocellulosic biomass [84, 95]. Among these pentose sugars, xylose is the most abundant in commonly studied biomass sources. Thus, the ability to improve xylose catabolism and conversion in a recombinant host such as *S. cerevisiae* would substantially improve the prospect of biofuels and biochemicals production.

Xylose catabolic pathway can be constructed in *S. cerevisiae* by expressing one of two main heterologous pathways[84]. The first pathway, with oxidoreductase-based chemistry, has been well-established in yeast through the heterologous expression of (at a minimum) a xylose reductase (XR) and xylitol dehyerogenase (XDH) leading to the ability of *S. cerevisiae* to utilize xylose [13, 40, 45, 80-82, 96]. However, this pathway is inherently limited by a cofactor imbalance with the xylose reductase utilizing NADPH and the xylitol dehydrogenase utilizing NAD^+ which leads a diversion of metabolic flux toward undesired products as a compensation reaction and decreases ethanol yield [84]. Recent work has focused on modifying the cofactor preference of these enzymes to make them more compatible and establish an oxidation-reduction cycle [97, 98]. However, even

with matching cofactor specificities, the oxidoreductase pathway requires cofactors which may limit overall pathway throughput. In all of these cases, the yield of ethanol from xylose still remains suboptimal when compared with native xylose utilizers.

A second, alternative pathway for xylose catabolism mainly exists in bacteria and rare yeasts. This isomerase-based pathway has no cofactor requirements and thus could lead to higher theoretical yields (0.51 g ethanol / g xylose) since no by-product is necessarily produced to compensate co-factor imbalance. By comparison, experimental ethanol yields using the oxidoreductase and xylose isomerase pathways in anaerobic conditions have been shown to be between 0.09 - 0.23 [95, 99]and near 0.43 g ethanol/g xylose, respectively [100]. For this reason, there is considerable interest in improving a xylose isomerase-based pathway in *S. cerevisiae* with a particular focus on improving both cell growth rate and xylose consumption rate. Recent reports of successful expression of xylose isomerase genes from *Piromyces sp.* [33], *Orpinomyces sp.* [34], and *Clostridium phytofermentans* [32] in *S. cerevisiae* raise the prospect of efficient xylose fermentation. Furthermore, researchers have applied adaptive evolutionary engineering [91], optimized metabolic flux by introducing/overexpressing xylose transporter and/or overexpressing downstream pathway [83, 101], and employed bioprospecting to identify other putative xylose isomerase enzymes [32, 102]. In all of these cases, extensive downstream overexpression and/or evolutionary engineering is required to improve cell growth and xylose consumption. Even still, these levels are not yet comparable with strains expressing an oxidoreductase pathway [2, 100, 103].

Beyond the assembly of xylose catabolic pathways, xylose isomerase is an important enzyme for the food industry, especially in the production of high-fructose corn syrup. For these applications, xylose isomerase has been extensively studied [104] to improve thermal stability [105, 106], pH optimum [107], and substrate preference [108]. However, these studies were mainly focused on obtaining a xylose isomerase that (i) has different optimum temperature and pH range, 60 – 80 °C and pH 7.0 - 9.0, respectively [109], than those for conventional ethanol fermentation, (ii) are expressed in *E. coli* rather than *S. cerevisiae* [78], and (iii) are found to be unsuccessfully expressed [77] or to be inactive at mesophilic temperature [110] in *S. cerevisiae* mainly due to protein misfolding [31]. Moreover, later attempts to improve the xylose isomerases for ethanol fermentation, such as cold-adaptation [111] and optimizing expression level [112], were unsatisfactory for use in constructing a functional xylose catabolic pathway in yeast.

Here, we report the first directed evolution study of a xylose isomerase gene (*xylA* from *Piromyces sp*.) for improved specific enzyme activity at the conditions tested, cell growth, xylose consumption rate, and ethanol production in the yeast *S. cerevisiae*. Directed evolution is an efficient approach for tailoring proteins that require refined functions such as higher stability, tolerance, substrate specificity, and product selectivity of protein. The iterative application of this method allows for proteins with significantly improved function to be easily obtained in a short period of time [79]. To this end, we subjected the *xylA* gene to iterative rounds of random mutagenesis (aided by error-prone PCR) followed by selection for increased cell growth on xylose as a sole carbon source. After three rounds of mutagenesis and selection, we obtained an improved mutant of

xylose isomerase that can offer a promising starting point for further strain engineering to improve xylose catabolism.

2.3 Results and Discussion

2.3.1 Identification of a xylose isomerase mutant conferring increased cell growth on xylose

In this study, we sought to utilize directed evolution to improve the poor growth and xylose consumption rates usually associated with xylose isomerase-based pathways in heterologous *S. cerevisiae*. First, the xylose isomerase gene (*xylA*) from *Piromyces sp.* was synthesized as a codon-optimized version for *S. cerevisiae* to maximize translational efficiency. This optimal version was then used as a template for random mutagenesis afforded by error-prone PCR using the GeneMorph II Random Mutagenesis kit. Specifically, a library size of 1×10^5 members (as measured by independent *E. coli* colonies post-transformation) was created using a range of mutation rates. This library was cloned into a yeast expression vector behind the strong GPD promoter for high level expression.

To create a suitable host strain for selection and testing of this xylose isomerase library, the standard haploid yeast BY4741 was chosen as a background. Two additional genetics changes were selected for this host strain. First, a *gre3* knockout strain of BY4741 was chosen as Gre3p is an aldose reductase that has been previously shown to inhibit xylose isomerase by nonspecifically producing xylitol from xylose [113]. Second, the overexpression of a heterologous transaldolase, *tal1* from *S. stipitis* was chosen since this

protein has been previously been reported to be beneficial for xylose utilization in *S. cerevisiae* [45]. The overexpression of downstream pathway enzymes to insure proper flux is advantageous for selection phenotypes [97]. To achieve overexpression of *tal1*, this gene was cloned behind a TEF promoter in the p415 plasmid. The resulting strain containing both the *gre3* knockout and the heterologous *tal1* expression was named BY4741-S1. Finally, the native chromosomal copy of *S. cerevisiae* xylulokinase (*XKS1*) was maintained as opposed to overexpressing a heterologous copy from *S. stipitis*. Previous work has shown that strong overexpression of heterologous xylulokinase can be toxic to cells, especially with non-optimal xylose catabolic pathways, thus we sought to initially avoid this improper balance of enzyme levels in this strain [41].

The mutant xylose isomerase library was transformed into *S. cerevisiae* BY4741-S1 and screening/selection was conducted on the basis of growth rate advantage. To this end, pools of cells were serially subcultured in xylose media for between 5 and 7 serial transfers. Following this selection, many mutant strains were isolated and tested for growth (a sampling of 20 from each round provided in **Figure 2.1**). The mutant conferring the most improved growth phenotype was subjected subsequent rounds of mutation and selections. This process was repeated for a total of three rounds of mutagenesis and selection. At each step, improved cell growth phenotypes of selected mutants were confirmed through retransformation of the xylose isomerase mutant gene into a fresh host strain to exclude any adaptive changes in the genomic DNA of isolated strains. It should be noted that cell adaptation can provide an additional means of improving the xylose utilizing cells created here. For example, the isolated strain after

third round of mutation showed almost double the growth rate compared to the retransformed strain. Thus, the ultimate production strain will utilize the synergy between pathway engineering and adaptive laboratory evolution.

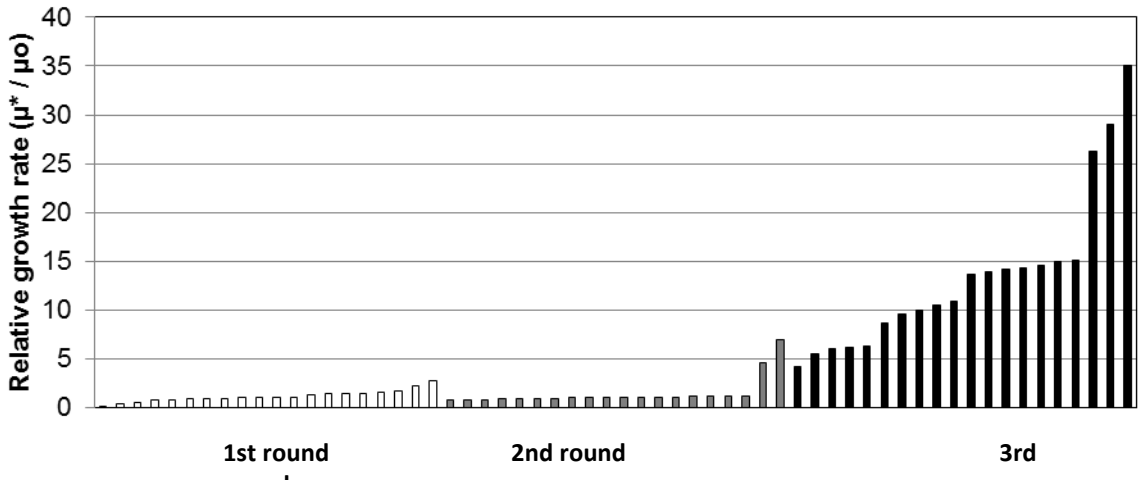


Figure 2.1: Comparison of relative growth rate of isolated mutants (selected). The growth rate of strains isolated during the first (white), second (grey), and third (black) rounds of mutagenesis (μ*) were compared to those of the wild type (μ^o). Twenty representative strains are shown out of the 40 tested for each round. The average growth rate of these isolated mutants increased progressively with the round of mutagenesis. The isolated strain with the highest growth rate in each round was selected.

Table 2.1: Comparison of literature based growth and fermentation rates with xylose as a sole carbon source. Fermentation performance of *S. cerevisiae* BY4741-S1A, S1A1, S1A2, S1A3 and S2A3 were compared with previously reported results. The strains developed here have 10 to 30 fold and 3 to 10 fold higher rates than previously reported engineered strains for xylose isomerase pathways and oxidoreducatase pathways, respectively. Legend: n.a.: not available.

strains	Strain description	Aerobic growth rate (h^{-1})	Xylose consumption rate ($g/g^{-1}\ h^{-1}$)	Ethanol production rate ($g/g^{-1}\ h^{-1}$)	Ethanol yield (g/g^{-1})	Reference
BY4741-S1A	*Piromyces* XI expressing strain overexpressing *tal1* and *Δgre3*	0.001	0.007	0.0032	0.46	This study
BY4741-S1A1	*Piromyces* XI mutant (*xylA*1*) strain overexpressing *tal1* and *Δgre3*	0.001	0.0082	0.0041	0.50	This study
BY4741-S1A2	*Piromyces* XI mutant (*xylA*2*) strain overexpressing *tal1* and *Δgre3*	0.003	0.0123	0.0060	0.49	This study
BY4741-S1A3	*Piromyces* XI mutant (*xylA*3*) strain overexpressing *tal1* and *Δgre3*	0.009	0.0126	0.0056	0.45	This study
BY4741-S2A3K	*Piromyces* XI mutant (*xylA*3*) strain overexpressing *tal1, XKS1* and Δ*gre3*	0.061	0.057	0.024	0.42	This study
CEN.PK-RWB 202	*Piromyces* XI expressing strain	0.005	n.a.	n.a.	n.a.	[33]

Table 2.1 (Continued)

strains	Strain description	Aerobic growth rate (h^{-1})	Xylose consumption rate ($g/g^{-1}\ h^{-1}$)	Ethanol production rate ($g/g^{-1}\ h^{-1}$)	Ethanol yield (g/g^{-1})	Reference
CEN.PK -RWB202-AFX	Evolved strain from CEN.PK- RWB 202 for 160days in xylose medium	0.18	0.340	0.140	0.42	[114]
CEN.PK- RWB217	*Piromyces* XI expressing strain overexpressing *XKS1, TKL1, RPE1, RKI1* and *Δgre3*	0.22	1.060	0.456	0.43	[83]
CEN.PK- TMB3066	*Piromyces* XI expressing strain overexpressing *XKS1, TKL1, RPE1, RKI1, HXT7* and *Δgre3*	0.02	0.005	0.002	0.43	[100]
BWY10Xyl	*Clostridium* XI expressing strain (evolved industrial strain in xylose medium)	0.04	0.07	0.03	0.43	[32]
CEN.PK2-1C -TMB 3424	XR/XDH expressing strain overexpressing *XKS1, GAL2, TKL1, RPE1, RKI1* and *Δgre3*	n.a.	0.06	0.007	0.12	[97]

The growth rates of the retransformed strains (S1A1, S1A2, and S1A3) expressing the best mutant xylose isomerase isolated from each round (*xylA*1, xylA*2, and xylA*3* from the first, second and third round, respectively) are shown in Table 2.1 compared to strains (S1A) harboring the wild type xylose isomerase gene (*xylA*). Growth on xylose is not observed in a control strain (S1C) harboring a blank plasmid as wild-type yeast cells lack a pathway to catabolize xylose. The growth rates on xylose progressively increased through the successive rounds of directed evolution (**Table 2.1**). The S1A3 imparted a 9 fold increase in growth rate compared to S1A.

2.3.2 Enzymatic assay of xylose isomerase mutants

As demonstrated above, these mutant *xylA* genes were able to increase the growth rate of yeast cells on xylose, thus we sought to understand the underlying mechanism of this improvement using *in vitro* kinetics assays. To do so, total protein was extracted from these strains and enzyme assays were conducted using a spectrophotometric-based coupled enzyme system [115]. Similar to the growth rates, the enzyme activities of xylose isomerase mutants increased progressively with the rounds of mutagenesis and selection (**Table 2.2**). The *xylA*3* had a 77% increase in V_{max} compared to *xylA* (0.094 µmol min^{-1} mg $protein^{-1}$ compared to 0.053 µmol min^{-1} mg $protein^{-1}$ for *xylA*3* and *xylA*, respectively). It should be noted that the V_{max} value measured here for *xylA* is consistent with previously reported values [32]. In contrast, the K_m values did not show the expected trend of decreasing K_m values, which refers to high substrate affinity leading high enzyme activity, with increased cell growth (the K_m for *xylA*3* is significantly higher

than the wild-type value). Nevertheless, these K_m values (averaging around 100 mM) are quite high, highlighting the root cause of difficulties with the isomerase system. Therefore, the improved cell growth seems to be derived mainly by increased V_{max} rather than decreased K_m of mutant xylose isomerase.

Table 2.2: Enzyme Kinetics and sequence analysis of *xylA* mutants. The wild type (*xylA*) and mutant xylose isomerase: first (*xylA*1*), second (*xylA*2*), and third (*xylA*3*) round mutant were expressed in *S. cerevisiae* and *in vitro* xylose isomerase activity was measured and reported. Sequence analysis reveals the mutations identified in each round. Error bars represent the standard deviation of biological triplicates.

	V_{max} (μM mg protein^{-1} min^{-1})	K_m (mM)	Mutations
xylA	0.053 ± 0.007	86.97 ± 6.56	-
*xylA*1*	0.064 ± 0.001	89.98 ± 3.93	E129D, V433I
*xylA*2*	0.083 ± 0.002	50.18 ± 7.08	E129D, V433I, E15D, E114G
*xylA*3*	0.094 ± 0.006	168.39 ± 11.98	E129D, V433I, E15D, E114G, T142S, A177T

2.3.3 Identification of critical mutations in isolated xylose isomerase mutants

In addition to enzyme analysis, these improved xylose isomerase genes were sequenced to identify the mutant residues responsible for improved activity. In each round of mutagenesis, two amino acid substitutions occurred which resulted in six mutations in the *xylA*3* third round mutant: E15D, E114G, T142S, E129D, A177T, and V433I (**Table 2.2**).

The cumulative amino acid substitutions obtained in this third round mutant were next investigated to identify which were necessary and sufficient for the improved performance of *xylA*3*. To accomplish this, single back-mutations of each amino acid substitution in the *xylA*3* were conducted to identify the essential nature of each of these six mutated sites. The strains expressing *xylA*3* with single back mutations of amino acid at the position 15 and 142, which were obtained during the 2^{nd} and 3^{rd} round respectively, showed 46% and 32% decrease in aerobic growth rates on xylose compared to the strain expressing *xylA*3*. This suggests that both E15D and T142S are necessary mutations that strongly contribute to the improved performance of mutant xylose isomerase (**Figure 2.2**). The remaining four mutations appeared to be neutral mutations as their back substitution did not have an impact on enzyme performance as measured by resulting growth rate. The beneficial effect the E15D and T142S mutations were confirmed by single and double point mutations of using *xylA* as a template. This test was conducted to determine whether these mutations either alone or in conjunction were sufficient to obtain the same phenotype as the isolated mutant. As shown in **Figure 2.2**, single substitution of either amino acid at the position 15 or 142 was not enough to confer improved activity, but substitutions at both position resulted in nearly the same performance as *xylA3** in aerobic growth rates. As a result, the improved performance of *xylA3** seems to be the result of synergistic mutations at position 15 and 142.

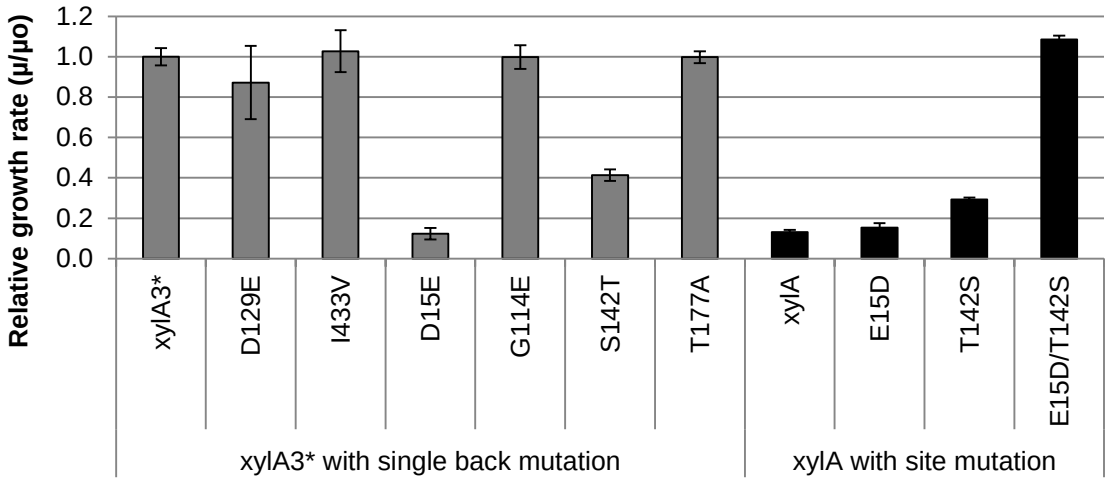


Figure 2.2: Identification of critical mutations by site-directed mutation. Aerobic cell growth rate (μ) of strains expressing the third round mutant with single back mutation (grey) and wild type with point mutation (black) were compared to that of the strain expressing the third round mutant xylose isomerase (μ^o). The back mutations at the amino acid position 15 and 142 resulted in 46% and 32% decrease in growth rates. Combining these two positions improved the performance of wild type xylose isomerase in aerobic growth on xylose. Error bars represent the standard deviation of biological triplicates.

2.3.4 Xylose isomerase mutants improve xylose consumption and ethanol production

We next evaluated the xylose consumption and ethanol production rates enabled by these xylose isomerase mutants using a series of high-cell density (OD = 20) batch fermentations in both oxygen-limited and micro-aerobic conditions. For the oxygen-limited conditions, 50 ml was cultured in a covered 125 ml flask. In these conditions, xylose consumption rates increased progressively through the rounds of directed evolution with S1A3 increasing xylose consumption rates by nearly 90% over the S1A (**Figure 2.3B**). As described above, the control strain (S1C) harboring a blank plasmid

does not grow and thus consumed no xylose. Interestingly, only S1A2 and S1A3 produced measurable ethanol levels in this condition (**Figure 2.3A**). The S1C, S1A, and S1A1 did not produce any ethanol (**Figure 2.3A**). Furthermore, S1A3 had a relatively short lag time for ethanol production whereas S1A2 started producing only after 95 hours of fermentation (**Figure 2.3A**). Thus, the identified mutant xylose isomerase (especially *xylA*3*) enables ethanol production capacity in these oxygen-limited conditions.

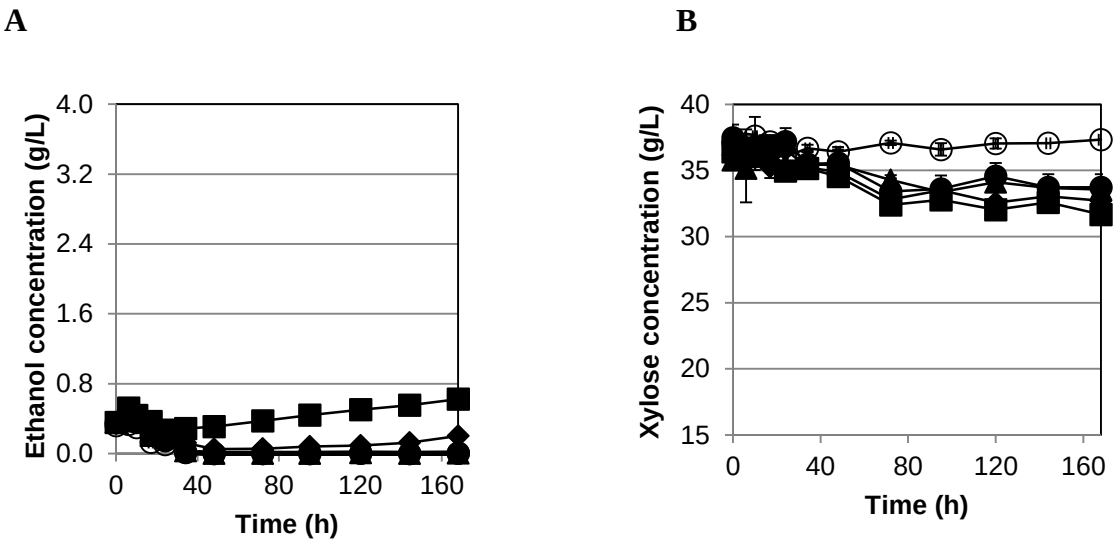


Figure 2.3: Oxygen-limited fermentation tests with *xylA* mutants. Ethanol production (A) and xylose consumption (B) profiles were measured in oxygen-limited conditions for *S. cerevisiae* BY4741-S1C (○), S1A (●), S1A1 (▲), S1A2 (◆), and S1A3 (■). Ethanol production was only present in the second and third round mutant. Xylose consumption was likewise increased in these strains. Error bars represent the standard deviation of biological triplicates.

As a second test for ethanol production, we utilized less aerobic conditions (specifically, a micro-aerobic condition) afforded by culturing 40 ml in a 50 ml sealed

vial. These conditions were expected to be more favorable for ethanol fermentations and likewise resulted in all strains producing ethanol at significantly increased rates (**Table 2.1**). Ethanol production capacity was greatly increased in strains harboring these mutant versions of xylose isomerase (**Figure 2.4A**) as was xylose utilization rates (**Figure 2.4B**). The S1A1 produced ethanol at the rate of 0.0041 g ethanol g cell^{-1} h^{-1} in this condition which represents an increase of 28% relative to that of S1A (0.0032 g ethanol g cell^{-1} h^{-1}). In addition, the ethanol yield in S1A1 (0.50 g ethanol / g xylose) was close to the theoretical value of 0.51 g/g. Subsequent mutants of xylose isomerase resulted in higher rates of ethanol production with an 88% and 75% increase over S1A for S1A2 and S1A3, respectively. It is interesting that in this more micro-aerobic condition, the performance of S1A2 and S1A3 were quite similar. Although S1A3 produced ethanol at a slightly lower rate (0.0056 g ethanol g cell^{-1} h^{-1}) than S1A2 (0.0060 g ethanol g cell^{-1} h^{-1}), the xylose consumption rate was slightly higher in S1A3 (0.0126 versus 0.0123 g xylose g cell^{-1} h^{-1}) (**Figure 2.4A and B**). Despite the increased rate, ethanol yields in these strains were slightly lower than S1A1 (0.49 and 0.45g ethanol /g xylose for S1A2 and S1A3 respectively). Collectively, these results along with the production profiles (**Figure 2.4**) suggest that downstream metabolic flux is limited in the cells harboring these improved xylose isomerase genes in these micro-aerobic conditions.

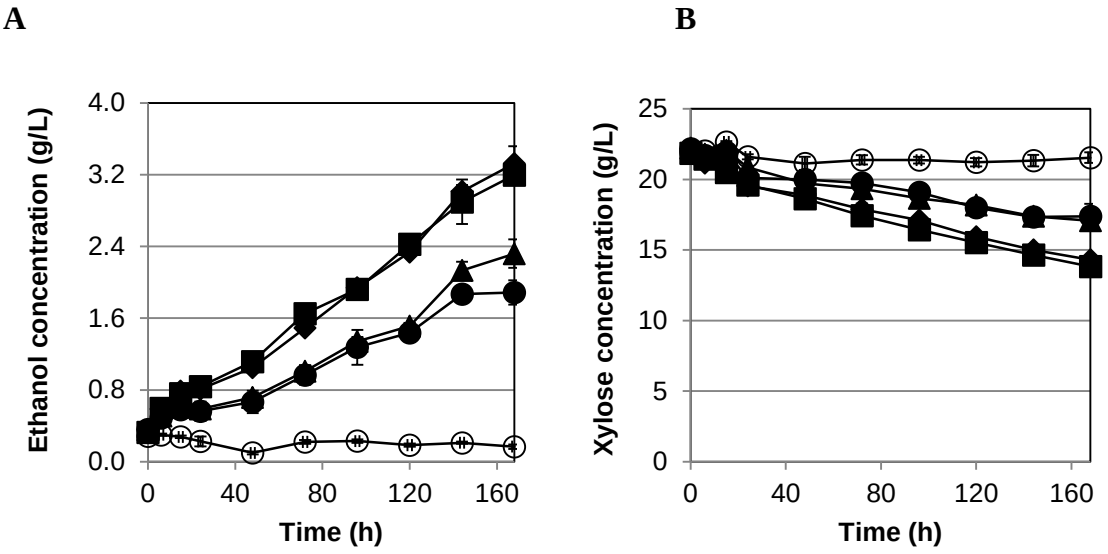


Figure 2.4: Micro-aerobic fermentation tests with *xylA* mutants. Ethanol production (A) and xylose consumption (B) profiles were measured in micro-aerobic conditions for *S. cerevisiae* BY4741-S1C (○), S1A (●), S1A1 (1), S1A2 (◆), and S1A3 (■). Ethanol production was increased over oxygen-limited conditions with both the second and third round having similar productivities. Total improvement in ethanol production was nearly 90% using these mutants. Xylose consumption was increased progressively with the round of directed evolution in these strains. Error bars represent the standard deviation of biological triplicates.

2.3.5 Improved performance of mutant xylose isomerase by additional engineering

The expression of the improved xylose isomerase in the minimally engineered strain of *S. cerevisiae* BY4741 still showed limited cell growth, xylose consumption, and ethanol production presumably due to insufficient downstream metabolic flux. To further improve the xylose fermentation performance, the xylose isomerase enzymes (mutant and wild-type) were cloned into a high copy plasmid p426-GPD [116] and co-transformed into

S. cerevisiae BY4741-S1 along with an additional plasmid expressing the xylulokinase *XKS1* driven by the TEF promoter [116]. The resulting strains were then named BY4741-S2AK, S2A2K, and S2A3K expressing *xylA*, *xylA*2*, and *xyl*3*, respectively. It should be noted that although these strains will have an expected higher downstream flux when compared to *S. cerevisiae* BY4741-S1, they are still relatively minimally engineered when compared with other strains commonly described in literature that contain many pentose phosphate pathway enzyme overexpressions (**Table 2.1**). Therefore, this experiment was intended to demonstrate the potential of the improved xylose isomerase enzyme as a promising starting point for further strain engineering.

As shown in **Figure 2.5**, the overexpression of xylulokinase significantly increased the aerobic growth rates in the strains expressing the various xylose isomerase enzymes (S2AK, S2A2K, and S2A3K expressing *xylA*, *xylA*2*, and *xylA*3*, respectively) compared to the strains without xylulokinase overexpression (S2A, S2A2, and S2A3 expressing *xylA*, *xylA*2*, and *xylA*3*, respectively). The improvement imparted by xylulokinase overexpression was most profound in strains harboring the mutant xylose isomerase variants. The aerobic growth rate of S2A3K (0.061 ± 0.00) was 20 fold higher than the strain without the xylulokinase overexpression and nearly 61 fold higher than a strain simply overexpressing wild-type xylose isomerase.

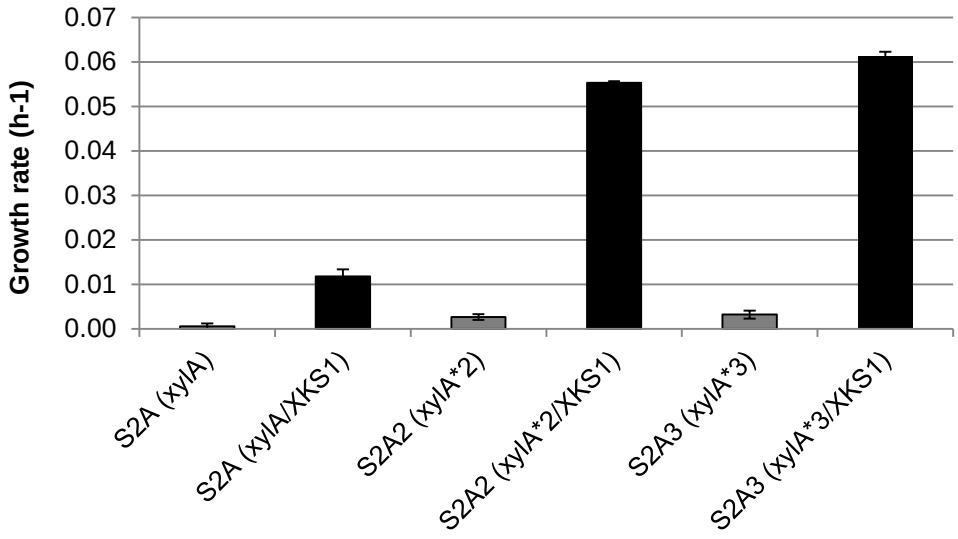


Figure 2.5: Aerobic growth rates of the strains expressing xylose isomerase with xylulokinase overexpression. Growth rates were measured for the strain expressing wild type (*xylA*), 2nd (*xylA*2*) and 3rd round (*xylA*3*) mutant xylose isomerase with xylulokinase overexpression (black) and compared to those without xylulokinase overexpression (grey). The improved growth rate of the strain expressing *xylA*3* (0.061 ± 0.001) was 5 fold higher than that of the strain expressing *xylA* (0.013 ± 0.001). Total improvement of the best strain compared with the strain simply expressing wild-type *xylA* is 61 fold. Error bars represent the standard deviation of biological triplicates.

Finally, we tested the performance of the strain expressing the 3rd round mutant xylose isomerase enzyme with xylulokinase overexpression (S2A3K) in a micro-aerobic xylose fermentation as described above. In these high-cell density batch fermentations, S2A3K produced ethanol at the rate of 0.024 g ethanol g cell^{-1} h^{-1} and consumed xylose a the rate of 0.057 g xylose g cell^{-1} h^{-1}, which represents an increase of nearly 8 folds relative to those of S1A (0.0032g ethanol g cell^{-1} h^{-1} and 0.007 g xylose g cell^{-1} h^{-1} S1A)

(**Figure 2.6** and **Table 2.1**). These production rates make this strain comparable to the rates obtained in strains harboring the traditional oxidoreductase pathway (**Table 2.1**). The ethanol yield in these fermentations was 0.42 g ethanol / g xylose which is significantly higher than the oxidoreductase pathway.

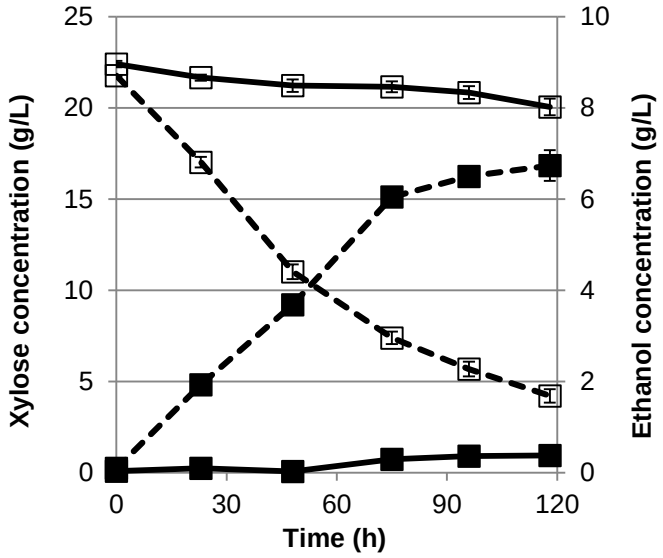


Figure 2.6: Micro-aerobic fermentation tests with xylulokinase overexpression. Ethanol production (closed square) and xylose consumption (open square) profiles were measured in micro-aerobic conditions for *S. cerevisiae* BY4741-S2A3 (solid line) and S2A3K (dotted line). Ethanol production was increased by xylulokinase overexpression. Total improvement in both ethanol production and xylose consumption rates were nearly 8 folds when the mutant xylose isomerase mutant was co-expressed with heterologous xylulokinase. Error bars represent the standard deviation of biological triplicates.

2.3.6 Discussion

This study demonstrates that a xylose isomerase pathway in *S. cerevisiae* can be improved by directed evolution. In doing so, the minimally engineered strain expressing the improved xylose isomerase outperformed the strain expressing wild type xylose isomerase in terms of ethanol production by nearly 90%, xylose consumption by 80%, and aerobic growth rate by 9 fold. These improvements were achieved by a 77% increase in *in vitro* catalytic activity for this enzyme. Combined with one additional downstream enzyme overexpression (xylulokinase), the strain expressing the mutant xylose isomerase exhibited a 61 fold increase in aerobic growth and an 8 fold increase in both ethanol production and xylose consumption rates. These improvements clearly show the potential of this mutant as a promising starting point for further strain engineering for efficient xylose fermentation.

The increased enzyme activity seems to be resulted from a series of identified mutations in both the active site and monomer-binding contacts. **Figure 2.7** shows three-dimensional protein structure modeling predicted based on the *Themotoga neapolitana* xylose isomerase [PDB code 1A0E [117]], which was available and shows high sequence similarity (65%) and identity (51%) to *xylA*, suggested that improved enzyme activity is likely due to two main factors: (i) increased substrate-enzyme interaction by mutations (Thr142 and Ala177) near the active site (His 102) and (ii) increased enzyme stability caused by mutations near the monomer-binding contacts (Glu15 and Val433). This theory is supported by the fact that T142S and E15D were both necessary for the improved

xylose isomerase performance of the *xylA*3* (**Figure 2.2**). It is of interest that mutations near the monomer-binding contacts occurred in the early rounds of mutagenesis and selection, prior to active site mutations. Previous reports suggest that xylose isomerase is only active when in the form of a dimer and tetramer [118]. Based on the evolutionary trajectory identified in this work, it seems that increased stability of xylose isomerase dimer/tetramer through mutations near the monomer-binding sites was a prerequisite for increased enzyme activity and active site mutations. The instability of dimerized xylose isomerase is one of the possible explanations for previously poor expression and activity of this enzyme in *S. cerevisiae* [109].

A **B**

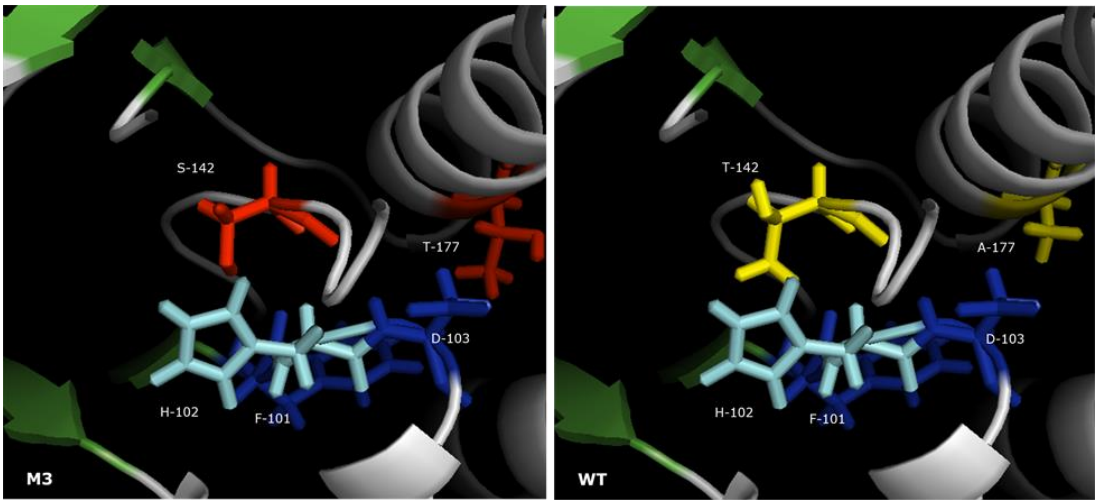


Figure 2.7: Structure predictions for xylose isomerase active site in wild-type and mutant (*xylA*3*). The smaller size of S142 in the xylose isomerase mutant could loosen structural inhibition. In addition, T177 in the xylose isomerase mutant seems to open the active site, thereby allowing xylose to more easily interact with this binding pocket. The mutations near active site of xylose isomerase in predicted three-dimensional structures of mutant (A) and wild type xylose isomerase (B). The mutation sites (S142 and T177) are shown in red and their wild type counterpart residues are shown in yellow. The F101, H102, and D103 residues, active site residues are shown in blue and sky-blue.

The performance of the identified xylose isomerase mutant was tested in two high-cell density fermentation assays— oxygen-limited and micro-aerobic conditions. In micro-aerobic conditions, the xylose consumption and ethanol production rates of S1A3 (0.0126 g xylose g cell^{-1} h^{-1} and 0.0056 g ethanol g cell^{-1} h^{-1}, respectively) were considerably higher than previously reported values for yeast cells harboring a xylose isomerase pathway (**Table 2.1**). As an example, when wild-type *Piromyces sp. xylA* was expressed in *S. cerevisiae* CEN.PK (TMB3066) background with *XKS1, TAL1, TKL1, RPE1, RKI1, HXT7* overexpression and *gre3* deletion, xylose consumption rate and ethanol production rates were reported as 0.005 g xylose g cell^{-1} h^{-1}and 0.002 g ethanol g cell^{-1} h^{-1}, respectively [100]. Thus, the results presented here—obtained strictly by overexpression of a mutant xylose isomerase along with *tal1* from *S. stipitis* and *gre3* deletion—are over 2 to 3 fold higher in these metrics. Moreover, ethanol production rate of S1A3 are also comparable to strains expressing an oxidoreductase pathway (**Table 2.1**). Most studies on ethanol fermentation with *S. cerevisiae* expressing oxidoreductase pathways have reported ethanol production rates in the range of 0.007 – 0.043 g ethanol g cell^{-1} h^{-1} in anaerobic condition, which depend on host strains, downstream gene overexpression, and reactor type [2, 97], though the highest reported ethanol production rate of 0.23 g ethanol g cell^{-1} h^{-1} was achieved by random mutagenesis of xylose reductase [97]. Based on the efforts reported here, the ethanol production rates of a xylose isomerase-based pathway are now comparable to early reports on the oxidoreductase pathway.

The improvements of xylose consumption and ethanol production of strain expressing mutant xylose isomerase are substantial especially since they were achieved

without the need for as extensive pathway and metabolic engineering in the host strain. This implies that xylose consumption rate and ethanol production rate are expected to be further increased by additional metabolic/evolutionary engineering. In previous research as shown in **Table 2.1**, xylose consumption and ethanol production rates of strains expressing wild type xylose isomerase from *Piromyces sp*. were improved by overexpressing downstream genes (3 fold increase in ethanol production rate) or evolutionary engineering (36 fold increased in aerobic growth rate) in a bioreactor. By overexpressing one additional enzyme, xylulokinase, we successfully demonstrated that the additional metabolic engineering would significantly improve xylose fermentation performances (**Table 2.1** and **Figure 2.5** and **6**). The S2A3K showed higher aerobic growth rate than the evolved-industrial strain expressing xylose isomerase from *Clostridium sp.* [32] and the strain expressing wild-type xylose isomerase from *Piromyces sp*. with extensive engineering of downstream genes and xylose transporter [83]. Moreover, the ethanol production rate of S2A3K (0.024 g ethanol g $cell^{-1}$ h^{-1}) was comparable to strains with an oxidoreductase pathway, which are usually in the range of 0.007 – 0.043 g ethanol g $cell^{-1}$ h^{-1} in anaerobic conditions [95, 97]. Given that the better performance was obtained without extensive strain engineering, it is expected that rates and yields could be further improved by the overexpression of other enzymes involved in pentose phosphate pathway or xylose transporters. Thus, the improved xylose isomerase reported here could serve as a critical starting point for further strain engineering to boost ethanol yields and productivity.

In oxygen limited conditions, the improvement of ethanol production is also important as this mutant pathway enabled ethanol production. *S. cerevisiae* strains S1A2 and S1A3 produced ethanol whereas S1A did not produce any ethanol. Moreover, xylose consumption rates in these conditions were increased by nearly 90% with S1A3. Ethanol fermentation is typically an anaerobic process that converts sugar to ethanol. Although *S. cerevisiae* can produce ethanol in aerobic condition by respiro-fermentation, ethanol is usually not the by-product in aerobic condition [119]. As a result, efficient industrial ethanol fermentors are mainly operated under anaerobic condition [120]; however, intermittent aerations are required to maintain *S. cerevisiae* viability and ethanol productivity during long period of fermentation. In the aerobic or oxygen-limited condition created by aeration, ethanol productivity decreases [120]. Hence, an idealized strain for ethanol fermentation will produce high concentrations of ethanol in anaerobic condition and is less vulnerable to lower yields under aerobic or oxygen-limited condition. To this end, the *S. cerevisiae* with improved xylose isomerase pathway reported here exhibited both of these properties with a rapid initiation of ethanol production under oxygen-limited condition and a strong ethanol production under more anaerobic conditions. Thus, this pathway could form the foundation of a promising host for industrial ethanol fermentation.

2.4. CONCLUDING REMARKS

This study presented a significantly improved xylose isomerase pathway in *S. cerevisiae* through the use of directed evolution without the need for extensive pathway

engineering. The xylose isomerase-based pathway is an advantage for yeast fermentation as it bypasses the imbalanced cofactor issues that plague the oxidoreductase pathway. The third round mutant created here has now alleviated xylose isomerase as the rate limiting step in this xylose catabolic pathway. Beyond this enzyme, it will be necessary to optimize the downstream genes involved in xylose metabolic pathway (downstream approach) and introduce efficient xylose transporter (upstream approach) to further improve xylose fermentation efficiency of engineered *S. cerevisiae* [84, 95]. In further strain engineering, it will be important to balance the upstream and downstream enzymatic activity since isomerization of xylose to xylulose is reversible and xylose formation is more favorable than xylulose formation at equilibrium by 80:20 [117]. Regardless, the xylose isomerase pathway has long-been considered an attractive alternative to the oxidoreductase pathway for expression in yeast. The results presented here demonstrate that this enzyme can be evolved for improved function in yeast leading to more efficient cell growth, xylose uptake rates, and ethanol fermentation rates—three phenotypes that were limited with the wild-type enzyme.

Chapter 3: Improving the xylose isomerase pathway-based catabolism

3.1 CHAPTER SUMMARY

Efficient xylose fermentation by yeast would improve the economical and sustainable nature of biofuels production from lignocellulosic biomass. However, the efficiency of xylose fermentation by the yeast *Saccharomyces cerevisiae* is suboptimal, especially in conversion yield, despite decades of research. Here, we present an improved performance of *S. cerevisiae* in xylose fermentation through systematic and evolutionary engineering approaches.

The engineering of *S. cerevisiae* harboring xylose isomerase-based pathway significantly improved the xylose fermentation performance without the need for intensive downstream pathway engineering. This strain contained two integrated copies of a mutant xylose isomerase, *gre3* and *pho13* deletion and *XKS1* and *S. stipitis tal1* overexpression. This strain was subjected to rapid adaptive evolution to yield the final, evolved strain (SXA-R2P-E) which could efficiently convert xylose to ethanol with a yield of 0.45 g ethanol / g xylose, the highest yield reported to date. The xylose consumption and ethanol production rates, 0.98 g xylose g cell^{-1} h^{-1} and 0.44 g ethanol g cell^{-1} h^{-1}, respectively, were also among the highest reported. During this process, the positive effect of a *pho13* deletion was identified for a xylose isomerase containing strain and resulted in up to 8.2 fold increase in aerobic growth rate on xylose. Moreover, these results demonstrated that low inoculum size and the cell transfer at exponential phase was found to the most effective adaptation strategy during a batch culture adaptation process.

These results suggest that the xylose isomerase pathway should be the pathway of choice for efficient xylose fermentation in *S. cerevisiae* as it can outperform strains with the oxidoreductase pathway in terms of yield and ethanol production and xylose consumption rates. Consequently, the strain developed in this study could significantly improve the prospect of biofuels production from lignocellulosic biomass.

3.2 INTRODUCTION

Lignocellulosic biomass can provide a sustainable feedstock for biofuels production. However, the inefficient fermentation of constituent pentose sugars such as xylose and arabinose limits the industrial scale conversion of lignocellulose by the yeast *Saccharomyces cerevisiae* [1, 84]. Several metabolic deficiencies in this yeast, including the lack of an endogenous pathway for xylose catabolism, require metabolic engineering. Many efforts have been reported that introduce heterologous xylose catabolism such as the oxidoreductase pathway from *Scheffersomyces stipitis* (encoded by *xyl1*, *xyl2*, and *xyl3*) and isomerase pathway from *Piromyces sp.*, (encoded by *xylA*) [2, 82]. While initial incorporation of these heterologous xylose catabolic pathways enable *S. cerevisiae* to convert xylose to ethanol, complementation is not enough and strains suffer from either low ethanol yield or productivity (or both) and thus require further improvement in xylose fermentation [84]. As a result, significant efforts have been made in modifying heterologous enzymes [35, 97, 98], optimizing metabolic flux through gene overexpression [41, 45, 83, 121] or deletion [85, 113, 122], evolving xylose utilizing strains by evolutionary engineering [81, 90, 91, 123], and identifying improved xylose transporter proteins [17, 61, 62].

As examples of additional strain engineering, the overexpression of xylulokinase [41] and downstream genes involved in pentose phosphate pathway [121], and the deletion of *gre3* [113] or *pho13* [85] genes have been shown to significantly improve xylose utilization rates. Individual enzyme modifications to alter cofactor preference of xylose reductase and xylitol dehydrogenase [97, 98] or improving enzyme activity of xylose isomerase by directed evolution [35] have improved heterologous xylose catabolism performance. Furthermore, whole-cell evolutionary engineering has been routinely applied to these rationally engineered strains to boost xylose fermentation efficiency [81, 90, 91]. Despite this array of attempts, the efficiency (especially yield) of xylose fermentation still remains suboptimal to achieve the goal of economical and sustainable biofuel production from lignocellulosic biomass, especially when compared with glucose conversion.

Until recently, the oxidoreductase pathway has been more intensely studied since consumption and growth rates with the xylose isomerase pathway have been extremely low [2]. However, the isomerase pathway does not require extensive cofactors as in the oxidoreductase pathway and thus has higher potential in terms of theoretical yield (0.51 g ethanol / g xylose) [35]. Toward this end, the experimental ethanol yields reported for the oxidoreductase pathway range from 0.09 to 0.39 for optimized strains [95, 124] whereas reports as high as 0.43 g ethanol / g xylose can be found for the isomerase pathway [90, 100]. Therefore, there is considerable interest in improving a xylose isomerase-based pathway in *S. cerevisiae* with a particular focus on improving both xylose consumption rates and yields.

In chapter 2, subjected the xylose isomerase enzyme, *xylA* from *Piromyces sp.*, to directed evolution [35]. In doing so, a xylose isomerase mutant was identified that improved aerobic growth rate by 61 fold and both ethanol production and xylose consumption rates by 8 fold. These results were obtained with a minimally engineered strain of *S. cerevisiae* and achieved rates that were comparable with the oxidoreductase pathway. Here, we further improve xylose-isomerase based catabolism of xylose in *S. cerevisiae* by using a combination of rational and evolutionary engineering in a rapid fashion. As a rational engineering approach, we integrated two copies of the xylose isomerase mutant gene (*xylA3**) into genome of *S. cerevisiae* along with genomic overexpressions of native xylulokinase (*XKS1*) and heterologous transaldolase (*tal1*) gene from *S. stipitis*. We demonstrated that the deletion of *pho13* can significantly improve the cell growth on xylose using the isomerase pathway only in high flux strains. Following this rational engineering, we subjected this strain to a rapid adaptation and obtained an evolved strain with the highest ethanol yield from xylose reported to date coupled with the second highest ethanol production and xylose consumption rates reported.

3.3 Results and Discussion

3.3.1 Rational construction of xylose isomerase-based strains

In this study, we sought to develop a *S. cerevisiae* strain with improved xylose catabolic rates and yields using the xylose isomerase pathway. To this end, we first established a genomic integration of the xylose isomerase pathway in *S. cerevisiae* by

expressing a mutant xylose isomerase, *xylA3**, developed by our group [35]. In our prior work, episomal expression of *xylA3** combined with overexpression of the native *XKS1* and *tal1* from *S. stipitis* resulted in improved growth rates and xylose consumption rates [35]. To achieve efficient xylose utilization, we first integrated *xylA3** into the genome of a *S. cerevisiae* BY4741 *gre3* knockout strain overexpressing *XKS1*, which was named as SXA-R1. In order to achieve catabolic levels commensurate with our previous plasmid-borne strain, an additional copy of *xylA3** and *tal1* from *S. stipitis* were integrated into SXA-R1 generating SXA-R2 (**Table 3.1**).

Table 3.1: *Saccharomyces cerevisiae* strains used in this study

strains	characteristics
BY4741 Δ*gre3*	*Mat a, his3Δ1, leu2Δ0, met15Δ0, ura3Δ0, YHR104w::kanMX4*
BY4741 Δ*pho13*	*Mat a, his3Δ1, leu2Δ0, met15Δ0, ura3Δ0, YDL236w::kanMX4,*
SXA-G	BY4741 Δ*gre3, URA::GPDp-xylA*3-CYC1t-TEFp- XKS1-CYC1t*, p413*TEF-TKL1-CYC1-TEF-tal1-CYC1*
SXA-P	BY4741 Δ*pho13,* p416*GPDp-xylA*3-CYC1t-TEFp- XKS1-CYC1t*, p413*TEF-TKL1-CYC1-TEF-tal1-CYC1*
SXA-GL	SXA-G, p415*GPDp-CYC1t*
SXA-GM	SXA-G, p415*GPDp-xylA3* -CYC1t*
SXA-GH	SXA-G, p425*GPDp-xylA3* -CYC1t*
SXA-PL	SXA-P, p415*GPDp-CYC1t*
SXA-PM	SXA-P, p415*GPDp-xylA3* -CYC1t*
SXA-PH	SXA-P, p425*GPDp-xylA3* -CYC1t*
SXA-R1	BY4741 Δ*gre3, URA::GPDp-xylA*3-CYC1t-TEFp- XKS1-CYC1t*
SXA-R2	BY4741 Δ*gre3, URA::GPDp-xylA*3-CYC1t-TEFp- XKS1-CYC1t, Leu:: GPDp-xylA*3-RPM1t-TEFp- tal1-CYC1t*

Table 3.1(Continued)

strains	characteristics
SXA-R1P	SXA-R1, *YDL236w :: His*
SXA-R2P	SXA-R2, *YDL236w:: His*
SXA-R2P-E	Evolved strain of SXA-R2P

3.3.2 The effect of *pho13* deletion with a xylose isomerase pathway

In an effort to further improve xylose consumption rates, we evaluated the impact of a *pho13* deletion on the xylose isomerase pathway. The *pho13* deletion has been reported to improve xylose fermentation and tolerance to toxic chemical in lignocellulosic hydrolysate [85, 122]. However, the improvement conferred by this deletion was mostly seen in strains harboring an oxidoreductase pathway with only marginally exhibited improvement with xylose isomerase based strains [85]. We hypothesized that the improvement of a *pho13* deletion could have been muted in prior xylose isomerase strains due to low catabolic rates. Therefore, we investigated the effect of deleting *pho13* in strains harboring the mutant xylose isomerase pathways, which has higher xylose catabolic rates compared to the wild-type xylose isomerase [35].

Here, the effect of *pho13* deletion was investigated in strains expressing different levels of xylose isomerase (thus conferring different levels of cell growth). To modulate the expression levels of xylose isomerase, we expressed single and double copies of *xylA3** in low copy plasmids and double copies of *xylA3** in a multi copy plasmids in the

pho13 knockout strain of *S. cerevisiae* BY4741 (**Figure 3.1**). As suspected, the impact of the *pho13* deletion was more pronounced in strains with higher expression of xylose isomerase (**Figure 3.2A**). As a comparison, the differences in cell growth for *gre3* knockout strains (a common deletion coupled with xylose isomerase pathways) expressing these same levels of *xylA3** was much more subtle (**Figure 3.2B**) compared to those in *pho13* knockout strains. This result suggests that the beneficial effect of the *pho13* deletion is only clear when xylose isomerase expression is sufficiently high. This result could explain prior reports of *pho13* deletion being insignificant in xylose-isomerase strains [92].

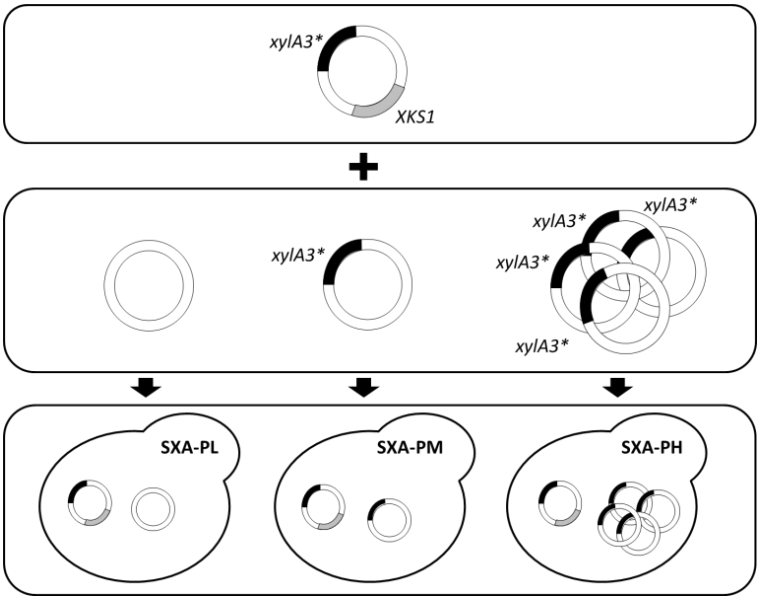


Figure 3.1: Schematic of the different *xylA3 expression plasmids utilized for *pho13* deletion evaluation.** The effect of *pho13* deletion was investigated in *S. cerevisiae* using plasmids to express a singular copy of *xylA3** (SXA-PL), double copy of *xylA3** in a low copy plasmid (SXA-PM), and a high copy plasmid (SXA-PH). These three arrangements enabled progressive increases in *xylA3** expression.

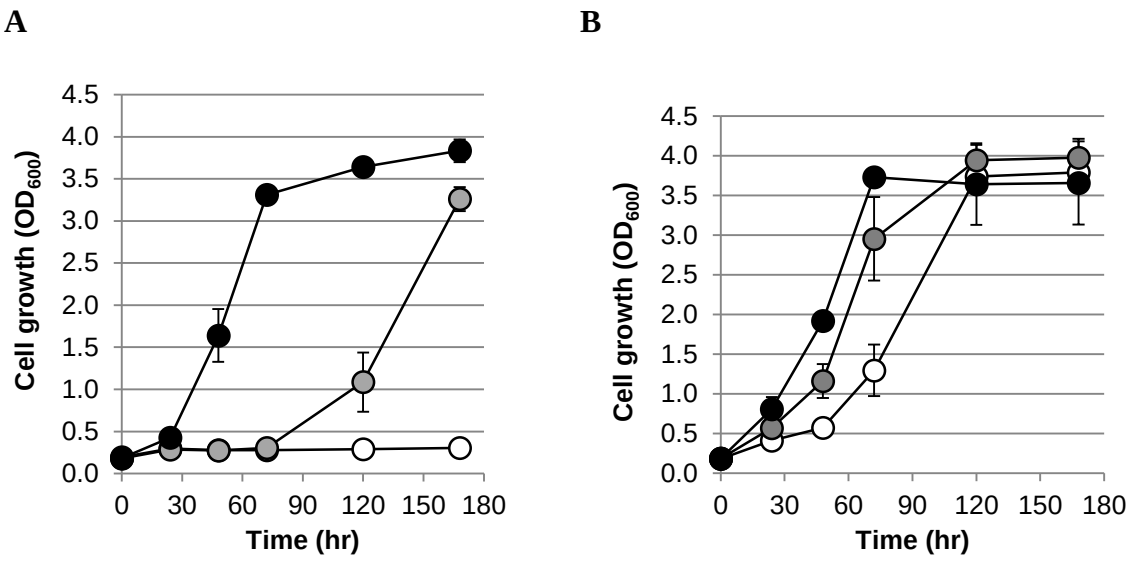


Figure 3.2: The impact of *pho13* and *gre3* deletions on xylose isomerase expressing strains. The cell grow of the knockout strains (Δ*pho13*: A, and Δ*gre3*: B) expressing different levels of the xylose isomerase gene (Figure 3.1) were compared: the strain expressing the singular copy plasmid of *xylA3** (white), the double copy plasmid of *xylA3** (grey) and a high copy plasmid (black). The beneficial effect of *pho13* deletion was clear only when xylose isomerase was sufficiently expressed. Error bars represent the standard deviation of biological triplicates.

Upon confirming the beneficial effect of *pho13* deletion in xylose isomerase strains, we deleted the *pho13* gene in the rationally engineered strains of SXA-R1 and SXA-R2 described above to make strains SXA-R1P and SXA-R2P respectively. By deleting *pho13*, both rationally engineered strains showed significant increases in cell growth on xylose (**Figure 3.3**). Specifically, the *pho13* deletion increased aerobic growth rates 2.5 fold and 8.2 fold in SXA1-R1P and SXA-R2P compared to their respective controls. These results further support the importance of coupling high xylose isomerase pathway flux with a *pho13* deletion.

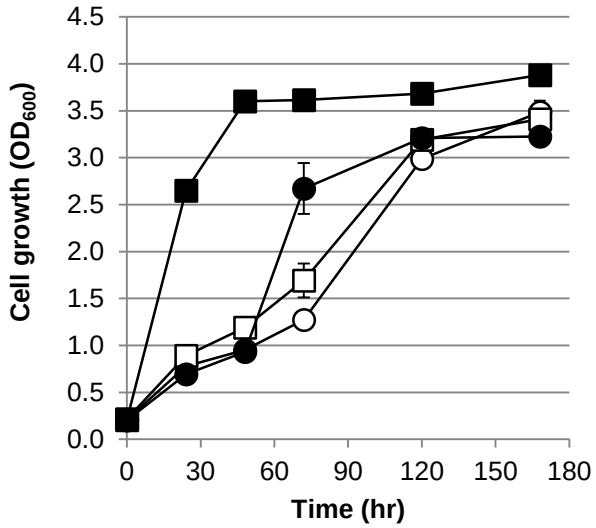


Figure 3.3: Rational engineering of *S. cerevisiae* expressing a xylose isomerase pathway. Rational strain engineering in a *gre3* background containing an integrated copy of *xylA*3* and *XKS1* overexpression (SXA-R1 strain) was evaluated. In addition, the expression of an additional copy of *xylA3** and *tal1* overexpression (square), and *pho13* deletion (black) gradually improved the cell growth on xylose. Strain identifications (SXA-R1: white circle, SXA-R2: white square, SXA-R1P: black circle, and SXA-R2P: black square) are described in the text. The expression of an additional copy of *xylA3** and *pho13* deletion showed synergistic effect on the cell growth on xylose. Error bars represent the standard deviation of biological triplicates.

3.3.3 Evolutionary engineering of the rationally engineered xylose utilizing strain

Upon developing the rationally engineered strain SXA-R2P containing two copies of *xylA3**, *gre3* and *pho13* deletions, and overexpression of XKS1 and *tal1*, we sought to further improve xylose catabolism in this strain by evolutionary engineering. Despite the effectiveness and simplicity of evolutionary engineering to improve desired phenotypes,

there are no standardized methods for an effective adaptation process. To this end, we sought to evolve the rationally engineered strain SXA-R2P using serial subculturing in xylose medium and evaluate the success rate using different inoculum sizes and different cell growth phases. The inoculum sizes used in this study were 0.5, 1, and 5% and cell transfers into fresh medium were conducted when the cell growth was at exponential and stationary phases.

During the rapid 24 days adaptation process investigated here, cultures were transferred 12 times in the case of the exponential phase transfer set and five times in the case of the stationary phase transfer set. Overall adaptation rate was determined on the basis of growth rate advantage. As adaptation progressed, the cell growth rate on xylose increased progressively for the exponential phase transfer set. Most specifically, the most significant increase in the cell growth was observed from the set using low inoculum size. In contrast, the cultures from the stationary phase transfer set showed no significant growth adaptation, and the differences among various inoculum sizes were negligible. This data suggests that the choice of cell growth phase plays a significant role in obtaining growth advantage in batch culture adaptation as well as defining success for an evolutionary engineering experiment. Following this adaptation process, 130 evolved strains from the exponential phase transfer set were isolated and tested for growth on xylose by using BioscreenC for high throughput analysis. The cell growth of 20 representative strains from each inoculum sizes were shown in **Figure 3.4**. The isolated strains from the pool with the low inoculum size, 0.5%, showed highest improvement in the growth rates on xylose compared to those from the medium and high inoculum sizes,

1% and 5%, respectively. As inoculum size increased, the efficiency in adaptation decreased, thus demonstrating the importance of this parameter in defining success for an evolutionary engineering experiment. Of these improved xylose utilizing strains, the isolate showing the highest growth on xylose was selected as the evolved strain and designated as SXA-R2P-E. When evaluated in 14ml culture tubes, SXA-R2P-E showed an aerobic growth rate of 0.128 h^{-1}, which was 22% higher than that of the initial strain, SXA-R2P, (0.105 h^{-1}) (**Figure 3.5**). This growth rate is among the highest for a xylose isomerase based strain reported.

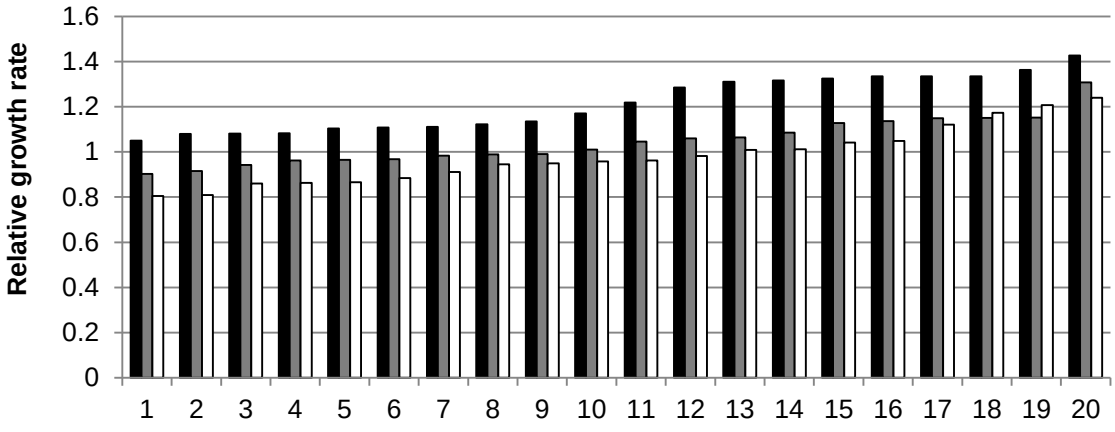


Figure 3.4: Adaptive evolution of the rationally engineered strain expressing the xylose isomerase pathway. The adaptive evolution of the rationally engineered strain (SXA-R2P) was performed in the exponential cell growth phase transfer set with different inoculum size of low (0.5%: black), medium (1%: grey), and high (5%: white). The graph shows representative strains with improved cell growth on xylose. The strains adapted in the culture with low inoculum size showed the highest improvement in the cell growth.

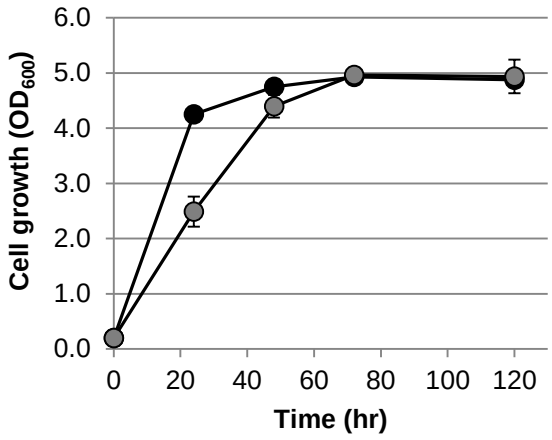


Figure 3.5: The aerobic cell growth of the evolutionary engineered strain of SXA-R2P-E on xylose. The evolutionary engineering improved the aerobic cell growth of the strain harboring a xylose isomerase-based pathway on xylose. The evolved strain of *SXA-R2P-E* (black) reached stationary growth phase faster than the rationally engineered strain of SXA-R2P (grey). Error bars represent the standard deviation of biological triplicates.

3.3.4 Improved xylose fermentation in the evolved strain SXA-R2P-E

To quantify the improvement of xylose fermentation performance of the evolved strain, SXA-R2P-E, we performed high-cell density (OD=20) batch fermentations in the micro-aerobic condition afforded by culturing 40 ml in a 50 ml sealed vial. Ethanol production capacity was greatly enhanced in SXA-R2P-E compared to the initial strain of SXA-R2P and the control strain which does not have a xylose catabolic pathway (**Figure 3.6A**). After 72 hours of fermentation, SXA-R2P-E produced nearly 8 g/L of ethanol from about 20g/L of xylose representing near complete consumption of xylose in this same timeframe (**Figure 3.6B**). The ethanol production and xylose consumption rates for

this evolved strain (0.054 ± 0.002 g ethanol g cell^{-1} h^{-1} and 0.141 ± 0.003 g xylose g cell^{-1} h^{-1}) were 3.9 fold and 4.3 fold higher than those of the starting strain, SXA-R2P (0.014 ± 0.002 g ethanol g cell^{-1} h^{-1} and 0.033 ± 0.003 g xylose g cell^{-1} h^{-1}). These increases were accompanied by an almost 20% increase in yield of ethanol approaching 0.39 g ethanol / g xylose for the evolved strain (**Table 3.2**). These results suggest that the adaptation process successfully improved the xylose catabolism in the rationally engineered strain and resulted in a strain with highly efficient xylose catabolism.

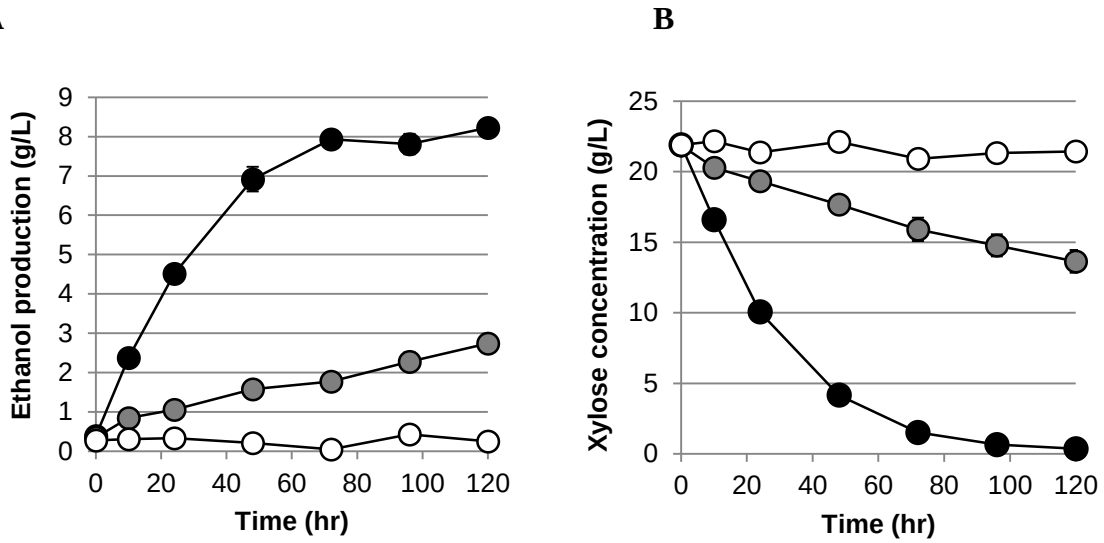


Figure 3.6: Micro-aerobic fermentation tests with the evolved strain. Ethanol production (A) and xylose consumption (B) profiles were measured in micro-aerobic conditions for wild type (white), the rationally engineered strain (grey), and evolved strain (black) of *S. cerevisiae*. Ethanol production and xylose consumption were significantly increased in the evolved strain. Total improvement in both ethanol production and xylose consumption rates were about 4 fold. Error bars represent the standard deviation of biological triplicates.

To further improve the xylose fermentation performance and mimic more industrial scale operation, we performed xylose fermentation in a bioreactor which offers more controlled fermentation condition. The bioreactor was operated with no inlet airflow and minimal stirring to create an anaerobic condition after cells consumed all oxygen initially dissolved in the medium (typically around 12 hours as monitored by a DO sensor). Ethanol production from xylose was significantly improved under these conditions along with the xylose consumption rate (**Figure 3.6**). The ethanol production rate of SXA-R2P-E in the bioreactor reached 0.44 ± 0.01 g ethanol g cell^{-1} h^{-1}, a value that is 8.1 fold higher than that obtained in the vial fermentation. The xylose consumption rate in the bioreactor was also increased by 7.0 fold (0.98 g xylose g cell^{-1} h^{-1} versus 0.14 g xylose g cell^{-1} h^{-1} in the vial fermentation). Finally, these anaerobic conditions and control resulted in a substantially increased ethanol yield of 0.45 g ethanol / g xylose, which is closed to the theoretical yield the theoretical value of 0.51 g/g [84]. Based on a survey of literature, the xylose fermentation performance of SXA-R2P-E is the highest in terms of ethanol yield and among the highest with respect to ethanol production and xylose consumption rates (**Table 3.2**).

Table 3.2: Comparison of representative, previously reported xylose fermentation performances. Fermentation performance of *S. cerevisiae* SXA-R2P-E was compared with previously reported results for representative isomerase and oxidoreductase pathways. Legend: n.a.: not available.

strains	Strain description	conditions	Xylose consumption rate (g/g^{-1} h^{-1})	Ethanol production rate (g/g^{-1} h^{-1})	Ethanol yield (g/g^{-1})	Reference
SXA-R2P-E	*xylA*3, tal1, XKS1, Δgre3, Δpho13* evolved (24 days)	Anaerobic Batch in bioreactor Synthetic medium (40g/L xylose)	0.98	0.44	0.45	This study
H131-A3-CS	*xylA, xyl3, TAL1,TKL1, RPE1, RKI1,* evolved (160 days)	Anaerobic Batch in bioreactor Synthetic medium (40g/L xylose)	0.94	0.40	0.43	[90]
H131-A3-ALCS	*xylA, xyl3, TAL1,TKL1, RPE1, RKI1,* evolved (160days)	Anaerobic Batch in bioreactor Synthetic medium (40g/L xylose)	1.87	0. 77	0.41	[90]

Table 3.2(Continued)

CEN.PK-RWB218	*xylA, XKS1, TKL1, RPE1, RKI1, Δgre3,* evolved (70days)	Anaerobic Batch in bioreactor Synthetic medium (20g/L xylose)	-	0.49 (projected)	0.41	[91]
SR8	Xyl1, xyl2, xyl3, *Δald6,* evolved (no detailed information)	Anaerobic Batch in flasks Synthetic medium (40g/L xylose)	0.65	0.25	0.39	[124]
CEN.PK2-1C -TMB 3424	Xyl1, xyl2, *XKS1, GAL2, TKL1, RPE1, RKI1, Δgre3* Selected (70days)	Anaerobic Batch in bioreactor Synthetic medium (60g/L xylose)	0.89	0.32	0.36	[97]
SR8	Xyl1, xyl2, xyl3, *Δald6,* evolved (no detailed information)	Oxygen-limited Batch in flasks complete medium (40g/L xylose)	0.87	0.28	0.31	[123]

3.3.5 Discussion

This study reports on the development of a superior strain for xylose conversion into ethanol on the basis of yield as well as a prominent strain with respect to rates. To accomplish this feat, we combined a minimal set of rational engineering targets along with short-term evolutionary engineering. The resulting ethanol yield (0.45 g ethanol / g xylose) was the highest reported to date, and ethanol production and xylose consumption rates were also among the highest reported.

Furthermore, we confirmed the beneficial effect of *pho13* deletion in a strain with a xylose isomerase pathway. In particular, we demonstrated that the *pho13* deletion was only beneficial when xylose isomerase was expressed at high enough levels. These results are in contrast to a prior report that investigated the effect of *pho13* deletion in both a strain expressing an isomerase pathway (wild-type *xylA* gene) as well as a strain harboring the oxidoreductase pathway [92]. The beneficial effect of the *pho13* deletion was clear in the strain expressing an oxidoreductase pathway, but not for the isomerase pathway strain. Due to the inefficiency of a wild-type isomerase pathway [35], it might have been hard to confirm the effect of *pho13* deletion in this background. Here, we clearly demonstrate the beneficial effect of *pho13* deletion with an improved xylose isomerase pathway, especially when xylose isomerase was highly expressed (**Figure 3.3**). The aerobic growth rate of SXA-R2P was 8.2 fold higher than that of SXA-R2, whereas SXA-R1P showed only 2.5 fold higher aerobic growth rate compared to SXA-R1.

A secondary aspect of this work is the comparison of conditions for evolutionary engineering. In particular, we confirmed that inoculum sizes and cell growth phases were important parameters influencing success of a rapid strain evolution project. Evolutionary engineering is a powerful approach to improve strains for a desired phenotype(s) [125]. For the case of biofuel production, evolutionary engineering has led to strains with improved tolerances to products and utilization of substrates [87, 90-92]. However, the detailed information about effective adaptation processes is not available, nor several conditions often compared. As a result, an adaptation process is usually conducted intuitively and takes a long period of time, from couple of months to several months. In this study, we evaluated the effectiveness of rapid batch culture evolutionary engineering with respect to the conditions of inoculum sizes (low, medium, and high) and cell growth phases (exponential and stationary phases). Of the tested conditions, the most effective combination was the low inoculum size (0.5%) and exponential growth phase transfer—an easily understandable combination for rapid growth selection. In the end, the isolated strain showed improved phenotype in a relatively short time frame of 24 days and resulting in 4.3 fold increase in xylose consumption rate and 3.9 fold increase in ethanol production rate compared to the initial strain. It should be mentioned that we utilized a very rapid and effective evolutionary engineering. In contrast, for alternative work to improve *S. cerevisiae* expressing a xylose isomerase pathway, Kuyper et al. [91] and Zhou et al. [90] spent about 70 and 160 days, respectively. Given the comparison between these strains (**Table 3.2**), it is evident that rapid evolutionary engineering combined with

rational strain engineering can be quite effective for developing strains with improved xylose utilization.

This study presented a significantly improved xylose catabolism in *S. cerevisiae*. To our knowledge, the highest ethanol productivity and xylose consumption rates were reported by Zhou et al. [90] using a combinatorially engineered strain of wild-type xylose isomerase expressing *S. cerevisiae* background with *xyl3, TAL1, TKL1, RPE1,* and *RKI1* overexpression (**Table 3.2**). Following the engineering of xylose catabolism, nutrients requirement of the strain by having auxotrophic markers were restored, and this led significant improvement in the xylose fermentation performance giving this strain the highest rates. Without the nutrient complementation, however, the performance of the combinatorially engineered strain itself (0.40 g ethanol g $cell^{-1}$ h^{-1} and 0.94 g xylose g $cell^{-1}$ h^{-1}, and 0.43 g ethanol / g xylose) falls behind the values reported in this study (0.44 g ethanol g $cell^{-1}$ h^{-1} and 0.98 g xylose g $cell^{-1}$ h^{-1}, and 0.45 g ethanol / g xylose). As a result, it can be presumed that optimizing culture condition or complementing nutrient requirement in SXA-R2P-E could enhance the xylose fermentation performance further. As an example, by changing the initial cell concentration and controlling to a different pH, SXA-R2P-E can achieve an even higher xylose consumption rate of 1.33 g xylose g $cell^{-1}$ h^{-1} (**Figure 3.7**). In addition, it is possible that further rational genetic changes including overexpression of downstream pentose phosphate pathway enzymes could further improve the ethanol production and xylose consumption rates.

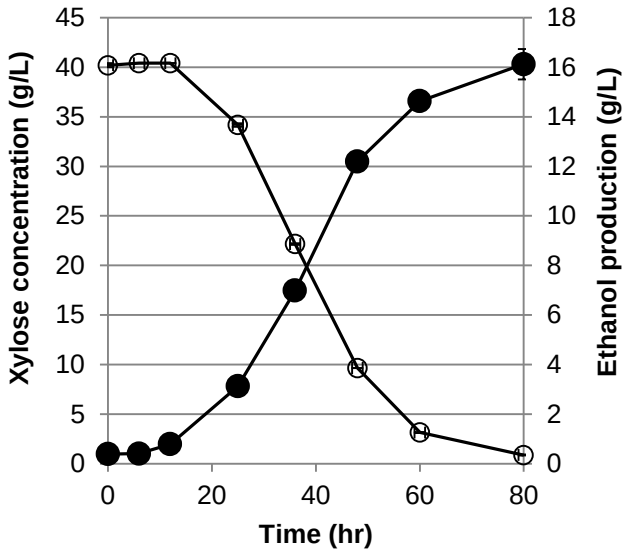


Figure 3.7: Anaerobic fermentation of xylose with the evolved strain with low initial OD. Ethanol production (black) and xylose consumption (white) profiles of the evolved strain were measured during anaerobic batch fermentation in a bioreactor. The evolved strain was inoculated at an initial OD of 1. Medium pH was maintained at 6.0 with 2.5 N NaOH. Error bars represent the standard deviation of technical duplicates.

Additional changes notwithstanding, the results achieved in this study clearly show that the xylose isomerase pathway has distinct functional and theoretical advantage over an oxidoreductase pathway for xylose fermentation in *S. cerevisiae*. The low ethanol productivity and xylose consumption rates that have been major challenges in the field are alleviated. In particular, the strain developed here outperforms evolved strains expressing oxidoreductase pathways in terms of ethanol productivity, xylose consumption rates, and yields (last three rows of **Table 3.2**). The result obtained in this study and representative data from previous similar studies (summarized in **Table 3.2**) show that the ethanol yields obtained with an isomerase pathway are in the range of 0.41

- 0.45 g ethanol / g xylose whereas those with an oxidoreductase pathway were in the range of 0.31 - 0.39 g ethanol / g xylose. The ethanol production and xylose consumption rates were also higher for isomerase pathway which were in the range of 0.40 – 0.77 g ethanol g cell^{-1} h^{-1} and 0.98 – 1.87 g xylose g cell^{-1} h^{-1}, respectively, compared to those for oxidoreductase pathway (0.25 – 0.32 g ethanol g cell^{-1} h^{-1} and 0.65 – 0.89 g xylose g cell^{-1} h^{-1}, respectively). Collectively, these results suggest that the xylose isomerase pathway should be the pathway of choice for efficient xylose fermentation in *S. cerevisiae*.

3.4. CONCLUDING REMARKS

This study presented a significantly improved xylose fermentation performance of *S. cerevisiae* expressing a xylose isomerase pathway along with minimal pathway engineering and evolutionary engineering. Without the need for extensive pathway engineering, the developed strain exhibits the highest ethanol yield (0.45 g/g), and the second highest ethanol production and xylose consumption rates ever reported. During this process, the positive effect of *pho13* deletion in a strain with the xylose isomerase pathway was clearly demonstrated. This experiment shows that the xylose isomerase-based pathway as developed here should be the pathway of choice over the oxidoreductase pathway for xylose utilization in *S. cerevisiae*.

Chapter 4: Blending bioprospecting with metabolic pathway engineering to improve xylose, arabinose, and xylan catabolism in *S. cerevisiae*

4.1 CHAPTER SUMMARY

Complete utilization of all available carbon sources in lignocellulosic biomass still remains a challenge in engineering *S. cerevisiae*. Here we demonstrate that a blended bioprospecting approach along with pathway engineering and directed evolution can be used to improve the catabolism of xylose, arabinose, and xylan in *S. cerevisiae*. Specifically, we perform whole genome sequencing based bioprospecting of a strain with remarkable pentose catabolic potential that we isolated and named *Ustilago bevomyces*. The expression of novel pentose catabolic genes for xylose, arabinose, and xylan was successful and enabled growth of each of these substrates in minimal medium by *S. cerevisiae*. The expression of alternative xylose catabolic genes improved cell growth by 85% compared to conventional heterologous genes from *S. stipitis*. Arabinose and xylan catabolic gene expressions enabled *S. cerevisiae* to grow on these sugars as a single carbon source in minimal medium. In addition, we suggest an alternative pathway for arabinose utilization in *S. cerevisiae* using a dual-functional reductase. Finally, we demonstrate that these pathways can be further improved by evolution.

4.2 INTRODUCTION

Bioconversion of lignocellulosic material is only sustainable and economical when all available carbon sources are completely utilized. However, the pentose sugars, which comprise a significant portion of lignocellulosic biomass, are poorly utilized by *Saccharomyces cerevisiae*. As a result, engineering pentose catabolism in yeast (esp. by importing pathways from native pentose utilizing organism) has been the subject of metabolic engineering for several decades [13, 35, 84, 97, 126-129]. Traditionally, only a handful of organisms have been sourced for pathways including, for the case of xylose catabolic genes, *Scheffersomyces stipitis* [10] and *Piromyces sp.*[33], and for the case of arabinose and xylan catabolic genes, *Aspergillus niger* [14], *Trichoderma reesei* [130], *Lactobacillus plantarum*[36], *E. coli,* and *Bacillus subtilis* [12]. After pathway complementation, pentose catabolism has been improved over using a variety of approaches [131] such as optimization of gene expression levels [41, 100], evolutionary engineering [123], directed evolution of pentose catabolic enzymes [35, 97], transporter engineering [62], and further bioprospecting searches for novel genes [130]. Even in light of these efforts, xylose fermentation remains much less efficient than glucose fermentation [2, 35, 132]. Moreover, the efficiencies of arabinose and xylan catabolic pathways are even worse resulting in insufficient cell growth, lower ethanol production, and inability to use minimal media formulations [2, 28, 36, 63, 84]. In particular, xylan utilization was possible only using surface display of xylan degrading enzymes and cultivating on complex medium [63]. The literature is still in disagreement over the actual proteins involved in an arabinose catabolic pathway [30]. Thus, more work still remains in the field to identify optimal pathways.

Despite difficulties in metabolically engineering *S. cerevisiae* for xylose utilization, many native organisms are innately adept at utilizing pentose sugars. Thus, bioprospecting, searching for novel genes from the strains with desired phenotypes in nature, remains a powerful method for importing desired cellular phenotypes [1]. Traditional bioprospecting relied on constructing cDNA libraries of an organism with desired phenotype and selecting for the responsible dominant allele [10, 133, 134]. However, recent bioinformatics advances and increases in the number of sequenced organisms have enabled homology-based bioprospecting [130, 135]. Even more so, the prospect of identifying an organism, conducting a whole genome sequence, and identifying novel pathways of interest has become more of a reality with the reduced cost of sequencing. Such an approach enabled the importation of succinate production pathways into *E. coli* through studying *Mannheimia succiniciproducens* [136].

Here, we demonstrate that a blended bioprospecting approach along with pathway engineering and directed evolution can be used to improve the catabolism of xylose, arabinose, and xylan in *S. cerevisiae*. In particular, we isolated a unique organism in our laboratory named *Ustilago bevomyces* that had superior pentose catabolism, conducted a whole genome sequence-based bioprospecting, imported novel pathways into *S. cerevisiae*, and improved them through directed evolution approaches. In doing so, we demonstrate that alternative pentose catabolic pathways can perform better than previously reported pathways. For the case of xylose, growth rates were improved up to 85% over a traditional *S. stipitis* pathway. We also identify novel sets of pathway

enzymes for both arabinose and xylan that enable growth with these substrates as the sole carbon sources in minimal medium.

4.3 Results and Discussion

4.3.1 Phenotypic and genotypic characteristics of *U. bevomyces*

During a culturing experiment with xylose as a sole carbon source, an excellent xylose utilizing organism was isolated. Remarkably, this isolated strain showed highly efficient xylose catabolism and even grow better on xylose than on glucose (**Figure 4.1A**). The growth of the isolated strain on xylose exceeded that of *S. stipitis*, the common source of xylose catabolic genes imported into *S. cerevisiae* (**Figure 4.1B**). In addition, the isolated strain showed growth on arabinose and xylan as sole carbon sources. These phenotypic traits suggested that this strain could be a promising source for catabolic genes to construct efficient pentose catabolism in *S. cerevisiae*.

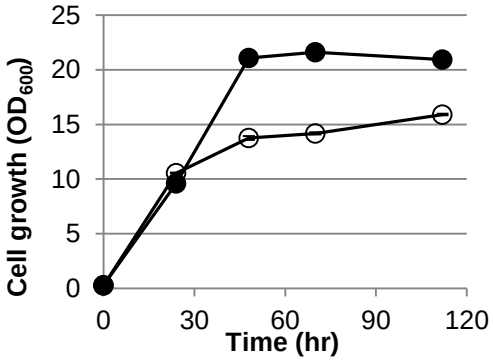
Initially, we sought to classify the isolated strain based on 18S rDNA. To do so, 18S rDNA was amplified from genomic DNA of the isolated strain using a fungus 18S rDNA specific primer set (EF4f and Fung5r) [137] resulting in a roughly 1kb fragment 18S rDNA fragment. Sanger sequencing of this fragment indicated 99% similarity with the plant pathogenic fungus *Ustilago maydis*. The morphology of the isolated strain was also similar to that of *U. maydis* in its yeast-like growth phase [138, 139] (**Figure 4.2**). As a result of the genomic similarity and location of isolation, we named this strain *Ustilago bevomyces*. Plant pathogens have recently received interest due to their potential to

harbor many plant cell-wall degrading enzymes [140]. To support this point further, a recent study reported that *U. maydis* could significantly enhance the performance of the *T. reesei* enzyme cocktail for the hydrolysis of lignocellulosic biomass [141]. Moreover, one of the hemicellulases, endo-1,4-beta-xylanase, from *U. maydis* was identified and confirmed to have activity on xylan [67]. Given the significant potential of *U. bevomyces*, we performed a whole genome sequence to aid the search for novel pentose catabolism pathways and genes.

The whole genome sequencing of *U. bevomyces* was conducted using Illumina technology and the obtained sequence reads were assembled by *de novo* genome assembler Velvet [142]. The N50 was 389,227 bp and the coverage of the assembly was over 10X. The full assembly size was 15.8Mbp, representing a genome which is roughly 80% the size of closely related *U. maydis* (**Table 4.1**) and similar to other yeast strains. Next, gene annotations were made using MAKER [143] and Blast2GO [144]. A structural genome annotation was carried out using the MAKER pipeline [143] to identify genomic elements. During this process, the transcripts of *U. maydis* were used as Expressed Sequence Tag (EST) evidence. After three rounds of annotation, 4,465 protein-coding genes were annotated with an average coding sequences (CDS) size of 665 bp and a maximum CDS of 5,423 bp. The average GC contents of the whole genome was 61%, consistent with the range for fungi in the basidiomycetes [145]. The structurally annotated protein-coding genes were then processed with Blast2GO [144] to identify Gene Ontology (GO) terms and conduct functional annotation. This resulted in a total of 3,367 protein-

coding genes which were annotated with GO terms and ultimately used as a pool to search for pentose catabolic genes.

A



B

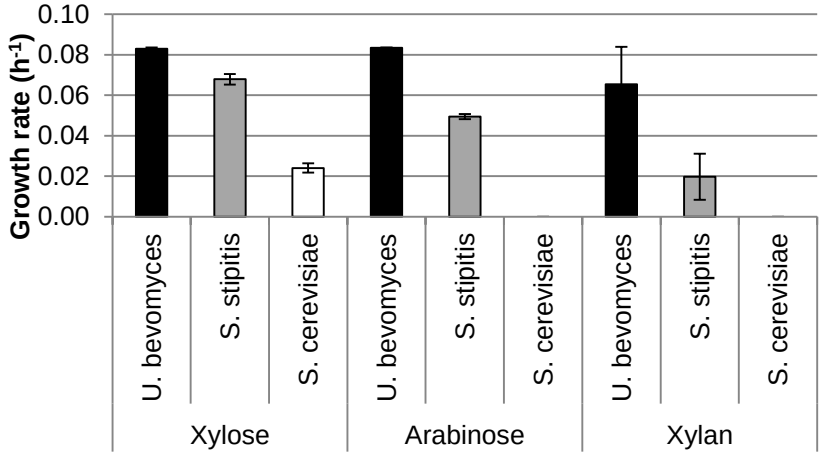


Figure 4.1: Growth phenotype of *U. bevomyces.* (A) Comparison of cell growth of *U. bevomyces* on glucose (open circle) and xylose (closed circle). (B) Comparison of cell growth of *U. bevomyces* on pentose sugars (monomers and polymers). Error bars represent the standard deviation of biological duplicates (A) or triplicates (B).

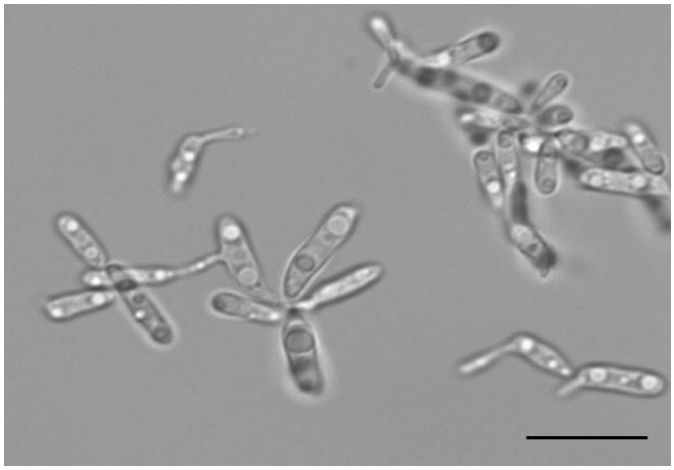


Figure 4.2: Morphology of *U. bevomyces.* Light microscopy image of *U. bevomyces.* Vacuoles are shown as circles inside cell. Scale bars: 10μm.

Table 4.1: General characteristic of several fungi genomes. General characteristic of *U. bevomyces* genome was compared to those of several fungi. *U. bevomyces: Ustilago bevomyces, S. stipitis: Scheffersomyces stipitis, S. cerevisiae: Saccharomyces cerevisiae, D. hansenii: Debaryomyces hansenii, Y. lipolytica: Yarrowia lipolytica T. reesei: Trichoderma ressei, and U. maydis: Ustilago maydis*.

species	genome size(Mb)	Avg. G+C content (%)	Total CDS	Avg. gene density (%)	Avg. G+C in CDS (%)	Avg. CDS size (codons)	Maximum CDS size (codons)	References
U. bevomyces	15.8	61.0	4,465	56.7	63.0	665	5,423	This study
S. stipitis	15.4	41.1	5,841	55.9	42.7	493	4,980	[146]
S. cerevisiae	12.1	38.3	5,807	70.3	39.6	485	4,911	[146]
D. hansenii	12.2	36.3	6,906	79.2	37.5	389	4,190	[146]
Y. lipolytica	20.5	49.0	6,703	46.3	52.9	476	6,539	[146]
T. reesei	33.9	52.0	9,129	40.40	NA	NA	NA	[147]
U. maydis	19.68	54.0	6,522	NA	NA	NA	NA	[139]

4.3.2 Identifying putative pentose catabolic pathways in *U. bevomyces*

Using the functional annotation of *U. bevomyces* obtained after whole genome sequencing, homology based searches were used to identify potential xylose, arabinose, and xylan catabolic genes (**Table 4.2**). Similar to most fungi, *U. bevomyces* utilizes an oxidoreductase pathway for xylose catabolism involving a xylose reductase (XR), xylitol dehydrogenase (XDH), and xylulokinase (XK). Interestingly, *U. bevomyces* has eight putative xylose reductases, each exhibiting distinct expression level profiles when cells were grown on different pentose substrates (**Figure 4.3**). This observation suggests these putative xylose reductase may each have different substrate affinities and activities and function combinatorially to create multiple pentose catabolic modes within the cell. Similar to xylose catabolism, the arabinose catabolic pathway in *U. bevomyces* seemingly adopts an oxidoreductase pathway and has homologous genes of two additional enzymes of L-arabitol dehydrogenase (LAD) and L-xylulose reductase (LXR). For xylan utilization, *U. bevomyces* contains a mixture of hemicellulases including xylanases (XNA), L-arabinofuranosidase (ABF), D-glucuronidase (GD), and xylosidases (XD) (**Table 4.2**).

These identified pentose catabolic genes from *U. bevomyces* are all quite distinct phylogenetically from previously reported pentose catabolic genes obtained from other commonly sourced organisms. To demonstrate this, phylogenetic trees were constructed based on amino acid sequence alignment of concatenated pentose catabolic genes (i.e. pathways) from *U. bevomyces*, *S. stipitis*, *S. shehatae*, *T. ressei*, *Piromyces sp.*, *E. coli*, *A.*

kawachii, *A. niger*, *A. oryzae*, *Aureobasidium pullulans*, and *B. pumilus* (**Figure 4.4**). For each pentose pathway evaluated, *U. bevomyces* genes formed a unique and distant phylogenetic group implying that these enzymes are quite different from previously used pathways imported into *S. cerevisiae*. Thus, *U. bevomyces* not only exhibits unique growth physiology on its own right, the uniqueness of its pathways makes bioprospecting an intriguing option.

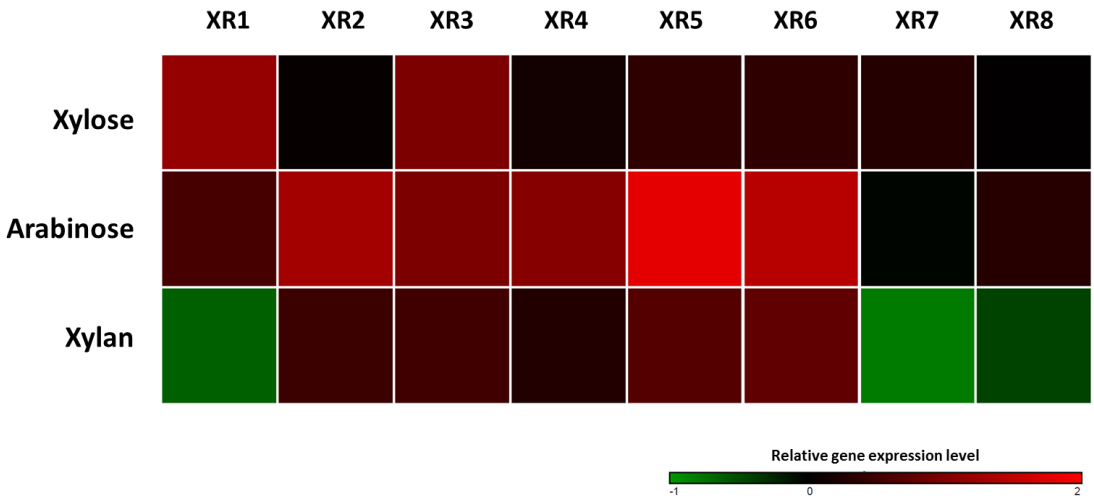


Figure 4.3: Expression of different xylose reductase in xylose, arabinose, and xylan cultures. The expression level of xylose reductases in xylose, arabinose, or xylan culture were measured by RT-qPCR. The relative expression levels were calculated based on the expression level in glucose culture. Error bars represent the standard deviation of biological triplicates.

Table 4.2: Identified pentose catabolic genes from *U. bevomyce*. Annotation carried out using MAKER and Blast2GO identified xylose, arabinose, and xylan catabolic genes from *U. vevomyces*. The sizes of the genes were ranged from 906 to 3219 bps. Mean similarity represents the average hsp (highest scoring pair)-similarity value for all the hits of a given sequence obtained from Blast2GO. *Mean similarity was not calculated by Blast2GO.

Pathway	Gene name	Gene description	Gene length (bp)	Mean similarity (%)
Xylose catabolic pathway	XR1	aldo-keto reductase	978	79.7
	XR2	aldo-keto reductase	1023	64.2
	XR3	aldo-keto reductase	906	63.3
	XR4	aldo-keto reductase	1431	59.1
	XR5	aldo-keto reductase	1410	59.6
	XR6	aldo-keto reductase	1113	61.9
	XR7	aldo-keto reductase	2052	59.1
	XR8	aldo-keto reductase	2142	*-
	XDH	xylitol dehydrogenase	1164	72.4
	XK	xylulokinase	1878	61.4
Arabinose catabolic pathway	LAD	L-arabitol dehydrogenase	1068	76.6
	LXR	L-xylulose reductase	825	86.1
Xylan catabolic pathway	XNA1	Xylanase	1026	79.4
	XNA2	Xylanase	672	73.9
	ABF	L-arabinofuranosidase	999	81.3
	GD	D-Glucuronidase	3219	59.3
	XD1	Xylosidase	1047	63.4
	XD2	Xylosidase/arabinosidase	1083	58.3

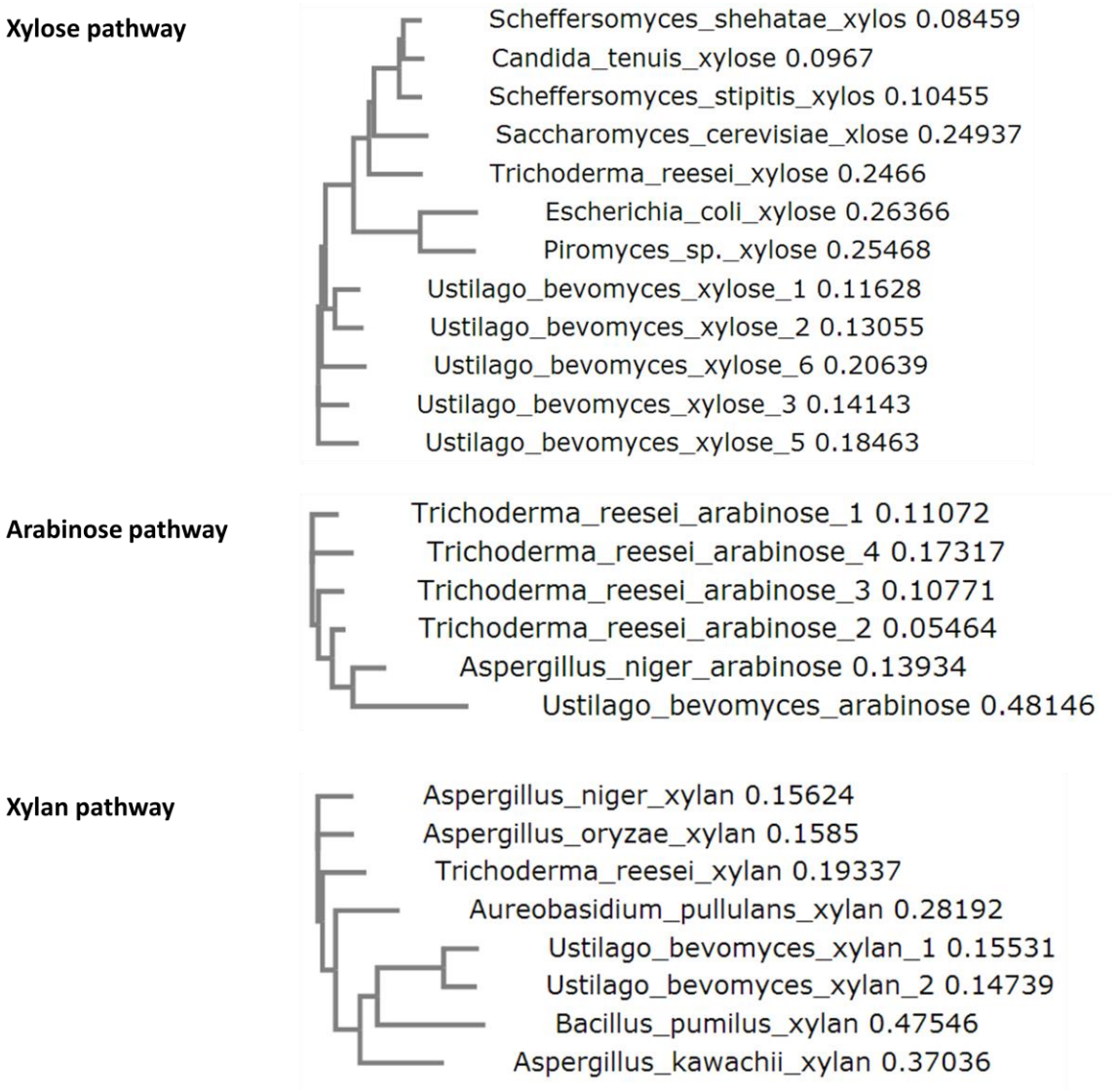


Figure 4.4: Phylogenetic trees of pentose catabolic genes from *U. bevomyces*. The phylogenetic trees were constructed from sequence alignment of the concatenated pentose catabolic genes (xylose pathway:XR, XDH, and XK, arabinose pathway: LAD and LXR, xylan pathway : XNA and XD) from *Ustilago bevomyces* and other organisms from which pentose catabolic genes have been imported. *Scheffersomyces stipitis*, *Scheffersomyces shehatae*, *Trichoderma ressei*, *Piromyces sp.*, *E. coli*, *Aspergillus kawachii*, *Aspergillus niger*, *Aspergillus oryzae*, *Aureobasidium pullulans*, and *Bacillus pumilus*. The number in each sequence stands for the different version of XR, LXR, and XNA. As example, Ustilago_bevomyces_xylose_1: the concatenated genes using XR1, XDH, and XK.

4.3.3 Importing pentose catabolic genes from *U. bevomyces* into *S. cerevisiae* to establish pathways

To evaluate the performance of these identified pathways, we sought to import these various genes into *S. cerevisiae*. To do so, genes were amplified from gDNA of *U. bevomyces* with specific primers (**Appendix B**), cloned into a standard yeast plasmid set [116] (**Table 4.3**), and transformed into *S. cerevisiae* BY4741 strains.

4.3.3.1 Xylose catabolic pathway

A heterologous xylose catabolic pathway was constructed in *S. cerevisiae* by importing various concatenations of a putative xylose reductase (XR) along with the xylitol dehydrogenase (XDH), and xylulokinase (XK) from *U. bevomyces*. Among the eight different xylose reductases from *U. bevomyces*, only five were selected for expression in *S. cerevisiae* xylose gene expression patterns (**Figure 4.3**). As a comparison, *xyl1* from *S. stipitis* was expressed with XDH and XK from *U. bevomyces*. Consistent with the gene expression profile (**Figure 4.3**), the strain expressing XR1 showed the highest growth rate on xylose, a value that was 52% higher than the strain expressing *xyl1* (**Figure 4.5A**).

To compare this *U. bevomyces* based pathway to the traditionally used *S. stipitis* pathway, we tested strains containing all possible permutations of *U. bevomyces* and *S. stiptitis* pathway enzymes. As a whole, expression of genes from *U. bevomyces* resulted in significantly higher xylose growth rates (**Figure 4.5B**). Moreover, the complete

pathway from *U. bevomyces* exhibited an 85% increase in growth rate relative to the strain expressing the complete pathway from *S. stipitis*.

Encouraged by the growth results, we next evaluated the xylose consumption and ethanol production enabled by *U. bevomyces* genes using a higher-cell density (OD = 15) batch fermentation. Unexpectedly, the ethanol production was slightly lower with *U. bevomyces* genes (2.7 g/L ethanol) relative to that with *S. stipitis* (3.0 g/L ethanol) and had a lower rate of xylose consumption (**Figure 4.6**). To understand the discrepancy between growth rate and fermentation capacity, the *in vitro* kinetics of XR1 were measured and compared with *xyl1* from *S. stipitis*. The NADPH-dependent enzyme activity (V_{max}) of XR1 (0.174 ± 0.044 μmol min^{-1} mg $protein^{-1}$) was similar to that of *xyl1* (0.189 ± 0.011 μmol min^{-1} mg $protein^{-1}$) (**Figure 4.7**). However, unlike *xyl1*, XR1 had a strict requirement for only NADPH coupled with a rather poor affinity for this cofactor (**Figure 4.7**). The cofactor imbalance in the oxidoreductase pathway caused by NADPH-preferred XR and NAD^+ -preferred XDH is a well-known challenge limiting ethanol production [84] and is thus the cause of this fermentation behavior. Nevertheless, the *U. bevomyces* based xylose pathway exhibits superior growth properties compared to commonly used pathways.

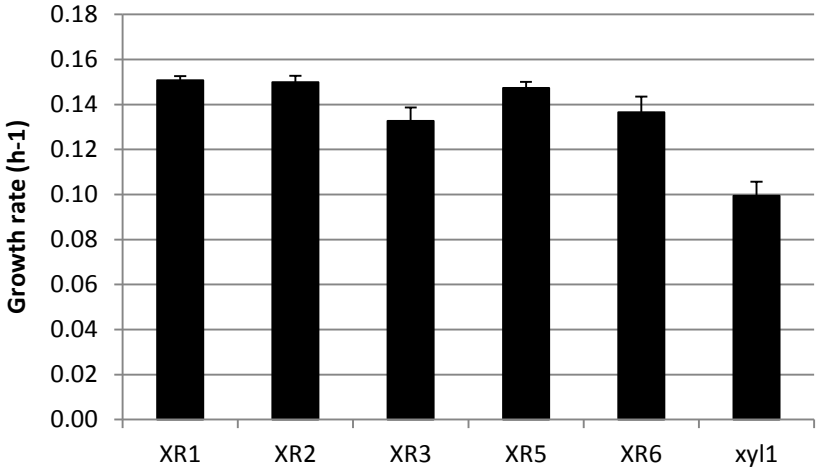
Table 4.3: Strains and plasmids used in this study

Strain/Plasmid	Descrption
U. bevomyces	Isolated strain from xylose culture
S. stipitis CBS 6054	Xylose fermenting yeast
S. cerevisiae BY4741_*trp5*	*Mat a; his3Δ1; leu2Δ0; met15Δ0; ura3Δ0; trp5; YHR104w::kanMX4*
S. cerevisiae YXS3[41]	*MATα; leu2::LEU2-TDH3P-PsXYL1-TDH3T; ura3::URA3-TDHP-PsXYL2-TDH3T; Ty3::G418R-PsXYL3; trp1-298 can1 cyn1 gal+*
P416-TEF-UbXR1	Putative xylose reductase 1 from *U. bevomyces*
P416-TEF-UbXR2	Putative xylose reductase 2 from *U. bevomyces*
P416-TEF-UbXR3	Putative xylose reductase 3 from *U. bevomyces*
P416-TEF-UbXR4	Putative xylose reductase 4 from *U. bevomyces*
P416-TEF-UbXR5	Putative xylose reductase 5 from *U. bevomyces*
P416-TEF-UbXR6	Putative xylose reductase 6 from *U. bevomyces*
P416-TEF-UbXR1-GPD-UbXDH	Putative xylose reductase 1 and xylitol dehydrogenase from *U. bevomyces*
P416-TEF-UbXR1-GPD-xyl2	Putative xylose reductase 1 from *U. bevomyces* and xylitol dehydrogenase from *S. stipitis*
P416-TEF-xyl1-GPD-UbXDH	Xylose reductase from *S. stipitis* and xylitol dehydrogenase from *U. bevomyces*

Table 4.3: Continued

P416-TEF-xyl1-GPD-xyl2	Xylose reductase and xylitol dehydrogenase from *S. stipitis*
P415-GPD-UbXDH-TEF-UbXK	Xylitol dehydrogenase and xylulokinase from *U. bevomyces*
P423-GPD-UbLAD	L-arabinitol dehydrogenase from *U. bevomyces*
P424-GPD-UbLXR	L-xylulose reductase from *U. bevomyces*
P423-GPD-UbXNA1-GPD-ABF	Xylanase 1 and arabinofuranosidase from *U. bevomyces*
P423-GPD-Ub XNA2-GPD-ABF	Xylanase 2 and arabinofuranosidase from *U. bevomyces*
P424-GPD-UbXD1	Xylosidase from *U. bevomyces*
P424-GPD-UbXD2	Xylosidase/arabinosidase from *U. bevomyces*

A



B

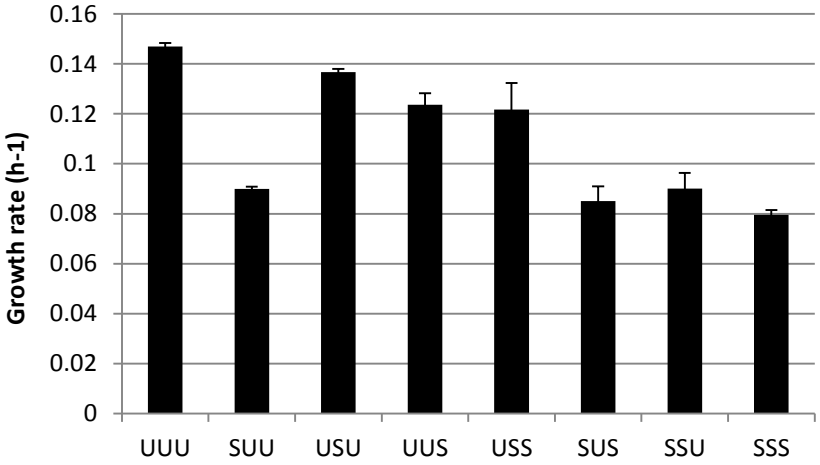


Figure 4.5: Growth rates of *S. cerevisiae* expressing xylose catabolic pathway genes from *U. bevomyces*. (A) Different putative XRs were expressed with XDH and XK from *U. bevomyces*. Xylose reductase gene (*xyl1*) from *S. stipitis* was also expressed as a control. (B) XR, XDH, and XK were expressed with different combinations from *U. bevomyces* and *S. stipitis*. Three letters in legend stands for the source organism for XR, XDH, and XK. UUU: all genes from *U. bevomyces*, USU: XR and XK from *U. bevomyces* and XDH from *S. stipitis*, SSS: all genes from *S. stipitis*. Error bars represent the standard deviation of biological triplicates**.**

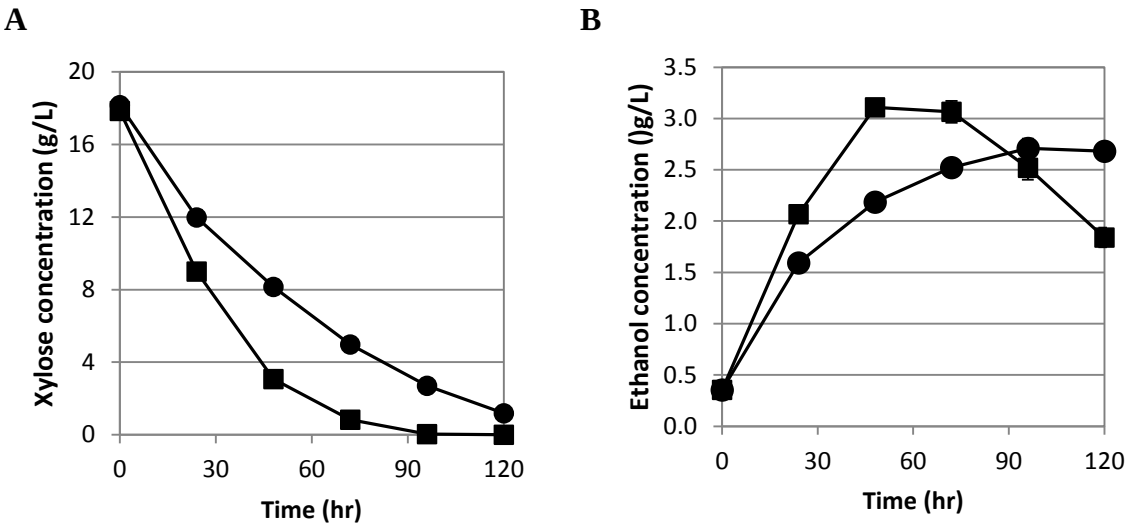


Figure 4.6: Ethanol fermentation of *S. cerevisiae* expressing xylose catabolic pathway genes from *U. bevomyces* and *S. stipitis*. Xylose consumption (A) and ethanol production (B) were measured during high cell density ethanol fermentation by *U. bevomyces* (●) and *S. stipitis* (■) in oxygen-limited condition. Error bars represent the standard deviation of biological triplicates.

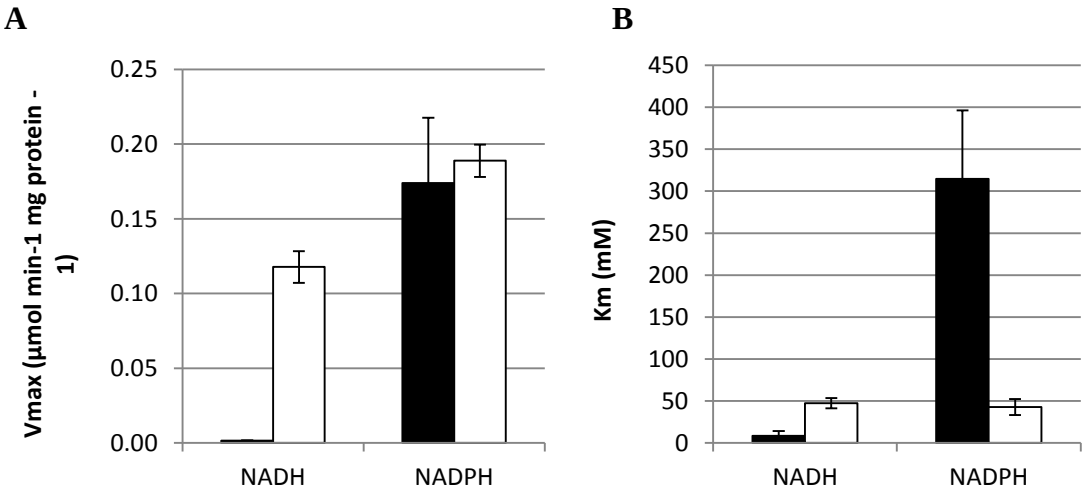


Figure 4.7: Enzyme activity of xylose reductase from *U. bevomyces* and *S. stipitis*. The in vitro enzyme activities of xylose reductase from S. cerevisiae expressing XR1 from *U. bevomyces* (black) and *xyl1* from *S. stipitis* (white) were measured spectrophotometrically in the presence of NADPH as a cofactor. Error bars represent the standard deviation of biological triplicates.

4.3.3.2 Arabinose catabolic pathways

An initial arabinose catabolic pathway was created by importing identified genes encoding XRs, LAD, LXR, XDH, and XK into *S. cerevisiae*. The heterologous expression of these five genes enabled *S. cerevisiae* to grow on arabinose as a single carbon source in rich medium but was not sufficient for obtaining robust growth in minimal medium suggesting that the cell growth might be resulted from nutrient sources in rich medium. This result is not surprising since there has been no report demonstrating the ability for *S. cerevisiae* to grow on arabinose as a single carbon source in minimal medium.

Interestingly, a few control strains expressing incomplete arabinose pathways (lacking the LXR) enabled cells to grow on arabinose. The LXR enzyme is attributed to the third step of a canonical arabinose oxidoreductase pathway to convert L-xylulose to xylitol (**Figure 4.8**). Since *U. bevomyces* has multiple xylose reductases and xylose reductases from *S. stipitis* and *T. reesei* have an enzyme activity for both D-xylose and L-arabinose [130], it is possible that a dually functional xylose reductase may exist. To confirm the dual function of a *U. bevomyces* XR as L-arabinose and L-xylulose reductase, we expressed XRs in *S. cerevisiae* complemented with XDH, XK, and LAD. Out of five different XRs tested, XR3 exhibited this dual functionality and enabled cells to grow on arabinose as a sole carbon source in both rich (**Figure 4.9**) and minimal medium formulations (**Figure 4.10**). A control strain is unable to grow on arabinose in minimal medium whereas a strain expressing XR3, XDH, XK and LAD from *U.*

bevomyces can grow at a rate of 0.011 ± 0.0033 h^{-1}. This suggests that XR3 is the preferred reductase for constructing arabinose catabolic pathway in the same way that XR1 was best for xylose. Although cell growth was not sufficiently high to perform a meaningful fermentation test, these results present the first documentation of recombinant *S. cerevisiae* on arabinose in minimal media.

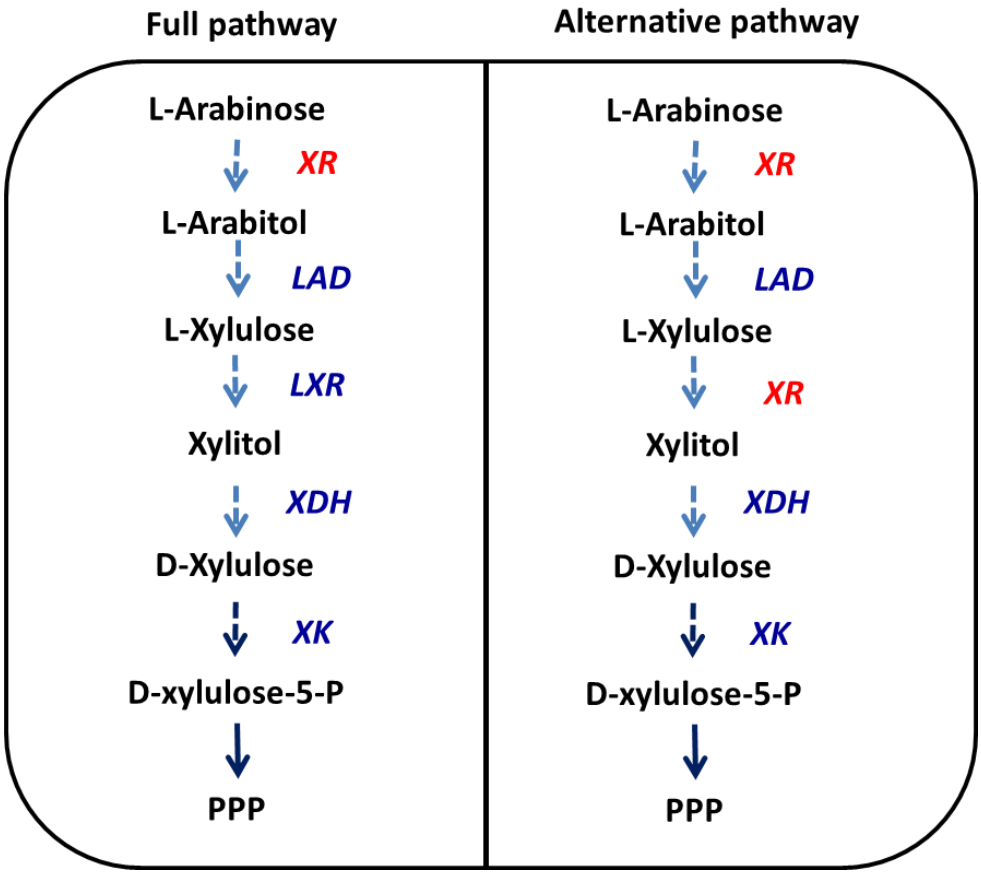


Figure 4.8: Conventional and alternative pathway for arabinose catabolism heterologously expressed in *S. cerevisiae.* Arabinose oxidoreductase pathway involves xylose reductase (XR), L-arabitol dehydrogenase (LAD), L-xylulose reductase (LXR), xylitol dehydrogenase (XDH), and xylulokinase (XK). In alternative pathway, XR functions as both XR and LXR for arabionose catabolism.

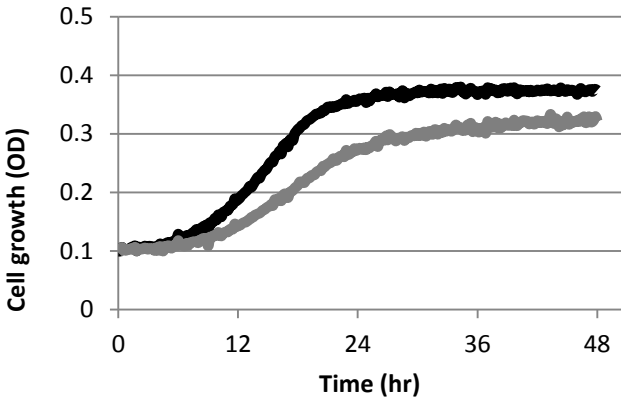


Figure 4.9: The cell growth of *S. cerevisiae* expressing alternative arabinose pathway. The cell growth of *S. cerevisiae* expressing XR3, LAD, XDH, and XK (black line) was compared to the control strain expressing empty plasmids (gray line) on arabinose as a single carbon source in rich medium (CSM + 1g/L of yeast extract and 2g/L of peptone). The minimal growth of the control strain seems to be resulted from nutritional source in the rich medium other than arabinose.

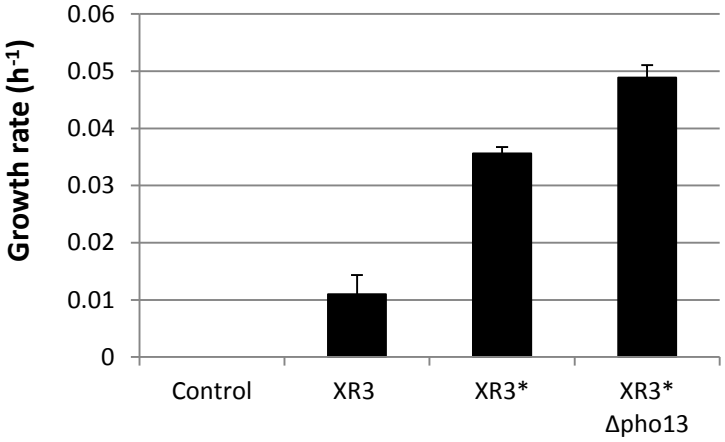
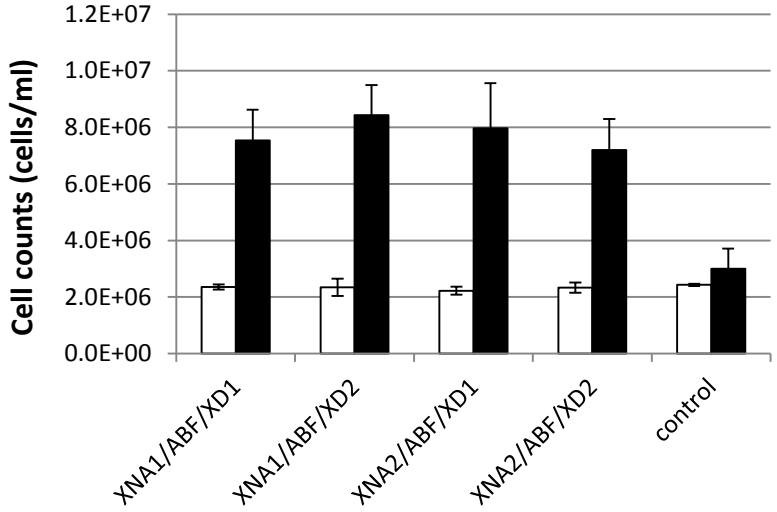


Figure 4.10: The cell growth of *S. cerevisiae* expressing the alternative pathway. The cell growth of *S. cerevisiae* expressing the alternative arabinose pathway was compared to the control strain harboring empty plasmids. Directed evolution of XR3 (XR3*) and *pho13* deletion (XR3* *Δpho13*) further improved the alternative pathway increasing growth rates. The growth test was performed on arabinose as a single carbon source in minimal medium. Error bars represent the standard deviation of biological triplicates.

4.3.3.3 Xylan catabolic pathways

Xylan hydrolysis is thought to require a complex mixture of hemicellulases including xylanase, xylosidase, and arabinofuranosidase. To create a simple, heterologous xylan catabolic pathway, we imported minimal combinations of hemicellulases into *S. cerevisiae* YSX3 strain [45], an engineered xylose utilizing strain that has integrated xylose catabolic pathway genes from *S. stipitis*. Strains expressing combinations of hemicellulases were grown on xylan-based minimal medium for five days and cell counts were compared with day zero (**Figure 4.11A**). In each of the combinations, the xylan catabolic pathways from *U. bevomyces* enabled growth on xylan whereas the control strain was unable to grow. We next evaluated ethanol production using a higher-cell density (OD = 20) batch fermentation. The expression of XNA1, ABF, and XD2 from *U. bevomyces* into the *S. cerevisiae* YSX3 strain enabled ethanol production of up to 0.26 ± 0.008 g/L using xylan as a single carbon source in minimal medium (**Figure 4.11C**). It should be noted that this cell growth and ethanol production from xylan was obtained by simply expressing only three xylan catabolic genes in *S. cerevisiae* without the need for any cell surface display unlike prior literature [63] and is the first report of growth and fermentation of xylan in minimal medium.

A



B

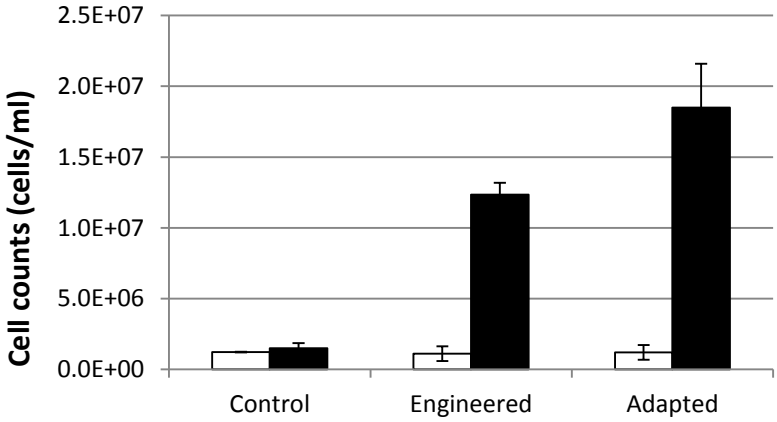


Figure 4.11: The performance of *S. cerevisiae* expressing xylan catabolic genes from *U. bevomyces*. Cell growth (A) of *S. cerevisiae* expressing a different set of xylan catabolic genes from *U. bevomyces* was measured before (white) and after five days (black) of cultivation. Xylan catabolic genes were expressed with minimal combinations which include xylanase (XNA1 or XNA2), xylosidase (XD1 or XD2), and arabinofuranosidase (ABF). Cell growth (B) and ethanol production (C) of the wild type, the engineered strain, and adapted strains of *S. cerevisiae* were compared after five days of fermentation. Error bars represent the standard deviation of biological triplicates.

C

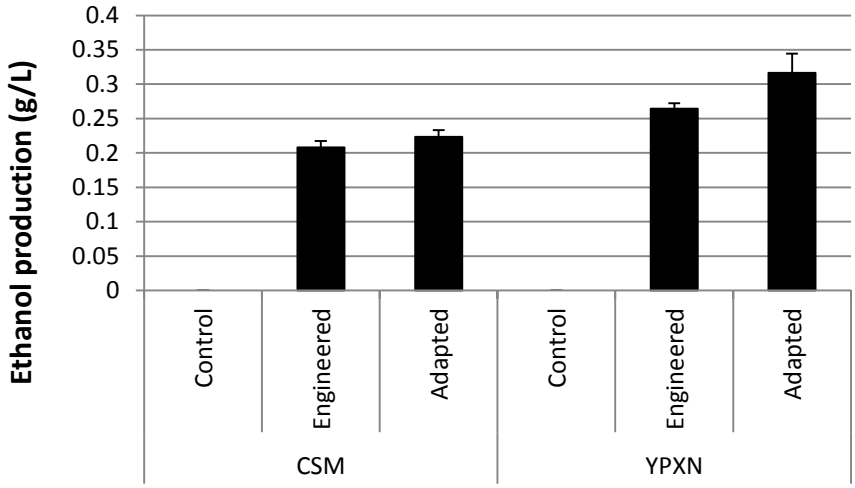


Figure 4.11(Continued)

4.3.4 Using evolution to improve the imported pentose catabolic pathways

Finally, we sought to blend bioprospecting with pathway engineering construction and evolution to create further optimized pathways. The simple, heterologous expression of these pathways from *U. bevomyces* without extensive downstream engineering provides a unique starting point for strain engineering. As a rapid test to improve these pathways further, we employed directed evolution of XR1, XR3, and LAD, and whole-cell adaptive / evolutionary engineering of xylan catabolism.

4.3.4.1 Evolution of the xylose catabolic pathway

To improve the fermentation capacity of the XR1-based xylose pathway, we sought to use directed evolution on XR1 with the hope of mitigating the strict cofactor

preference issues. To this end, a 1 x 10^5 library of XR1 was generated by error-prone PCR, transformed into *S. cerevisiae* BY4741, and screened on the basis of growth rate advantage in the micro-aerobic conditions of a sealed vial to increase the selection pressure against the cofactor imbalance. After three 3 serial transfers, hundreds of mutant were plated onto xylose plates supplemented by TTC solution to screen for ethanol production [148]. After testing many variants in test tubes, the most promising mutant, designated as XR1*, was retransformed into a fresh strain. The ethanol production from the strain expressing the XR1* based xylose pathway was improved about 20% compared to those with wild-type XR1, but still remained lower than the strain expressing the xylose catabolic genes from *S. stipitis*. These results suggest that the directed evolution of XR1 could partially resolve the issues of cofactor imbalance and further directed evolution of XR1 may result in even higher ethanol production. Nevertheless, these results demonstrate an improvement in the imported xylose catabolic pathway via directed evolution.

4.3.4.2 Evolution and strain engineering for the arabinose catabolic pathway

Expression of the bioprospected *U. bevomyces* XR3, XDH, XK and LAD enabled *S. cerevisiae* to grow on arabinose for the first time as a single carbon source in minimal medium. To improve this pathway performance further, we employed a similar directed evolution scheme for both XR3 and LAD. Specifically, 1 x 10^5 error-prone PCR based libraries were independently transformed into *S. cerevisiae* BY4741 with the remainder of the arabinose pathway and subcultured for growth on arabinose. The best mutants from

each library were designated as XR3* and LAD*, and transformed into a fresh host to confirm the improve cell growth on arabinose. The expression of XR3* and LAD* resulted in 3.2 and 1.6 fold increases, respectively, in growth rates on arabinose compared to a wild-type pathway counterpart (**Figure 4.10 and 4.12**). The coexpression of XR3* and LAD* only improved arabinose growth by 2 fold indicating that these mutants were specific for the wild-type version of their counterpart (**Figure 4.12**).

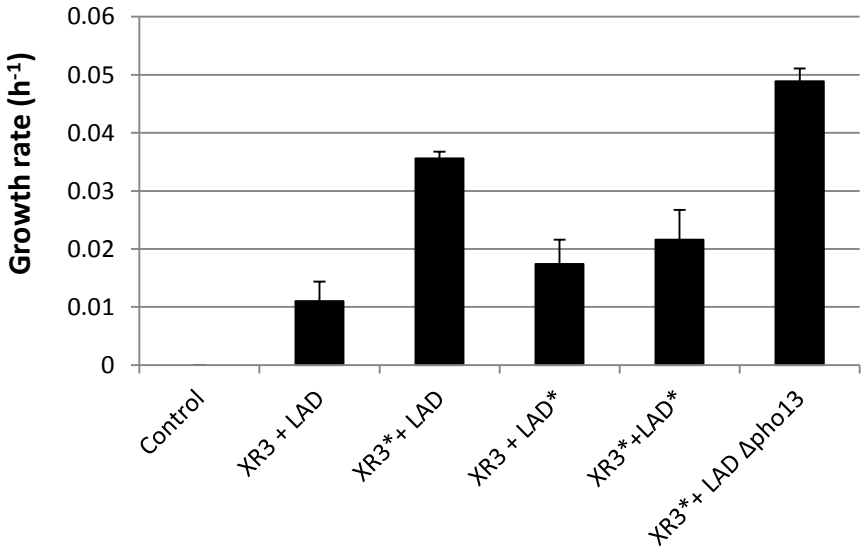


Figure 4.12. Directed evolution of XR3 and LAD for improved cell growth on arabinose. The cell growth of *S. cerevisiae* expressing the different combination of the XR and LAD with their wild-type (XR and LAD) and mutant (XR3* and LAD*) versions were compared to the control strain harboring empty plasmids(control). The *pho13* knockout strain expressing XR3* (XR3* *Δpho13*) showed significantly improved cell growth on arabinose. The growth test was performed on arabinose as a single carbon source in minimal medium. Error bars represent the standard deviation of biological triplicates.

Although cell growth was still too low to conduct ethanol fermentations (nearly 0.036 h^{-1}), these results clearly show the potential of directed evolution (in even just a single round) for improving the arabinose catabolic pathway. Beyond the pathway, additional strain engineering may be useful to improve arabinose catabolic pathways in *S. cerevisiae*. As an example, the deletion of *pho13* (para-nitrophenyl phosphatase) gene known to be beneficial in xylose utilization[85] further improved the cell growth on arabinose to obtain a final growth rate of nearly 0.05 h^{-1} on minimal medium (**Figure 4.10**) suggesting downstream engineering can also significantly improve the arabinose pathway.

4.3.4.3 Evolutionary evolution to improve the xylan catabolic pathway

Due to the number of enzymes required for the xylan catabolic pathway and no immediate rate limiting target, we opted to utilize an adaptive, evolutionary engineering approach to improve this pathway performance. To this end, selected the strain containing XNA1 (xylanase 1) and XD2 (xylosidase 2) with ABF (arabinofuranosidase) from *U. bevomyces* and performed exponential phase, low inoculum serial subcultures of this strain in minimal xylan medium. The xylan utilizing strain was serially-subcultured in xylan medium when the cells were exponential phase using low inoculum size of 0.5%. The best mutant to emerge from this selection showed nearly 50% improvements in growth rate on xylan minimal medium (**Figure 4.11B**). Moreover, this evolved strain was able to produce 20% more ethanol in fermentations and produced up to 0.32 g/L on

xylan (**Figure 4.11C**). Thus, these results demonstrate the ability to use evolution to improve the bioprospected xylan pathay.

4.3. 5 Discussion

This study uses a blended bioprospecting and pathway engineering and evolution approach to generate novel and functional pathways for xylose, arabinose, and xylan in *S. cerevisiae*. These pathways performed better than existing pathways with respect to growth rate on xylose by more than 85% and enabled the first reports of growth on minimal medium with arabinose and xylan as sole carbon sources and a novel pathway with a dual-functional reductase for arabinose catabolism. Moreover, these genes provide promising starting points for further engineering and evolution as evinced by rapid improvement by both directed and evolutionary engineering. These results demonstrate that the pentose catabolic genes from *U. bevomyces* offer promising starting points for developing pathways in *S. cerevisiae*. Moreover, this study provides a straightforward workflow of isolating a non-model organism, sequencing, bioprospecting, importing, and evolving pathways in yeast.

U. bevomyces is closely related to a plant pathogen *U. maydis* based on 99% 18S rDNA sequence similarity; however, these strains are quite distinct with respect to pentose catabolic enzymes. When compared with the counterpart sequences of *U. maydis* [139], pentose catabolic enzymes only ranged from 61% - 93% similarity. The most similar sequence was LXR involved in arabinose pathway, and the most dissimilar sequence was XD2 involved in xylan degradation. It is interesting to note the strong sequence

homology with LXR given its dispensability in the arabinose pathway. A recently identified endo-1,4-beta xylanase from *U. maydis* [67] was also compared to the counterpart of *U. bevomyces*, but the identify value of these two protein sequences were only 72 % (**Appendix C**). Thus, while clearly part of the same genus, *U. bevomyces* is distinct from *U. maydis* and has intriguing physiological and genetic properties, especially with respect to pentose catabolism.

The xylose catabolic genes from *U. bevomyces* could offer alternative choices to construct efficient xylose catabolic pathways in *S. cerevisiae*. While these enzymes provided an 85% growth rate advantage over the traditionally used *xyl1*, *xyl2*, and *xyl3* from *S. stipitis*, ethanol production levels were not improved. The *S. stipitis* pathway has been intensively studied and optimized for almost 30 years since its first expression in *S. cerevisiae* [10]. In particular, the optimization of gene expression levels of *xyl1*, *xyl2*, and *xyl3* significantly improved xylose fermentation [110, 149-151]. In this study, we compared ethanol production by expressing the XR, XDH, and XK from both *U. bevomyces* and *S. stipitis* under expression conditions which were optimal for the *S. stipitis* pathway. Thus, further expression optimization as well as protein engineering may help improve the *U. bevomyces* pathway performance and cofactor preference.

This study suggests an alternative pathway for arabinose catabolism in *S. cerevisiae* (**Figure 4.8**) in which a singular reductase enzyme can dually function to convert both L-arabinose to L-arabitol and L-xylulose to xylitol. In *T. reesei*, L-arabinose is converted to L-arabitol by a xylose reductase which has dual functions as xylose and arabinose reductases[130]. The reductase XR3 from *U. bevomyces* seems to be even more

promiscuous than the *T. reesei* reductase and can use D-xylose, L-arabinose, and L-xylulose as substrates. The expression profiling of *U. bevomyces* (**Figure 3**) indicates that XR3 is highly expressed in both xylose and arabinose culture which is consistent with the hypothesis that this enzyme has dual functions. The efforts to find efficient LAD and LXR from various organisms have been reported for last decades [9, 14, 15, 135]. However, there has been some disagreement in the literature over whether some previously reported LXRs actually have LXR activity. Recently, only *lxrA* from *A. niger* and *LXR3* from *T. reesi* have validated with activity [30, 130]. The expression of the alternative *U. bevomyces* arabinose pathway enabled *S. cerevisiae* to grow on arabinose as a sole carbon source in minimal medium. While there have been other reports of *S. cerevisiae* growth on arabinose using heterologous pathways [12, 28, 36, 59], all of these reports use either rich medium or strains adapted for long periods of time. To our knowledge, this report is the first to demonstrate actual cell growth of *S. cerevisiae* on arabinose minimal medium as a result of simple pathway importation.

The arabinose pathway discovered here is a promising starting point for constructing an efficient utilizing strain. This test was conducted in a minimally engineered strain which does not have any overexpression of downstream genes or sugar transporters. Therefore, further downstream pathway engineering and/or evolutionary engineering can further improve this catabolic pathway. As demonstrated here, the expression of the mutant reductase XR3* along with a *pho13* knockout increased growth 4.4 fold (**Figure 4.10**).

A functional xylan catabolic pathway was achieved here without the need for surface display like prior research in the field [63, 67]. Here, simple expression of a minimal set of hemicellulases enabled growth on xylan and growth rates were increased 50% by evolutionary engineering. Naturally, these hemicellulaes are secreted as mixtures in *Trichoderma* and *Aspergillus* or expressed as a multi-enzyme cellulosome type complex as in *Clostridia* [152]. Most studies have reported only enzyme activity of these hemicellulases when expressed in engineered *S. cerevisiae* or have demonstrated conversion of xylan to reduced sugars [66, 67]. To our knowledge, there is only one report fully reporting ethanol production from xylan, but this strain used a surface display technique and was conducted in rich medium [63]. As a result, the minimal medium based xylan growth reported here is highly significant in the area of xylan catabolism engineering. Moreover, the engineering conducted here was with a minimal number of enzymes which indicates more possibility as further hemicellulases are imported.

A recent study addressed the potential of *U. maydis* for lignocellulosic biomass degradation and showed that *U. maydis* outperformed *T. reesei* in the saccharification of wheat straw [141]. Given the high similarity to *U. maydis*, *U. bevomyces* would be a good source organism for similar lignocellulolytic enzymes. An analysis of cellulases and hemicellulases from whole genome-sequenced fungal plant pathogens revealed that the number of these enzymes ranged from 2 to 60 and 4 to 43, respectively [140]. In particular, *T. reesei* has 26 hemicellulases and *U. maydis* has only 10 hemicellulases. From our whole genome sequencing, *U. bevomyces* has only 7 hemicellulases and we have confirmed that three of these are functional in *S. cerevisiae.* Therefore, the moderate

number of lignocellulolytic enzymes in *U. bevomyces* and ease of transfer makes this source of genes quite promising for further lignocellulose engineering efforts in yeast.

4.4. CONCLUDING REMARKS

This study presents a complete workflow from isolating an organism to whole genome sequencing and bioprospecting to pathway engineering and evolution. This blended approach was highly successful in engineering novel phenotypes of xylose, arabinose, and xylan growth by *S. cerevisiae*. The decreasing cost of whole genome sequencing offers more opportunities to understand organisms in nature and perform targeted bioprospecting of entire pathways and bypasses the need for more laborious cDNA libraries. These results both establish a blended methodology and provide excellent starting points for future engineering efforts to improve xylose, arabinose, and xylan catabolism.

Chapter 5: Conclusions and Major Findings

Production of biofuels and biochemicals from lignocellulosic biomass is expected to offer promising alternatives in future energy markets and manufacturing industries. With the advances in metabolic engineering, microorganisms are now able to produce biofuels and biochemicals from lignocellulosic biomass, but the conversion efficiency of lignocellulosic biomass into valued products remains suboptimal. The improvement of efficiency, therefore, requires a new metabolic engineering approach.

In this study, therefore, *S. cerevisiae*, the platform organism for biofuels and biochemical production, was engineered in a combinatorial way for efficient pentose catabolism to facilitate economical bioethanol production from lignocellulosic biomass. To this end, the directed evolution of xylose isomerase, optimization of xylose catabolic pathway, evolutionary engineering for the whole cell system were performed to develop the strain with efficient xylose catabolism as a model combinatorial engineering. This approach was then expanded to the initial studies of combinatorial engineering of oxidoreductase-based xylose, arabinose and xylan catabolism using novel genes identified from a xylose utilizing non-model organism of *U. bevomyces* through whole genome sequencing.

Directed evolution of heterologous enzymes improved essential enzymes that can be adapted for efficient pentose catabolism including improved xylose isomerase, xylose reductase, and L-arabinose dehydrogenase for efficient xylose and arabinose utilization and ethanol production. The directed evolution of xylose isomerase increased xylose

isomerase activity by 77% which resulted in 61 fold increase in aerobic growth rate and 8 folds increase in xylose consumption and ethanol production rates when expressed in *S. cerevisiae*. In addition, the directed evolution of xylose reductase and L-araabinose dehydrogenase improved ethanol production from xylose and cell growth on arabinose, respectively. These improved enzymes served as key components to develop the strain with efficient pentose catabolism.

Optimizing metabolic pathway by rational engineering offered better metabolic flux through the newly introduced metabolic pathway, thus significantly improved pentose catabolic pathway, specifically xylose catabolic pathway, in *S. cerevisiae*. Though the rational engineering was conducted in minimal manner by overexpressing downstream genes, *XKS1* and *tal1*, and deleting of *gre3* and *pho13*, the strain expressing evolved xylose isomerase pathway showed significantly improved xylose catabolism. The rational engineering resulted in the improvement in aerobic growth on xylose by over 100 fold, and showed increased xylose consumption and ethanol production rates by 24 and 14 folds, respectively, compared to wild-type strain. This result demonstrates that rational engineering can optimize metabolic flux through a new pathway introduced by a heterologous gene expression and enhance the performance of the pathway.

Evolutionary engineering further improve pentose catabolism in the rationally engineered strain throughout the whole cell metabolism. The evolutionary engineering of xylose and xylan utilizing strains showed the aerobic growth rate on xylose and xylan increased by 22% and 50%, respectively. The increased aerobic growth rates resulted in higher ethanol production from xylose and xylan. The evolved xylose utilizing strain

(SXA-R2P-E) showed the highest ethanol yield (0.45g ethanol / g xylose) and the second highest xylose consumption and ethanol production rates reported to date. The evolutionary engineering of xylan utilizing strain enabled ethanol production from xylan as a single carbon source in minimal medium for the first time.

The identification of novel pentose catabolic genes offered key components to construct efficient xylose, arabinose, and xylan catabolic pathways. The expression of alternative arabinose pathway genes and xylan catabolic genes enabled *S. cerevisiae* to grow on arabinose and xylan as a single carbon source in minimal medium for the first time. The successful expression of these enzymes and further improvement of these pathways by combinatorial engineering approach demonstrated that these genes could be a good starting point to construct efficient pentose catabolic pathways in *S. cerevisiae* and could be significantly improved as xylose isomerase pathway.

In conclusion, this study 1) developed *S. cerevisiae* with efficient xylose isomerase-based pathway that converts xylose to ethanol with high yield and productivity, 2) suggested rational engineering targets to construct efficient xylose isomerase-based pathway, 3) identified novel xylose, arabinose, and xylan catabolic genes for efficient pentose catabolic pathway construction in *S. cerevisiae*. In doing so, this work established 1) effective evolutionary engineering strategies and 2) a new approach to transfer a desired phenotype from a non-model organism to a model-organism. Therefore, this study represents a novel tool for combinatorial engineering of a new metabolic pathway. This study improves the prospect of biofuels and biochemicals production from a wide variety of feedstocks in ‘advanced cell factories’.

Chapter 6: Proposals for Future Works

This study demonstrated combinatorial engineering of pentose catabolic pathways for efficient biofuels and biochemicals production from lignocellulosic biomass. Collectively, the pentose utilizing strains and the strain engineering strategies developed in this study significantly advanced the field of strain engineering for biofuels and biochemical productions. Yet, there remain unanswered questions and several follow-up studies that could expand upon these results.

Through combinatorial engineering, we developed an efficient xylose utilizing strain (SXA-R2P-E) showing the highest ethanol yield and the second highest ethanol production and xylose consumption rates from xylose ever reported. Yet, carbon flux through xylose catabolic pathway could be limited even with the efficient xylose catabolism of SXA-R2P-E since native sugar transporters in *S. cerevisiae* are inefficient in pentose transport [84]. Here, transporter engineering will be necessary to further improve carbon flux through xylose catabolism and thus xylose fermentation performance. In our lab, xylose specific transporters has been developed significantly increasing xylose fermentation performance in a couple of strains of *S. cerevisiae*[17, 61, 62]. By introducing these xylose specific transporters, therefore, the xylose fermentation performance of SXA-R2P-E is expected to be further improved.

Evolutionary engineering significantly improved xylose fermentation performance in SXA-R2P-E compared to its parent strain of the rationally engineering strain, SXA-R2P. While evolutionary engineering of xylose utilizing strains of SXA-

R2P-E, adaptive changes in the genomic DNA are expected to be happened. However, the detailed information of the changes was not elucidated in this study. Therefore, genomic and/or transcriptomic analysis of SXA-R2P-E will be necessary to offer an idea of the important genomic components which lead significantly improved performance in xylose fermentation.

The construction of arabinose catabolic pathway complemented by *U. bevomyces* genes enabled *S. cerevisiae* to grow on arabinose as a single carbon source in minimal medium for the first time. Thought the directed evolution of XR3 and *pho13* deletion improved the cell growth on arabinose, applicable ethanol production was not observed requiring further engineering of the arabinose utilizing strain. Here, directed evolution of LAD in the strain expressing XR3*, downstream gene overexpressions, and evolution evolutionary engineering could be applied to increase the cell growth and ethanol production of arabinose utilizing strains. Once developed, the arabinose utilizing strain could be used as a base strain for developing arabinose transporters to improve arabinose uptake. To date, few arabinose transporters are available limiting arabinose fermentation in *S. cerevisiae*[84].

The initial effort in the combinatorial engineering of xylan utilizing strain made ethanol production possible from xylan without surface display of hemicellulases. Yet, the fermentation performance is still insufficient. Here, combinatorial engineering of xylan utilizing strain could be performed by using a different background strain and/or expressing additional hemicellulases. *S. cerevisiae* requires heterologous xylose catabolic pathway to utilize xylan, a xylose polymer. To support sufficient xylose catabolism,

therefore, SXA-R2P-E developed during xylose utilizing strain engineering could be used as a background strain to develop a xylan utilizing strain. The expression of multiple xylanases and/or xylosidases would also be necessary to support higher catalytic activity for depolymerization of xylan.

This work engineered *S. cerevisiae* for efficient pentose catabolism and resulted in satisfactory outcomes. Combined with downstream engineering for advanced biofuel or biochemical production, the engineered strains developed in this study could improve the prospect of economical biofuels and biochemical production from lignocellulosic biomass.

Chapter 7: Materials and Methods

7.1 Common Materials and Methods

7.1.1 Culture conditions and media

Yeast strains were routinely propagated at 30 ℃ in yeast synthetic complete (YSC) medium composed of 6.7 g/L yeast nitrogen base, 20g/L glucose, and appropriate dropout complete supplement mixture (CSM) (MP Biomedicals, Solon, OH). Mutant selection, growth characterizations, and ethanol fermentation were conducted in the identical medium except that 20 g/L or 40 g/L of xylose, arabinose, or xylan was added as a carbon source. *Escherichia coli* strain *DH10β* for all cloning and plasmid propagation was grown at 37 ℃ in Luria-Bertani (LB) Broth supplemented with 50 μg/mL of ampicillin. All strains were cultivated with 225 RPM orbital shaking. Yeast and bacterial strains were stored at -80℃ in 15% glycerol.

7.1.2 Cloning and transformation

The heterologous genes for pentose catabolic pathways were routinely cloned into Mumberg plasmids with appropriate promoter and selection marker [116]. The amplification of heterologous genes was conducted by PCR reactions using Phusion DNA polymerase (Thermo Fisher Scientific, Waltham, MA). Cloning and bacterial transformations were conducted according to standard methods[153]. Restriction enzymes were purchased from New England Biolabs (Ipswich, MA) and Thermo Fisher Scientific (Waltham, MA). The

plasmid was isolated from *E.coli* using Zippy Plasmid Miniprep Kit (Zymo Research, Irvine, CA). Yeast transformation was conducted using Frozen EZ Yeast Transformation II Kit (Zymo Research, Irvine, CA) according to the manufacturer's instructions.

7.1.3 Fermentation assays and growth analysis

Ethanol fermentations were performed at high yeast optical density in oxygen-limited conditions and micro-aerobic conditions. For oxygen-limited conditions, cells were grown at 30℃ in 50 ml of YSC medium containing 40g/L of xylose/arabinose/xylan as a sole carbon source in a 125 ml flask. For micro-aerobic conditions, cells were grown at 30℃ in 40 ml of YSC medium with 20g/L of xylose/arabinose/xylan in a sealed 50 ml falcon tube. For both of these fermentations, a one day old pre-culture (grown on glucose) was pelleted and re-suspended in xylose/arabinose/xylan medium and inoculated at an initial OD of 20. Aerobic growth rate analysis was performed at low yeast optical density while culturing cells at 30℃ in 5ml of YSC medium with 20g/L of xylose in a 14 ml culture tubes. Xylose concentrations were measured using an YSI 7100 Multiparameter Bioanalytical System (YSI Life Sciences, Yellow Springs, OH) and ethanol concentrations were measured using Ethanol Assay, UV-method Kit (R-Biopharm, Darmstadt, Germany). Cell density was measured spectrophotometrically at 600 nm absorbance, from which cell dry weight was calculated for the ethanol production and xylose consumption rates. One OD_{600} unit was considered as 0.17 g cells/L [45]. Fermentation and growth assays were performed in biological triplicate.

7.1.4 Mutant library construction

A library of randomly mutated pentose catabolic genes was generated by error-prone PCR using GeneMorph II Random Mutagenesis Kit (Stratagene, La Jolla, CA). According to the directions, libraries using low (0-4.5 mutations/kb), medium (4.5-9 mutations/kb), and high (9-16 mutations/kb) mutagenesis rates were cloned to achieve a library size of over 1 x 10^5. The randomly amplified pentose catabolic genes were digested with appropriate enzymes and ligated into yeast cloning vectors. The mutant plasmid library was transformed into *E. Coli DH10β* using standard electroporation protocols [153]. The resulting *E. coli* library was harvested from the petri-dishes, isolated and retransformed into *S. cerevisiae*. This process and procedure for mutagenesis was repeated if necessary. These yeast libraries were allowed to grow for three days on plates before being scraped off and harvested for selection. These cells were then pooled and subjected to growth-based enrichment.

7.2 Materials and Methods for Chapter 2

7.2.1 Strain construction

S. cerevisiae strain BY4741-S1 deleted in *gre3* (*Mat a; his3Δ1; leu2Δ0; met15Δ0; ura3Δ0; YHR104w::kanMX4, p415$_{TEF}$-tal1*) was used as a host strain in this study. This strain was obtained by transforming *S. cerevisiae* BY4741 *gre3* knockout strain (supplied by Zhihua Li, The University of Texas at Austin) with *tal1* from *S. spititis* cloned into a p415 vector under the control of the TEF promoter [116]. A yeast

codon-optimized version of xylose isomerase gene (*xylA*) from *Piromyces sp.* was synthesized by Blue Heron (Bothell, WA) and cloned into the Mumberg plasmid p416-GPD [116] to form p416-GPD-XI. The xylose isomerase mutant library was then constructed in the *S. cerevisiae* strain BY4741-S1 as described in common methods.

7.2.2 Mutant selection

To enrich fast-growing transformants, cells from the yeast library were cultured and serial-transferred into 20 ml of fresh YSC medium with xylose as a sole carbon source in a loosely closed 50 ml falcon tube. Serial transfer was repeated every 4 to 5 days using 10% inoculums. After seven rounds of serial transfers, cells from the final culture were plated onto YSC medium with xylose and the largest colonies were isolated. In total, 40 cells were selected from these plates for further characterization. Cell growth of isolated variants was compared in 5 ml of YSC medium with xylose in a 14 ml culture tube and the fastest growing variants were initially selected. The vectors from these promising mutants were isolated, sequenced, and retransformed. A second growth rate measurement against control confirmed that the growth rate increase was due to the mutant xylose isomerase and not background adaptation of the host strain. The mutant xylose isomerase conferring the highest growth rate on xylose was used as the template for the next round of mutagenesis and selection. In total, two additional rounds of iterative random mutagenesis and the selection were conducted. The best performing mutant was termed *xylA*1* (1st round mutnat), *xylA*2* (2nd round mutant), and *xylA*3* (3rd round mutant), respectively, and the re-transformed strains expressing these mutant

xylose isomerase are termed S1A1, S1A2, and S1A3, respectively. The number of serial transfers was reduced from seven to five in the second and the third round of the selection processes.

7.2.3 Site-directed mutagenesis of identified mutations

Site-directed mutations were conducted to confirm the beneficial mutations in the xylose isomerase mutant. A series of six individual back mutations were made using the third round mutant xylose isomerase (*xylA*3*) by using Quikchange II kit (Stratagene, La Jolla, CA) and transformed into *S. cerevisiae* to identify beneficial mutations on the basis of cell growth. Once identified, the beneficial mutations were introduced to wild type xylose isomerase to confirm the improved cell growth phenotype on xylose.

7.2.4 *In vitro* xylose isomerase activity measurements

Xylose isomerase activities from cell extracts were assayed by measuring the decrease of NADH in a 1 ml of reaction mixture at 340 nm using a spectrophotometer [115]. The reaction mixture contained 100 mM Tris-HCl buffer (pH 7.5), 0.15 mM NADH, 10 mM $MgCl_2$, and 2 U sorbitol dehydrogenase (Roche, Mannheim, Germany). The cell extracts were prepared from the yeast transformants cultivated until early exponential growth phase in selective medium by using YPER Plus Dialyzable Yeast Protein Extraction Reagent (Thermo Scientific, Rockford, IL). The protein contents of cell extracts were determined by the method of Bradford assay using Pierce BCA Protein

Assay Kit (Thermo Scientific, Rockford, IL). Enzyme assays were performed in biological triplicate and 25 to 500 mM xylose was used to determine kinetic parameters.

7.2.5 Protein structure prediction for xylose isomerase

The three-dimensional structure of the xylose isomerase protein was predicted by Swiss-Model (http://swissmodel.expasy.org), a web-based server for automated comparative modeling of three dimensional protein structures [154]. The crystal structure of *Thermotoga neapolitana* [PDB code 1A0E [117]] was used as a template. The active site of xylose isomerase was predicted using the computer program PHYRE (a protein fold recognition server; http://www.sbg.bio.ic.ac.uk/phyre/[155]). The predicted protein model and mutations were visualized using Pymol version 1.3.

7.3 Materials and Methods for Chapter 3

7.3.1 Strain construction

S. cerevisiae strains used in this are summarized in **Table 3.1.** To construct xylose utilizing strain, the xylose isomerase mutant (*xylA3**) under the control of GPD promoter and an additional copy of *XKS1* under the control of TEF promoter were integrated into genome of *S. cerevisiae* BY4741 *gre3* knockout strain using integration vectors [61] resulting in SXA-R1. For the higher expression of xylose isomerase expression, an additional copy of *xylA3** under the control of GPD promoter and *tal1* from *S. stipitis* under the control of TEF promoter were also integrated into genome of SXA-R1 strain

resulting in SXA-R2. The integration of the target genes were conducted by linearizing integration vectors by proper restriction enzymes and transforming into the target strains. In each of SXA-R1 and SXA-R2 strains, *pho13* gene was deleted to further improve the xylose utilization by homologous recombination. *Pho13* knockout strain of SXA-R1 and SXA-R2 were named as SXA-R1P and SXA-R2P, respectively.

7.3.2 Evolutionary engineering of xylose utilizing strain

To improve xylose utilization of the rationally engineered strain, SXA-R2P was cultured and serial-transferred into 20 ml of fresh YSC medium with 20 g/L of xylose as a sole carbon source in a closed 50 ml falcon tube. The cells were transferred into fresh medium when they are at the exponential and stationary phase using 0.5, 1, 5% inoculums. After 12 and five rounds of exponential and stationary phase transfer, cells were plated onto YSC medium with 20 g/L xylose and the largest colonies were isolated. In total, 130 cells were selected from these plates for further characterization. Cell growth of isolated variants was compared using Bioscreen C and confirmed in 5ml of YSC medium with 20 g/L of xylose in a 14 ml culture tube. The fastest growing strain was selected and named as SXA-R2P-E.

7.3.3 Large scale ethanol fermentation

Ethanol fermentations were performed in sealed 50 ml falcon tubes and a bioreactor. The small scale fermentation was conducted as described in common methods. For a large scale experiment, ethanol fermentation was conducted in 3 L

Bioreactor with 1.8 L of YSC medium with 40 g/L of xylose. The medium was supplemented with the anaerobic growth factors ergosterol (0.01 g/L) and Tween 80 (0.42 g/L) [90]. Medium pH was maintained at 5.0 with 2.5 N NaOH. Though the reactor was not purged with nitrogen gas, DO in the vessel was dropped after 12 hours of operation and anaerobic condition was maintained throughout the fermentation. Fermentation in a bioreactor was performed in technical duplicate.

7.4 MATERIALS AND METHODS FOR CHAPTER 4

7.4.1 *U. bevomyces* culture condition and growth test

U. bevomyces was isolated during a xylose culture experiment. *U. bevomyces* was routinely propagated at 30 ℃ in Yeast synthetic complete (YSC) medium composed of 6.7 g/L yeast nitrogen base, 20 g/L glucose, and CSM (MP Biomedicals, Solon, OH). Growth rate analysis was performed at low optical density while culturing cells at 30 ℃ in 50ml of YSC medium with 20 g/L of glucose, xylose, arabinose or xylan in 125 ml flasks. *S. cerevisiae* BY4741 and *S. stipitis* CBS 6054 were used as control strains for growth analysis (**Table 4.3**). To calculate growth rate, cell density was measured spectrophotometrically at 600 nm absorbance. The growth rate on xylan was calculated from cell counts based on plating.

7.4.2 Whole genome sequencing

Genomic DNA of *U. bevomyces* was prepared from an overnight culture by using Wizard Genomic DNA purification Kit (Promega Corporation, Wisconsin, USA) according to the manufacturer's instruction. Initial classification was conducted using 18S rDNA amplified from the genomic DNA of the *U. bevomyces* using a fungus 18S rDNA specific primer set [137] and Sanger sequencing. The National Center for Biotechnology Information (NCBI) Basic Local Alignment Search Tool (BLAST)[156] was used to identify closely related species. Whole genome sequencing was carried out using Illumina HiSeq 2000 at the UT Genomic Sequencing and Analysis Facility. The genome sequences were assembled by *de novo* genome assembler Velvet [142]. Structural genome annotation was carried by using the MAKER pipeline [143]. The structurally annotated genes were processed by Blast2GO [144] for functional annotation.

7.4.3 Gene expression analysis by RT-qPCR

Relative gene expressions were quantified using real-time qPCR. Cells grown in YSC media with glucose, xylose, arabinose or xylan were used for whole cell RNA extraction using YeaStar RNA Kit (Zymo Research, Irvine, CA) followed by DNase treatment using Ambion DNA-free kit (Life Technologies, Carlsbad, CA). cDNA synthesis was performed using the Applied Biosystems High Capacity Reverse Transcription Kit (Life Technologies, Carlsbad, CA). Primers used for RT-qPCR reaction are as shown in **Appendix B**. RT-qPCR was performed on a ViiA7 Real Time PCR

System (Life Technologies) using Fast Start SYBR Green Master Mix (Roche, Penzberg, Germany) following the manufacturer's instructions. The RNA expressions from xylose, arabinose, or xylan culture were compared to those from glucose culture.

7.4.4 Heterologous expression of *U. bevomyces* genes in *S. cerevisiae*

To express pentose pathway genes from *U. bevomyces* into *S. cerevisiae*, the target genes were PCR amplified and cloned into Mumberg plasmids, p416, p415, p423, and p424 under the control of TEF or GPD promoter [116]. *S. cerevisiae* BY4741 (*Mat a; his3Δ1; leu2Δ0; met15Δ0; ura3Δ0; trp5::kanMX4*) *trp5* knockout strain from the Yeast Deletion Database and *S. cerevisiae* YXS3[41] were used as host strains in this study [157] (**Table 4.2**). *S. cerevisiae* was routinely propagated at 30 ℃ in YSC with medium composed of 6.7 g/L yeast nitrogen base, 20g/L glucose, and CSM, CSM-Leu-Ura, CSM-Leu-His-Ura or CSM-His-Leu-Trp-Ura (MP Biomedicals, Solon, OH). Growth characterizations were conducted in the identical medium except that 20 g/L xylose, arabinose, or xylan was added as a carbon source.

7.4.5 *In vitro* xylose reductase activity measurement

Xylose reductase from cell extracts were assayed by measuring the decrease of NADH or NADPH in a 1ml of reaction mixture at 340 nm using a spectrophotometer [123]. The reaction mixture contained 0.7 ml of 50 mM potassium phosphate buffer (pH 6.5), 0.1 ml of the cell extracts, 0.1 ml of 2 mM NADH or NADPH, and 0.1 ml of xylose. The cell extracts were prepared from the yeast transformants cultivated until early exponential

growth phase in selective medium by using YPER Plus Dialyzable Yeast Protein Extraction Reagent (Thermo Scientific, Rockford, IL). The protein content of cell extracts was determined by a Bradford assay using Pierce BCA Protein Assay Kit (Thermo Scientific, Rockford, IL). Enzyme assays were performed in biological triplicate and 0.1 to 1 M xylose was used to determine kinetic parameters.

7.4.6 Directed evolution of XR3 and LAD

The libraries of randomly mutated XR3 and LAD were generated by error-prone PCR. Error-prone PCR of XR3 and LAD genes were conducted with the primer set XR3-EP-f, XR3-EP-r, LAD-EP-f, and LAD-EP-r (Appendix B) using GeneMorph II Random Mutagenesis Kit (Stratagene, La Jolla, CA). Libraries with the size of over 1 x 10^5 were generated as previously reported [35]. The resulting libraries were transformed into *S. cerevisiae* BY4741 (*Mat a; his3Δ1; leu2Δ0; met15Δ0; ura3Δ0; trp5::kanMX4*) *trp5* knockout strain. Growth-based enrichment and selection was conducted to obtain the best mutants. The isolated mutants were sequenced and retransformed into fresh host. Cell growth of isolated variants was compared using Bioscreen C and confirmed in 5ml of YSC medium with 20 g/L of arabinose in a 14 ml culture tube. The best performing mutants were termed XR3* and LAD*.

7.4.7 Evolutionary engineering of xylan utilizing strain

To improve xylan utilization, the xylan utilizing strain expressing XNA1, XD2, and ABF was cultured and serially-transferred into 5 ml of fresh YSC medium with 20

g/L of xylan as a sole carbon source in a 14 ml culture tube. The cells were transferred into fresh medium when they are at the exponential phase using 0.5% inoculums. After three weeks of serial subculturing, cells were plated onto YSC medium with 20 g/L xylan and the largest colonies were isolated. In total, 80 cells were selected from these plates for further characterization. Cell growth of isolated variants was compared using Bioscreen C and confirmed by measuring cell counts after 5 days of incubation in 5ml of YSC medium with 20 g/L of xylan in a 14 ml culture tube. The fastest growing strain was selected as the evolved strain.

Appendix A: Protein sequences for pentose catabolic enzymes from *U. bevomyces*

>XR1
MSLNKYVTLNDGNKIPQIGLGTWLSKPGEVANAVEVAVKAGYRHLDLARIYQNQEEIGGA
FKKIIPSIVKREELFITSKLWNNSHRPENVQAAYEETLQQLGLDYLDLYLIHWPVAFKPG
QDLVPKIADGKQTDIDREVSIVDTWKALIELQKAGKVKSIGVSNFTQNDLDAIINATGVV
PAVNQIEAHPLLPQDDLVAYSKAKNIHLTAYSPLGNNLSGKTKIVDYPQVSEVAKKYNAD
PAQVLIAWGVKRGYSVIPKSVTESRIKSNFEQIELKDEDYETVTSLYKELGKVRFNIPYT
YAPQWDINIFGEDAEKPAKLTVKTQ

>XR2
MSLKNHHFTLNTGAKMPAIGLGTWLAKPDEVRQAVSVALKQGYRHIDAARVYGNEHEVGQ
ALSEALSSGVCSRSDVWITSKLWNTDHTRAEAAINRTLSDIKTDYLDLYLVHWPVPFTPA
EDANMLFPRNSDGDVDIDTKTSLLDTWKQMIALKKKGLTKAIGLSNVSKDAVEFIVENSD
EIPSSVQVEFHPLIAPLQKELKDYCESKGIVVTAYSPLGNNRQSLPRIIDLDNVKALAAK
NNRTEAQLCINWCVSNGVCCIPKSVTPHRLQSNLESLNFDLTDEERKTLNVLPNTLGTKR
FNIPIQYPAPNGPWLIDLFNTEHESCAIWKAQSNGSVVVA

>XR3
MPIPTVRLPSGVELPRLAFGTGTALYQSAAAKQVTMALNAGFRFVDGAEVYANEESAGEG
IQAFLKSSGLKRDDIYVLTKVGKDGMKDLQSAVKEEMRKLQVDYLDSYLLHFPPRGKDGL
PSNVDAWRQLEKIKAAGLTKSIGVSNWLASDIQQLLDAGLSVPDINQIEFHPYLYANAEY
TKLLALQKKHGIVTMTYAALAPFYKSTLPEDGPLHKALTAIAAKNDKRTTAAVLLRWALQ
RSEGIITTTTSKESRANELLDQLDSNNDAQLHLDEDELKQIDQAGKEHGYEKYYMAPFMQ
P

>XR5
MPVYPSVEPSNADFPQQPALSSLAHDELKHGLAPWRKDAINTAENLARGRGDLATDASPL
VFGGGVFGQDMYNNADTLASDLPIRTVRLALRYGINAFDTSPYYFPSEFTLGNILTALRP
EFPRESYYIITKCGRYGPEHKHFDYSAAKIDSSVRGSCKRLGTEYLDVVLTHDAEFVCDK
VGRSHHDGWESGIVSGLVDPQDVGLQQTREQVIESLGLAPNVEAASKVHGKGDEQFLEAI
AALFKLKDQGIVRRVGISGYPLPVLLRLSRLVATTAPYRPLDAVLSYSNRCLHSDVLVGW
KELFAADPRGDSTNALPELQWSAPLLMNGSPFSMGLLTDGTPPAWHPASDALKAATKEAS
RNLVAQGDSLTMTALTYGLRGSEVAHPSGAGPQLRTLVGLSHPDHVHSAVEAYRVLCSGA
TEAGDALFPATDAANDRRDAYTKQAKNEKAVRDLFAQRGVREWSWSSGL

>XR6
MSSSPIPTIALGGSANHINVGRIAFGCMGMTWTEPAKMTPDAEAFKTIKAAVDAGSNFLN
TGAFYGPPSDPYANLKLLRRFYDAHPEYKDKTVLSVKGGMDTPAYQAKGMAGLKADASVE
ALEIDLRAIREHLGTDQGGKNVDVYEVARRDTSMSVTQAMLNMLSLSTQTYTDAEGNKVT
GKGLFDHISLSELGLASIQEAVKAAPVACVELEVSPWELEVFTSGIVEFCEANNLPILAY
SPVGKGLLTGTIKSAADIPEGDVRSHMDRLNKDNIAKNLELANEFVQLAEKQSPKVTPAQ
LGLAWLVASSKVMIPLPGTSKASRAKENAEAAAIKLDDQTKNELDQKVKQFKVAGGRYNE

AARNNHALMG

>XDH
MSAQVAPNNSLATTPAEGNVSFVLQEIEKVSFEDRPVVAPKRGQVQVNIRQTGLCASDCH
YLHHGRIGDFVVRKPMVLGHESSGIVTAVGEGVTTHKVGDRVALEPGVPCGGCHSCLQGS
YNHCPDLEFAATPPYDGTLCTYYNIMAPFAHHVPDTMSLEEASLMEPLSVAVYSAAIRGQ
VKAMENVLVFGAGPIGLLNAAVCKAYSAKRVVVVDVVDSKLEFAKGFCATSTFKPSLPKE
GEAKIDSANRNAQDLIKSIGDDVGAHEGFDLVLECTGAEPCIQMGIQALRPKGRFVQVGM
GRSEVEFPITRVCVKEIDVTGSFRYGAGAYKTSISLVSTGLIDVTKMVTHRFLFKDAIKA
FDTTTKGVGEDGKTAIKVQISQGEGKH

>XK
MQSTEALFLGLDASTQALKASLLDAHLTVLGELEVRFDPDLPHYGTNGGVSAPTADDDQG
TVVAPVMLYVEALDMLGDKMREATWPLARIRAISAAGQQHASVYFSRAAPHILTTLASDK
TLTAQVERAFSRKVVPNWQDSSTVEACRAFEDAMGGAEALAEVTGSKAHTRFTGPQIYKF
RKQQPEAYKDTERIGLVSSFVTSMLCVGEGEDAESVIKGIDESDACGMNLLDMRPASSRS
ADESVGKIEPGWCQKLLALASGETDGDGSLGGAEELERKLGVVYRDAGASVGKIGSWWKQ
RYGFSDDCHIFPGTGDNPATFLAFSLAQRQAIISLGTSDTVMVATDQYVPDPDFHAFFHP
AQSPDGGQRFFNMLVYKSGSLAREWVRDQYCSADWDTFNADVEKHRIESTSHKRVGFYWL
RPEIIPAGASGVHRYTTASSQSDWNKVDDFDEAGMNASAILETQFLNYRYRSSSIVAGTS
SSPDTSKLPLASIYAVGGASSNTTITQTMADVFGCDIVKPVQPSDDGKGWTAANYNFCSV
GAAYKAVWGWSRTQLPQGEASPQFDDFVHATKSRQAAKAGLTAQGQDKVEGGVQVICRPR
DGRTQIYTDIVQEWKTLEERSQKQA

>LXR
MSMTIEDCAAVRQPRSVPDTPTNVLEQLSMKGKVVVVTGASDGLGYAAVEAVAEAGADVA
LWYNSNDVAIQKAEQLAKTHSIRAKAYQCEVSQVDHVRACIDKVVADFGRIDVFVANAGM
AISKPILEQTVEEYKRQMEVNVDGVFYCAKYVGEVFKRQGKGNLIITSSMSAHIVNVPVD
QPVYNSTKAAVTHMGKSLAREWREFARVNIVSPGFFDTKLGASPRCINEAYRMSALGRQG
HVKEIKGLFLYLASDASSYQTGSDTLIDGGYTLP

>LAD
MLAPTARLAARTSARPSAVARAFASSSRSMSQDNKGKAKPLDTAPFASTKDDVVVPYPED
SSRHPADDVGHNTTGGRVGRHTQRTLASFSMEGKVCVVTGAARGIGNLIARTFVESGANH
VAIVDLNEDECQHAAREVEEWFTQHGGVKPGELDIQGYGCDISDEAQVQDVIGRIHKRFG
KIHVAVNSAGIVENYPATEYPTPKLKKLFDININGSYFVAREVAKCMMQDQTKGSIVMIA
SMSGSVVNVPQAQAPYNASKAAVKHLASSMAVEWAKAGIRVNSLSPGYMLTSLSRAVLEN
SPNGKELRTTWENLTPMGRLGDPEDLKGAVVYLASDASAFTTGSDLIVDGGYTSV

>XNA1
KTNFAALLSATLVVASAVGASVIPAEAHTLHERAATTLNAAIKADGRRYFGTATDQGTF
SISQVSNIIKSEMGCVTPENSMKWDATQPSRGQFTFSGADALVNYATSNGKMIRGHTLVW
HSQLPSWVQAITDPTDLTNVLKQRISTLVGRYKGKVYAWDVVNEIFNEDGTMRKSVFYNV
LGESYVKIAFEAARAADPNAKLYINDYNLDVSTYPKLTGLVSKVKQWRSEGIPIDGIGSQ
SHLAAPGSFGDASGVGAAMKAVCSAAPECAMTELDIAGAAASDYAKATSACLAQSNCVGI
TVWGVSDNLSWRSDKSPLLWDGSYNKKAAYSAVLNTLNQGH

>XNA2
MKLSALLAFASVAGVAVSSPIASALEADKDAGSLFKRQSINYVQNYNGNLANFKYNEGAG
TYSGSWNNPGDFVIGLGWSQGTSSRIITFGGNYQSNQGSYYAVYGWLNSPLTEYYVVENY
SYDPCTASNTQQVGTVTSDGSSYKICKHTQVNQPSIQGTKTFGQYFSVRQSKRSSGSVTL
ANHFNAWKKYGFAIGAANPDFNYQVFATEAFAGQGSVSVQVSG

>ABF
MKISSSLVAISVALAATLVGASPTAERRQTCALPTSYKWIPSAPLAQPANGWVALKDFTH
VPYNGKHLVYASTHDQGTKYGSMGFAPFSTWADMATATQTAMKTSAVAPTIFYFSPKKTW
ILASQWGVASFFYMTSSDPTNPNGWSEEKPLFTGKIADSDTGPIDQTLIGDSSKMCLFFA
GDNGKIYRSCMPIGNFPGSFGTASEVVMSDTKAKLFEAVQVYTLEGQNKYLMIVEAIGSG
GRYFRSFTADSLEGKWTVNAGDESAPFAGKANSGASWTDDVSHGDLIRSNPDQTMTVDPC
RLQLLYQGRDKSVATPSYDLAPYRPGLLTLQR

>XC1
WNVLKATLVGASVGLLSATWAISAPVGMVERAETYVGYGFYYFIGNAAGEERIFAAVSQ
GNSPTGWDLVNGGQPILTSTVGTQGVRDPSIVRSADGSKFYLLATDLNIGSGTSFGEAAT
LGSRSIVVWESSDDLQSWSEPRLVEVIGKEGGSAWAPEALYNPATQKYDVYFSAQLYPDS
DTNHTGSSYFRIMRSSTTDFTSFSPAEVYVDRKGDSVLDMTFLQTNTGLYRFIKNENPTS
DPHPLTVYQERSDGGVDGSWTKVTENIGAGVIGANEGPTAFVDNTDAGKSWLWVDEYTNR
GYVALNTNNTQQGTWTFDSAANEPSNHARHGTIIPLTQTQYDNLRAQI

>XC2
MLSVQVKLLLLVQLCLAACVASHPVSSEPSLLKRASDKVGYLFIHFYDDYKEPGVYTTYP
AGEQVFGHLSNGNDALSYKALKGGAPLLTSTVGTKGVRDMYLVSKGDESQHYIIATDLNQ
TAVGGFSSPFLSRSLVIWESEGASLTKWKPSRLVEVVPDTFRMAWAPEAIWLDDEQRFLV
YWSSNKYADASHSGTPDYDKIYSSYTTDFVTFTDPEVFMDLGNNVGVIDLTIGRGPTSGS
NQYVRFFKDESVYKVRGQVSNSGIQGPWVDIGSATEYVDNNNQAEAPIYFKENTSNRWFV
FLDQYGRSPAGYYPYSADNGIDQYGYTDLGFPSGMPTQLKHGSIKPLTQSQYDEINAAWA

Appendix B: Primers used to amplify the genes from *U. bevomyces*

Primer	Sequence (5' → 3')
PCR primers	
UbXR1-f	actgtatctagaacatcatcgtcatgtcgctcaa
UbXR1-r	actgtcgaattctgtgcctctcagccagacta
UbXR2-f1	atcgctctagacgataatgtctctcaagaatcaccacttca
UbXR2-r1	agcgtcaatgtgacggtagccctgctt
UbXR2-f2	ggctaccgtcacattgacgctgcccgcgtatac
UbXR2-r2	atcgcgaattctgcgagctaaaaatgaatcacacaaacta
UbXR3-f	atcgctctagagaccatgcccatccctact
UbXR3-r	atcgcgaattccgcaactagggctgcat
UbXR5-f	atcgctctagatggtcaacgatgccggtgtat
UbXR5-r	atcgcgaattccttcccgtaaacacaatagacacaatga
UbXR6-f	atcgctctagagatccattcgccatgtcgtctt
UbXR6-r	atcgcgaattcgagtttgtgagagtctatcccatga
UbXDH-f	actgtatctagaagtccgataagcaaccatgtct
UbXDH-r	actgtacccgggcggacagaagcctagtgctt
UbXK-f	actgtatctagacctccttcagcaacagagatttcataa
UbXK-r	actgtaactagtgtgatcggtgagacacctcag
UbLAD-f	atcgacactagtccgcgttctcccaatgct
UbLAD-r	atcgacgaattcagtaatgcattacaccgacgtgta
UbLXR-f	atcgacactagtcaagatgtcgatgaccatcgaaga

UbLXR-r	atcgacgaattctgacacaatcaagggagtgtgta
UbXNA1-f1	atcgacactagtcgatgaagaccaactttgca
UbXNA1-r1	atacgtcgaaacatccaaattgtagtcgttgatgtagagctt
UbXNA1-f2	acaatttggatgtttcgacgtatccgaagttga
UbXNA1-r2	atcgacgaattcatcaatggccctggttga
UbXNA2-f	atcgacactagtcatcatgaagctttccgctcttct
UbXNA2-r	atcgacgaattcgcaacctcacccgctcactt
UbABF-f	atcgacactagtcgttagactctctccgatatgaagatct
UbABF-r	atcgacgaattcagagtgaaggctttacctctgga
UbXD1-f	atcgtaactagtgcatcatgtggaacgtgctcaa
UbXD1-r	atcgtagaattcgttgcggctgttgctttagatct
UbXD2-f	atcgtaactagtatgctgagcgtgcaagtcaa
UbXD2-r	atcgtagaattccctgttggccggtacgacta
RT-qPCR primers	
UbXR1-f	gctcaaggacgaggactacg
UbXR1-r	tactcggtcatccccaagag
UbXR2-f	gtgacgtcgacattgacacc
UbXR2-r	ttaccaaagccatcggtctc
UbXR3-f	cagatcgagttccaccccta
UbXR3-r	cacccctacctctacgcaaa
UbXR4-f	ctgatgacccaaggtggact
UbXR4-r	gagatcttccaagcgctgtc

UbXR5-f	caacagaccagggaacaggt
UbXR5-r	gcatcaacgcatttgatacg
UbXR6-f	catcaagctcgacgatcaaa
UbXR6-r	aagaacgagctcgaccaaaa
UbXR7-f	cgacgaaaggaatccctaca
UbXR7-r	cgcacatcgctacactctgt
UbXR8-f	acccagaacacgtcgaaaag
Error-prone PCR primers	
XR3-EP-f	tacgaatctaatctaagttttctagagaccatg
XR3-EP-r	atcgatctcgagaagcttgatatcgaattccgcaacta
LAD-EP-f	atcgatactagtccgcgttctcccaatg
LAD-EP-r	tcgataagcttgatatcgaattcagtaatgcatta

Appendix C: The comparison of pentose catabolic enzymes from *U. bevomyces* and *U. maydis*

The amino acid sequences of identified pentose catabolic genes from *U. bevomyces* were blasted to genome sequence of *U. maydis*. The sequences with the highest match from *U. maydis* genome were named as their counterpart from *U. bevomyces*. The most dissimilar sequence was LXR (A), and the most dissimilar sequence was XD2 (B). Recently identified endo-1,4-beta xylanase from *U. maydis*[67] was compared to the counterpart of *U. bevomyces* (C).

A

```
U.bevomyces_LXR      MSMTIEDCAAVRQPRSVPDTPTNVLEQLSMKGKVVVVTGASDGLGYAAVEAVAEAGADVA
U.maydis_LXR         MSMTIQDCSAFREPRPVPDTPTNVLEQLSMKGKVVVVTGASEGLGYAAIEAVAEAGADVA
                     *****:**:*.*:** *************************:******:***********

U.bevomyces_LXR      LWYNSNDVAIQKAEQLAKTHSIRAKAYQCEVSQVDHVRACIDKVVADFGRIDVFVANAGM
U.maydis_LXR         LWYNSNDSAIEKAQVLAKTHAIRAKAYQCEVSLVDKVRATIDKVIADFGRIDVFVANAGM
                     ******* **:**: *****:********** **:*** ****:***************

U.bevomyces_LXR      AISKPILEQTVEEYKRQMEVNVDGVFYCAKYVGEVFKRQGKGNLIITSSMSAHIVNVPVD
U.maydis_LXR         AISKPILEQTVEEYKRQMEVNVDGVFYCAKYVGEVFKRQGKGNLIITSSMSAHIVNVPVD
                     ************************************************************

U.bevomyces_LXR      QPVYNSTKAAVTHMGKSLAREWREFARVNIVSPGFFDTKLGASPRCINEAYRMSALGRQG
U.maydis_LXR         QPVYNSTKAAVTHMGKSLAREWREFARVNIVSPGFFDTKLGASPRVVNEAYRMSALGRQG
                     ******************************************* :*************

U.bevomyces_LXR      HVKEIKGLFLYLASDASSYQTGSDTLIDGGYTLP
U.maydis_LXR         HVKEIKGLFLYLASDASSYQTGSDTIIDGGYTLP
                     *************************:********
```

B

```
U.bevomyces_XD2      MLSVQVKLLLLVQLCLAACVASHPVSSEPSLLKRASDKVGYLFIHFYDDYKEPGVYTTYP
U.maydis_XD2         -----MHLPSL--LSLALWFSLVSLVTWSPLRKHVRDKVGYLFIHFYDNYSSPGVYETYP
                          ::*  *  *.**  .:   : :   * *:. ************:*..**** ***

U.bevomyces_XD2      AGEQVFGHLSNGNDALSYKALKGGAPLLTSTVGTKGVRDMYLVSKGDESQHYIIATDLNQ
U.maydis_XD2         AG---------------------------------VRDMYIVSRADESQHFIIGTDLNQ
                     **                                  *****:**:.*****:**.*****

U.bevomyces_XD2      TAVGGFSSPFLSRSLVIWESEGASLTKWKPSRLVEVVPDTFRMAWAPEAIWLDDEQRFLV
U.maydis_XD2         TAAGGFGGKFVSRSLVIWDSKKASLTQWNEPRLVTVVPEEYRMAWAPEAIWLDNDEHFFV
                     **.***.. *:*******:*: ****:*:  *** ***: :************::::*:*

U.bevomyces_XD2      YWSSNKYADASHSGTPDYDKIYSSYTTDFVTFTDPEVFMDLGNNVGVIDLTIGRGPTSGS
U.maydis_XD2         YWSSNKFSDASHTGDADYDKIYASYTTDFVTFTEPHVYLDLGSHNGVIDLTLGHGPIDGD
                     ******::****:*  ******:**********:*.*::***.. ******:*:** .*.

U.bevomyces_XD2      NQYVRFFKDESVYKVRGQVSNSGIQGPWVDIGSATEYVDNNNQAEAPIYFKENTSNRWFV
U.maydis_XD2         SQYVRFFKDESVYKVCGQVSNNGIHADWQDIESASECVDNHHRAEAANLFAG--------
                     .************** *****.**:. * ** **:* ***..:***   *

U.bevomyces_XD2      FLDQYGRSPAGYYPYSADNGIDQYGYTDLGFPSGMPTQLKHGSIKPLTQSQYDEINAAWA
U.maydis_XD2         ------------------------------------------------------------
```

C

```
U.bevomyces_XNA2       MKLSALLAFASVAGVAVSSPIASALEADKDAGSLFKRQSINYVQNYNGNLANFKYNEGAG
U.maydis_XNA2          MKFATVLAFATAAGAAFASPLASSE--TTEAGQLSKRQSINYVQNYNGNAANFKYDQHAG
                       **::::****:.**.*.:**:**:    .:**.* ************** *****:: **

U.bevomyces_XNA2       TYSGSWNNPGDFVIGLGWSQGTSSRIITFGGNYQSNQGSYYAVYGWLNSPLTEYYVVENY
U.maydis_XNA2          TYSTRWTNPPDFVVGLGWSPGNSYRTIKFSGSYSSSSSSYSAVYGWLNNPLTEYYVVENY
                       ***  *.** ***:***** *.* * *.*.*.*.*...** *******.***********

U.bevomyces_XNA2       SYDPCTASNTQQVGTVTSDGSSYKICKHTQVNQPSIQGTKTFGQYFSVRQSKRSSGSVTL
U.maydis_XNA2          SYDPCSNSGAQVVGSVTSDGSNYKICKHTQYDQPSIQGTKTFGQYFSVRANKRNSGSVTL
                       *****: * :* **:******.******** :***************** .**.******

U.bevomyces_XNA2       ANHFNAWKKYGFAIGAANPDFNYQVFATEAFAGQGSVSVQVSG
U.maydis_XNA2          SKHFNAWKQHGFANGAANPDFNYQVFATEAFGGTGSASMSVSG
                       ::******::*** *****************.* **.*:.***
```

References

1. Alper, H. & Stephanopoulos, G. Engineering for biofuels: exploiting innate microbial capacity or importing biosynthetic potential? *Nature Reviews Microbiology* **7**, 715-723 (2009).
2. Van Vleet, J.H. & Jeffries, T.W. Yeast metabolic engineering for hemicellulosic ethanol production. *Current Opinion in Biotechnology* **20**, 300-306 (2009).
3. Fischer, C.R., Klein-Marcuschamer, D. & Stephanopoulos, G. Selection and optimization of microbial hosts for biofuels production. *Metabolic Engineering* **10**, 295-304 (2008).
4. Hamelinck, C.N., Hooijdonk, G.v. & Faaij, A.P.C. Ethanol from lignocellulosic biomass: techno-economic performance in short-, middle- and long-term. *Biomass and Bioenergy* **28**, 384-410 (2005).
5. Himmel, M.E. et al. Biomass Recalcitrance: Engineering Plants and Enzymes for Biofuels Production. *Science* **315**, 804-807 (2007).
6. Hahn-HäGerdal, B., Lindén, T., Senac, T. & Skoog, K. Ethanolic fermentation of pentoses in lignocellulose hydrolysates. *Appl Biochem Biotechnol* **28-29**, 131-144 (1991).
7. Olsson, L. & Hahn-Hägerdal, B. Fermentation of lignocellulosic hydrolysates for ethanol production. *Enzyme and Microbial Technology* **18**, 312-331 (1996).
8. Lee, J. Biological conversion of lignocellulosic biomass to ethanol. *J Biotechnol* **56**, 1-24 (1997).
9. Richard, P., Verho, R., Putkonen, M., Londesborough, J. & Penttilä, M. Production of ethanol from l-arabinose by *Saccharomyces cerevisiae* containing a fungal l-arabinose pathway. *FEMS Yeast Research* **3**, 185-189 (2003).
10. Kötter, P., Amore, R., Hollenberg, C. & Ciriacy, M. Isolation and characterization of the *Pichia stipitis* xylitol dehydrogenase gene, XYL2, and construction of a xylose-utilizing *Saccharomyces cerevisiae* transformant. *Curr Genet* **18**, 493-500 (1990).
11. Walfridsson, M. et al. Ethanolic fermentation of xylose with *Saccharomyces cerevisiae* harboring the *Thermus thermophilus xylA* gene, which expresses an active xylose (glucose) isomerase. *Appl Environ Microbiol* **62**, 4648 - 4651 (1996).
12. Becker, J. & Boles, E. A Modified *Saccharomyces cerevisiae* Strain That Consumes l-Arabinose and Produces Ethanol. *Applied and Environmental Microbiology* **69**, 4144-4150 (2003).
13. Nevoigt, E. Progress in metabolic engineering of *Saccharomyces cerevisiae*. *Microbiol. Mol. Biol. Rev.* **72**, 379-412 (2008).
14. Richard, P., Putkonen, M., Väänänen, R., Londesborough, J. & Penttilä, M. The Missing Link in the Fungal l-Arabinose Catabolic Pathway, Identification of the l-Xylulose Reductase Gene. *Biochemistry* **41**, 6432-6437 (2002).

15. Richard, P., Londesborough, J., Putkonen, M., Kalkkinen, N. & Penttila, M. Cloning and expression of a fungal L-arabinitol 4-dehydrogenase gene. *J. Biol. Chem.* **276**, 40631-40637 (2001).
16. van Maris, A. et al. Alcoholic fermentation of carbon sources in biomass hydrolysates by *Saccharomyces cerevisiae*: current status. *Antonie van Leeuwenhoek* **90**, 391-418 (2006).
17. Young, E., Poucher, A., Comer, A., Bailey, A. & Alper, H. Functional survey for heterologous sugar transport proteins, using *Saccharomyces cerevisiae* as a host. *Appl. Environ. Microbiol.* **77**, 3311-3319 (2011).
18. Ho, N.W.Y., Lin, F.P., Huang, S., Andrews, P.C. & Tsao, G.T. Purification, characterization, and amino terminal sequence of xylose reductase from *Candida shehatae*. *Enzyme and Microbial Technology* **12**, 33-39 (1990).
19. Zhao, X., Gao, P. & Wang, Z. The production and properties of a new xylose reductase from fungus *Neurospora crassa*. *Appl Biochem Biotechnol* **70-72**, 405-414 (1998).
20. Woodyer, R., Simurdiak, M., van der Donk, W.A. & Zhao, H. Heterologous Expression, Purification, and Characterization of a Highly Active Xylose Reductase from *Neurospora crassa*. *Applied and Environmental Microbiology* **71**, 1642-1647 (2005).
21. Berghäll, S., Hilditch, S., Penttilä, M. & Richard, P. Identification in the mould *Hypocrea jecorina* of a gene encoding an NADP(+): d-xylose dehydrogenase. *FEMS Microbiology Letters* **277**, 249-253 (2007).
22. Zhang, F. et al. Cloning, expression, and characterization of xylose reductase with higher activity from *Candida tropicalis*. *J Microbiol.* **47**, 351-357 (2009).
23. Tran, L.H., Kitamoto, N., Kawai, K., Takamizawa, K. & Suzuki, T. Cloning and expression of a NAD+-dependent xylitol dehydrogenase gene (*xdhA*) of *Aspergillus oryzae*. *J Biosci Bioeng* **97**, 419-422 (2004).
24. Verho, R., Putkonen, M., Londesborough, J., Penttila, M. & Richard, P. A novel NADH-linked l-xylulose reductase in the l-arabinose catabolic pathway of yeast. *J Biol Chem* **279**, 14746-14751 (2004).
25. Nair, N. & Zhao, H. Biochemical Characterization of an l-Xylulose Reductase from *Neurospora crassa*. *Applied and Environmental Microbiology* **73**, 2001-2004 (2007).
26. Sullivan, R. & Zhao, H. Cloning, characterization, and mutational analysis of a highly active and stable l-arabinitol 4-dehydrogenase from *Neurospora crassa*. *Applied Microbiology and Biotechnology* **77**, 845-852 (2007).
27. Bae, B., Sullivan, R.P., Zhao, H. & Nair, S.K. Structure and engineering of L-arabinitol 4-dehydrogenase from *Neurospora crassa*. *J Mol Biol* **402**, 230-240 (2010).
28. Bera, A., Sedlak, M., Khan, A. & Ho, N.Y. Establishment of l-arabinose fermentation in glucose/xylose co-fermenting recombinant *Saccharomyces cerevisiae* 424A(LNH-ST) by genetic engineering. *Applied Microbiology and Biotechnology* **87**, 1803-1811 (2010).

29. Suzuki, T. et al. Cloning and expression of NAD+-dependent L-arabinitol 4-dehydrogenase gene (ladA) of *Aspergillus oryzae*. *J Biosci Bioeng* **100**, 472-474 (2005).
30. Mojzita, D., Vuoristo, K., Koivistoinen, O.M., Penttila, M. & Richard, P. The 'true' L-xylulose reductase of filamentous fungi identified in *Aspergillus niger*. *FEBS Lett* **584**, 3540-3544 (2010).
31. Gardonyi, M. & Hahn-Hagerdal, B. The *Streptomyces rubiginosus* xylose isomerase is misfolded when expressed in *Saccharomyces cerevisiae*. *Enzyme Microb Technol* **32**, 252 - 259 (2003).
32. Brat, D., Boles, E. & Wiedemann, B. Functional expression of a bacterial xylose isomerase in *Saccharomyces cerevisiae*. *Appl. Environ. Microbiol.* **75**, 2304-2311 (2009).
33. Kuyper, M. et al. High-level functional expression of a fungal xylose isomerase: the key to efficient ethanolic fermentation of xylose by *Saccharomyces cerevisiae*. *FEMS Yeast Research* **4**, 69-78 (2003).
34. Madhavan, A. et al. Xylose isomerase from polycentric fungus *Orpinomyces*: gene sequencing, cloning, and expression in *Saccharomyces cerevisiae* for bioconversion of xylose to ethanol. *Applied Microbiology and Biotechnology* **82**, 1067-1078 (2009).
35. Lee, S.-M., Jellison, T. & Alper, H.S. Directed evolution of xylose isomerase for improved xylose catabolism and fermentation in the yeast *Saccharomyces cerevisiae*. *Applied and Environmental Microbiology* **78**, 5708-5716 (2012).
36. Wisselink, H.W. et al. Engineering of *Saccharomyces cerevisiae* for Efficient Anaerobic Alcoholic Fermentation of l-Arabinose. *Applied and Environmental Microbiology* **73**, 4881-4891 (2007).
37. Ho, N.W.Y., Chen, Z. & Brainard, A.P. Genetically Engineered *Saccharomyces* Yeast Capable of Effective Cofermentation of Glucose and Xylose. *Applied and Environmental Microbiology* **64**, 1852-1859 (1998).
38. Ho, N.Y., Chen, Z., Brainard, A. & Sedlak, M. in Recent Progress in Bioconversion of Lignocellulosics, Vol. 65. (eds. G.T. Tsao et al.) 163-192 (Springer Berlin Heidelberg, 1999).
39. Eliasson, A., Christensson, C., Wahlbom, C. & Hahn-Hagerdal, B. Anaerobic xylose fermentation by recombinant *Saccharomyces cerevisiae* carrying *XYL1*, *XYL2*, and *XKS1* in mineral medium chemostat cultures. *Appl Environ Microbiol* **66**, 3381 - 3386 (2000).
40. Richard, P., Toivari, M.H. & Penttilä, M. The role of xylulokinase in *Saccharomyces cerevisiae* xylulose catabolism. *FEMS Microbiology Letters* **190**, 39-43 (2000).
41. Jin, Y.-S., Ni, H., Laplaza, J.M. & Jeffries, T.W. Optimal growth and ethanol production from xylose by recombinant *Saccharomyces cerevisiae* require moderate D-xylulokinase activity. *Appl. Environ. Microbiol.* **69**, 495-503 (2003).
42. Matsushika, A. & Sawayama, S. Efficient bioethanol production from xylose by recombinant *Saccharomyces cerevisiae* requires high activity of xylose reductase

and moderate xylulokinase activity. *Journal of Bioscience and Bioengineering* **106**, 306-309 (2008).
43. Metzger, M.H. & Hollenberg, C.P. Isolation and characterization of the *Pichia stipitis* transketolase gene and expression in a xylose-utilising Saccharomyces cerevisiae transformant. *Applied Microbiology and Biotechnology* **42**, 319-325 (1994).
44. Walfridsson, M., Hallborn, J., Penttila, M., Keranen, S. & Hahn-Hagerdal, B. Xylose-metabolizing *Saccharomyces cerevisiae* strains overexpressing the *TKL1* and *TAL1* genes encoding the pentose phosphate pathway enzymes transketolase and transaldolase. *Appl Environ Microbiol* **61**, 4184-4190 (1995).
45. Jin, Y.-S., Alper, H., Yang, Y.-T. & Stephanopoulos, G. Improvement of xylose uptake and ethanol production in recombinant *Saccharomyces cerevisiae* through an inverse metabolic engineering approach. *Appl. Environ. Microbiol.* **71**, 8249-8256 (2005).
46. Lagunas, R. Sugar transport in *Saccharomyces cerevisiae*. *FEMS Microbiol Rev* **10**, 229-242 (1993).
47. Andre, B. An overview of membrane transport proteins in *Saccharomyces cerevisiae*. *Yeast* **11**, 1575-1611 (1995).
48. Boles, E. & Hollenberg, C.P. The molecular genetics of hexose transport in yeasts. *FEMS Microbiol Rev* **21**, 85-111 (1997).
49. Pao, S.S., Paulsen, I.T. & Saier, M.H., Jr. Major facilitator superfamily. *Microbiol Mol Biol Rev* **62**, 1-34 (1998).
50. Ozcan, S. & Johnston, M. Function and regulation of yeast hexose transporters. *Microbiol Mol Biol Rev* **63**, 554-569 (1999).
51. Hamacher, T., Becker, J., Gardonyi, M., Hahn-Hagerdal, B. & Boles, E. Characterization of the xylose-transporting properties of yeast hexose transporters and their influence on xylose utilization. *Microbiology* **148**, 2783-2788 (2002).
52. Weierstall, T., Hollenberg, C.P. & Boles, E. Cloning and characterization of three genes (*SUT1-3*) encoding glucose transporters of the yeast *Pichia stipitis*. *Mol Microbiol* **31**, 871-883 (1999).
53. Leandro, M.J., Goncalves, P. & Spencer-Martins, I. Two glucose/xylose transporter genes from the yeast *Candida intermedia*: first molecular characterization of a yeast xylose-H+ symporter. *Biochem J* **395**, 543-549 (2006).
54. Saloheimo, A. et al. Xylose transport studies with xylose-utilizing *Saccharomyces cerevisiae* strains expressing heterologous and homologous permeases. *Applied Microbiology and Biotechnology* **74**, 1041-1052 (2007).
55. Hector, R.E., Qureshi, N., Hughes, S.R. & Cotta, M.A. Expression of a heterologous xylose transporter in a *Saccharomyces cerevisiae* strain engineered to utilize xylose improves aerobic xylose consumption. *Applied Microbiology and Biotechnology* **80**, 675-684 (2008).
56. Du, J., Li, S. & Zhao, H. Discovery and characterization of novel d-xylose-specific transporters from *Neurospora crassa* and *Pichia stipitis*. *Mol Biosyst* **6**, 2150-2156 (2010).

57. Galazka, J.M. et al. Cellodextrin transport in yeast for improved biofuel production. *Science* **330**, 84-86 (2010).
58. Runquist, D., Hahn-Hagerdal, B. & Radstrom, P. Comparison of heterologous xylose transporters in recombinant *Saccharomyces cerevisiae*. *Biotechnol Biofuels* **3**, 1754-6834 (2010).
59. Subtil, T. & Boles, E. Improving L-arabinose utilization of pentose fermenting *Saccharomyces cerevisiae* cells by heterologous expression of L-arabinose transporting sugar transporters. *Biotechnol. Biofuels* **4**, 38 (2011).
60. Verho, R., Penttila, M. & Richard, P. Cloning of two genes (*LAT1,2*) encoding specific L: -arabinose transporters of the L: -arabinose fermenting yeast *Ambrosiozyma monospora*. *Appl Biochem Biotechnol* **164**, 604-611 (2011).
61. Young, E.M., Comer, A.D., Huang, H. & Alper, H.S. A molecular transporter engineering approach to improving xylose catabolism in *Saccharomyces cerevisiae*. *Metabolic Engineering* **14**, 401-411 (2012).
62. Young, E.M., Tong, A., Bui, H., Spofford, C. & Alper, H.S. Rewiring yeast sugar transporter preference through modifying a conserved protein motif. *PNAS* **111**, 131-136 (2014).
63. Sun, J., Wen, F., Si, T., Xu, J.H. & Zhao, H. Direct conversion of xylan to ethanol by recombinant *Saccharomyces cerevisiae* strains displaying an engineered minihemicellulosome. *Appl Environ Microbiol* **78**, 3837-3845 (2012).
64. La Grange, D.C., Pretorius, I.S., Claeyssens, M. & van Zyl, W.H. Degradation of Xylan to d-Xylose by Recombinant Saccharomyces cerevisiae Coexpressing theAspergillus niger β-Xylosidase (xlnD) and the Trichoderma reesei Xylanase II (xyn2) Genes. *Applied and Environmental Microbiology* **67**, 5512-5519 (2001).
65. la Grange, D.C., Pretorius, I.S. & van Zyl, W.H. Expression of a Trichoderma reesei beta-xylanase gene (XYN2) in Saccharomyces cerevisiae. *Applied and Environmental Microbiology* **62**, 1036-1044 (1996).
66. Zhang, J., Siika-aho, M., Tenkanen, M. & Viikari, L. The role of acetyl xylan esterase in the solubilization of xylan and enzymatic hydrolysis of wheat straw and giant reed. *Biotechnology for Biofuels* **4**, 60 (2011).
67. Geiser, E., Wierckx, N., Zimmermann, M. & Blank, L. Identification of an endo-1,4-beta-xylanase of *Ustilago maydis*. *BMC Biotechnology* **13**, 59 (2013).
68. Steen, E. et al. Metabolic engineering of *Saccharomyces cerevisiae* for the production of n-butanol. *Microbial Cell Factories* **7**, 36 (2008).
69. Kondo, T. et al. Genetic engineering to enhance the Ehrlich pathway and alter carbon flux for increased isobutanol production from glucose by *Saccharomyces cerevisiae*. *J Biotechnol* **159**, 32-37 (2012).
70. Brat, D., Weber, C., Lorenzen, W., Bode, H. & Boles, E. Cytosolic re-localization and optimization of valine synthesis and catabolism enables inseased isobutanol production with the yeast *Saccharomyces cerevisiae*. *Biotechnology for Biofuels* **5**, 1-16 (2012).
71. Shen, C.R. et al. Driving forces enable high-titer anaerobic 1-butanol synthesis in *Escherichia coli*. *Appl Environ Microbiol* **77**, 2905-2915 (2011).

72. Nielsen, J. & Jewett, M.C. Impact of systems biology on metabolic engineering of *Saccharomyces cerevisiae*. *FEMS Yeast Research* **8**, 122-131 (2008).
73. Blount, B.A., Weenink, T. & Ellis, T. Construction of synthetic regulatory networks in yeast. *FEBS Letters*.
74. Santos, C.N.S. & Stephanopoulos, G. Combinatorial engineering of microbes for optimizing cellular phenotype. *Current Opinion in Chemical Biology* **12**, 168-176 (2008).
75. Zotchev, S.B., Sekurova, O.N. & Katz, L. Genome-based bioprospecting of microbes for new therapeutics. *Current Opinion in Biotechnology* **23**, 941-947 (2012).
76. Krivoruchko, A., Siewers, V. & Nielsen, J. Opportunities for yeast metabolic engineering: Lessons from synthetic biology. *Biotechnology Journal* (2011).
77. Sarthy, A. et al. Expression of the *Escherichia coli* xylose isomerase gene in *Saccharomyces cerevisiae*. *Appl Environ Microbiol* **53**, 1996 - 2000 (1987).
78. Amore, R., Wilhelm, M. & Hollenberg, C. The fermentation of xylose - an analysis of the expression of *Bacillus* and *Actinoplanes* xylose isomerase genes in yeast. *Appl Microbiol Biotechnol* **30**, 351 - 357 (1989).
79. Bloom, J.D. & Arnold, F.H. In the light of directed evolution: Pathways of adaptive protein evolution. *Proceedings of the National Academy of Sciences* **106**, 9995-10000 (2009).
80. Bera, A., Ho, N., Khan, A. & Sedlak, M. A genetic overhaul of *Saccharomyces cerevisiae* 424A(LNH-ST) to improve xylose fermentation. *J Ind Microbiol Biotechnol* **38**, 617-626 (2011).
81. Garcia Sanchez, R. et al. Improved xylose and arabinose utilization by an industrial recombinant *Saccharomyces cerevisiae* strain using evolutionary engineering. *Biotechnology for Biofuels* **3**, 13 (2010).
82. Hahn-Hagerdal, B., Karhumaa, K., Jeppsson, M. & Gorwa-Grauslund, M. Metabolic engineering for pentose utilization in *Saccharomyces cerevisiae*. *Adv Biochem Eng Biotechnol* **108**, 147 - 177 (2007).
83. Kuyper, M. et al. Metabolic engineering of a xylose-isomerase-expressing *Saccharomyces cerevisiae* strain for rapid anaerobic xylose fermentation. *FEMS Yeast Research* **5**, 399-409 (2005).
84. Young, E., Lee, S.-M. & Alper, H. Optimizing pentose utilization in yeast: the need for novel tools and approaches. *Biotechnology for Biofuels* **3**, 24 (2010).
85. Van Vleet, J.H., Jeffries, T.W. & Olsson, L. Deleting the para-nitrophenyl phosphatase (pNPPase), *PHO13*, in recombinant *Saccharomyces cerevisiae* improves growth and ethanol production on d-xylose. *Metabolic Engineering* **10**, 360-369 (2008).
86. Çakar, Z.P., Turanlı-Yıldız, B., Alkım, C. & Yılmaz, Ü. Evolutionary engineering of *Saccharomyces cerevisiae* for improved industrially important properties. *FEMS Yeast Research* **12**, 171-182 (2012).
87. Hong, K.-K., Vongsangnak, W., Vemuri, G.N. & Nielsen, J. Unravelling evolutionary strategies of yeast for improving galactose utilization through

integrated systems level analysis. *Proceedings of the National Academy of Sciences* **108**, 12179-12184 (2011).
88. Argueso, J.L. et al. Genome structure of a *Saccharomyces cerevisiae* strain widely used in bioethanol production. *Genome Research* **19**, 2258-2270 (2009).
89. Alper, H., Moxley, J., Nevoigt, E., Fink, G.R. & Stephanopoulos, G. Engineering Yeast Transcription Machinery for Improved Ethanol Tolerance and Production. *Science* **314**, 1565-1568 (2006).
90. Zhou, H., Cheng, J.-s., Wang, B.L., Fink, G.R. & Stephanopoulos, G. Xylose isomerase overexpression along with engineering of the pentose phosphate pathway and evolutionary engineering enable rapid xylose utilization and ethanol production by *Saccharomyces cerevisiae*. *Metabolic Engineering* **14**, 611-622 (2012).
91. Kuyper, M. et al. Evolutionary engineering of mixed-sugar utilization by a xylose-fermenting *Saccharomyces cerevisiae* strain. *FEMS Yeast Research* **5**, 925-934 (2005).
92. Shen, Y. et al. An efficient xylose-fermenting recombinant *Saccharomyces cerevisiae* strain obtained through adaptive evolution and its global transcription profile. *Applied Microbiology and Biotechnology* **96**, 1079-1091 (2012).
93. Hahn-Hägerdal, B., Karhumaa, K., Fonseca, C., Spencer-Martins, I. & Gorwa-Grauslund, M. Towards industrial pentose-fermenting yeast strains. *Applied Microbiology and Biotechnology* **74**, 937-953 (2007).
94. Sonderegger, M. & Sauer, U. Evolutionary Engineering of *Saccharomyces cerevisiae* for Anaerobic Growth on Xylose. *Appl. Environ. Microbiol.* **69**, 1990-1998 (2003).
95. Matsushika, A., Inoue, H., Kodaki, T. & Sawayama, S. Ethanol production from xylose in engineered *Saccharomyces cerevisiae* strains: current state and perspectives. *Applied Microbiology and Biotechnology* **84**, 37-53 (2009).
96. Lu, C. & Jeffries, T. Shuffling of promoters for multiple genes to optimize xylose fermentation in an engineered *Saccharomyces cerevisiae* strain. *Appl. Environ. Microbiol.* **73**, 6072-6077 (2007).
97. Runquist, D., Hahn-Hagerdal, B. & Bettiga, M. Increased ethanol productivity in xylose-utilizing *Saccharomyces cerevisiae* via a randomly mutagenized xylose reductase. *Appl. Environ. Microbiol.* **76**, 7796-7802 (2010).
98. Watanabe, S. et al. Ethanol production from xylose by recombinant *Saccharomyces cerevisiae* expressing protein-engineered NADH-preferring xylose reductase from *Pichia stipitis*. *Microbiology* **153**, 3044-3054 (2007).
99. Bikard, D., Julié-Galau, S., Cambray, G. & Mazel, D. The synthetic integron: an in vivo genetic shuffling device. *Nucleic Acids Research* **38**, e153 (2010).
100. Karhumaa, K., Sanchez, R., Hahn-Hagerdal, B. & Gorwa-Grauslund, M.-F. Comparison of the xylose reductase-xylitol dehydrogenase and the xylose isomerase pathways for xylose fermentation by recombinant *Saccharomyces cerevisiae*. *Microbial Cell Factories* **6**, 5 (2007).

101. Madhavan, A. et al. Alcoholic fermentation of xylose and mixed sugars using recombinant *Saccharomyces cerevisiae* engineered for xylose utilization. *Applied Microbiology and Biotechnology* **82**, 1037-1047 (2009).
102. Parachin, N. & Gorwa-Grauslund, M. Isolation of xylose isomerases by sequence- and function-based screening from a soil metagenomic library. *Biotechnology for Biofuels* **4**, 9 (2011).
103. Bettiga, M., Hahn-Hagerdal, B. & Gorwa-Grauslund, M. Comparing the xylose reductase/xylitol dehydrogenase and xylose isomerase pathways in arabinose and xylose fermenting *Saccharomyces cerevisiae* strains. *Biotechnology for Biofuels* **1**, 16 (2008).
104. Bhosale, S., Rao, M. & Deshpande, V. Molecular and industrial aspects of glucose isomerase. *Microbiol. Rev.* **60**, 280-300 (1996).
105. Meng, M., Bagdasarian, M. & Zeikus, J.G. Thermal stabilization of xylose isomerase from *Thermoanaerobacterium thermosulfurigenes*. *Nat Biotech* **11**, 1157-1161 (1993).
106. Sriprapundh, D., Vieille, C. & Zeikus, J.G. Molecular determinants of xylose isomerase thermal stability and activity: analysis of thermozymes by site-directed mutagenesis. *Protein Engineering* **13**, 259-265 (2000).
107. Lambeir, A.M. et al. Protein engineering of xylose (glucose) isomerase from *Actinoplanes missouriensis*. 2. Site-directed mutagenesis of the xylose binding site. *Biochemistry* **31**, 5459-5466 (1992).
108. Meng, M., Lee, C., Bagdasarian, M. & Zeikus, J.G. Switching substrate preference of thermophilic xylose isomerase from D-xylose to D-glucose by redesigning the substrate binding pocket. *Proceedings of the National Academy of Sciences* **88**, 4015-4019 (1991).
109. van Maris, A. et al. Development of efficient xylose fermentation in *Saccharomyces cerevisiae*: xylose isomerase as a key component. *Advances in Biochemical Engineering/Biotechnology* **108**, 179-204 (2007).
110. Walfridsson, M., Anderlund, M., Bao, X. & Hahn-Hagerdal, B. Expression of different levels of enzymes from the *Pichia stipitis XYL1* and *XYL2* genes in *Saccharomyces cerevisiae* and its effects on product formation during xylose utilisation. *Appl Microbiol Biotechnol* **48**, 218 - 224 (1997).
111. Lönn, A., Gárdonyi, M., van Zyl, W., Hahn-Hägerdal, B. & Otero, R.C. Cold adaptation of xylose isomerase from Thermus thermophilus through random PCR mutagenesis. *European Journal of Biochemistry* **269**, 157-163 (2002).
112. Parachin, N.S., Bergdahl, B., van Niel, E.W.J. & Gorwa-Grauslund, M.F. Kinetic modelling reveals current limitations in the production of ethanol from xylose by recombinant *Saccharomyces cerevisiae*. *Metabolic Engineering* **13**, 508-517 (2011).
113. Traff, K., Otero Cordero, R., van Zyl, W. & Hahn-Hagerdal, B. Deletion of the *GRE3* aldose reductase gene and its influence on xylose metabolism in recombinant strains of *Saccharomyces cerevisiae* expressing the *xylA* and *XKS1* genes. *Appl Environ Microbiol* **67**, 5668 - 5674 (2001).

114. Kuyper, M., Winkler, A.A., van Dijken, J.P. & Pronk, J.T. Minimal metabolic engineering of *Saccharomyces cerevisiae* for efficient anaerobic xylose fermentation: a proof of principle. *FEMS Yeast Research* **4**, 655-664 (2004).
115. Kersters-Hilderson, H., Callens, M., Van Opstal, O., Vangrysperre, W. & De Bruyne, C.K. Kinetic characterization of d-xylose isomerases by enzymatic assays using d-sorbitol dehydrogenase. *Enzyme and Microbial Technology* **9**, 145-148 (1987).
116. Mumberg, D., Müller, R. & Funk, M. Yeast vectors for the controlled expression of heterologous proteins in different genetic backgrounds. *Gene* **156**, 119-122 (1995).
117. Chayen, N.E., Conti, E., Vielle, C. & Zeikus, J.G. Crystallization and initial X-ray analysis of xylose isomerase from *Thermotoga neapolitana*. *Acta Crystallographica Section D* **53**, 229-230 (1997).
118. Lavie, A., Allen, K.N., Petsko, G.A. & Ringe, D. X-ray crystallographic structures of D-xylose isomerase-substrate complexes position the substrate and provide evidence for metal movement during catalysis. *Biochemistry* **33**, 5469-5480 (1994).
119. Wiebe, M.G. et al. Central carbon metabolism of *Saccharomyces cerevisiae* in anaerobic, oxygen-limited and fully aerobic steady-state conditions and following a shift to anaerobic conditions. *FEMS Yeast Research* **8**, 140-154 (2008).
120. Sánchez, Ó.J. & Cardona, C.A. Trends in biotechnological production of fuel ethanol from different feedstocks. *Bioresource Technology* **99**, 5270-5295 (2008).
121. Karhumaa, K., Hahn-Hägerdal, B. & Gorwa-Grauslund, M.-F. Investigation of limiting metabolic steps in the utilization of xylose by recombinant *Saccharomyces cerevisiae* using metabolic engineering. *Yeast* **22**, 359-368 (2005).
122. Fujitomi, K., Sanda, T., Hasunuma, T. & Kondo, A. Deletion of the *PHO13* gene in *Saccharomyces cerevisiae* improves ethanol production from lignocellulosic hydrolysate in the presence of acetic and formic acids, and furfural. *Bioresource Technology* **111**, 161-166 (2012).
123. Kim, S.R. et al. Rational and evolutionary engineering approaches uncover a small set of genetic changes efficient for rapid xylose fermentation in *Saccharomyces cerevisiae*. *PLoS One* **8**, 26 (2013).
124. Wei, N., Quarterman, J., Kim, S.R., Cate, J.H.D. & Jin, Y.-S. Enhanced biofuel production through coupled acetic acid and xylose consumption by engineered yeast. *Nature communications* **4** (2013).
125. Portnoy, V.A., Bezdan, D. & Zengler, K. Adaptive laboratory evolution — harnessing the power of biology for metabolic engineering. *Current Opinion in Biotechnology* **22**, 590-594 (2011).
126. Bae, J., Laplaza, J. & Jeffries, T. Effects of gene orientation and use of multiple promoters on the expression of *XYL1* and *XYL2* in *Saccharomyces cerevisiae*. *Appl Biochem Biotechnol* **145**, 69-78 (2008).

127. Karhumaa, K., Påhlman, A.-K., Hahn-Hägerdal, B., Levander, F. & Gorwa-Grauslund, M.-F. Proteome analysis of the xylose-fermenting mutant yeast strain TMB 3400. *Yeast* **26**, 371-382 (2009).
128. Krahulec, S., Klimacek, M. & Nidetzky, B. Analysis and prediction of the physiological effects of altered coenzyme specificity in xylose reductase and xylitol dehydrogenase during xylose fermentation by Saccharomyces cerevisiae. *Journal of Biotechnology* **158**, 192-202 (2012).
129. Scalcinati, G. et al. Evolutionary engineering of Saccharomyces cerevisiae for efficient aerobic xylose consumption. *FEMS Yeast Research* **12**, 582-597 (2012).
130. Metz, B. et al. A Novel l-Xylulose Reductase Essential for l-Arabinose Catabolism in *Trichoderma reesei*. *Biochemistry* **52**, 2453-2460 (2013).
131. Jeffries, T.W. Engineering yeasts for xylose metabolism. *Current Opinion in Biotechnology* **17**, 320-326 (2006).
132. Wei, N., Quarterman, J., Kim, S.R., Cate, J.H. & Jin, Y.S. Enhanced biofuel production through coupled acetic acid and xylose consumption by engineered yeast. *Nat. Commun.* **4** (2013).
133. Takuma, S. et al. Isolation of xylose reductase gene ofPichia stipitis and its expression in *Saccharomyces cerevisiae*. *Appl Biochem Biotechnol* **28-29**, 327-340 (1991).
134. Harhangi, H.R. et al. Xylose metabolism in the anaerobic fungus *Piromyces sp.* strain E2 follows the bacterial pathway. *Archives of Microbiology* **180**, 134-141 (2003).
135. Young, E., Poucher, A., Comer, A., Bailey, A. & Alper, H. Functional survey for heterologous sugar transport proteins, using *Saccharomyces cerevisiae* as a host. *App.l Environ. Microbiol.* **77**, 3311-3319 (2011).
136. Lee, S.J. et al. Metabolic engineering of *Escherichia coli* for enhanced production of succinic acid, based on genome comparison and in silico gene knockout simulation. *Appl Environ Microbiol* **71**, 7880-7887 (2005).
137. Smit, E., Leeflang, P., Glandorf, B., Dirk van Elsas, J. & Wernars, K. Analysis of Fungal Diversity in the Wheat Rhizosphere by Sequencing of Cloned PCR-Amplified Genes Encoding 18S rRNA and Temperature Gradient Gel Electrophoresis. *Applied and Environmental Microbiology* **65**, 2614-2621 (1999).
138. Carbo, N. & Perez-Martin, J. Activation of the cell wall integrity pathway promotes escape from G2 in the fungus *Ustilago maydis*. *PLoS Genet* **6**, 1001009 (2010).
139. Kamper, J. et al. Insights from the genome of the biotrophic fungal plant pathogen *Ustilago maydis*. *Nature* **444**, 97-101 (2006).
140. Gibson, D.M., King, B.C., Hayes, M.L. & Bergstrom, G.C. Plant pathogens as a source of diverse enzymes for lignocellulose digestion. *Current Opinion in Microbiology* **14**, 264-270 (2011).
141. Couturier, M. et al. Post-genomic analyses of fungal lignocellulosic biomass degradation reveal the unexpected potential of the plant pathogen *Ustilago maydis*. *BMC Genomics* **13**, 57 (2012).

142. Zerbino, D.R. & Birney, E. Velvet: Algorithms for de novo short read assembly using de Bruijn graphs. *Genome Research* **18**, 821-829 (2008).
143. Holt, C. & Yandell, M. MAKER2: an annotation pipeline and genome-database management tool for second-generation genome projects. *BMC Bioinformatics* **12**, 491 (2011).
144. Conesa, A. & Gotz, S. Blast2GO: A comprehensive suite for functional analysis in plant genomics. *Int J Plant Genomics* **619832**, 619832 (2008).
145. Storck, R. Nucleotide Composition of Nucleic Acids of Fungi II. Deoxyribonucleic Acids. *Journal of Bacteriology* **91**, 227-230 (1966).
146. Jeffries, T.W. et al. Genome sequence of the lignocellulose-bioconverting and xylose-fermenting yeast *Pichia stipitis*. *Nat Biotech* **25**, 319-326 (2007).
147. Martinez, D. et al. Genome sequencing and analysis of the biomass-degrading fungus *Trichoderma reesei* (syn. *Hypocrea jecorina*). *Nat Biotechnol* **26**, 553-560 (2008).
148. Jingping, G., Hongbing, S., Gang, S., Hongzhi, L. & Wenxiang, P. A genome shuffling-generated *Saccharomyces cerevisiae* isolate that ferments xylose and glucose to produce high levels of ethanol. *J Ind Microbiol Biotechnol* **39**, 777-787 (2012).
149. Tantirungkij, M., Nakashima, N., Seki, T. & Yoshida, T. Construction of xylose-assimilating *Saccharomyces cerevisiae*. *J Ferment Bioeng* **75**, 83 - 88 (1993).
150. Johansson, B., Christensson, C., Hobley, T. & Hahn-Hagerdal, B. Xylulokinase Overexpression in Two Strains of *Saccharomyces cerevisiae* Also Expressing Xylose Reductase and Xylitol Dehydrogenase and Its Effect on Fermentation of Xylose and Lignocellulosic Hydrolysate. *Appl. Environ. Microbiol.* **67**, 4249-4255 (2001).
151. Jin, Y.-S., Jones, S., Shi, N.-Q. & Jeffries, T.W. Molecular Cloning of *XYL3* (D-Xylulokinase) from *Pichia stipitis* and Characterization of Its Physiological Function. *Appl. Environ. Microbiol.* **68**, 1232-1239 (2002).
152. Shallom, D. & Shoham, Y. Microbial hemicellulases. *Current Opinion in Microbiology* **6**, 219-228 (2003).
153. Sambrook, J., Fritch, E. & Maniatis, T. Molecular cloning: a laboratory manual. 2nd ed. (1989).
154. Schwede, T., Kopp, J., Guex, N. & Peitsch, M.C. SWISS-MODEL: an automated protein homology-modeling server. *Nucleic Acids Research* **31**, 3381-3385 (2003).
155. Kelley, L.A. & Sternberg, M.J.E. Protein structure prediction on the Web: a case study using the Phyre server. *Nat. Protocols* **4**, 363-371 (2009).
156. Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. Basic local alignment search tool. *J Mol Biol* **215**, 403-410 (1990).
157. Giaever, G. et al. Functional profiling of the *Saccharomyces cerevisiae* genome. *Nature* **418**, 387-391 (2002).