

Flux balance analysis predicts NADP phosphatase and NADH kinase are critical to balancing redox during xylose fermentation in *Scheffersomyces stipitis*

Kevin Correia¹, Anna Khusnutdinova¹, Peter Yan Li¹, Jeong Chan Joo¹, Greg Brown¹, Alexander F. Yakunin¹, and Radhakrishnan Mahadevan^{1,2,*}

¹Department of Chemical Engineering and Applied Chemistry, University of Toronto, Canada, ON

²Institute of Biomaterials and Biomedical Engineering, University of Toronto, Ontario, Canada

Corresponding author:
Radhakrishnan Mahadevan*

Email address: krishna.mahadevan@utoronto.ca

ABSTRACT

Xylose is the second most abundant sugar in lignocellulose and can be used as a feedstock for next-generation biofuels by industry. *Saccharomyces cerevisiae*, one of the main workhorses in biotechnology, is unable to metabolize xylose natively but has been engineered to ferment xylose to ethanol with the xylose reductase (XR) and xylitol dehydrogenase (XDH) genes from *Scheffersomyces stipitis*. In the scientific literature, the yield and volumetric productivity of xylose fermentation to ethanol in engineered *S. cerevisiae* still lags *S. stipitis*, despite expressing of the same XR-XDH genes. These contrasting phenotypes can be due to differences in *S. cerevisiae*'s redox metabolism that hinders xylose fermentation, differences in *S. stipitis*' redox metabolism that promotes xylose fermentation, or both. To help elucidate how *S. stipitis* ferments xylose, we used flux balance analysis to test various redox balancing mechanisms, reviewed published omics datasets, and studied the phylogeny of key genes in xylose fermentation. *In vivo* and *in silico* xylose fermentation cannot be reconciled without NADP phosphatase (NADPase) and NADH kinase. We identified eight candidate genes for NADPase. *PHO3.2* was the sole candidate showing evidence of expression during xylose fermentation. *Pho3.2p* and *Pho3p*, a recent paralog, were purified and characterized for their substrate preferences. Only *Pho3.2p* was found to have NADPase activity. Both NADPase and NAD(P)H-dependent XR emerged from recent duplications in a common ancestor of *Scheffersomyces* and *Spathaspora* to enable efficient xylose fermentation to ethanol. This study demonstrates the advantages of using metabolic simulations, omics data, bioinformatics, and enzymology to reverse engineer metabolism.

1 INTRODUCTION

Xylose is the second most abundant sugar in lignocellulose and can be used as a carbon source for biofuels and biochemicals by industry (Jeffries and Jin, 2004). Known catabolic pathways include xylose isomerase (XI) (Schellenberg et al., 1984), the xylose reductase (XR)-xylitol dehydrogenase (XDH) pathway (Horitsu et al., 1968), the Weimberg pathway (Weimberg, 1961), and the Dahms pathway (Dahms, 1974). *Saccharomyces cerevisiae*, one of the main workhorses of industrial biotechnology, has evolved to ferment glucose to ethanol rapidly but cannot natively grow on xylose or ferment it to ethanol

(Jeffries and Jin, 2004), even though it has XR and XDH genes. Screening of various yeasts has found that ethanol fermentation with xylose in yeasts is rare (Toivola et al., 1984) and requires NADH-linked XR to alleviate the cofactor imbalance in the XR-XDH pathway during oxygen limitation (Schneider et al., 1981; Slininger et al., 1982; Bruinenberg et al., 1984).

Early research on xylose fermentation in yeasts focused on process optimization of native xylose fermenters (Slininger et al., 1985, 1990; du Preez, 1994), but advancements in genetic engineering enabled the expression of bacterial XI (Sarthý et al., 1987; Amore et al., 1989) and the XR-XDH pathway in *S. cerevisiae* (Kötter and Ciriacy, 1993). *S. cerevisiae* engineered with the XR-XDH pathway from *S. stipitis*, encoded by *XYL1* and *XYL2*, have typically outperformed strains expressing bacterial XI. More recently, evolved strains of *S. cerevisiae* expressing fungal XI has led to faster growth rates and fewer byproducts (Zhou et al., 2012; Verhoeven et al., 2017). Kwak and Jin (2017) provide a review of engineering xylose fermentation in *S. cerevisiae*. After more than 30 years of engineering xylose fermentation in *S. cerevisiae*, the yield and volumetric productivity of engineered *S. cerevisiae* with XI or XR-XDH have lagged native xylose fermenters like *S. stipitis* and *S. passalidarum* in the scientific literature (van Vleet and Jeffries, 2009; Kim et al., 2013a). The inability of *S. cerevisiae*, engineered the XR-XDH pathway from *S. stipitis*, to anaerobically ferment xylose to ethanol at high yields is especially puzzling because the same genes appear to enable xylose fermentation in wild-type *S. stipitis* (Wahlbom et al., 2003). These contrasting phenotypes can be due to differences in the transcriptome, proteome, and metabolome of *S. cerevisiae* that hinder xylose fermentation, differences in the transcriptome, proteome, and metabolome of *S. stipitis* that promote xylose fermentation, or a combination of both. The yeast research community has largely focused on targets in *S. cerevisiae* (van Vleet et al., 2008; Wei et al., 2013; Kim et al., 2013b), while few studies have probed for targets in *S. stipitis* beyond central metabolism (Jeppsson et al., 1995; Freese et al., 2011; Wohlbach et al., 2011). Although the XI pathway has less technical challenges for industrial fermentation than the XR-XDH pathway in *S. cerevisiae*, the redox balancing of the XR-XDH pathway in *S. stipitis* has not been fully elucidated and may offer insight into new redox balancing strategies.

Flux balance analysis (FBA) is a computational method often used to gain insight into metabolism (Österlund et al., 2013; McCloskey et al., 2013), and is well suited to study redox metabolism in yeasts (Pereira et al., 2016). To date, there are four genome-scale network reconstructions (GENRE's) of *S. stipitis*, the most widely studied xylose fermenting yeast: *iBB814* (Balagurunathan et al., 2012), *iSS884* (Caspeta et al., 2012), *iTL885* (Liu et al., 2012), and *iPL912* (Li, 2012). Our xylose fermentation simulations with the four models led to xylitol accumulation when we forced the *in vitro* XR cofactor selectivity (60% NADPH), removed cytosolic NADP-dependent acetaldehyde dehydrogenase from the metabolic models, which is not encoded in *S. stipitis*' genome (Correia et al., 2017), prevented flux

through degradation pathways and considered alternative optima. Failing to reconcile *in silico* and *in vivo* xylose fermentation in *S. stipitis*, we created a consensus GENRE for *S. stipitis* to analyze redox balancing mechanisms during xylose fermentation (Figure 1), and reviewed published omics datasets to guide our understanding of potential flux constraints.

The only mechanisms that were able to eliminate *in silico* xylitol accumulation during anaerobic xylose fermentation were phosphorylating NADP-dependent glyceraldehyde 3-phosphate dehydrogenase (GAPDH), non-phosphorylating NADP-dependent GAPDH, or NADP phosphatase (NADPase) and NADH kinase. The expression of phosphorylating NADP-dependent GAPDH, encoded by *K. lactis*' *GDPI* (Verho et al., 2002), in XR-XDH engineered *S. cerevisiae* decreased its xylitol yield and increased its ethanol yield (Verho et al., 2003); however, there is no strong biochemical or bioinformatic evidence of either form of GAPDH in *S. stipitis*. Expression of cytosolic NADH kinase led to an increase in the xylitol yield of XR-XDH engineered *S. cerevisiae* (Hou et al., 2009), but our simulations indicate that NADPase and NADH kinase are both required to completely balance redox cofactors.

The presence of NADPase in *S. cerevisiae* or *S. stipitis* is unknown since it is a eukaryotic orphan enzyme. Eight NADPase candidate genes were identified in *iSS885* and *iPL912*, which had functional annotations from KEGG Orthology (Kanehisa et al., 2011) and PathwayTools (Karp et al., 2009), respectively. The genes encoding the NADPase candidates, the XR-XDH pathway, and enzymes related to NADPH regeneration are outlined in Table 1, along with expression data (Yuan et al., 2011; Huang and Lefsrud, 2012) and orthology information (Correia et al., 2017). *PHO3.2* was the most promising candidate since *S. cerevisiae* does not encode any homologs (Table 1), and its expression in *S. stipitis* was confirmed during xylose fermentation via shotgun proteomics (Huang and Lefsrud, 2012). Therefore, we expressed Pho3.2p and Pho3p, its paralog with 77% identity, in *Komagataella phaffii*, purified them via 6xHis tags, and characterized their substrate preferences. We also studied the phylogenetic origin of our proposed mechanism in the *Scheffersomyces-Spathaspora* clade. We propose NADPase and NADH kinase are critical to xylose fermentation in *S. stipitis* since they can balance redox cofactors in the absence of oxygen (Figure 2). In contrast, *S. cerevisiae* requires oxygen to balance redox, and loses CO₂ when the oxidative pentose phosphate pathway regenerates NADPH.

2 METHODS

Genome-scale network reconstruction and analysis. The *S. stipitis* GENRE was obtained from a pan-fungal GENRE that combined the GENRE's of *S. stipitis* (Balagurunathan et al., 2012; Caspeta et al., 2012; Li, 2012; Liu et al., 2012), *Schizosaccharomyces pombe* (Sohn et al., 2012), *Aspergillus niger* (Andersen et al., 2008), *Yarrowia lipolytica* (Pan and Hua, 2012; Loira et al., 2012), *Komagataella phaffii* (Caspeta

et al., 2012), *Kluyveromyces lactis* (Dias et al., 2014), and *S. cerevisiae* (Heavner et al., 2013). Model simulations were carried out using CONstraints Based Reconstruction and Analysis for Python version 0.9.1 (COBRApy) (Ebrahim et al., 2013). The xylose uptake rate was set to $10 \text{ mmol} \cdot \text{gDCW}^{-1} \cdot \text{h}^{-1}$. The growth associated maintenance (GAM) and non-growth associated maintenance (NGAM) were set to $60 \text{ mmol} \cdot \text{gDCW}^{-1}$ and $0 \text{ mmol} \cdot \text{gDCW}^{-1} \cdot \text{h}^{-1}$, respectively. Flux variability analysis (FVA) was used to evaluate alternative optimum solutions (Mahadevan and Schilling, 2003). The exchange bounds for erythritol, ribitol, arabitol, sorbitol, and glycerol were all set to zero to simplify the solution space for polyols. The cofactor selectivities of NADH and NADPH were varied in a single XR reaction (Balagurunathan et al., 2012). XR solely driven by NADH or NADPH were blocked. The ethanol yield or biomass growth rate were maximized depending on the simulation. We reviewed the data from all published transcriptomics and proteomics studies for *S. stipitis* to guide our understanding of its metabolism (Jeffries et al., 2007; Jeffries and Van Vleet, 2009; Wohlbach et al., 2011; Yuan et al., 2011; Huang and Lefsrud, 2012; Papini et al., 2012; Huang and Lefsrud, 2014).

Cloning *PHO3/PHO3.2*, and transformation in *K. phaffii*. *PHO3* and *PHO3.2* were amplified from *Scheffersomyces stipitis* CBS 5773 genomic DNA with Phusion polymerase (NEB), without their native signal peptides, as predicted by SignalP (Petersen et al., 2011). The genes were inserted into pPICZ α ,B using restriction enzyme digestion and ligation. Primer sequences can be found in Table S1. The plasmids were transformed into *Escherichia coli* BL21 using electroporation, selected on low salt LB (Lennox)-ZeocinTM ($10 \mu\text{g}/\text{mL}$), and sequence verified (the Centre for Applied Genomics, Toronto). The plasmids were linearly digested by PmeI, and $5 \mu\text{g}$ of each digestion were transformed into *K. phaffii* KM71H using electroporation (EasySelect *Pichia* Expression Kit, Invitrogen). Cells were recovered in 4 mL of YPD for 2 hours and plated on YPD agar with ZeocinTM ($100 \mu\text{g}/\text{mL}$).

Agar acid phosphatase assay. An acid phosphatase assay was used to screen for *K. phaffii* colonies with the highest expression of Pho3p and Pho3.2p (Dorn, 1965). Colonies were plated on BMMY agar and incubated at 30°C. $100 \mu\text{L}$ of 100% methanol was dispensed on the lid of each inverted Petri dish after 24 and 48 hours of growth. 20 mg of Fast Garnet GBC sulphate salt (MilliporeSigma) and 2 mg of 1-naphthyl phosphate disodium salt (MilliporeSigma) were dissolved in 4 mL of 0.6 M acetate buffer (pH 4.8). 4 mL of the solution was flooded into the Petri dish and examined for five minutes (Figure S1).

General phosphatase assay with para-nitrophenyl phosphate (pNPP). In addition to the agar acid phosphatase assay, a general phosphatase assay was used to monitor the activity of various *K. phaffii* clones for secreted phosphatase in liquid media (Kuznetsova et al., 2005). $100 \mu\text{L}$ of the fermentation broth from wild-type *K. phaffii* and mutants expressing Pho3p and Pho3.2p were assayed for pNPP phosphatase after 24 hours of induction. The supernatant enriched with phosphatase from each culture were prepared

by treatment with 0.1% of Triton X100 or Tween 20 on a rotator at +4°C for 30 minutes or sonication in 1 mL volume during 10 seconds on ice. The supernatant was separated from the cells by centrifugation at 13 000 rpm and a benchtop Eppendorf 5424 centrifuge. The cells were resuspended in an equal volume of BMMY media. The pNPP phosphatase assay was carried out in 200 µL reactions in sealed 96 well-plates incubated overnight at 30°C. The reaction mixture consisted of the phosphatase-enriched supernatant collected from the fermentation, 4 mM pNPP, 0.5 mM MnCl₂, 5 mM MgCl₂, and 100 mM HEPES pH 7.5. pNPP phosphatase activity was estimated due to the increasing absorbance of para-nitrophenyl at 410 nm; the absorbance of wild-type *K. phaffii* cultures was subtracted as background.

Protein expression and purification. The clones with the highest expression of Pho3p and Pho3.2p were grown in BMGY as described in EasySelect *Pichia* Expression Kit (Invitrogen™). 5 mL of 100% methanol was added at 24 and 48 hours after inoculation in 1 L of broth in a 4 L baffled flask. We failed to purify any secreted protein in the fermentation broth after 24, 48 and 72 hours after methanol induction; the final pH was adjusted to 7.5, buffered with 50 mM HEPES, 0.4 M NaCl and 5 mM imidazole for Ni-NTA binding. Most of the phosphatase activity was found to be associated with the cells (Figure S2). The cells were collected and stored at -20°C until purification. The cells were sonicated (Qsonica, dual horn probe, 2.5 minutes, 80% of maximal amplitude) to detach the acid phosphatase from the cell surface. The phosphatases relative sizes were estimated to be 55 kDa with glycosylation on 12% PAAG (Figure S3) and were sequence confirmed by mass spectrometry (Supplementary Information). The phosphatases were further purified with a Ni-NTA agarose (Quiagen) by their 6xHis tag according to manufacturer's protocol. It should be noted that longer sonication time or Y-PER Yeast Protein extraction reagent (Thermo) application increased contamination with an alcohol dehydrogenase (40 kDa band), identified as *K. phaffii* Adh2p (A0A1B2JBQ8) (Figure S3). Size exclusion chromatography (Superdex 10/300 GL) failed to separate the phosphatases from *K. phaffii* Adh2p. The proteins formed a tight dimer, with a relative size of 110 kDa (data not shown). We attempted to purify Utr1p from *S. cerevisiae* and *S. stipitis* using *E. coli* and *K. phaffii* as expression hosts, but were unable to collect soluble protein.

Alcohol dehydrogenase enzyme assay. The *K. phaffii* Adh2p contaminant was purified from wild-type *K. phaffii* and assayed for NADH and NADPH oxidation at 340 nm in 96 well-plates at 30 °C, with 0.01-10 mM butyraldehyde, 1 mM ZnCl₂, 50 mM HEPES pH -7.5. Both NADH and NADPH reduced butyraldehyde, but no butanol oxidation with NAD or NADP was detected.

NADPase phosphatase (NADPase) enzyme assay. NADPase activity was assayed in a coupled reaction with formate dehydrogenase (P33160). 2.9 µg of Pho3p and Pho3.2p were incubated at 30°C in 200 µL reactions with NADP, 50 mM sodium formate, 20 µg of strictly NAD-dependent formate dehydrogenase, 50 mM HEPES pH 7.5, 5 mM MgCl₂ and 0.5 mM MnCl₂. The reduction of NAD was

monitored at 340 nm in 96-well microplates. 10 mM butyraldehyde was used to estimate the concentration of the contaminated Pho3p and Pho3.2p Ni-NTA eluted samples via NADH-dependent butyraldehyde reductase activity.

Phosphatase screen with natural substrates using the malachite green assay. Pho3p and Pho3.2p were assayed for their substrate preferences with the malachite green assay (Kuznetsova et al., 2015). The assay was performed in 96-well microplates in 160 μ L reactions containing of 100 mM HEPES pH 7.5, 5 mM $MgCl_2$, 0.5 mM $MnCl_2$, and 3 μ g the acid phosphatase. 3.125 mM was the final substrate concentration of cytidine 2' monophosphate, inosine triphosphate, NADP, FMN, coenzyme A, glyphosate, adenosine 3',5' diphosphate. The remaining substrates were assayed at 6.25 mM: AMP, CMP, GMP, IMP, UMP, XMP, 2'AMP, 2'CMP, 3'AMP, 3'CMP, dAMP, dCMP, dGMP, dIMP, dTMP, dUMP, ADP, CDP, GDP, IDP, TDP, UDP, dADP, dCDP, dGDP, ATP, CTP, GTP, ITP, TTP, UTP, dATP, dCTP, dGTP, dITP, dUTP, NADP, FMN, PEP, CoA, phosphocholine, α -glucose 1-phosphate, β -glucose 1-phosphate, β -glucose 6-phosphate, fructose 1-phosphate, fructose 6-phosphate, ribose 5-phosphate, mannose 1-phosphate, mannose 6-phosphate, galactose 1-phosphate, fructose 1,6-bisphosphate, erythrose 4-phosphate, trehalose 6-phosphate, glucose 1,6-bisphosphate, sucrose 6-phosphate, 2-deoxy-D-glucose 6-phosphate, 2-D-ribose 5-phosphate, glucosamine 6-phosphate, 6-phospho-D-gluconate, L-2-phosphoglycerate, 3-phosphoglycerate, glyceraldehyde 3-phosphate, phytic acid, thiamine monophosphate, thiamine diphosphate, phosphoserine, phosphothreonine, phosphotyrosine, 2-phosphoascorbate, pyridoxal 5-phosphate, polyphosphate, glycerol 1-phosphate, glycerol 2-phosphate, glycerol 3-phosphate, ribulose 1,5-bisphosphate, lactose 1-phosphate, N-acetyl- α -D-glucosamine 1-phosphate, α -D-glucosamine 1-phosphate, N-acetyl- α -D-glucosamine 6-phosphate, dihydroxyacetone phosphate, phosphonoacetate, phosphonoformate, phosphono-methyl-glycine, AMP-ramidate, D-sorbitol 6-phosphate, glyphosate, phosphoryl-ethanolamine, NMN, diphosphate, 2',5'-ADP, PAP (3', 5' - ADP), PAPS, pNPP, 5-methyl dCMP. After 30 minutes of incubation at 30°C, free phosphate concentration was estimated by mixing a reaction aliquot with 40 μ L of the malachite green solution to a final volume of 200 μ L volume. After 1 minute of 1000 rpm orbital shaking, the production of phosphate was measured according to the optical density at 630 nm. The malachite green solution was prepared fresh by mixing 1 mL of 7.5% NH_4MoO_4 with 80 μ L 11% Tween-20 and malachite stock solution (1 L: 1.1 g malachite green, 150 mL H_2SO_4 , 750 mL water).

Syntenic and phylogenetic analysis of *XYL1* and *PHO3.2* homologs. AYbRAH was used to annotate proteins in the *XYL1* and *PHO3.2* loci, since genome annotations are lacking for several *Scheffersomyces* and *Spathaspora* species (Correia et al., 2017). The protein sequences of Xyl1p and Pho3.2p from *S. stipitis* were queried against the genomic nucleotide sequences of *Debaryomyces hansenii*, *Suhyomyces*

tanzawaensis, *Scheffersomyces* species, and *Spathaspora* species using TBLASTN; assembly accessions are listed in Table 2. The genomic loci 50 kbp upstream and downstream of the *XYL1* and *PHO3.2* hits were then queried with a protein sequence of the organism's closest relative for each Fungal Ortholog Group (FOG) in AYbRAH using BLASTP (Correia et al., 2017). The gene coordinates were manually reviewed to remove spurious open reading frames (see Supplementary Information). Biopython's GenomeDiagram was used to illustrate the synteny of the *XYL1* and *PHO3.2* loci (scripts available at <https://github.com/kcorreia/>) (Cock et al., 2009). The *XYL1* and *PHO3.2* homologs were aligned with MAFFT version 7.24. PhyML version 3.2.0 was used to reconstruct the phylogeny of *XYL1* and *PHO3.2* homologs with 1000 bootstrap replicates (Correia et al., 2017). These trees were used to help distinguish between the *XYL1/XYL1.2* and *PHO3/PHO3.2* paralogs.

3 RESULTS & DISCUSSION

***In vivo* NADPH source in *S. stipitis* during xylose fermentation.** *S. stipitis* Xyl1p prefers NADPH to NADH *in vitro*, but its NADPH source and XR selectivity for NADPH are unknown *in vivo*. We first used FVA to analyze the impact of each NADPH source on the maximum ethanol yield from xylose, assuming the *in vitro* NADPH selectivity for XR, in presence and absence of NADPase (Figure 3). The oxidative pentose phosphate pathway is able to ferment xylose to ethanol *in silico*, but its ethanol yields are less than the *in vivo* yields when oxygen uptake rate (OUR) is less than $3 \text{ mmol} \cdot \text{gDCW}^{-1} \cdot \text{h}^{-1}$; the addition of NADPase is unable to resolve the redox imbalance with NADPH regenerated from the oxidative pentose phosphate pathway. Jeffries et al. (2007) proposed the succinate bypass to regenerate NADPH, but it is unable to ferment xylose to ethanol below $10 \text{ mmol} \cdot \text{gDCW}^{-1} \cdot \text{h}^{-1}$ in our simulations. NADP-dependent isocitrate dehydrogenase can also enable ethanol fermentation, but not below $10 \text{ mmol} \cdot \text{gDCW}^{-1} \cdot \text{h}^{-1}$. Only NADPase and NADH kinase, phosphorylating NADP-dependent GAPDH, and non-phosphorylating NADP-dependent GAPDH were able to anaerobically ferment xylose to ethanol at the maximum theoretical yield. There is no strong evidence for NADPH regeneration from NADP-dependent GAPDH in *S. stipitis*, since it does not have any genes orthologous to *K. lactis*' phosphorylating NADP-dependent GAPDH (Correia et al., 2017), and there is no significant upregulation of *UGA2*, the most likely source of nonphosphorylating GAPDH in *S. stipitis* (Brunner et al., 1998), in its xylose-fermenting transcriptome (Jeffries and Van Vleet, 2009; Yuan et al., 2011). Therefore, these simulations and omics data support NADPase/NADH kinase balancing redox cofactors during xylose fermentation in *S. stipitis* if XR prefers NADPH *in vivo*.

Impact of XR cofactor preference on xylose fermentation. The *in vivo* XR cofactor preference is not known, so we used FVA to explore the maximum xylitol yield with the ethanol yield as the objective

function, with varying OUR's and NADPH selectivities, in the presence and absence of NADPase/NADH kinase (Figure 4A). Anaerobic xylose fermentation leads to xylitol accumulation when any amount of NADPH drives XR flux *in silico*. The *in silico* xylitol yield falls within the experimental polyol range when the NADPH selectivity is less than 10% NADPH, far less than the XR's *in vitro* selectivity (60% NADPH). *In silico* anaerobic xylose fermentation to ethanol is infeasible when the XR selectivity for NADPH is greater than 80% in the absence of NADPase and NADH kinase. The *in vitro* cofactor selectivity predicts the xylitol yield to be greater than 0.3 g/g when the OUR is less than $2 \text{ mmol} \cdot \text{gDCW}^{-1} \cdot \text{h}^{-1}$, which is more than the typical polyol yield of 0.1 g/g observed with *S. stipitis* (Ligthelm et al., 1988b; Su et al., 2015). The presence of NADPase/NADH kinase in the model enables the fermentation of xylose to ethanol at all OUR's and all NADPH selectivities. There is no impact on the ethanol yield when they are present individually. These simulations indicate that either XR can change its cofactor selectivity to NADH during oxygen-limiting conditions, NADPase/NADH kinase plays a role in redox cofactor balancing in *S. stipitis* or additional enzymes are absent in the metabolic model.

Previous attempts to determine the XR cofactor preference are summarized in Table 3. *In vitro* enzyme characterization and crude enzyme assays for XR at varying aeration rates with *S. stipitis* both show higher selectivity for NADPH than NADH. Verduyn et al. (1985b) found that XR preferred NADPH as a cofactor even in the presence of NADH. The accumulation of xylitol in the first engineered *S. cerevisiae* strain with the XR-XDH pathway (Kötter and Ciriacy, 1993) also suggests NADPH is the preferred cofactor *in vivo*. In contrast, Ligthelm et al. (1988c) used nuclear magnetic resonance (NMR) spectroscopy and ^{13}C labelling to infer that NADH is the preferred XR cofactor during anaerobic conditions. Dellweg et al. (1990) corroborated their conclusions via metabolic flux analysis and predicted that NADPH is preferred by XR aerobically. Dellweg et al. (1990) suggested the concentration of redox cofactors may exert metabolic control over the *in vivo* XR cofactor preference. The conflicting conclusions for the *in vivo* XR preference inferred from *in vitro* enzyme activities and model-based analysis require a reexamination of the assumptions made during the pre-genomic age.

Although metabolic control over the XR cofactor preference may exist to an extent, the discrepancy between the *in vitro* and *in vivo* XR cofactor preferences is likely due to an incomplete metabolic network reconstruction. *In vivo* flux cannot be directly measured in most cases but can be inferred using accurate metabolic network reconstructions. Ligthelm et al. (1988c) demonstrated anaerobic xylose fermentation did not use NADPH regenerated from the oxidative pentose phosphate pathway and concluded that XR likely used NADH regenerated from glycolysis during anaerobic xylose fermentation. These results do not prove XR uses NADH as a cofactor anaerobically since the presence of cytosolic NADPase/NADH kinase or NADP-dependent GAPDH in *S. stipitis*' metabolic network would obfuscate the ability to

resolve the XR cofactor preference using D-1-¹³C-xylose and NMR. The increased expression of *ZWF1* and *UTR1* in the xylose-fermenting transcriptome (Table 1), the lack of evidence for NADP-dependent GAPDH in *S. stipitis*, and the complete anaerobic fermentation of xylose to ethanol supports our proposed redox balancing mechanism of NADPase and NADH kinase.

Our previous simulations demonstrate xylitol accumulates when we maximize the ethanol yield in the absence of NADPase/NADH kinase, but we wanted to consider the impact of the growth rate on xylose fermentation. Growth-coupled xylose fermentation is especially relevant to *S. stipitis* since it has suboptimal growth with xylose, but not glucose (Ligthelm et al., 1988c; Shi et al., 2002). FVA was used to explore the xylitol yield solution space by assuming the *in vitro* XR cofactor preference in the wild-type background, a bypassed Complex I (Shi et al., 2002), and in the presence and absence of NADPase/NADH kinase. The minimum xylitol yield was predicted to be 0.52 g/g and 0.27 g/g, at the maximum and minimum growth rates, respectively. These yields are greater than the maximum 0.10 g/g polyol yield typically seen in xylose fermentation with *S. stipitis* (Cadete et al., 2012). Bypassing Complex I (type I NADH dehydrogenase) only has a minor impact on reducing the xylitol yield at the maximum growth rate. The addition of NADPase *or* NADH kinase does not change the solution space; however, the presence of NADPase *and* NADH kinase enables the *in silico* xylitol yield to overlap with the experimental range (Ligthelm et al., 1988c; Shi et al., 2002; Wahlbom et al., 2003; Jeffries et al., 2007; Jeffries and Van Vleet, 2009). These results demonstrate no growth-coupled mechanism balances redox cofactors given our metabolic constraints and further support NADPase and NADH kinase as critical to balancing redox cofactors anaerobically.

***PHO3* and *PHO3.2* characterization.** We used *K. phaffii* to express and purify Pho3p and Pho3.2p. Purified Pho3.2p showed Michaelis-Menten kinetics with NADP (Figure 5); no activity was detected with Pho3p. The assay conditions were not optimized, and therefore the *in vitro* NADPase activity may not reflect its *in vivo* activity. Pho3.2p was found to be more promiscuous than Pho3p using the malachite green phosphate assay (Figure S4). Although Pho3.2p's NADPase activity may balance redox cofactors under oxygen limitation, its broad activity may have suboptimal effects if the *in vitro* activities are relevant *in vivo*. These include creating futile cycles with its phosphatase activity on ribose 5-phosphate, erythrose 4-phosphate, and fructose 6-phosphate. *PHO3.2* is not highly expressed, and is not significantly upregulated during xylose fermentation (Table 1). Pho3.2p's low expression has allowed it to evade top-down omics approaches. Our characterization of Pho3.2p confirmed it has NADPase activity, in addition to a broader range of activities.

The phylogenetic origin of xylose fermentation to ethanol in *Scheffersomyces-Spathaspora*.

NADP phosphatase. *PHO3.2* is derived from a tandem duplication of *PHO3* in a common ancestor of

Scheffersomyces, *Spathaspora*, and *Suhomyces tanzawaensis* (Figures S5 and S6). *PHO3* was subsequently lost in several *Scheffersomyces* and *Spathaspora* species. The presence of *PHO3.2* in yeasts unable to ferment xylose to ethanol indicates that it may offer a fitness advantage beyond balancing redox cofactors; one possibility is fine-tuning the concentrations or ratios of redox cofactors (Kawai et al., 2005). The absence of *PHO3.2* homologs in wild-type *S. cerevisiae* prevents adaptive laboratory evolution from climbing the redox imbalance hurdle created by expressing the XR-XDH pathway from *S. stipitis*.

PHO3 and *PHO3.2* belong to the acid phosphatase family and contain the survival E protein motif, which was first characterized in *Thermotoga maritima* (Zhang et al., 2001; Lee et al., 2001). *Candida albicans*' Pho100p is the only eukaryotic homolog of Pho3p to be characterized (MacCallum et al., 2009), which scavenges for phosphate in the extracellular. The gain of NADPase activity by Pho3.2p is an example of neofunctionalization, which would have required several changes in its regulation and sequence. First, Pho3.2p would have required changes in its expression from phosphate limitation to at least oxygen limitation to enable redox cofactor balancing during xylose fermentation. Second, a change in its signal peptide or mature protein sequence would have redirected it from the extracellular to the cytoplasm or embedded in the membrane of an organelle. Lastly, Pho3.2p would have likely required sequence mutations to gain NADPase activity, since it is absent in *S. stipitis* Pho3p; as a consequence, it may have gained greater enzyme promiscuity. These changes are consistent with the view that duplications are critical to the evolution of metabolism (Correia et al., 2017), and it enabled the *PHO3.2* paralog to gain NADPase activity in the *Scheffersomyces-Spathaspora* clade.

NAD(P)H-dependent xylose reductase. *XYL1* is part of the large aldo-keto reductase family, which has a broad range of activities and generally prefers NADPH as a cofactor (Bennett et al., 1997). Several yeasts maintain *XYL1* in their genome despite the inability to grow on xylose. Yeasts that have NAD(P)H-dependent XR are able to ferment xylose to ethanol under oxygen limitation (Bruinenberg et al., 1984). The best xylose fermenters belong in the *Scheffersomyces-Spathaspora* clade, and have NAD(P)H-dependent XR encoded by *XYL1.2* (Table 1) (Mamoori et al., 2013). *XYL1.2* originated from a tandem duplication of *XYL1*, in a common ancestor of *Spathaspora* and *Scheffersomyces* (Figures S7 and S8). Other yeasts that can ferment xylose to ethanol, although at lower yields than *Scheffersomyces-Spathaspora* species, have independently evolved NAD(P)H-dependent XR. These include *Pachysolen tannophilus*, from a recent duplication of *XYL1* (Ditzelmüller et al., 1985; Correia et al., 2017), and likely convergent evolution in the *XYL1* ortholog for *Spathaspora hagerdaliae*, *Candida tropicalis*, and *Candida tenuis* (Bruinenberg et al., 1984). NAD(P)H-dependent XR evolved independently in yeasts, but Xyl1.2p's higher preference for NADH enables superior xylose fermentation to ethanol in the *Scheffersomyces-Spathaspora* clade.

NADH kinase. *UTR1* encodes a cytoplasmic NAD kinase in *S. cerevisiae* but has been shown to have

slight activity with NADH (Mori et al., 2005). Its physiological role is to synthesize *de novo* NADP, and not *de novo* NADPH or NADPH regeneration (Kawai et al., 2001; Mori et al., 2005). We were unable to purify Utr1p from *S. cerevisiae* or *S. stipitis* to compare their enzyme activities, but there is some indirect evidence Utr1p in *S. stipitis* regenerates NADPH. Swapping *S. stipitis* *UTR1* in place for *S. cerevisiae* *UTR1* led to a decrease in *S. cerevisiae*'s growth rate on glucose; its growth rate was restored to near wild-type by further swapping *S. stipitis* *NDE1* in place for *S. cerevisiae* *NDE1*, which encode NAD(P)H and NADH dehydrogenase, respectively (unpublished results). Furthermore, the amino acid alignment of *UTR1* in budding yeasts reveals conserved motifs found exclusively in the CTG clade, but outside the conserved NAD kinase domain (see Supplementary Information). Many of these yeasts grow well on xylose and can ferment it to xylitol and ethanol (Papon et al., 2014). NAD kinase may have evolved to NADH kinase in the CTG clade and replaced NADP-dependent acetaldehyde dehydrogenase as an alternative NADPH source (Correia et al., 2017). Additional characterization of Utr1p orthologs can trace its preference for phosphorylating NAD or NADH in budding yeasts.

Cofactor balancing in metabolic pathways. NADP phosphatase and NADH kinase appear to have evolved millions of years ago to balance redox cofactors during xylose fermentation, in a possible symbiotic relationship between *Scheffersomyces-Spathaspora* species and wood-ingesting beetles (Suh et al., 2003), but these enzymes are relevant to today's metabolic engineers. Gevo filed a patent which proposed the use of NADH kinase and NADP(H) phosphatase to resolve the redox imbalance between NAD-dependent GAPDH and NADPH-dependent isobutanol reductase in the isobutanol pathway; no results were provided in their patent to confirm its feasibility (Buelter et al., 2008). Our xylose fermentation simulations show that the sole expression of NADP phosphatase *or* NADH kinase cannot increase the ethanol yield from xylose, but the expression of NADP phosphatase *and* NADH kinase may balance redox cofactors anaerobically. This is not consistent with our bioinformatic and physiological analysis of *Scheffersomyces-Spathaspora* yeasts (Table 1). These yeasts demonstrate that NAD(P)H-dependent XR enables xylose fermentation to ethanol, but NADPH-dependent XR and NADPase/NADH kinase still accumulate xylitol (Bruinenberg et al., 1984). Further investigation is needed to determine if our proposed mechanism can support *sustained* anaerobic fermentation with an entirely imbalanced pathway, or if other factors, such as inhibition of XR or XDH by NAD(P)(H) (Verduyn et al., 1985b; Dellweg et al., 1990), limit its ability *in vivo*. Our analysis of xylose fermentation in *S. stipitis* highlights the need to understand and engineer metabolism at a systems-level, which has been advocated by Kim et al. (2011) and Meadows et al. (2016).

Next steps in understanding xylose fermentation in yeasts. Comparative genetic using CRISPR-Cas9 can be used to validate further or refute our proposed mechanism (Figure 2A). This can be achieved

by perturbing the mechanism in part or whole for *S. stipitis*, or introducing it in yeasts unable to ferment xylose to ethanol, such as *Meyerozyma guilliermondii* (Nolleau et al., 1995). For example, *PHO3.2*'s role in metabolism can be demonstrated by a knockout in *S. stipitis* or expression in *Meyerozyma guilliermondii*. Swapping *Meyerozyma guilliermondii*'s NADPH-dependent XR for *S. stipitis*' NAD(P)H-dependent XR can test *S. stipitis*' ability to ferment xylose to ethanol with a completely imbalanced pathway in the presence of NADPase and a putative NADH kinase; the corollary tests the ability for *M. guilliermondii* to ferment xylose to ethanol with NAD(P)H-dependent XR but without NADPase. Furthermore, the putative NADH kinase in *S. stipitis* can be swapped with NAD kinase in *S. cerevisiae* to test its impact on redox metabolism and xylose fermentation, especially during anaerobic fermentation. The emancipation of non-conventional yeast genetics by CRISPR-Cas9 provides an exciting future to study the evolution of metabolism in yeasts.

4 CONCLUSION

In this study, we sought to understand how *S. stipitis* balances redox cofactors during xylose fermentation by integrating transcriptomics and proteomics with FBA from a consensus *S. stipitis* GENRE. We could not reconcile flux simulations with *S. stipitis*' near theoretical ethanol yields during oxygen limitation without the presence of NADPase and NADH kinase in the metabolic model, when NADPH was the dominant cofactor driving XR flux. Our proposed mechanism of NADPase and NADH kinase can balance redox cofactors independent of oxygen, but requires ATP. In contrast, most yeasts generate NADPH from the oxidative pentose phosphate pathway but require oxygen to reoxidize NADH. NADPase activity was confirmed from purified Pho3.2p using *K. phaffii*; it has a broader activity than Pho3p, its recent paralog. NADH-linked XR activity and NADPase from Pho3.2p both originated from tandem gene duplications in a common ancestor of *Scheffersomyces* and *Spathaspora* species. This study demonstrates the advantages of using a bottom-up approach that combines metabolic modelling, omics analysis, bioinformatics, and enzymology to reverse engineer metabolism.

5 ACKNOWLEDGMENTS

The authors acknowledge Tracy Chan, Dr. Julie-Anne Gandier, and Prof. Emma Master for their help with protein expression in *K. phaffii*; Dr. Andrew Quaile for his help with protein sequencing by mass spectrometry; Dr. Goutham Vemuri for encouraging the consensus *S. stipitis* GENRE to be built; Prof. Amy Caudy for her suggestions on improving the figures and tables; Prof. Uwe Sauer for his discussion on the ATP balance in anaerobic xylose fermentation.

6 FUNDING

K.C. was supported by Bioconversion Network and NSERC CREATE M3.

7 CONTRIBUTIONS

R.M., P.Y.L., and K.C. conceived the study. P.Y.L. and K.C. reconstructed the metabolic network of *S. stipitis*. G.B., J.C.J., K.C. and A.K. performed cloning, purification, and characterization of Pho3p and Pho3.2p. K.C. performed the bioinformatic analysis. K.C. and A.K. wrote the manuscript. All the authors reviewed the manuscript.

8 ABBREVIATIONS

BMGY	Buffered complex Medium containing Glycerol
BMMY	Buffered complex Medium containing Methanol
FBA	Flux Balance Analysis
FVA	Flux Variability Analysis
GENRE	GEome-scale Network REconstruction
GAPDH	GlycerAldehyde 3-Phosphate DeHydrogenase
HEPES	(4-(2-HydroxyEthyl)-1-PiperazineEthaneSulfonic acid)
NADPase	NADP phosphatase
PAAG	PolyAcrylAmide Gel
RPKM	Reads Per Kilobase of transcript per Million mapped reads
XI	Xylose Isomerase
XDH	Xylitol DeHydrogenase
XR	Xylose Reductase
YPD	Yeast extract-Peptone-Dextrose

REFERENCES

- Amore, R., Wilhelm, M., and Hollenberg, C. P. (1989). The fermentation of xylose—an analysis of the expression of *Bacillus* and *Actinoplanes* xylose isomerase genes in yeast. *Applied Microbiology and Biotechnology*, 30(4):351–357.
- Andersen, M. R., Nielsen, M. L., and Nielsen, J. (2008). Metabolic model integration of the bibliome, genome, metabolome and reactome of *Aspergillus niger*. *Molecular Systems Biology*, 4(1):178.
- Balagurunathan, B., Jonnalagadda, S., Tan, L., and Srinivasan, R. (2012). Reconstruction and analysis of a genome-scale metabolic model for *Scheffersomyces stipitis*. *Microbial Cell Factories*, 11(1):1.
- Bennett, M. J., Schlegel, B. P., Lewis, M., Trevor, M., et al. (1997). Comparative anatomy of the aldo–keto reductase superfamily. *Biochemical Journal*, 326(3):625–636.
- Bruinenberg, P. M., de Bot, P. H., van Dijken, J. P., and Scheffers, W. A. (1984). NADH-linked aldose reductase: the key to anaerobic alcoholic fermentation of xylose by yeasts. *Applied Microbiology and Biotechnology*, 19(4):256–260.

- Brunner, N. A., Brinkmann, H., Siebers, B., and Hensel, R. (1998). NAD⁺-dependent glyceraldehyde-3-phosphate dehydrogenase from *Thermoproteus tenax*: The first identified archaeal member of the aldehyde dehydrogenase superfamily is a glycolytic enzyme with unusual regulatory properties. *Journal of Biological Chemistry*, 273(11):6149–6156.
- Buelter, T., Meinhold, P., Feldman, R. M. R., Eckl, E., Hawkins, A., Aristidou, A., Dundon, C. A., Bastian, D. L., Arnold, F., and Urano, J. (2008). Engineered microorganisms capable of producing target compounds under anaerobic conditions. *US Patent*, page WO2010051527A2.
- Cadete, R. M., Alejandro, M., Sandström, A. G., Ferreira, C., Gírio, F., Gorwa-Grauslund, M.-F., Rosa, C. A., and Fonseca, C. (2016). Exploring xylose metabolism in *Spathaspora species*: *XYL1*. 2 from *Spathaspora passalidarum* as the key for efficient anaerobic xylose fermentation in metabolic engineered *Saccharomyces cerevisiae*. *Biotechnology for Biofuels*, 9(1):167.
- Cadete, R. M., Melo, M. A., Dussán, K. J., Rodrigues, R. C., Silva, S. S., Zilli, J. E., Vital, M. J., Gomes, F. C., Lachance, M.-A., and Rosa, C. A. (2012). Diversity and physiological characterization of D-xylose-fermenting yeasts isolated from the Brazilian Amazonian forest. *PLoS One*, 7(8):e43135.
- Caspeta, L., Shoaie, S., Agren, R., Nookaew, I., and Nielsen, J. (2012). Genome-scale metabolic reconstructions of *Pichia stipitis* and *Pichia pastoris* and *in silico* evaluation of their potentials. *BMC Systems Biology*, 6(1):1.
- Cock, P. J., Antao, T., Chang, J. T., Chapman, B. A., Cox, C. J., Dalke, A., Friedberg, I., Hamelryck, T., Kauff, F., Wilczynski, B., et al. (2009). Biopython: freely available Python tools for computational molecular biology and bioinformatics. *Bioinformatics*, 25(11):1422–1423.
- Correia, K., Yu, S. M., and Mahadevan, R. (2017). Reconstructing the evolution of metabolism in budding yeasts. *bioRxiv*, page 237974.
- Dahms, A. S. (1974). 3-Deoxy-D-pentulosonic acid aldolase and its role in a new pathway of D-xylose degradation. *Biochemical and Biophysical Research Communications*, 60(4):1433–1439.
- Dellweg, H., Klein, C., Prahl, S., Rizzi, M., and Weigert, B. (1990). Kinetics of ethanol production from D-xylose by the yeast *Pichia stipitis*. *Food Biotechnology*, 4(1).
- Dias, O., Pereira, R., Gombert, A. K., Ferreira, E. C., and Rocha, I. (2014). iOD907, the first genome-scale metabolic model for the milk yeast *Kluyveromyces lactis*. *Biotechnology Journal*, 9(6):776–790.
- Ditzelmüller, G., Kubicek-Pranz, E., Röhr, M., and Kubicek, C. (1985). NADPH-specific and NADH-specific xylose reduction is catalyzed by two separate enzymes in *Pachysolen tannophilus*. *Applied Microbiology and Biotechnology*, 22(4):297–299.
- Dorn, G. (1965). Genetic analysis of the phosphatases in *Aspergillus nidulans*. *Genetics Research*, 6(1):13–26.

- du Preez, J. (1994). Process parameters and environmental factors affecting D-xylose fermentation by yeasts. *Enzyme and Microbial Technology*, 16(11):944–956.
- Dujon, B., Sherman, D., Fischer, G., Durrens, P., Casaregola, S., Lafontaine, I., De Montigny, J., Marck, C., Neuvéglise, C., Talla, E., et al. (2004). Genome evolution in yeasts. *Nature*, 430(6995):35.
- Ebrahim, A., Lerman, J. A., Palsson, B. O., and Hyduke, D. R. (2013). COBRApy: CONstraints-Based Reconstruction and Analysis for Python. *BMC Systems Biology*, 7(1):74.
- Freese, S., Vogts, T., Speer, F., Schäfer, B., Passoth, V., and Klinner, U. (2011). C-and N-catabolic utilization of tricarboxylic acid cycle-related amino acids by *Scheffersomyces stipitis* and other yeasts. *Yeast*, 28(5):375–390.
- Haselbeck, R. J. and McAlister-Henn, L. (1993). Function and expression of yeast mitochondrial NAD- and NADP-specific isocitrate dehydrogenases. *Journal of Biological Chemistry*, 268(16):12116–12122.
- Heavner, B. D., Smallbone, K., Price, N. D., and Walker, L. P. (2013). Version 6 of the consensus yeast metabolic network refines biochemical coverage and improves model performance. *Database*, 2013.
- Ho, N., Lin, F., Huang, S., Andrews, P., and Tsao, G. (1990). Purification, characterization, and amino terminal sequence of xylose reductase from *Candida shehatae*. *Enzyme and Microbial Technology*, 12(1):33–39.
- Horitsu, H., Tomoeda, M., and Kumagai, K. (1968). Pentose metabolism in *Candida utilis*: Part IV. NADP specific polyol dehydrogenase. *Agricultural and Biological Chemistry*, 32(4):514–517.
- Horne, R. N., Anderson, W. B., and Nordlie, R. C. (1970). Glucose dehydrogenase activity of yeast glucose 6-phosphate dehydrogenase. Inhibition by adenosine 5'-triphosphate and other nucleoside 5'-triphosphates and diphosphates. *Biochemistry*, 9(3):610–616.
- Hou, J., Vemuri, G. N., Bao, X., and Olsson, L. (2009). Impact of overexpressing NADH kinase on glucose and xylose metabolism in recombinant xylose-utilizing *Saccharomyces cerevisiae*. *Applied Microbiology and Biotechnology*, 82(5):909–919.
- Huang, E. and Lefsrud, M. (2014). Fermentation monitoring of a co-culture process with *Saccharomyces cerevisiae* and *Scheffersomyces stipitis* using shotgun proteomics. *Journal of Bioprocessing and Biotechniques*, 144:1–7.
- Huang, E. L. and Lefsrud, M. G. (2012). Temporal analysis of xylose fermentation by *Scheffersomyces stipitis* using shotgun proteomics. *Journal of Industrial Microbiology & Biotechnology*, 39(10):1507–1514.
- Jeffries, T. and Jin, Y.-S. (2004). Metabolic engineering for improved fermentation of pentoses by yeasts. *Applied Microbiology and Biotechnology*, 63(5):495–509.
- Jeffries, T. W., Grigoriev, I. V., Grimwood, J., Laplaza, J. M., Aerts, A., Salamov, A., Schmutz, J.,

- Lindquist, E., Dehal, P., Shapiro, H., et al. (2007). Genome sequence of the lignocellulose-bioconverting and xylose-fermenting yeast *Pichia stipitis*. *Nature Biotechnology*, 25(3):319–326.
- Jeffries, T. W. and Van Vleet, J. R. H. (2009). *Pichia stipitis* genomics, transcriptomics, and gene clusters. *FEMS Yeast Research*, 9(6):793–807.
- Jeppsson, H., Alexander, N., and Hahn-Hagerdal, B. (1995). Existence of cyanide-insensitive respiration in the yeast *Pichia stipitis* and its possible influence on product formation during xylose utilization. *Applied and Environmental Microbiology*, 61(7):2596–2600.
- Kanehisa, M., Goto, S., Sato, Y., Furumichi, M., and Tanabe, M. (2011). KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Research*, 40(D1):D109–D114.
- Karp, P. D., Paley, S. M., Krummenacker, M., Latendresse, M., Dale, J. M., Lee, T. J., Kaipa, P., Gilham, F., Spaulding, A., Popescu, L., et al. (2009). Pathway Tools version 13.0: integrated software for pathway/genome informatics and systems biology. *Briefings in Bioinformatics*, 11(1):40–79.
- Kawai, S., Fukuda, C., Mukai, T., and Murata, K. (2005). MJ0917 in archaeon *Methanococcus jannaschii* is a novel NADP phosphatase/NAD kinase. *Journal of Biological Chemistry*, 280(47):39200–39207.
- Kawai, S., Suzuki, S., Mori, S., and Murata, K. (2001). Molecular cloning and identification of *UTR1* of a yeast *Saccharomyces cerevisiae* as a gene encoding an NAD kinase. *FEMS Microbiology Letters*, 200(2):181–184.
- Kim, J., Reed, J. L., and Maravelias, C. T. (2011). Large-scale bi-level strain design approaches and mixed-integer programming solution techniques. *PLoS One*, 6(9):e24162.
- Kim, S. R., Park, Y.-C., Jin, Y.-S., and Seo, J.-H. (2013a). Strain engineering of *Saccharomyces cerevisiae* for enhanced xylose metabolism. *Biotechnology Advances*, 31(6):851–861.
- Kim, S. R., Skerker, J. M., Kang, W., Lesmana, A., Wei, N., Arkin, A. P., and Jin, Y.-S. (2013b). Rational and evolutionary engineering approaches uncover a small set of genetic changes efficient for rapid xylose fermentation in *Saccharomyces cerevisiae*. *PloS one*, 8(2):e57048.
- Kötter, P. and Ciriacy, M. (1993). Xylose fermentation by *Saccharomyces cerevisiae*. *Applied Microbiology and Biotechnology*, 38(6):776–783.
- Kuznetsova, E., Nocek, B., Brown, G., Makarova, K. S., Flick, R., Wolf, Y. I., Khusnutdinova, A., Evdokimova, E., Jin, K., Tan, K., et al. (2015). Functional diversity of haloacid dehalogenase superfamily phosphatases from *Saccharomyces cerevisiae*: biochemical, structural, and evolutionary insights. *Journal of Biological Chemistry*, pages jbc–M115.
- Kuznetsova, E., Proudfoot, M., Sanders, S. A., Reinking, J., Savchenko, A., Arrowsmith, C. H., Edwards, A. M., and Yakunin, A. F. (2005). Enzyme genomics: Application of general enzymatic screens to discover new enzymes. *FEMS Microbiology Reviews*, 29(2):263–279.

- Kwak, S. and Jin, Y.-S. (2017). Production of fuels and chemicals from xylose by engineered *Saccharomyces cerevisiae*: a review and perspective. *Microbial Cell Factories*, 16(1):82.
- Lee, J. Y., Kwak, J. E., Moon, J., Eom, S. H., Liong, E. C., Pedelacq, J.-D., Berendzen, J., and Suh, S. W. (2001). Crystal structure and functional analysis of the SurE protein identify a novel phosphatase family. *Nature Structural and Molecular Biology*, 8(9):789.
- Li, P. Y. (2012). *In silico* metabolic network reconstruction of *Scheffersomyces stipitis*. Master's thesis, University of Toronto.
- Ligthelm, M. E., Prior, B. A., and du Preez, J. C. (1988a). The effect of respiratory inhibitors on the fermentative ability of *Pichia stipitis*, *Pachysolen tannophilus* and *Saccharomyces cerevisiae* under various conditions of aerobiosis. *Applied Microbiology and Biotechnology*, 29(1):67–71.
- Ligthelm, M. E., Prior, B. A., and du Preez, J. C. (1988b). The oxygen requirements of yeasts for the fermentation of D-xylose and D-glucose to ethanol. *Applied Microbiology and Biotechnology*, 28(1):63–68.
- Ligthelm, M. E., Prior, B. A., du Preez, J. C., and Brandt, V. (1988c). An investigation of D-{1-13 C} xylose metabolism in *Pichia stipitis* under aerobic and anaerobic conditions. *Applied Microbiology and Biotechnology*, 28(3):293–296.
- Liu, T., Zou, W., Liu, L., and Chen, J. (2012). A constraint-based model of *Scheffersomyces stipitis* for improved ethanol production. *Biotechnology for Biofuels*, 5(1):72.
- Lobo, F. P., Gonçalves, D. L., Alves, S. L., Gerber, A. L., de Vasconcelos, A. T. R., Basso, L. C., Franco, G. R., Soares, M. A., Cadete, R. M., Rosa, C. A., et al. (2014). Draft genome sequence of the D-xylose-fermenting yeast *Spathaspora arborariae* UFMG-HM19. 1AT. *Genome Announcements*, 2(1):e01163–13.
- Loira, N., Dulermo, T., Nicaud, J.-M., and Sherman, D. J. (2012). A genome-scale metabolic model of the lipid-accumulating yeast *Yarrowia lipolytica*. *BMC Systems Biology*, 6(1):35.
- Lopes, D. D., Cibulski, S. P., Mayer, F. Q., Siqueira, F. M., Rosa, C. A., Hector, R. E., and Ayub, M. A. Z. (2017). Draft genome sequence of the d-xylose-fermenting yeast *Spathaspora xylofermentans* UFMG-HMD23. 3. *Genome Announcements*, 5(33):e00815–17.
- Lopes, M. R., Batista, T. M., Franco, G. R., Ribeiro, L. R., Santos, A. R., Furtado, C., Moreira, R. G., Goes-Neto, A., Vital, M. J., Rosa, L. H., et al. (2018). *Scheffersomyces stambukii* fa, sp. nov., a D-xylose-fermenting species isolated from rotting wood. *International Journal of Systematic and Evolutionary Microbiology*.
- Lopes, M. R., Morais, C. G., Kominek, J., Cadete, R. M., Soares, M. A., Uetanabaro, A. P. T., Fonseca, C., Lachance, M.-A., Hittinger, C. T., and Rosa, C. A. (2016). Genomic analysis and D-xylose fermentation

- of three novel *Spathaspora* species: *Spathaspora girioi* sp. nov., *Spathaspora hagerdaliae* fa, sp. nov. and *Spathaspora gorwiae* fa, sp. nov. *FEMS Yeast Research*, 16(4):fow044.
- MacCallum, D. M., Castillo, L., Nather, K., Munro, C. A., Brown, A. J., Gow, N. A., and Odds, F. C. (2009). Property differences among the four major *Candida albicans* strain clades. *Eukaryotic Cell*, 8(3):373–387.
- Mahadevan, R. and Schilling, C. (2003). The effects of alternate optimal solutions in constraint-based genome-scale metabolic models. *Metabolic Engineering*, 5(4):264–276.
- Mamoori, Y. I., Yahya, A. G. I., and AL-Jelawi, M. H. (2013). Expression of xylose reductase enzyme from *Spathaspora passalidarum* in *Saccharomyces cerevisiae*. *Iraqi Journal of Science*, 54:316–323.
- McCloskey, D., Palsson, B. Ø., and Feist, A. M. (2013). Basic and applied uses of genome-scale metabolic network reconstructions of *Escherichia coli*. *Molecular Systems Biology*, 9(1):661.
- Meadows, A. L., Hawkins, K. M., Tsegaye, Y., Antipov, E., Kim, Y., Raetz, L., Dahl, R. H., Tai, A., Mahatdejkul-Meadows, T., Xu, L., et al. (2016). Rewriting yeast central carbon metabolism for industrial isoprenoid production. *Nature*, 537(7622):694–697.
- Morais, C. G., Batista, T. M., Kominek, J., Borelli, B. M., Furtado, C., Moreira, R. G., Franco, G. R., Rosa, L. H., Fonseca, C., Hittinger, C. T., et al. (2017). *Spathaspora boniae* sp. nov., a D-xylose-fermenting species in the *Candida albicans*/Lodderomyces clade. *International Journal of Systematic and Evolutionary Microbiology*, 67(10):3798–3805.
- Mori, S., Kawai, S., Shi, F., Mikami, B., and Murata, K. (2005). Molecular conversion of NAD kinase to NADH kinase through single amino acid residue substitution. *Journal of Biological Chemistry*, 280(25):24104–24112.
- Nolleau, V., Preziosi-Belloy, L., and Navarro, J. (1995). The reduction of xylose to xylitol by *Candida guilliermondii* and *Candida parapsilosis*: incidence of oxygen and pH. *Biotechnology Letters*, 17(4):417–422.
- Okada, N., Tanimura, A., Hirakawa, H., Takashima, M., Ogawa, J., and Shima, J. (2017). Draft genome sequences of the xylose-fermenting yeast *Scheffersomyces shehatae* NBRC 1983T and a thermotolerant isolate of *S. shehatae* ATY839 (JCM 18690). *Genome Announcements*, 5(20):e00347–17.
- Österlund, T., Nookaew, I., Bordel, S., and Nielsen, J. (2013). Mapping condition-dependent regulation of metabolism in yeast through genome-scale modeling. *BMC Systems Biology*, 7(1):1.
- Pan, P. and Hua, Q. (2012). Reconstruction and *in silico* analysis of metabolic network for an oleaginous yeast, *Yarrowia lipolytica*. *PLoS One*, 7(12):e51535.
- Papini, M., Nookaew, I., Uhlén, M., and Nielsen, J. (2012). *Scheffersomyces stipitis*: a comparative systems biology study with the Crabtree positive yeast *Saccharomyces cerevisiae*. *Microbial Cell*

- Factories*, 11(1):1.
- Papon, N., Courdavault, V., and Clastre, M. (2014). Biotechnological potential of the fungal CTG clade species in the synthetic biology era. *Trends in Biotechnology*, 32(4):167–168.
- Parrish, W. R., Stefan, C. J., and Emr, S. D. (2004). Essential role for the myotubularin-related phosphatase Ymr1p and the synaptojanin-like phosphatases Sjl2p and Sjl3p in regulation of phosphatidylinositol 3-phosphate in yeast. *Molecular Biology of the Cell*, 15(8):3567–3579.
- Pereira, R., Nielsen, J., and Rocha, I. (2016). Improving the flux distributions simulated with genome-scale metabolic models of *Saccharomyces cerevisiae*. *Metabolic Engineering Communications*, 3:153–163.
- Petersen, T. N., Brunak, S., von Heijne, G., and Nielsen, H. (2011). SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nature Methods*, 8(10):785.
- Ramos, F., el Guezzer, M., Grendon, M., and Wiame, J.-M. (1985). Mutations affecting the enzymes involved in the utilization of 4-aminobutyric acid as nitrogen source by the yeast *Saccharomyces cerevisiae*. *European Journal of Biochemistry*, 149(2):401–404.
- Riley, R., Haridas, S., Wolfe, K. H., Lopes, M. R., Hittinger, C. T., Göker, M., Salamov, A. A., Wisecaver, J. H., Long, T. M., Calvey, C. H., et al. (2016). Comparative genomics of biotechnologically important yeasts. *Proceedings of the National Academy of Sciences*, 113(35):9882–9887.
- Rizzi, M., Harwart, K., Erlemann, P., Bui-Thanh, N.-A., and Dellweg, H. (1989). Purification and properties of the NAD⁺-xylitol-dehydrogenase from the yeast *Pichia stipitis*. *Journal of Fermentation and Bioengineering*, 67(1):20–24.
- Roseiro, J. C., Peito, M. A., Gírio, F. M., and Amaral-Collaco, M. (1991). The effects of the oxygen transfer coefficient and substrate concentration on the xylose fermentation by *Debaryomyces hansenii*. *Archives of Microbiology*, 156(6):484–490.
- Sarthy, A., McConaughy, B., Lobo, Z., Sundstrom, J., Furlong, C., and Hall, B. (1987). Expression of the *Escherichia coli* xylose isomerase gene in *Saccharomyces cerevisiae*. *Applied and Environmental Microbiology*, 53(9):1996–2000.
- Schellenberg, G., Sarthy, A., Larson, A., Backer, M., Crabb, J., Lidstrom, M., Hall, B., and Furlong, C. (1984). Xylose isomerase from *Escherichia coli*. Characterization of the protein and the structural gene. *Journal of Biological Chemistry*, 259(11):6826–6832.
- Schneider, H., Wang, P., Chan, Y., and Maleszka, R. (1981). Conversion of D-xylose into ethanol by the yeast *Pachysolen tannophilus*. *Biotechnology Letters*, 3(2):89–92.
- Shi, F., Kawai, S., Mori, S., Kono, E., and Murata, K. (2005). Identification of ATP-NADH kinase isozymes and their contribution to supply of NADP(H) in *Saccharomyces cerevisiae*. *The FEBS Journal*, 272(13):3337–3349.

- Shi, N.-Q., Cruz, J., Sherman, F., and Jeffries, T. W. (2002). SHAM-sensitive alternative respiration in the xylose-metabolizing yeast *Pichia stipitis*. *Yeast*, 19(14):1203–1220.
- Skoog, K. and Hahn-Hägerdal, B. (1990). Effect of oxygenation on xylose fermentation by *Pichia stipitis*. *Applied and Environmental Microbiology*, 56(11):3389–3394.
- Slininger, P., Bothast, R., Okos, M., and Ladisch, M. (1985). Comparative evaluation of ethanol production by xylose-fermenting yeasts presented high xylose concentrations. *Biotechnology Letters*, 7(6):431–436.
- Slininger, P., Bothast, R., Van Cauwenberge, J., and Kurtzman, C. (1982). Conversion of D-xylose to ethanol by the yeast *Pachysolen tannophilus*. *Biotechnology and Bioengineering*, 24(2):371–384.
- Slininger, P., Branstrator, J., Lomont, J., Dien, B., Okos, M., Ladisch, M., and Bothast, R. (1990). Stoichiometry and kinetics of xylose fermentation by *Pichia stipitis*. *Annals of the New York Academy of Sciences*, 589(1):25–40.
- Sohn, S. B., Kim, T. Y., Lee, J. H., and Lee, S. Y. (2012). Genome-scale metabolic model of the fission yeast *Schizosaccharomyces pombe* and the reconciliation of *in silico*/*in vivo* mutant growth. *BMC Systems Biology*, 6(1):1.
- Su, Y.-K., Willis, L. B., and Jeffries, T. W. (2015). Effects of aeration on growth, ethanol and polyol accumulation by *Spathaspora passalidarum* NRRL Y-27907 and *Scheffersomyces stipitis* NRRL Y-7124. *Biotechnology and Bioengineering*, 112(3):457–469.
- Suh, S.-O., Marshall, C. J., Mchugh, J. V., and Blackwell, M. (2003). Wood ingestion by passalid beetles in the presence of xylose-fermenting gut yeasts. *Molecular Ecology*, 12(11):3137–3145.
- Toivola, A., Yarrow, D., Van Den Bosch, E., Van Dijken, J. P., and Scheffers, W. A. (1984). Alcoholic fermentation of D-xylose by yeasts. *Applied and Environmental Microbiology*, 47(6):1221–1223.
- van Vleet, J. and Jeffries, T. W. (2009). Yeast metabolic engineering for hemicellulosic ethanol production. *Current Opinion in Biotechnology*, 20(3):300–306.
- van Vleet, J. H., Jeffries, T. W., and Olsson, L. (2008). Deleting the para-nitrophenyl phosphatase (pNPPase), PHO13, in recombinant *Saccharomyces cerevisiae* improves growth and ethanol production on D-xylose. *Metabolic Engineering*, 10(6):360–369.
- Veras, H. C. T., Parachin, N. S., and Almeida, J. R. M. (2017). Comparative assessment of fermentative capacity of different xylose-consuming yeasts. *Microbial Cell Factories*, 16(1):153.
- Verduyn, C., Jzn, J. F., van Dijken, J. P., and Scheffers, W. A. (1985a). Multiple forms of xylose reductase in *Pachysolen tannophilus* CBS 4044. *FEMS Microbiology Letters*, 30(3):313–317.
- Verduyn, C., Van Kleef, R., Frank, J., Schreuder, H., Van Dijken, J., and Scheffers, W. (1985b). Properties of the NAD(P)H-dependent xylose reductase from the xylose-fermenting yeast *Pichia stipitis*.

- Biochemical Journal*, 226(3):669–677.
- Verho, R., Londesborough, J., Penttilä, M., and Richard, P. (2003). Engineering redox cofactor regeneration for improved pentose fermentation in *Saccharomyces cerevisiae*. *Applied and Environmental Microbiology*, 69(10):5892–5897.
- Verho, R., Richard, P., Jonson, P. H., Sundqvist, L., Londesborough, J., and Penttilä, M. (2002). Identification of the first fungal NADP-GAPDH from *Kluyveromyces lactis*. *Biochemistry*, 41(46):13833–13838.
- Verhoeven, M. D., Lee, M., Kamoen, L., Van Den Broek, M., Janssen, D. B., Daran, J.-M. G., Van Maris, A. J., and Pronk, J. T. (2017). Mutations in *PMRI* stimulate xylose isomerase activity and anaerobic growth on xylose of engineered *Saccharomyces cerevisiae* by influencing manganese homeostasis. *Scientific Reports*, 7:46155.
- Wahlbom, C. F., van Zyl, W. H., Jönsson, L. J., Hahn-Hägerdal, B., and Otero, R. R. C. (2003). Generation of the improved recombinant xylose-utilizing *Saccharomyces cerevisiae* TMB 3400 by random mutagenesis and physiological comparison with *Pichia stipitis* CBS 6054. *FEMS Yeast Research*, 3(3):319–326.
- Wei, N., Xu, H., Kim, S. R., and Jin, Y.-S. (2013). Deletion of *FPS1*, encoding aquaglyceroporin Fps1p, improves xylose fermentation by engineered *Saccharomyces cerevisiae*. *Applied and Environmental Microbiology*, 79(10):3193–3201.
- Weimberg, R. (1961). Pentose oxidation by *Pseudomonas fragi*. *Journal of Biological Chemistry*, 236(3):629–635.
- Wohlbach, D. J., Kuo, A., Sato, T. K., Potts, K. M., Salamov, A. A., LaButti, K. M., Sun, H., Clum, A., Pangilinan, J. L., Lindquist, E. A., et al. (2011). Comparative genomics of xylose-fermenting fungi for enhanced biofuel production. *Proceedings of the National Academy of Sciences*, 108(32):13212–13217.
- Yuan, T., Ren, Y., Meng, K., Feng, Y., Yang, P., Wang, S., Shi, P., Wang, L., Xie, D., and Yao, B. (2011). RNA-Seq of the xylose-fermenting yeast *Scheffersomyces stipitis* cultivated in glucose or xylose. *Applied Microbiology and Biotechnology*, 92(6):1237–1249.
- Zhang, R.-G., Skarina, T., Katz, J., Beasley, S., Khachatryan, A., Vyas, S., Arrowsmith, C., Clarke, S., Edwards, A., Joachimiak, A., et al. (2001). Structure of *Thermotoga maritima* stationary phase survival protein SurE: a novel acid phosphatase. *Structure*, 9(11):1095–1106.
- Zhou, H., Cheng, J.-s., Wang, B. L., Fink, G. R., and Stephanopoulos, G. (2012). 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(6):611–622.

Table 1. Xylose fermentation related genes in *Scheffersomyces stipitis*, including the xylose reductase(XR)-xylitol dehydrogenase (XDH) pathway, NADPH regeneration, and NADP phosphatase candidates. AYbRAH annotations, transcriptomics, proteomics, and functional characterization are outlined for all the genes.

	Gene name	Locus tag	Protein name	AYbRAH annotation ^a			Transcriptomics ^b		Proteomics ^c average spectral hits	Characterized function
				HOG	FOG	<i>S. cerevisiae</i> homolog	glucose	xylose		
XR-XDH pathway	<i>XYL1</i>	PICST_89614	NAD(P)H-dependent xylose reductase	HOG00232	FOG00421	paralog	67.1	11 175.4	126	NADH and NADPH-linked D-xylose reductase (Verduyn et al., 1985a)
	<i>XYL2</i>	PICST_86924	NAD-linked xylitol dehydrogenase	HOG00428	FOG00908	ortholog	40.0	2555.3	136	NAD-xylitol dehydrogenase (Rizzi et al., 1989)
	<i>XKS1</i>	PICST_68734	D-xylulose 5-kinase	HOG00427	FOG00907	ortholog	11.5	2156.4	42	phosphorylates D-xylulose for entry into the pentose phosphate pathway
NADPH regeneration	<i>ZWF1</i>	PICST_85065	glucose-6-phosphate 1-dehydrogenase	HOG00414	FOG00879	ortholog	250.0	1772.6	19	first committed step in the oxidative pentose phosphate pathway, regenerates NADPH (Horne et al., 1970)
	<i>GND1</i>	PICST_69500	6-phosphogluconate dehydrogenase, decarboxylating	HOG00417	FOG00889	ortholog	730.4	1658.3	50	last step in the oxidative pentose phosphate pathway, regenerates NADPH
	<i>IDP2</i>	PICST_43870	NADP-linked isocitrate dehydrogenase, cytosolic	HOG00262	FOG00618	ortholog	4.6	16.8	5	NADPH source during growth on non-fermentable carbon sources (Haselbeck and McAlister-Henn, 1993)
	<i>UGA2</i>	PICST_40468	NADP-linked succinate semialdehyde dehydrogenase, cytosolic	HOG00216	FOG00365	ortholog	12.2	15.56	0	involved in the utilization of gamma-aminobutyrate (Ramos et al., 1985)
	<i>UTR1</i>	PICST_87580	NAD(H) kinase, cytosolic	HOG00407	FOG00863	ortholog	4.0	18.0	0	de novo synthesis of NAD(H) pool (Shi et al., 2005)
NADPase candidates	<i>PHO3</i>	PICST_32593	acid phosphatase	HOG00624	FOG01534	absent	6.6	8.4	0	secreted acid phosphatase by <i>Candida albicans</i> under low-phosphate conditions (MacCallum et al., 2009)
	<i>PHO3.2</i>	PICST_47650	acid phosphatase	HOG00624	FOG01535	absent	2.4	2.9	4	uncharacterized ortholog group
	<i>PHO3.3</i>	PICST_83142	acid phosphatase-like protein	HOG00624	FOG01626	absent	2.1	0.9	0	uncharacterized ortholog group
	<i>PHO5</i>	PICST_46975	acid phosphatase	HOG00677	FOG01619	paralog	1.5	3.0	0	secreted acid phosphatase
	<i>PHO6</i>	PICST_61096	acid phosphatase	HOG00677	FOG01619	paralog	16.5	23.1	0	secreted acid phosphatase
	<i>PHO12</i>	PICST_46121	acid phosphatase	HOG00677	FOG01619	paralog	8.7	11.9	0	secreted acid phosphatase
	<i>YBU4</i>	PICST_74933	predicted tubulin-tyrosine ligase	HOG00671	FOG01612	ortholog	12.7	9.8	0	uncharacterized ortholog group
	<i>YMR1</i>	PICST_82137	myotubularin-related dual specificity phosphatase	HOG00670	FOG01611	ortholog	4.5	6.4	0	myotubularin phosphatase family, pathway signalling (Parrish et al., 2004)

^aHomolog Group (HOG) and Fungal Ortholog Group (FOG) identifications from AYbRAH (Correia et al., 2017).

^bRNA-seq from glucose and xylose fermentation (Yuan et al., 2011).

^cAverage spectral hits sampled from three time points in two independent xylose fermentation runs (Huang and Lefsrud, 2012).

^dReads Per Kilobase Million (RPKM).

Table 2. Xylose fermentation genotypes and phenotypes for *Debaryomyces hansenii*, *Suhomyces tanzawaensis*, *Spathaspora*, and *Scheffersomyces* species. GenBank Assembly Accessions were used to analyze the synteny of *XYL1* and *PHO3.2* loci.

Species	Organism code	Genotype		XR NADPH selectivity	Phenotype		Genome sequence	
		<i>XYL1.2</i>	<i>PHO3.2</i>		Xylitol yield (g/g)	Ethanol yield (g/g)	GenBank Assembly Accession	Reference
<i>Debaryomyces hansenii</i> CBS 767	dha				0.14-0.45 ^a	0.02-0.16 ^a	GCA_000006445.2	Dujon et al. (2004)
<i>Suhomyces tanzawaensis</i> NRRL Y-17324	ctz		✓				GCA_001661415.1	Riley et al. (2016)
<i>Scheffersomyces stipitis</i> CBS 6054	pic	✓	✓	0.58 ^b	0.01-0.06 ^{c,d}	0.45-0.47 ^{c,d}	GCA_000209165.1	Jeffries et al. (2007)
<i>Scheffersomyces stambukii</i> UFMG-CM-Y427	sheu		✓		0.58-0.66 ^e	0.12-0.17 ^e	GCA_002245345.1	Lopes et al. (2018)
<i>Scheffersomyces lignosus</i> JCM 9837	shel	✓	✓				GCA_001599395.1	
<i>Scheffersomyces shehatae</i> ATY839	she	✓	✓	0.71 ^f	0.18 ^c	0.37 ^c	GCA_002118035.1	Okada et al. (2017)
<i>Spathaspora passalidarum</i> NRRL Y-27907	spa	✓	✓	0.36 ^g	0.02-0.05 ^{g,d}	0.44-0.48 ^{g,d}	GCA_000223485.1	Wohlbach et al. (2011)
<i>Spathaspora arborariae</i> UFMG-19.1A	spaa		✓	0.57 ^g	0.04-0.18 ^g	0.31-0.32 ^g	GCA_000497715.1	Lobo et al. (2014)
<i>Spathaspora xylofermentans</i> UFMG-HMD23.3	spax		✓	1.00 ^g	0.51 ^g	0.08 ^g	GCA_002105455.1	Lopes et al. (2017)
<i>Spathaspora boniae</i> UFMG-CM-Y306	spau			0.96 ^h	0.46 ^h	0.19 ^h	GCA_002094185.1	Morais et al. (2017)
<i>Spathaspora girioi</i> UFMG-CM-Y302	spagi		✓	1.00 ⁱ	0.35 ⁱ	0.22 ⁱ	GCA_001657455.1	Lopes et al. (2016)
<i>Spathaspora hagerdaliae</i> UFMG-CM-Y303	spah		✓	1.00 ⁱ	0.21-0.24 ⁱ	0.25-0.28 ⁱ	GCA_001655755.1	Lopes et al. (2016)
<i>Spathaspora gorwiae</i> UFMG-CM-Y312	spago		✓	0.49 ⁱ	0 ⁱ	0.1 ⁱ	GCA_001655765.1	Lopes et al. (2016)

^aRoseiro et al. (1991)

^bVerduyn et al. (1985b)

^cLigthelm et al. (1988a)

^dVeras et al. (2017)

^eLopes et al. (2018)

^fHo et al. (1990)

^gCadete et al. (2016)

^hMorais et al. (2017)

ⁱLopes et al. (2016)

Table 3. Estimated xylose reductase cofactor selectivity using various techniques in the literature.

Method	NADH	NADPH	Reference
<i>In vitro</i> : purified enzyme characterization from <i>S. stipitis</i>	40%	60%	Verduyn et al. (1985b)
<i>In vitro</i> : crude enzyme assay in <i>S. stipitis</i> at varying aeration rates	40%	60%	Skoog and Hahn-Hägerdal (1990)
<i>In vivo</i> : <i>XYL1</i> and <i>XYL2</i> expression in <i>S. cerevisiae</i>	Minor	Major	Kötter and Ciriacy (1993)
<i>In vivo</i> : ¹³ C xylose (anaerobic)	100%	0%	Ligthelm et al. (1988c)
<i>In silico</i> : MFA (aerobic)	5%	95%	Dellweg et al. (1990)
<i>In silico</i> : MFA (anaerobic)	85%	15%	Dellweg et al. (1990)

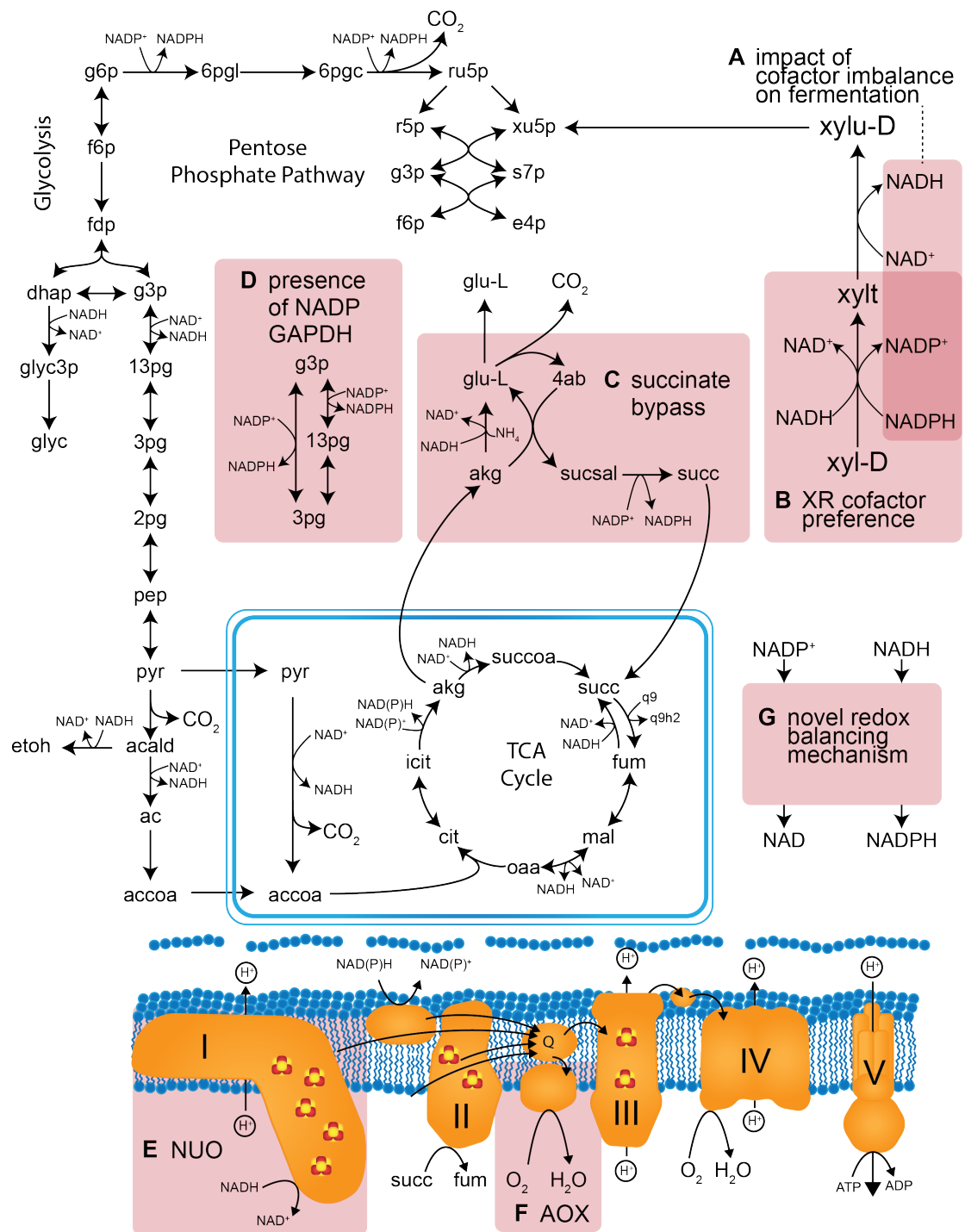


Figure 1. Simplified map of xylose fermentation in *Scheffersomyces stipitis* and potential redox balancing mechanisms. Uncertainties in xylose fermentation are highlighted in red squares: (A) the impact of the redox cofactor imbalance on metabolism, (B) the *in vivo* XR cofactor preference, (C) the use of the succinate bypass to regenerate NADPH, (D) the presence of non or phosphorylating glyceraldehyde 3-phosphate dehydrogenase (GAPDH) to regenerate NADPH, (E) the impact of bypassing Complex I (NUO) during xylose fermentation, (F) the ability of alternative oxidase (AOX) to oxidize NADH during xylose fermentation, and (G) the presence of novel redox balancing mechanisms.

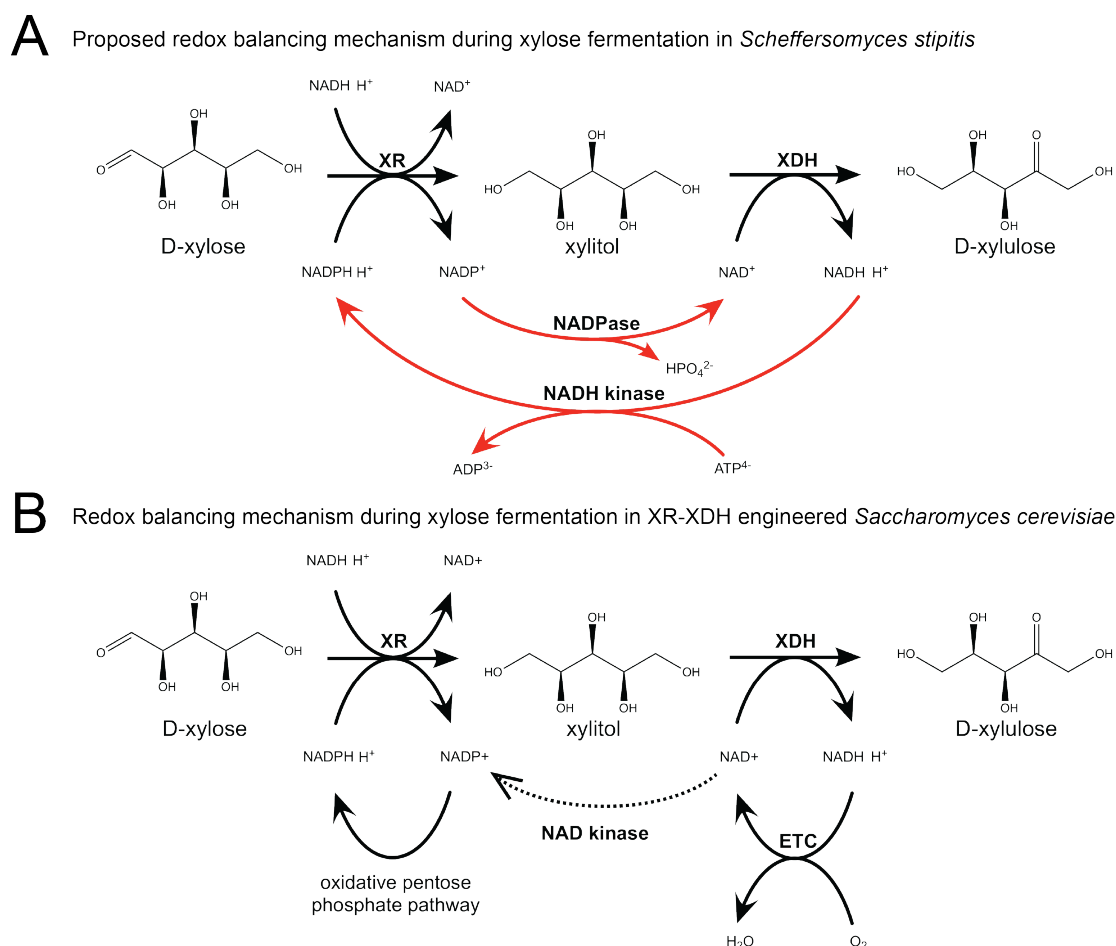


Figure 2. (A) Proposed redox balancing during xylose fermentation in *Scheffersomyces stipitis*. NADH kinase regenerates NADPH; NAD(P)H flux drives xylose reductase (XR); NADP phosphatase (NADPase) dephosphorylates NADP to NAD; NAD is reduced to NADH by xylitol dehydrogenase (XDH). This redox balancing scheme is consistent with the ^{13}C results from (Ligthelm et al., 1988c), independent of oxygen availability, does not have a loss of CO_2 from the oxidative pentose phosphate pathway, but requires ATP. **(B)** Redox balancing during xylose fermentation in engineered *S. cerevisiae* with the XR-XDH pathway from *S. stipitis*. NAD kinase phosphorylates a fraction of the NAD pool for *de novo* NADP synthesis (dotted line); the oxidative pentose phosphate pathway regenerates NADPH; NAD(P)H drives XR. XDH regenerates NADH; NADH is reoxidized to NAD by the electron transport chain (ETC). Under this redox balancing scheme, there is a loss of CO_2 from the oxidative pentose phosphate, oxygen is required to reoxidize NADH, and therefore xylose cannot be anaerobically fermented to ethanol at the maximum theoretical yield.

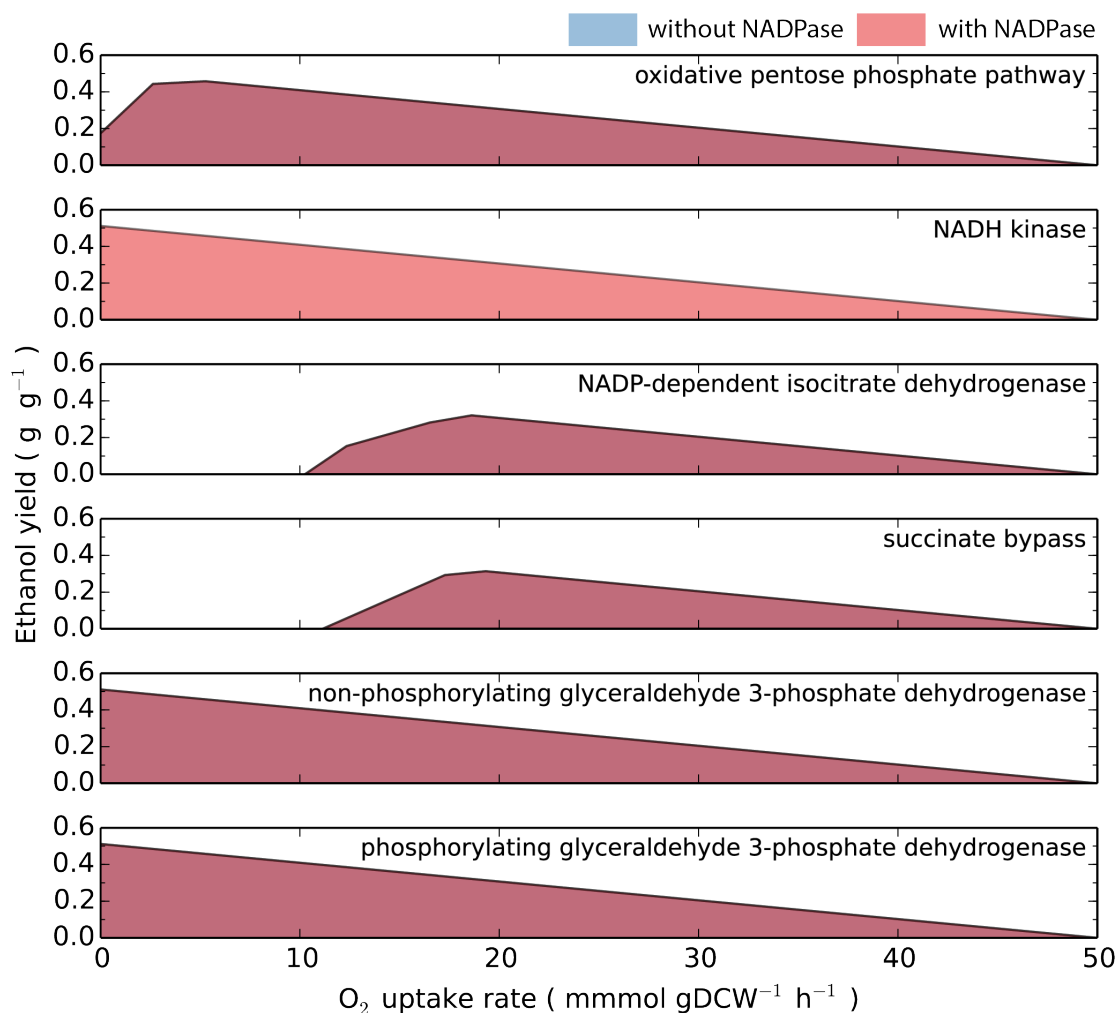


Figure 3. Ethanol yield as a function of NADPH source and oxygen uptake rate (OUR). There is a drop in the ethanol yield when the oxidative pentose phosphate regenerates NADPH at OUR's close to anaerobic levels. The highest ethanol yields were obtained with NADP phosphatase/NADH kinase, phosphorylating glyceraldehyde 3-phosphate dehydrogenase (GAPDH), and non-phosphorylating GAPDH. NADP-dependent isocitrate dehydrogenase and the succinate bypass were unable to ferment xylose to ethanol below 10 mmol · gDCW⁻¹ · h⁻¹.

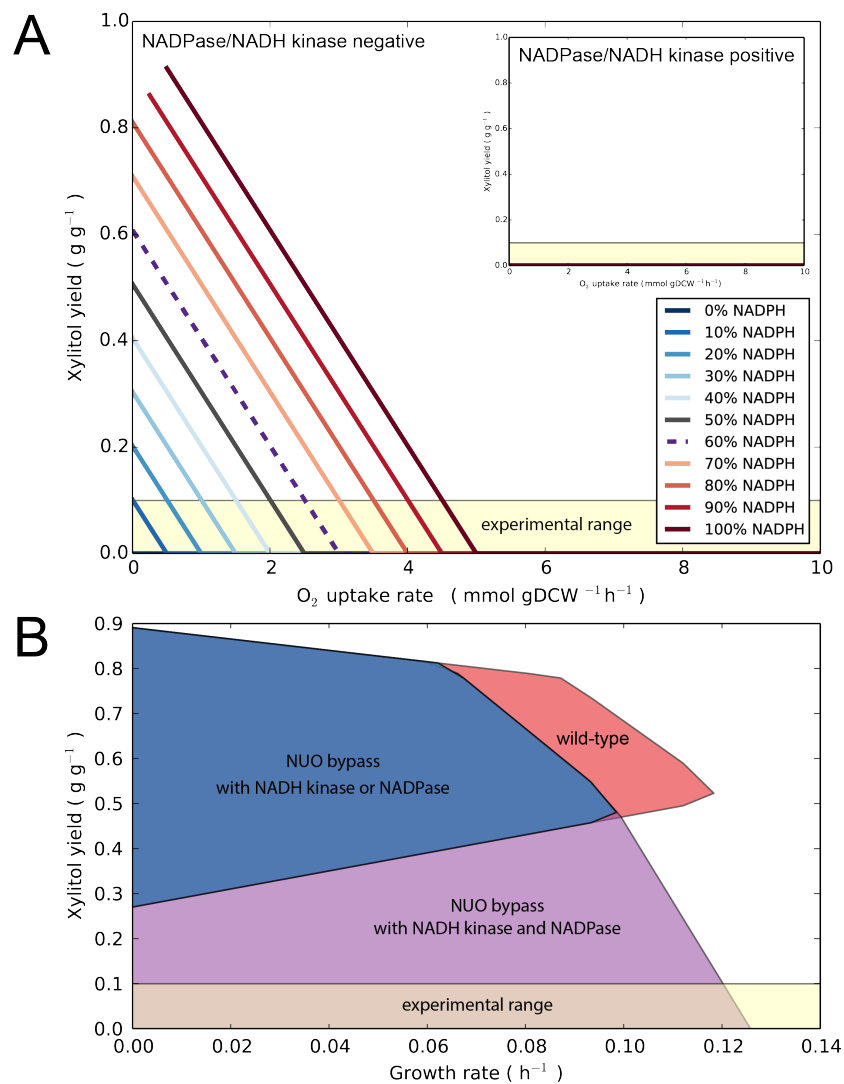


Figure 4. Xylitol yield sensitivity to oxygen uptake rate (OUR) and growth rate. **(A)** Simulations maximized xylose to ethanol with and without NADP phosphatase (NADPase) and NADH kinase. Anaerobic xylose fermentation is only feasible *in silico* when xylose reductase (XR) is driven by more than 80% NADPH. The *in silico* xylitol yield exceeds the 10% polyol yield typically observed *in vivo* when OUR is less than $2 \text{ mmol} \cdot \text{gDCW}^{-1} \cdot \text{h}^{-1}$. The presence of NADPase and NADH kinase in the metabolic model eliminates xylitol yield at all OUR's and XR cofactor selectivities. **(B)** Xylitol production envelope with and without NADPase and NADH kinase when XR is driven by 60% NADPH. OUR was constrained to $1 \text{ mmol} \cdot \text{gDCW}^{-1} \cdot \text{h}^{-1}$. Simulations without NADPase/NADH kinase lead to xylitol accumulation at the optimal growth rate and at all suboptimal growth rates. Bypassing Complex I (NUO) only reduces the maximum growth rate and has a marginal decrease in the xylitol yield at the optimal growth rate. The presence of NADH kinase or NADPase does not reduce the xylitol yield to the experimental polyol range; however, the addition of NADPase and NADH kinase enables the xylitol yield to fall within the experimental polyol range.

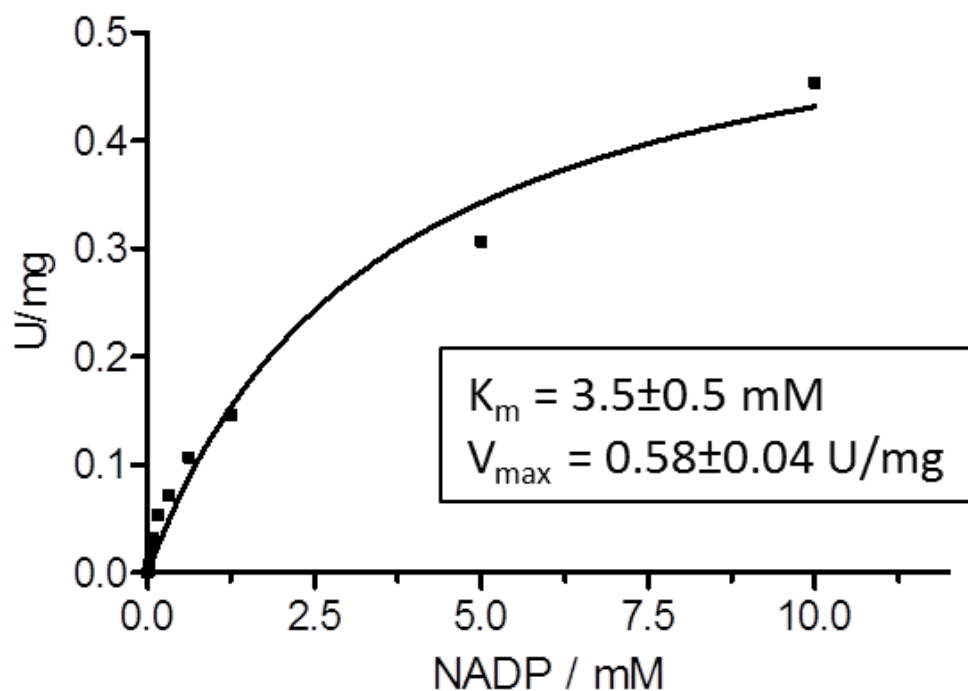


Figure 5. Pho3.2p Michaelis–Menten kinetics with NADP as a substrate. Reaction conditions: 50 mM HEPES pH 7.5, 50 mM NaFormate, 20 μg formate dehydrogenase, 5 mM MgCl_2 , 0.5 mM MnCl_2 , and 2.9 μg pure phosphatase added. The enzyme assay was not optimized for K_m or V_{max} .