Universidade de Brasília
Instituto de Ciências Biológicas
Departamento de Biologia Celular
Programa de Pós-Graduação em Biologia Molecular

BÁRBARA GOMES PAES

MOLECULAR AND PHYSIOLOGICAL RESPONSES OF *Pichia pastoris* TO LIGNOCELLULOSIC HYDROLYSATE COMPONENTS.

SUPERVISOR: PROF. DRA. ILDINETE SILVA PEREIRA

CO-SUPERVISOR: DR. JOÃO RICARDO MOREIRA DE ALMEIDA

SUPERVISOR DURING THE EXCHANGE: PROF. DR. DIETHARD MATTANOVICH

Brasília

2021

Universidade de Brasília
Instituto de Ciências Biológicas
Departamento de Biologia Celular
Programa de Pós-Graduação em Biologia Molecular

BÁRBARA GOMES PAES

RESPOSTAS MOLECULARES E FISIOLÓGICAS DE *Pichia pastoris* A COMPONENTES DE HIDROLISADO LIGNOCELULÓSICO.

ORIENTADORA: PROF. DRA. ILDINETE SILVA PEREIRA

COORIENTADOR: DR. JOÃO RICARDO MOREIRA DE ALMEIDA

COORIENTADOR NO ESTÁGIO SANDUÍCHE: PROF. DR. DIETHARD MATTANOVICH

Brasília

2021

BÁRBARA GOMES PAES

RESPOSTAS MOLECULARES E FISIOLÓGICAS DE *Pichia pastoris* A COMPONENTES DE HIDROLISADO LIGNOCELULÓSICO. / MOLECULAR AND PHYSIOLOGICAL RESPONSES OF *Pichia pastoris* TO LIGNOCELLULOSIC HYDROLYSATE COMPONENTS.

Tese apresentada ao programa de Pós-graduação em Biologia Molecular do Instituto de Ciências Biológicas da Universidade de Brasília para obtenção do título de Doutor em Biologia Molecular.

Área de concentração: Fisiologia e engenharia genética de microrganismos de interesse industrial.

Orientadora: Prof. Dra. Ildinete Silva Pereira
Coorientador: Dr. João Ricardo Moreira de Almeida
Coorientador no estágio sanduíche: Prof. Dr. Diethard Mattanovich

Brasília
2021

O desenvolvimento deste projeto de pesquisa foi apoiado financeiramente pela EMBRAPA e CNPq (Projeto YEASTACID), e executado na Embrapa Agroenergia. A Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) apoiou através da concessão de bolsa de estudos. A Universidade de Brasília (UnB) e Fundação de Apoio à Pesquisa do Distrito Federal (FAP/DF) apoiaram a participação em congressos e cursos. Doutorado sanduíche realizado através da bolsa Ernst Mach Worldwide 2019-2020 em parceria com Prof. Dr. Diethard Mattanovich e a Universidade de Recursos Naturais e Ciências da Vida (BOKU, Viena, Áustria).

Paes, Bárbara Gomes

RESPOSTAS MOLECULARES E FISIOLÓGICAS DE *Pichia pastoris* A COMPONENTES DE HIDROLISADO LIGNOCELULÓSICO. / MOLECULAR AND PHYSIOLOGICAL RESPONSES OF *Pichia pastoris* TO LIGNOCELLULOSIC HYDROLYSATE COMPONENTS.

Tese de Doutorado - Programa de Pós-graduação em Biologia Molecular, Instituto de Ciências Biológicas da Universidade de Brasília/UnB –

Orientadora: Ildinete Silva-Pereira

Respostas moleculares e fisiológicas de *Pichia pastoris* a componentes de hidrolisado lignocelulósico. / Molecular and physiological responses of *Pichia pastoris* to lignocellulosic hydrolysate components.

Bárbara Gomes Paes

Tese de Doutorado

Banca Examinadora

Membro Interno
Helson Mario Martins do Vale

Membro Externo
Leandro Vieira

Membro Externo
Betania Ferraz Quirino

Suplente
Fernando Araripe G. Torres

Brasília, 24 de maio 2021

Aos meus pais.

AGRADECIMENTOS

À minha família, que me apoiou incondicionalmente em todos os momentos, mas principalmente naqueles que eu estava caída. As vezes até com uma árvore, ou duas em cima. Agradeço aos meus pais por terem me dado herança mais valiosa (d)o mundo. Agradeço a minha irmã por ser minha companheira de aventura. As vezes até com uma árvore, ou duas em cima.

À minha orientadora Ildinete Silva-Pereira, que juntamente à Cynthia Kyaw são minhas fadas madrinhas acadêmicas, e por entre outras coisas, terem me apresentado o João Ricardo no começo da minha vida acadêmica. Ao querido (co)orientador João Ricardo, por todos os ensinamentos, paciência, exemplos e lições importantes, algumas até relacionadas com academia, pesquisa e ciência. Desejo a todos que quero bem, mentores como os que eu tive.

To Diethard Mattanovich, for giving me the opportunity to join his lab in Vienna for part of my Ph.D. research, which brought me joy, good times, and great people and experiences to my life. To all the people in the Mattanovich-Gasser lab, for sharing not only thoughts and laminas, but stories and moments: Marina Jecmenica, Jennifer Staudacher, Özge Ata, Donny Rudinatha, Thomas Gaßler, and many others. A special thanks to Lina Heistinger for so much support, and Nora Tir for putting up with me at home too. Danke meine Lieben!

Agradeço muito aos meus amigos, e ao me sentar para escrever, percebi quanta gente incrível colecionei na minha jornada. Aos que achei nos laboratórios do Brasil, Nathália Vilela, Débora Trichez, Clara Vida, Henrique Veras, Letícia Malmann, Luana Serra, Victor Mendes, Fernanda Guilhelmelli. Aos que viraram minha família na Áustria: Gica Trierweiler e Andreas Jurdak, Apolinário Passos, e Ale Burger. Aos amigos que me faziam lembrar de mim, do que gosto, e quem eu sou para além da pós-graduação, os do RPG, dos Femivinho, do que o sol toca, da divulgação científica, da Garagem. Aos que tropecei e catei para mim: Bruno Amui, Adriana Miranda e Álvaro Campos, Veronica Slobodian e Igor Drummond, Débora Azzi-Nogueira, Thaís Ziober, Tupá Guerra, Marcelo Moussallem, Luciana Oliveira, Oscar Oliveira, Milla Góes, Ana Noronha, Lígia Amoroso, Flávia Sant'Anna, Gabriela Sobral e Lucas Camargos, e outros tantos outros que foram sorrisos, lágrimas, conselhos e momentos inesquecíveis. Muitíssimo obrigada, membros do Dragões de Garagem,

que foram verdadeiras luzes na escuridão. Na Europa me disseram que eu uso a palavra “amigo” muito livremente. Eu só me considero uma pessoa de sorte mesmo.

Agradeço aos artistas, músicos, aos podcasters, e tantos outros produtores de conteúdo que foram companhia e inspiração, descanso, informação e trilha sonora em tantos e tantos e tantos momentos. Agradeço a todos os cientistas brasileiros que resistiram, se expuseram e lutaram contra o obscurantismo, a desinformação, e o negacionismos, em um dos momentos mais difíceis como sociedade.

À EMBRAPA Agroenergia que me deu todo apoio, oportunidade e estrutura de desenvolver o projeto. À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e ao Austrian Federal Ministry of Education, Science and Research (BMBWF) pelo apoio financeiro.

À sociedade brasileira.

E a todos que, de uma forma ou de outra, se sintam parte deste processo (ou não), o meu sincero muito obrigada.

I don’t know what I’m looking for.”

“What not?”

“Because … because … I think it might be because if I knew I wouldn’t be able to look for them.”

“What, are you, crazy?”

“It's a possibility I haven't ruled out yet.”

Douglas Adams

The Hitchhiker's Guide to the Galaxy

“I didn’t want to just know names of things. I remember really wanting to know how it all worked.”

Elisabeth Blackburn

Nobel prize winner in physiology/ medicine 2009

RESUMO

Komagataella phaffii, anteriormente conhecida como *Pichia pastoris*, é um organismo modelo comummente usado na pesquisa e na indústria. Tem sido considerada para a produção de produtos de alto valor agregado, especialmente proteínas. Suas aplicações biotecnológicas envolvem principalmente a utilização de fontes clássicas de carbono como glicose, glicerol, sorbitol e metanol, embora muitos outros substratos também tenham sido pesquisados. As biomassas lignocelulósicas são ricas em açúcares fermentáveis, como a xilose, o segundo açúcar mais abundante na natureza, que pode ser usado para a produção de produtos químicos renováveis. No entanto, a utilização microbiana da xilose depende da capacidade inata de utilização desta pelo microrganismo, e de sua tolerância aos compostos inibidores presentes nos hidrolisados lignocelulósicos. Atualmente, entende-se que *P. pastoris* não é capaz de utilizar xilose como fonte de carbono a menos que seja modificada para isso, e sua sensibilidade a tais inibidores é pouco compreendida. Neste estudo, apresentamos as respostas moleculares e fisiológicas de *P. pastoris* aos principais componentes do hidrolisado lignocelulósico: inibidores derivados da lignocelulose, e a xilose. A análise fisiológica e transcricional de *P. pastoris* X33 para ácido acético, furaldeídos e hidrolisado de bagaço de cana-de-açúcar mostrou que eles afetam o metabolismo celular de forma dose-dependente, e se correlacionam positivamente com a quantidade de genes diferencialmente expressos. Ao contrário de outras leveduras industriais como *Saccharomyces cerevisiae*, *P. pastoris* pode consumir concomitantemente ácido acético com glicose como fonte de carbono, possivelmente ajudando a reduzir sua toxicidade para as células. A tolerância ao inibidor, e a capacidade de utilização da xilose de 25 isolados diferentes naturais e um de laboratório de *Komagataella*, de seis espécies diferentes, também foram avaliadas. Nenhum isolado com tolerância aos inibidores distintamente maior foi identificado, entretanto linhagens capazes de crescer em xilose foram. Três linhagens de melhor crescimento foram selecionadas e submetidas à engenharia adaptativa de laboratório (ALE) para otimização do consumo de xilose. A caracterização detalhada da assimilação da xilose pela via de oxirredução foi confirmada por ensaios utilizando marcação com isótopos de carbono ^{13}C, apesar de precisar de mais de dez dias para duplicar. Por fim, uma estratégia de engenharia genética foi empregada para melhorar a tolerância de *P. pastoris* X33 ao ácido acético. Para isso, o gene homólogo *HAA1*,

previamente descrito em *S. cerevisiae* como fator de transcrição envolvido na resposta ao estresse com ácido acético, foi identificado e superexpresso em *P. pastoris* X33. Isso melhorou o crescimento da levedura na presença de 2g.L^{-1} 4,9 vezes após 24 h de cultivo. Juntos, os resultados apresentados aqui abrem caminho para a compreensão do metabolismo de *P. pastoris* na presença de hidrolisados lignocelulósicos, ácido acético, furaldeídos e xilose.

Palavras-chave: *Pichia pastoris. Komagataella phaffii.* Hidrolisado lignocelulósico. Xilose. Ácido acético. Furaldeídos. Transcritoma.

ABSTRACT

Komagataella phaffii, previously known as *Pichia pastoris,* is a common model organism used in research and industry. It has been considered for the production of high value-added products, especially proteins. Its biotechnological applications mostly involve the utilization of classical carbon sources like glucose, glycerol, sorbitol, and methanol, although many other substrates have also been researched. Lignocellulosic biomasses are rich in fermentable sugars, like xylose, the second most abundant sugar in nature, which can be used for the production of renewable chemicals. The microbial utilization of xylose is dependent on the innate utilization capacity by the microorganism, and its tolerance to inhibitory compounds present in lignocellulosic hydrolysates. The current understanding is that *P. pastoris* is not capable of utilizing xylose as a carbon source unless engineered towards it, and its sensitivity to such inhibitors is poorly understood. In this study, we present the molecular and physiological responses of *P. pastoris* to the major lignocellulosic hydrolysate components: lignocellulose-derived inhibitors, and xylose. The physiological and transcriptional analysis of *P. pastoris* X33 to acetic acid, furaldehydes, and sugarcane bagasse hydrolysate showed they affect cell metabolism in a dose-dependent way, and it positively correlates with the amount of differentially expressed genes. Unlike other industrial yeasts like *Saccharomyces cerevisiae*, *P. pastoris* can co-consume acetic acid with glucose as carbon source, possibly helping it reduce its toxicity to the cells. The inhibitor tolerance and xylose utilization capacity of 25 different natural and one laboratory *Komagataella* isolates of six different species was also evaluated. No isolates with higher tolerance towards inhibitors were identified, however, strains capable of growing on xylose were. The three best growing strains were selected and underwent adaptative laboratory evolution (ALE) for xylose consumption optimization. Detailed characterization of xylose assimilation via the oxidoreductase pathway was confirmed using carbon isotope ^{13}C labeling, despite it needing more than ten days to duplicate. Finally, a genetic engineering strategy was employed to improve *P. pastoris* X33 tolerance towards acetic acid. For this, the homologous *HAA1* gene previously described in *S. cerevisiae* as a transcriptional factor involved in acetic acid stress response was identified and overexpressed in *P. pastoris* X33. These improved yeast's growth in the presence of $2 g.L^{-1}$ 4.9-fold after 24 h of cultivation. Altogether the results presented here paves the way to understanding

P. pastoris' metabolism in presence of lignocellulosic hydrolysates, acetic acid, furaldehydes, and xylose.

Keywords: *Pichia pastoris. Komagataella phaffii.* Lignocellulosic hydrolysate. Xylose. Acetic acid. Furaldehydes. Transcriptome.

LIST OF FIGURES

Figure 1 – Primary and secondary metabolites produced in *P. pastoris.* 23

Figure 2 – Structure of lignocellulosic biomass and its constitutive polymers; cellulose, hemicellulose, and lignin. 24

Figure 3 – Lignocellulosic biomass composition, its main constitutive monomers, and pre-treatment derived inhibitors. 26

Figure 4 – Xylose metabolic pathways of xylose in various microorganisms.. 28

Figure 5 – Schematic representation of furaldehydes, phenolics, and organic acid effects on yeast cells. 30

Figure 6 – *P. pastoris* growth in microtiter plates containing YNB supplemented with different dilutions of lignocellulosic hydrolysate and 5,4 $g.L^{-1}$ glucose. 44

Figure 7 – *P. pastoris* growth profile in the presence of different inhibitors. 46

Figure 8 – Principal component analysis (PCA) of RNA-seq data of *P. pastoris* cultivated in different inhibitors.. 48

Figure 9 – Venn diagrams representing *P. pastoris* differentially expressed genes (DEGs) in response to lignocellulose-derived inhibitors. 49

Figure 10 – Gene expression of central carbon metabolism pathways for *P. pastoris*. Possible carbon sources are presented in colored squares: glucose; glycerol, ethanol, acetate. Bar charts represent the transcriptional changes (log2 fold) of genes in acetic acid 2g.L-1 (light blue), 6g.L-1 (dark blue), FH 0.9/0.15g.L-1 (yellow), FH 1.5/0.25g.L-1 (orange), hydrolysate 10% (light green) and 30% (dark green) with P-values <0.05 or 0.1 (* on top of bar chart). 51

Figure 11 – Representativeness of GO categories for the DEGs of *P. pastoris* in presence of different inhibitors. 54

Figure 12 – Hierarchical clustering heat map of the 630 DEGs common in all categories. 60

Figure 13 – Schematic representation of the 96-deep-well plate for dilutions (a) and agar plate (b) of the spotting assays. 68

Figure 14 – Growth pattern of *Komagataella* strains in media with or without the addition of lignocellulose-derived inhibitors. 71

Figure 15 – Spotting assay of 26 strains of *Komagataella* strains in YNB supplemented with glucose-containing $2g.L^{-1}$ acetic acid in three different time points: 48 h, 120 h, and 168 h. 72

Figure 16 – Spotting assay of 26 strains of *Komagataella* in strains in media containing 0.15 g.L^{-1} HMF and 0.9 g.L^{-1} of furfural in three different time points: 48 h, 120 h, and 168 h. 74

Figure 17 – Stacking experiment of *Komagataella* strains plated and stacked in the presence or absence of furaldehydes. 75

Figure 18 – Spotting assay of 26 *Komagataella* strains in YNB media with glucose and xylose after 120 h of incubation. 77

Figure 19 – Spotting assay of selected strains on xylose. From top to bottom, strains *K. phaffii X33*, *C. lignohabitans*, *K. phaffii CBS2612*, *K. phaffii CBS7435*, *K. pastoris,* and *K. populi* in different time points of growth.....78

Figure 20 – Culture of strains *K. phaffii X33*, *K. pastoris*, and *K. populi* in YNB supplemented with 2% Xylose (continuous line) and YNB with no addition of carbon source (dotted line, -C)......79

Figure 21 – *K. phaffii X33* cultivations in xylose isotopes ^{12}C and ^{13}C.....80

Figure 22 – Comparison of the general growth of strains before (dotted line) and after (continuous lines) undergoing adaptive laboratory evolution process. 82

Figure 23 – Cloning strategy for *HAA1* overexpression. Source: Author's collection.96

Figure 24 – Overview of plasmids for HAA1 expression in *P. pastoris* assembled....97

Figure 25 – Confirmation of the constructed systems for HAA1 overexpression.....98

Figure 27 – Growth curves of strains overexpressing HAA1 under the promoters *pGAP (red triangle)* or p *RPP1b (blue diamond) in YNB media with or without supplementation of 2 g.L^{-1} of acetic acid (AA2)*. 101

Figure 28 – Intracellular amounts of acetic acid on X33 (gray-left), p *RPP1b*-HAA1(orange-middle), and p*GAP*-HAA1(blue-right) strains.....102

LIST OF TABLES

Table 1 – Composition of different lignocellulosic feedstock after physicochemical pretreatment. 27

Table 2 – Strategies for enhancing microbial tolerance to inhibitors. 33

Table 3 – Selected differentially expressed genes (log 2-fold) of *P. pastoris* cultivated in presence of acetic acid, furaldehydes (HMF and furfural), and sugarcane biomass hydrolysate.. 56

Table 4 – List of *Komagataella* strains used in Chapter 2, their codes, origin of isolation, location, and collection source. 65

Table 5 – List of primers used in Chapter 3. Source: Author's collection. 89

Table 6 – List of plasmids used in Chapter 3. Source: Author's collection. 90

LIST OF ABBREVIATIONS AND SYMBOLS

%	Percent
ALE	Adaptative laboratory evolution
Bp	Base pairs
CDS	Coding Site
Cl	*Candida lignohabitans*
DCW	Dry cell weight
DEG	Differentially expressed gene
DNA	Desoxyribonucleic acid
DTT	DL-dithiothreitol
EA-IRMS	Elemental analysis with isotope ratio mass spectrometry
EDTA	Ethylenediamine tetraacetic acid
FB	FB Buffer
G	Grams
GGA	Golden Gate assembly
GO	Gene ontology
H	Hours
HCl	Chloridoid acid
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HMF	Hydroximethylfurfural
HPLC	High liquid throughput chromatography
KH_2PO_4	Potassium dihydrogen phosphate
LB	Luria Bertani media
M	Molar
Min	Minutes
mL	Milliliter
mM	Millimolar
mRNA	Messenger RNA
Na_2HPO_4 *$2H_2O$	Disodium hydrogen phosphate dihydrate
NaCl	Sodium Chloride
NaOH	Sodium hydroxide
Nm	Nanometers

ºC	Degrees Celsius
OD_{600}	Optical density at 600 nanometers
PBS	Phosphate-buffered saline
PCA	Principal component analysis
PCR	Polymerase chain reaction
pH	Potassium hydrogenionic
QS	Quenching solution
RNA	Ribonucleic acid
RNA-Seq	RNA sequencing
Rpm	Rotations per minute
TAE	Tris base, acetic acid, EDTA buffer
TCA	Tricarboxylic acid cycle
w/v	weigh / volume
WT	Wild type
X*g*	times gravity force
YNB	Yeast nitrogen base
YPD	Yeast peptone dextrose
YPG	Yeast peptone glycerol
µL`	Microliter
µM	Micromolar
LDI	Lignocellulose derived inhibitors

CONTENTS

1 INTRODUCTION 21
1.1 *Pichia pastoris* 21
1.1.1 Industrial and biotechnological applications 22
1.2 LIGNOCELLULOSIC BIOMASS 23
1.2.1 Xylose 28
1.2.2 Lignocellulose derived inhibitors (LDI) 29
1.3 INCREASED TOLERANCE OF YEAST TOWARD LIGNOCELLULOSE DERIVED INHIBITORS 31
1.4 JUSTIFICATION 34
2 OBJECTIVES OF THE WORK 35
2.1 GENERAL OBJECTIVE 35
2.2 SPECIFIC OBJECTIVES 35
CHAPTER 1 – *PICHIA PASTORIS* RESPONSE TO LIGNOCELLULOSE-DERIVED INHIBITORS 36
3 MATERIALS AND METHODS 39
3.1 STRAIN AND MEDIA 39
3.2 *P. pastoris* GROWTH ASSAY IN LIGNOCELLULOSIC HYDROLYSATE 39
3.3 PHYSIOLOGICAL CHARACTERIZATION OF *P pastoris* IN PRESENCE OF LIGNOCELLULOSE DERIVED INHIBITORS (LDI) 40
3.4 RNA EXTRACTION AND QUALITY ANALYSIS 40
3.5 RNA SEQUENCING AND DATA ANALYSIS 40
3.6 QUANTIFICATION OF METABOLITES 42
4 RESULTS AND DISCUSSION 43
4.1 EFFECT OF SUGARCANE BAGASSE HYDROLYSATE ON *P. pastoris* GROWTH 43
4.2 *P. pastoris* FERMENTATIVE PROFILE IN PRESENCE OF LIGNOCELLULOSE-DERIVED INHIBITORS 45
4.3 TRANSCRIPTIONAL RESPONSE OF *P. pastoris* TOWARDS LIGNOCELLULOSE-DERIVED INHIBITORS 47
4.4 CENTRAL CARBON METABOLISM 50
4.5 GENE ONTOLOGY ANALYSIS 52
5 CONCLUSIONS 61

CHAPTER 2 – CHARACTERIZATION OF NEW *KOMAGATAELLA* STRAINS FOR TOLERANCE TO ACETIC ACID, FURALDEHYDES, AND XYLOSE CONSUMPTION 62
6 MATERIALS AND METHODS.....65
6.1 YEAST STRAINS AND MAINTENANCE.....65
6.2 CULTURE MEDIA AND SOLUTIONS.....67
6.2.1 PBS 10X (Stock) 67
6.2.2 YNB.....67
6.3 SPOTTING ASSAY 67
6.4 FLASK GROWTH ASSAY.....68
6.5 DRY CELL WEIGHT (DCW).....68
6.6 ^{13}C XYLOSE CULTURE ASSAY 69
6.7 BIOMASS LABELING (ISOTOPE RATIO MASS SPECTROMETRY) 69
6.8 ADAPTIVE LABORATORY EVOLUTION (ALE) 69
7 RESULTS AND DISCUSSION.....70
7.1 TOLERANCE OF *Komagataella* STRAINS TO ACETIC ACID AND FURALDEHYDES 70
7.2 XYLOSE METABOLISM BY *Komagataella* STRAINS 77
7.2.1 Spotting assays.....77
7.2.2 Characterization of selected *Komagataella* strains on xylose 79
7.2.3 Confirmation of xylose metabolism by *K. phaffii* using labeled xylose. ...80
7.2.4 Adaptive laboratory evolution (ALE) 81
8 CONCLUSIONS.....83
CHAPTER 3 – OVEREXPRESSION OF THE GENE *HAA1* ENHANCES ACETIC ACID TOLERANCE IN *PICHIA PASTORIS* 85
8 MATERIALS AND METHODS.....87
8.1 CELL STRAINS AND MAINTENANCE 87
8.1.1 *Escherichia coli* 87
8.1.2 *Pichia pastoris X33* (Invitrogen).....87
8.2 CULTURE MEDIA AND SOLUTIONS.....87
8.2.1 PBS 10X (Stock) 87
8.2.2 TAE Buffer 50x (Stock).....88
8.2.3 FB buffer (sterile filtrated) 88
8.2.4 HEPES 1M pH 8,0 (Stock) 88

8.2.5 DTT 1M (Stock) 88
8.2.6 Sorbitol 1M (Stock) 88
8.2.7 0,5 M EDTA 88
8.2.8 YNB 88
8.3 *HAA1* IDENTIFICATION 89
8.4 GOLDEN GATE ASSEMBLY (GGA) 89
8.5 PCR – POLYMERASE CHAIN REACTION 89
8.6 PRIMERS UTILIZED IN CHAPTER 3 89
8.7 PLASMIDS ARE UTILIZED IN CHAPTER 3 90
8.8 CONSTRUCTION OF *HAA1* EXPRESSION VECTORS 91
8.9 AGAROSE GEL ELECTROPHORESIS 91
8.10 CHEMICALLY COMPETENT *E. coli* CELL PREPARATION AND TRANSFORMATION 91
8.11 ELECTRO-COMPETENT *P. pastoris* CELLS AND TRANSFORMATION 92
8.12 SCREENING FOR ACETIC ACID-TOLERANT CLONES 93
8.13 SPOTTING ASSAY 93
8.14 SEQUENCING 93
8.15 INTRACELLULAR ACETIC ACID CONTENT ASSAY 93
9 RESULTS AND DISCUSSION 95
9.1 *HAA1* OVEREXPRESSION 95
9.2 SPOTTING ASSAYS 98
9.3 GROWTH ASSAYS IN LIQUID MEDIA 100
9.4 INTRACELLULAR ACETIC ACID ASSAY 101
10 CONCLUSIONS 104
11 FUTURE PROSPECTS 105
REFERENCES 107
APPENDIX 127
Appendix A – Published papers 127
Appendix B – Supplementary material 128
B1 - Supplementary Tables 128
B2 - *Pichia pastoris* (*Komagataella phaffii*) *HAA1* complete sequence 128
B3 – *Pichia pastoris* (*Komagataella phaffii*) mutated *HAA1* sequence 129

1 INTRODUCTION

1.1 *Pichia pastoris*

The yeast *P. pastoris* was isolated in the 1920s (1) and extensively studied for a distinctive and not so common feature: the capacity of oxidizing methanol to produce energy, assimilate it as the sole carbon source for growth, and generation of products of interest. Today, *P. pastoris* is established as one of the most common eukaryotic expression systems both in academia and industry and is mostly used for the production of recombinant proteins, as reviewed by (2–4).

Initially described as *Zygosaccharomyces pastori* (1), it had its name changed into the *Pichia pastoris* by Herman Phaff when other related strains were isolated in the 1950s (5). More recently, thanks to more modern sequencing techniques and phylogenetic analysis, *Pichia* was reclassified to the new genus *Komagataella* (6), and the previous *P. pastoris* was split into two different species: *K. phaffii* and *K. pastoris* (7). Their genome was sequenced in 2009, still before the separation into two species (8,9). Until the genome sequencing, not much besides physiological characteristics (fast growth, cheap and easy cultivation conditions, methylotrophy, efficient heterologous protein production) was known (3,4). Many information gaps still exist, however in recent decades, technological advances in sequencing, annotation, genomic editing techniques, characterization, and metabolic models have expanded our understanding and accumulated knowledge of these yeasts (10,11).

Currently, the genus *Komagataella* consists of seven species based on the divergence of marker gene sequences: *K. phaffii, K. pastoris, K. ulmi, K. kurtzmanii, K. mondaviorum, K. populi,* and *K. pseudopastoris* (12). All the strains were isolated from trees in either North America or Europe. Few natural isolates are found in culture collections and these have been thoroughly characterized (13). Previous *P. pastoris* strains, now called *K. phaffii* and *K. pastoris,* have been the most studied among all, and commonly used as protein expression systems. The characterization of *Komagataella* strains and the discovery of new members of this family can contribute to a better understanding of speciation and its unique mechanisms, besides providing new biotechnological tools.

1.1.1 Industrial and biotechnological applications

When researchers discovered *P. pastoris* was capable of using methanol (14) it was expected to be used as a protein source for animal nutrition (15). This was based on *P. pastoris* ability to grow to high cell densities (>130 g/L dry cell weight on methanol in continuous culture) and use methanol synthesized from natural gas (16). The reduction of soy prices (the main feed alternative) and the oil crisis in the ’80s however, made the development of the process unviable.

In the 1980s, the isolation and characterization of *AOX1* and *AOX2* genes permitted the development of molecular tools and methods of *P. pastoris* genetic manipulations (17). The creation of a series of products for researchers to work with *P. pastoris,* like vectors, strains, and commercial kits made it one of the most used eukaryotic model organism, as reviewed by (4,15).

Currently, *P. pastoris* is a major recombinant protein expression system in both academia and industry as a result of its high levels of protein production, secretion, and type of post-translational modifications. It was engineered to have a human-like glycosylation pattern (18) making it a great model for medical research, and the development of, for example, antibodies through yeast fermentation (19). It is estimated that more than five thousand different proteins have used *P. pastoris* as a production host (2), and it also demonstrates a great potential for the production of different metabolites (Fig. 1). Through metabolic engineering *P. pastoris* was designed for producing lactic acid, lovastatin, monacolins, lycopene, xylitol, riboflavin, and others, (reviewed in 3).

Like most industrial microorganisms, *P. pastoris’* initial application was based on its natural distinctive ability of interest: methanol consumption. With tools for genetic manipulation, metabolic engineering, and synthetic biology, scientists and industries can design microorganisms based on key intrinsic characteristics towards new industrial traits of interest (20). Carbon source utilization is an essential limitation in the use of microorganisms in biotechnology. *P. pastoris* is known to use glucose, glycerol, and ethanol as sole carbon sources. Research including mannitol, alanine, trehalose (21), lactic acid (22) sorbitol, methanol, fructose, L-rhamnose, acetic acid, and others (23–25) as carbon sources has been done.

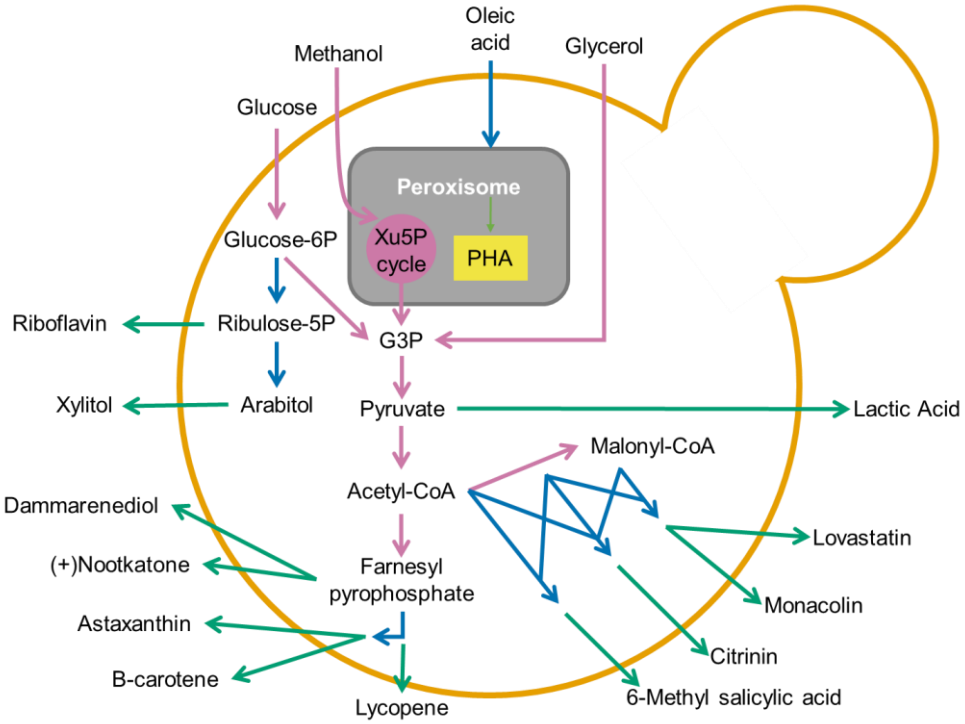

Figure 1 – Primary and secondary metabolites produced in *P. pastoris*. Specific pathways towards products already made in *P. pastoris*, and to their precursors are highlighted in green and blue, respectively. Common metabolic nodes at key precursors (pathways highlighted in pink) indicate targets for future engineering of chassis strains as platforms for the development of multiple production strains for different metabolic products. Source: Adapted from (3).

The increasing commercial demands for sustainable and competitive processes have driven research towards investigating even more unconventional carbon sources for *P. pastoris* utilization. Through metabolic engineering, strains capable of metabolizing xylose (26) and cellulose (27) have been created. More recently, *P. pastoris* was also converted into an autotroph capable of growing into CO_2 (28). The use of cheap and abundant substrates for microorganism utilization is viewed as a key strategy in the development of biotechnological processes.

1.2 LIGNOCELLULOSIC BIOMASS

Lignocellulosic biomass is the main structural component of plants, the largest source of terrestrial organic matter, and it has been cited as the most promising raw

material for bioprocesses because of its sustainability, availability, and low price (29). The term refers to a complex structure constituted by cellulose, hemicellulose, and lignin. Proportions of the three components vary depending on the species, age, environment, and tissue of the plant. Cellulose is a homogeneous polymer composed of D-glucose and represents between 30% to 70% of the total biomass (Fig. 2). Hemicellulose is formed by pentoses (xylose, arabinose) and hexoses (mannose, glucose, galactose), and makes up 25% to 35% of the biomass. Lignin, the smallest fraction of biomass (10% to 25%), is composed of long heterogeneous chains of aromatic alcohols that make it insoluble in water (30–32).

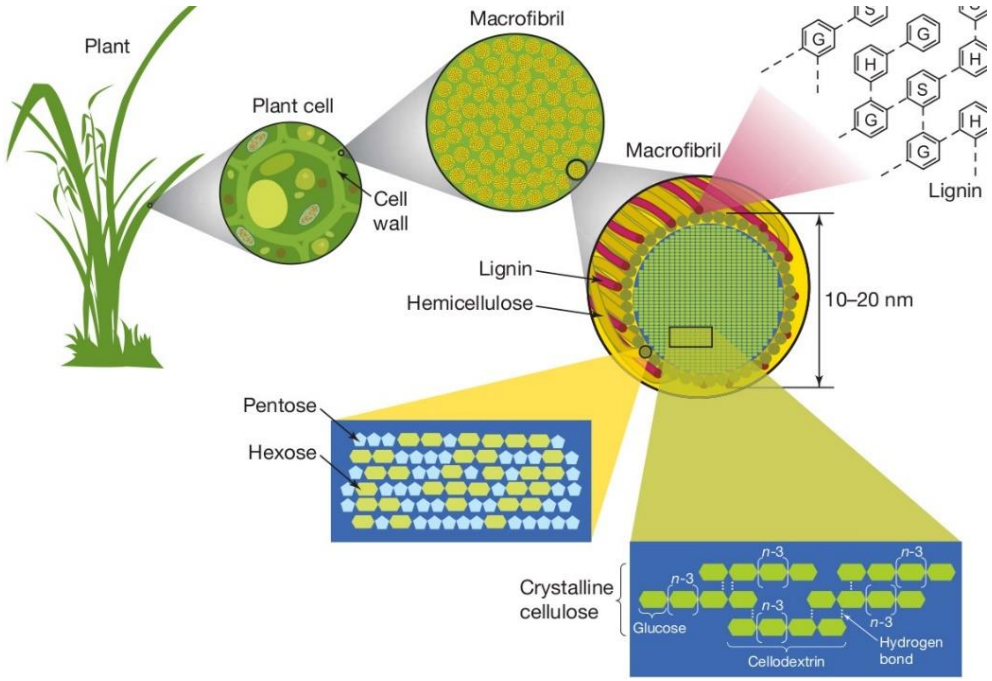

Figure 2 – Structure of lignocellulosic biomass and its constitutive polymers; cellulose, hemicellulose, and lignin. Source: (33).

The high sugar content from different hydrolysates can be sourced as a carbon source by microorganisms. These sugars however are not in an available fermentable form and before they can be used, the biomass needs to undergo pretreatment and hydrolysis processes. Pretreatment is the first step to make lignocellulose sugars available for microbial utilization. It aims to disrupt the tight structures of the main components of biomass and make cellulose and hemicellulose polymers more accessible for enzymatic digestion. Pretreatment methods can be chemical (acids,

base, chemical pulping, other solvents), physical (temperatures), oxidative (involve H_2O_2 and O_2 and O_3), biological (microorganisms), and/or a combination of these, like physicochemical (steam explosion, liquid hot water). Each technique will target a specific portion of the biomass and have different by-product formations (32,34). An ideal pretreatment should maximize sugar recovery and minimize inhibitor formation (35).

In general, acid pretreatments are capable of hydrolyzing part of the hemicellulose into soluble sugars. Residual hemicellulose and cellulose can then undergo enzymatic hydrolysis to convert leftover polymers into fermentable monomers. Hydrolysis of cellulose requires the synergistic activity of three main groups of enzymes: cellobiohydrolases (CBH, EC 3.2.1.91) that hydrolyze the exposed ends of cellulose chains; Endo-1,4- ß-D-glucanases (EB, EC 3.2.1.4) that randomly cleaves internal bonds in amorpha cellulose; and beta-glucosidase (EC 3.2.1.21) that hydrolyze cellobiose into glucose monomers. Other groups of proteins can act as auxiliary to cellulases on the cleavage and reduction of cellulose crystallinity, like swollenins, expansins, and lytic polysaccharide monooxygenases (LPMO) (29,36).

Hemicellulose heterogeneous structure results in a more complex mechanism necessary for hydrolysis when compared to cellulose. Hemicellulases include enzymes that will attack either internal bonds or oligosaccharides liberating monomers, respectively endo-1,4-β-d-xylanases (EC 3.2.1.8), and 1,4-β-d-xylosidases (EC 3.2.1.37) liberating xylose, endo-1,4-β-d- mannanases (EC 3.2.1.78), and 1,4-β-d-mannosidases (EC 3.2.1.25) liberating mannose. Other enzymes include α-d-galactosidases (EC 3.2.1.22), α-l-arabinofuranosidases (EC 3.2.1.55), α-glucuronidases (EC 3.2.1.139), acetyl xylan esterases (EC 3.1.1.72) and feruloyl and p-cumaric acid esterases (EC 3.1.1.73), and will, among others, liberate arabinose and galactose from hemicellulase (36,37).

Many microorganisms, including bacteria, fungi, and animals, have been reported to produce cellulases. Filamentous fungi like *Trichoderma reesei, Aspergillus orizae,* and *Penicillium funiculosum* are famous for producing efficient cellulases in high amounts and have been extensively studied (37,38). Currently, most commercial enzymatic cocktails employed in plant biomass degradation are fungal-derived. These are produced by companies like Novozymes, DuPont, Genencor, Dyadic, and AB

Vista, with many enzymes and combinations available to attend to different hemicellulose biomass types and compositions (36).

The lignocellulosic hydrolysate is a dark liquid containing sugars (pentoses and hexoses), organic acids, furaldehydes, and phenolic compounds from cellulose, hemicellulose, and lignin. Residual fibers, ashes, and some other solid particles can be found in small amounts (Fig. 3).

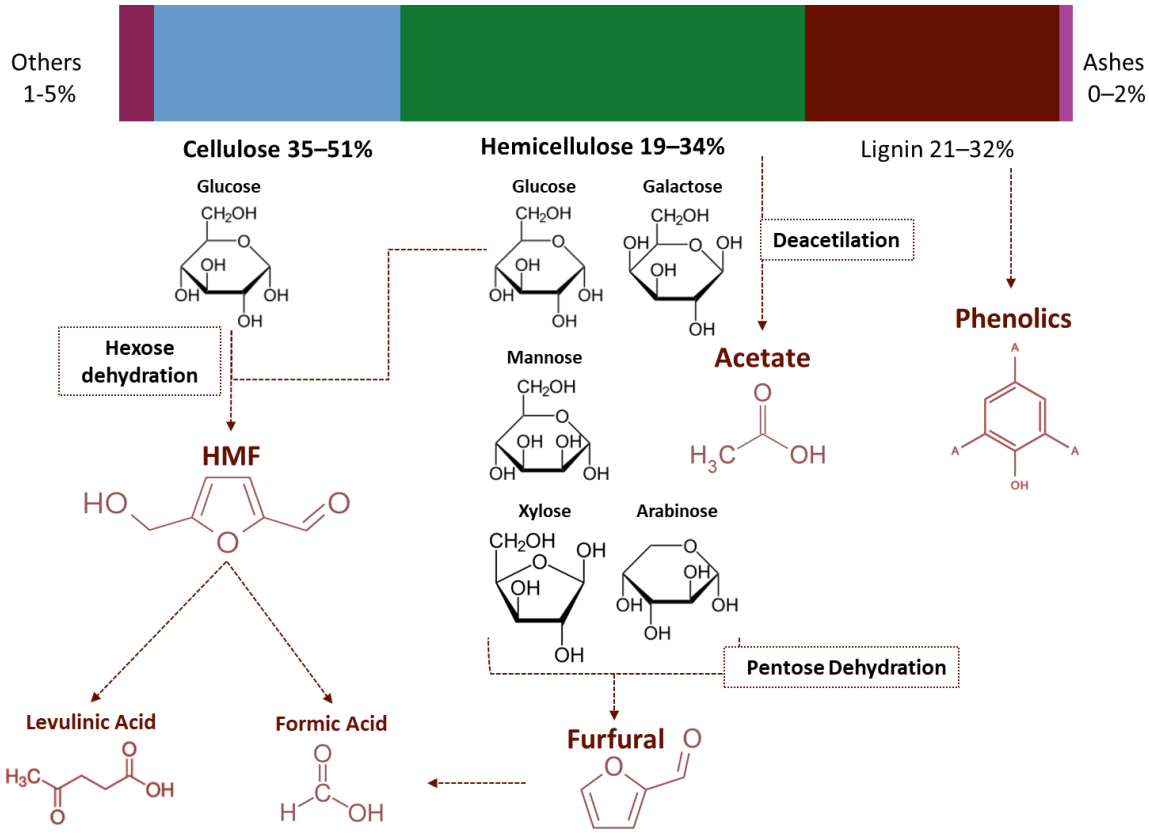

Figure 3 – Lignocellulosic biomass composition, its main constitutive monomers, and pre-treatment derived inhibitors. Source: Adapted from (39).

The proportions of each of the components in the lignocellulose hydrolysates will vary depending on the biomass, the pre-treatment, and the hydrolysis processes it underwent (Table 1).

Table 1 – Composition of different lignocellulosic feedstock after physicochemical pretreatment. Source: Adapted from (40).

	Composition (% dry basis)						Soluble Inhibitors in Pre-Hydrolysate (g/L)				
Biomass	Cellulose	Hemi-cellulose	Lignin	Solid concentration in the	Method	Condition	Phenols	Furans	Acetic Acid	Others	Reference
Corn stover	37	22.7	18.6	10-20%	Liquid hot water	Temp. 190 C, residence time 45 min	181–246 AU[2]	0.74–8.37	2.0–2.8	Xylo-oligomers 9.71–21.7	(41,42)
Wheat straw	30.2	21	17	30%	Steam explosion	Temp. 190–210 C, residence time 2–10 min, sulfuric acid 0.2%	nm	0.16–2.14	0.04–1.01	nm	(43,44)
Maple	41	15	29.1	23%	Liquid hot water	Temp. 180–200 C, residence time 24 min	1.3	4.1	13.1	Sugar oligomer 12.7, xylo-oligomers 11.2	(45,46)
Sugarcane bagasse	43.1	31.1	11.4	10%	Liquid hot water	Temp. 180–200 C, residence time 30 min	1.4–2.4	0.5–5.1	1.1–3.4	Gluco-oligomers 0.8, xylo-oligomers 6.5–12.5	(47,48)
Olive tree pruning	25	11.1	16.2	20%	Steam explosion	Temp. 190–240 C, residence time 5 min, sulfuric acid 0–2%	nm 1	0–3.2	0.4–4.2	Formic acid, 0.8–1.8	(49)

nm 1: not measured; AU 2: Absorbance Unit.

1.2.1 Xylose

Xylose is an aldopentose monomer and the main component of the hemicellulose xylan. After glucose, xylose is the second most abundant sugar present in nature, reaching up to 30% of lignocellulose hydrolysate sugars (50). Its immense availability makes it an important substrate for the production of value-added products. Consequently, the discovery and optimization of microorganisms able to use xylose as a carbon source are of special interest. Special attention has been given to microorganisms capable of efficiently fermenting them into products of interest, like bioethanol (51–53). Microorganisms naturally capable of fermenting xylose have three main pathways for xylose catabolism: isomerization, redox, and the α-ketoglutarate pathways (Fig. 4).

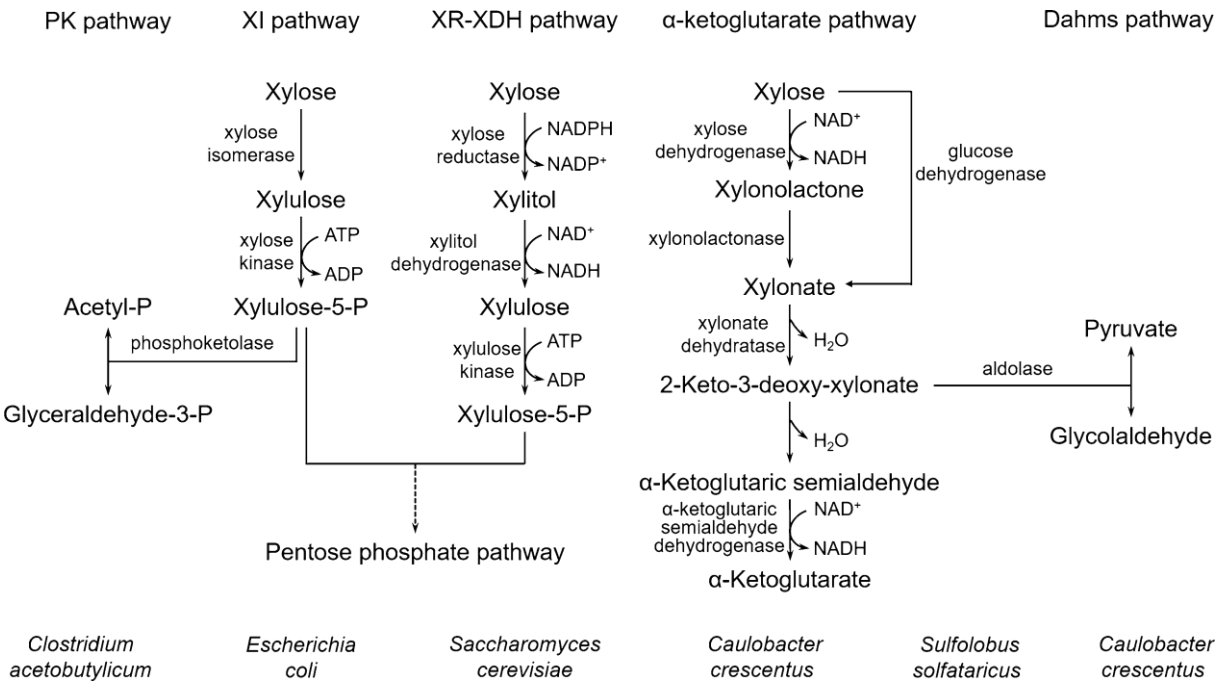

Figure 4 – Xylose metabolic pathways of xylose in various microorganisms. PK phosphoketolase, XI xylose isomerase, XR xylose reductase, XDH xylitol dehydrogenase. Source: (54).

In the isomerization pathway, xylose is isomerized to xylulose in one step by xylose isomerase (XI). It is found mainly in prokaryotes, but also present in filamentous fungi. The xylose oxidoreduction pathway is found mainly in yeasts and filamentous fungi, and consists of two stages: in the first, xylose is reduced to xylitol by xylose reductase (XR), and this one, oxidized to xylulose by xylitol dehydrogenase (XDH). In

both routes, xylulose is phosphorylated to xylulose-5-phosphate by xylulokinase (XK). Xylulose-5-phosphate is metabolized via the pentose phosphate pathway until it enters the glycolytic pathway. (54,55). Xylulose-5-phosphate can undergo the phosphoketolase (PK) pathway ins some bacteria like *Clostridia acetobutylicum* strains, being cleaved into acetylphosphate and glyceraldehyde-3-phosphate by phosphoketolase.

A more recent pathway was described for the bacteria *Caulobacter crescentus*, employing α-ketoglutarate as intermediate. Xylose is converted into xylonolactone by xylose dehydrogenase (XDH) and then to xylonate by xylonolactonase (XylC). Dehydration of xylonate into 2-keto-3-deoxy-xylonate and α-ketoglutaric semialdehyde, which is then oxidized into α-ketoglutarate by α-ketoglutaric semialdehyde dehydrogenase. The intermediate, 2-keto-3-deoxy-xylonate can undergo the Dahms pathway by an aldolase cleavage into glycolaldehyde and pyruvate, as reviewed by (54).

1.2.2 Lignocellulose derived inhibitors (LDI)

Pre-treatment processes facilitate the release of fermentable sugars from biomass; however, they end up also leading to the formation and release of compounds that inhibit microbial metabolism (Fig. 5). These inhibitors can be classified into three main groups: furaldehydes, such as 2-furaldeyde (furfural) and 5-hydroxymethyl-2-furaldehyde (HMF), weak acids (acetic acid, formic acid, and levulinic acid), and phenolic compounds (vanillin, syringaldehyde, coniferyl aldehyde, and others) (39,56). The pretreatment and hydrolysis processes, as well as biomass sources, influence the formation and concentrations of the aforementioned compounds in lignocellulosic hydrolysates (34,56,57).

Furaldehydes like furfural and 5-hydroxymethyl-2-furadehyde (HMF) are formed respectively from the dehydration of pentoses and hexoses in high temperatures and low pH. A series of phenolic compounds can be produced from lignin degradation. Organic acids can be released from hemicellulose deacetylation, like acetic acid, or the degradation of furaldehydes (Fig. 5), as reviewed by (39,58). Together, acetic acid, furfural, and HMF are frequently, in this order, the most abundant inhibitors found in some lignocellulosic hydrolysates, like sugarcane bagasse (59).

Physiological effects and resistance mechanisms to lignocellulose-derived inhibitors have been widely investigated on *Saccharomyces cerevisiae* (57,60–62) and to a minor extent in other yeasts, like *Zygosaccharomyces* (63), *Spathaspora passalidarum* (64), *Candida* spp (65,66), and others (62,67,68).

Furaldehydes have a negative effect on yeast growth, prolonging the lag phase, causing redox imbalance, damaging the cell wall, DNA and reducing cell viability. In microorganisms such as *S. cerevisiae, Pichia stitpitis, Escherichia coli,* and *Zymomona mobilis* furaldehydes have also been shown to reduce volumetric yield and ethanol productivity. HMF and furfural are highly toxic to cells, especially since they have a synergistic effect and are usually converted into less toxic compounds by the cell during the extended lag phase (39,70). Acetic acid is generally the most abundant inhibitor in most lignocellulose hydrolysates (59). It can inhibit the formation of biomass, affecting cell metabolism, since it can cross the plasma membrane, decreasing cytosolic pH. To regulate the intracellular pH, cells induce the activity of ATP-dependent proton pumps, leading to ATP depletion (39,56,58,71).

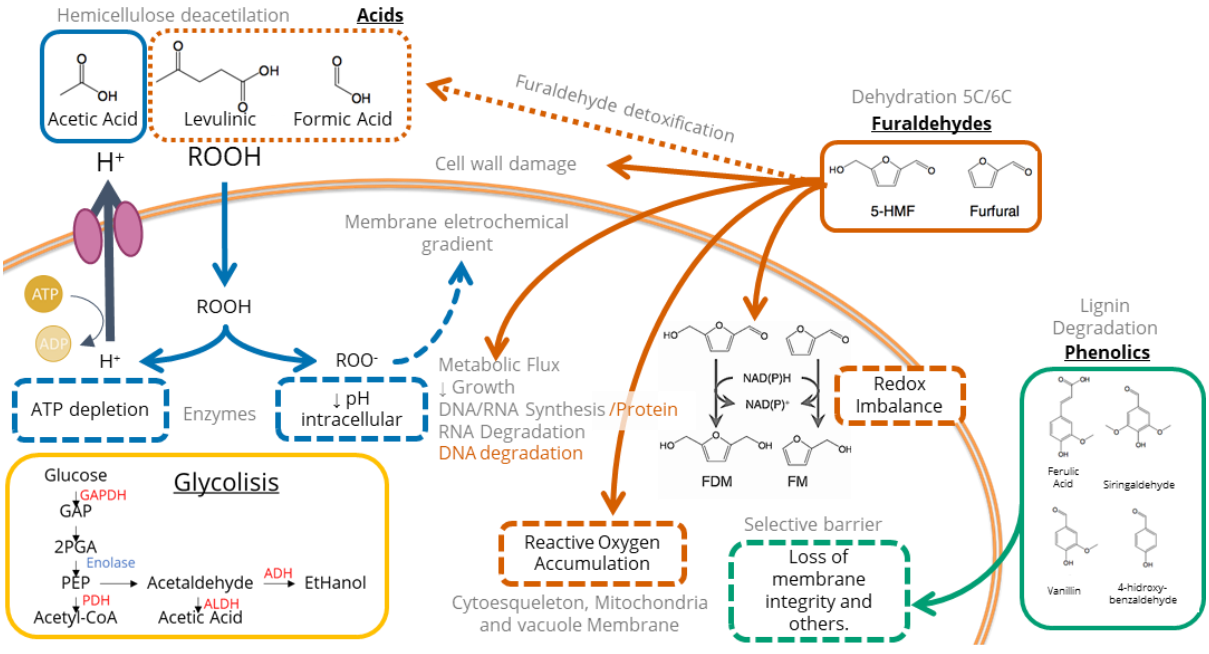

Figure 5 – Schematic representation of furaldehydes, phenolics, and organic acid effects on yeast cells. Source: Adapted from (56).

Overall, the global transcriptional and translational responses to lignocellulose-derived inhibitors are complex and involve the polygenetic modulation of various metabolic pathways, such as carbon, lipid, amino acid metabolism, and regulatory

pathways, among others. The differential gene expression redirects the yeast's metabolism to allow repair of damages caused by the inhibitors and increase the innate detoxification activities (66,70,72–75). These responses have been defined for *S. cerevisiae* (57,60–62) and other few organisms like *Zygosaccharomyces* (63), and *Spathaspora passalidarum* (64). There is a relatively small body of literature on this topic for methylotrophic yeasts, such as *Ogatae polymorpha* and *Pichia pastoris* (76). On this matter, recent observations pointed to *O. polymorpha* tolerance to wheat straw hydrolysate containing different concentrations of acetic acid, formic acid, furaldehydes, and phenolic compounds. The results demonstrated that the sugar uptake by the yeast was reduced in the presence of inhibitors. The yeast was still able to consume some xylose and produce xylitol in presence of 12.24 $g.L^{-1}$ of acetic acid and 4.17 $g.L^{-1}$ of total phenolics, (68). Although *P. pastoris* growth in lignocellulose-derived media has been reported, its global response mechanisms to the inhibitors were not previously described (77).

1.3 INCREASED TOLERANCE OF YEAST TOWARD LIGNOCELLULOSE DERIVED INHIBITORS

A fundamental trait for lignocellulose hydrolysate utilization by microorganisms is their robustness against the inhibitors formed during pretreatment of biomass. When naturally tolerant strains are not available for the desired process, adaptative, genetic, and metabolic engineering strategies may be used to improve strain response for inhibitory compounds.

Adaptative laboratory evolution (ALE) takes advantage of natural evolutionary mechanisms for strain optimization. Through an unbiased selection of naturally occurring mutations, a desirable trait becomes more prevalent in the population (78,79). This is typically done through a series of culture transfers, or continuous cultivation with a selective pressure agent (80). Mutagenic agents like UV light and EMS (Ethyl methanesulfonate) can be used to increase the genetic variation in the population (20). In yeasts, this strategy has been most commonly used in strains that have a naturally higher tolerance to inhibitors or are of great industrial interest like *S. cerevisiae* (81), *P. stipitis* (82), *S. passalidarum* (64).

The combination of ALE methods with direct engineering is also a common approach for yeast optimization. For example, the ALE of xylose utilizing *S. cerevisiae* strain in a cocktail of different lignocellulose inhibitors resulted in the increase of growth rate from 0.18 h^{-1} to 0.33^{-1} h and a 24 h reduction from the lag phase (83).

Understanding cellular response to lignocellulose-derived inhibitors became more convenient and accessible with the advancement and consequently, cheapening of sequencing technologies and multi-omics analysis (genomics, transcriptomics, proteomics, and metabolomics). Consequently, many studies have been dedicated to better understand the genes and mechanisms involved in stress tolerance in different organisms and the development of more robust strains (Table 2).

Table 2 – Strategies for enhancing microbial tolerance to inhibitors. Source: Adapted from (84).

Strategies	Approach	Inhibitors	Microorganism	Effect	References
In situ detoxification	Expression of dehydrogenases	Furfural, HMF, others	*S. cerevisiae*	Ethanol production improvement +20–30%	(70)
Efflux pumps	Expression of *eilAR* module	Ionic liquids	*E. coli*	Engineered strains could grow and produce biofuel in the presence of inhibitors.	(85)
	Expression of *rcdA*	Ionic liquids	*E. coli*	Limonene production improvement	(86)
Stress responses	Expression of *groESL*	n-Butanol	*Clostridium acetobutylicum*	n-Butanol production improvement 40%	(87)
	Expression of *TPS1*	Ethanol	*S. cerevisiae*	Ethanol production improved 8.7%	(88)
	Overexpression of transcriptional factor *HAA1*	Acetic acid	*S. cerevisiae*	Better growth in engineered strains and less intracellular acid content.	(89)
Membrane engineering	Expression of *fabD*	n-Butanol	*E. coli*	n-Butanol production improved 200%	(90)
	Expression of *mgtA*	Succinic acid	*E. coli*	succinic acid production improved 50%	(91)
Adaptive laboratory evolution	Adapted laboratory	Hydrolysate	*S. cerevisiae*	Improvement of cell viability of 86%and 95% ethanol	(92)
	Adapted laboratory	Ionic liquids	*E. coli*	Engineered strains could grow in the presence of inhibitors.	(93)
Random mutagenesis	Plasma mutagenesis	Hydrolysate	*Rhodosporidium toruloides*	Improvement of 18% in growth and 13% in lipid production.	(94)
	UV radiation	Ethanol	*S. cerevisiae*	Ethanol 40%	(95)

1.4 JUSTIFICATION

The yeast *P. pastoris* has established value as a biotechnological chassis organism for the production of proteins and chemicals (96). Another element of considerable biotechnological potential, the lignocellulosic hydrolysate is an abundant, cheap, and sustainable resource. The great promise of lignocellulosic residues utilization as an environmentally sustainable option still finds two major bottlenecks: xylose utilization and tolerance to lignocellulosic-derived inhibitors. Xylose is the second most abundant sugar present in nature (50), and currently underutilized in biorefineries because many industrial microorganisms are unable to efficiently use it (29). At the same time, the use of any sugar present in the hydrolysate is conditioned to the robustness and capacity of the microorganism in question to tolerate the toxic effects of the inhibitors formed during the pre-treatment process of the biomass (59). Few previous studies exploring lignocellulosic hydrolysate as a carbon source for *P. pastoris* have been found.

Physiological and transcriptional characterization in lignocellulose hydrolysate and its derived inhibitors were globally investigated here (Chapter 1). By looking into different *Komagataella* strains, evidence that *P. pastoris* can consume xylose as only carbon source was found. This was further validated and characterized, and selected strains underwent an adaptative laboratory evolution process towards faster growth and xylose consumption (Chapter 2). Lastly, the accumulated knowledge for this yeast and cutting-edge technology allowed easy genetic manipulation and the construction of strains more robust against acetic acid effects through overexpression of the *HAA1* gene (Chapter 3). By combining different strategies, we hereby present a holistic characterization of *P. pastoris* responses to lignocellulosic hydrolysate components.

2 OBJECTIVES OF THE WORK

2.1 GENERAL OBJECTIVE

Describe the molecular and physiological responses of *Pichia pastoris* to xylose and lignocellulose-derived inhibitors.

2.2 SPECIFIC OBJECTIVES

- Evaluate the physiological and transcriptional response of *P. pastoris* in the presence of lignocellulose-derived inhibitors (chapter 1).
- Characterize and compare the performance of 26 different *Komagataella* strains regarding xylose utilization and tolerance to lignocellulose-derived inhibitors (chapter 2).
- Improve *P. pastoris* tolerance to acetic acid through *HAA1* overexpression (chapter 3).

CHAPTER 1 – *Pichia pastoris* RESPONSE TO LIGNOCELLULOSE-DERIVED INHIBITORS

This chapter is based on the article 'Physiological characterization and transcriptome analysis of *Pichia pastoris* reveals its response to lignocellulose-derived inhibitors.'

Lignocellulosic biomass is an abundant raw material that can be converted by physicochemical and microbial processes into different products, such as biofuels, building-block chemicals, and high added-value chemicals (20,29). Before microbial fermentation, the biomass needs to undergo pretreatment and hydrolysis to release the monosaccharides present in the biomass. During pretreatment, compounds that inhibit microbial metabolism are also released or formed during dehydration of pentoses and hexoses, hemicellulose deacetylation, or lignin breakdown (34,39). These inhibitors can be classified into three main groups: furaldehydes, such as 2-furaldeyde (furfural) and 5-hydroxymethyl-2-furaldehyde (HMF), weak acids (acetic acid, formic acid, and levulinic acid), and phenolic compounds (vanillin, syringaldehyde, coniferyl aldehyde, and others) (39,56). The pretreatment and hydrolysis processes, as well as biomass source influence the formation and concentrations of the aforementioned compounds in lignocellulosic hydrolysates" (34,56,57).

Effects of lignocellulose-derived inhibitors on yeast physiology and resistance mechanisms have been extensively investigated for *Saccharomyces cerevisiae* (57,60–62) and to a minor extent for other yeasts, like *Zygosaccharomyces* (63), *Spathaspora passalidarum* (64), *Candida* spp (65,66), and others (62,67,68). Inhibitory effects and mechanisms vary depending on the chemical structure of the specific inhibitor and its concentration. Generally, they are cytotoxic and hinder microbial growth, reduce cell vitality, and fermentation efficiency. Main mechanisms involve inhibition of essential enzymes related to cell metabolism, DNA replication, RNA, and protein synthesis and redox imbalance, and damaging cellular membranes (39,97–101).

The *S. cerevisiae* response to inhibitors is complex and involves a polygenetic modulation of various metabolic pathways, such as carbon, lipid, amino acid metabolism, and regulatory pathways, among others. The differential gene expression redirects the yeast's metabolism to allow repair of damages caused by the inhibitors and increase the innate detoxification activities (66,70,72–75). The translation of such

complex mechanisms to yeasts of industrial interest is, therefore, crucial. Among those, methylotrophic yeast, such as *Ogatae polymorpha* and *Pichia pastoris*, can be highlighted given their role in the production of fuels and chemicals (76). On this matter, recent observations pointed to *O. polymorpha* tolerance to wheat straw hydrolysate when containing different concentrations of acetic acid, formic acid, furaldehydes, and phenolic compounds. The results demonstrated that the sugar uptake by the yeast was reduced in the presence of inhibitors. The yeast was still able to consume some xylose and produce xylitol in presence of 12.24 g.L-1 of acetic acid and 4.17 g.L-1 of total phenolics, (68). However, the methylotrophic yeast response mechanisms to the inhibitors were not previously reported.

As previously mentioned, the yeast *Komagataella phaffii*, previously known and here referred to as *Pichia pastoris* (96) is a methylotrophic yeast extensively used in the production of heterologous proteins and metabolites both in industry and academia (10). To this date, more than five thousand different proteins have been heterologously expressed in this yeast (2). The biotechnological potential of *P. pastoris* has been amplified by its use in metabolic engineering programs and the production of many other compounds besides proteins have been considered (3,102), including alcohols, acids, vitamins, and others (96,103–106). The increasing interest in *P. pastoris* has led to the construction of recombinant strains capable of metabolizing cellulose (27), glucose, and glycerol (103). Both xylose, (26,106) and acetic acid have also been studied as an alternative carbon source for this yeast (22,107).

Here we unveil the potential of *P. pastoris* for the conversion of sugars present in lignocellulosic hydrolysates. More specifically, we evaluate the yeast's physiological response to acetic acid, furaldehydes, and sugarcane biomass hydrolysate. Using RNA-seq based transcriptome analysis, we investigated the global response of *P. pastoris* in the presence of different concentrations of those compounds. Lastly, the physiological and transcriptional dose-dependent response of *P. pastoris* to the inhibitors are presented and discussed.

3 MATERIALS AND METHODS

3.1 STRAIN AND MEDIA

The yeast *P. pastoris X33* was used in this work (Invitrogen, USA). Stock cultures of yeast grown in YPD medium (1% w/v yeast extract, 2% w/v peptone, 2% w/v glucose) were preserved in 30% glycerol and maintained at -80°C.

To evaluate the effect of inhibitors on yeast metabolism, the medium employed was composed of (w/v): YNB (yeast nitrogen base, Sigma Aldrich Y0626) without amino acids (0.68% YNB, 2% ammonium sulfate), glucose 2%, 4% xylose, 4.10^{-5}% biotin, buffered to pH 5.5 with phthalate buffer (1.02% potassium hydrogen phthalate with 0.22% potassium hydroxide w/v). For each culture condition, inhibitory compounds were added to the media in the following concentrations 2 $g.L^{-1}$ and 6 $g.L^{-1}$ of acetic acid; a mixture of 0.9 $g.L^{-1}$ furfural and 0.15 $g.L^{-1}$ HMF (FH 0.9/0.15 $g.L^{-1}$) and 1.5 $g.L^{-1}$ furfural and 0.25 $g.L^{-1}$ HMF (FH 1.5/0.25 $g.L^{-1}$) for furaldehydes; and sugarcane bagasse hydrolysate diluted to 10% and 30% of the initial concentration. The sugarcane bagasse hydrolysate was obtained by steam explosion of sugarcane bagasse, then for the breakdown of the oligomers in the hemicellulose-rich fraction, the liquid fraction of the steam explosion was subjected to hydrolysis with 0.5% H_2SO_4 (w/w) at 130 °C for 100 min (108). The final composition of the sugarcane bagasse hydrolysate was: 5.4 $g.L^{-1}$ glucose, 90.3 $g.L^{-1}$ xylose, 19.4 $g.L^{-1}$ acetic acid, 2.9 $g.L^{-1}$ furfural, 0.55 $g.L^{-1}$ HMF. In the media containing diluted hydrolysate, the amount of glucose and xylose present in the hydrolysate was taken into account to keep the final glucose and xylose concentration at 2% and 4%, respectively.

3.2 *P. pastoris* GROWTH ASSAY IN LIGNOCELLULOSIC HYDROLYSATE

Colonies grown in YPD medium (yeast extract 1%, peptone 2%, glucose 2%) from glycerol stocks were inoculated in 5 mL YPD and incubated overnight at 28ºC, 200 rpm in a rotary shaker. Cells were harvested, washed twice with sterile water, and inoculated to an initial OD_{600} of 0.1 in 200 µl of media, in a 96-well plate. As media, YNB was supplemented with dilutions of lignocellulosic hydrolysate (from zero to 90%) and glucose concentration was adjusted to 5.4 $g.L^{-1}$ in all dilutions. Each dilution was

repeated in technical triplicate. Plates were incubated for 26h in the Epoch2 microplate reader (Biotek), at 28ºC and 210 rpm double orbital, and D_{600} was measured every two hours.

3.3 PHYSIOLOGICAL CHARACTERIZATION OF *P pastoris* IN PRESENCE OF LIGNOCELLULOSE DERIVED INHIBITORS (LDI)

Cells plated in YPD medium were initially inoculated in 5 mL YPD and grown overnight (28 °C, 200 rpm on a rotary shaker). The next day cells were transferred to 200 mL YPD in 1 L shake flasks and grown overnight at the same conditions. The culture was washed twice with distilled water and diluted down to an initial optical density (OD) at 600 nm of 5 in 50 mL of medium in 250 mL shake flasks. The culture was incubated for 30 h at 28 °C and 200 rpm. Samples for transcriptome were withdrawn after 4h of incubation and samples for metabolite analysis were withdrawn regularly. All experiments were carried out in biological triplicate.

3.4 RNA EXTRACTION AND QUALITY ANALYSIS

RNA was extracted using TRIZOL (Thermo Fisher Scientific, USA) reagent following the manufacturer´s protocol with few modifications. A culture of 5 mL was harvested for 1 min at 14.000 x *g* at 4 °C. The supernatant was discarded and 1 mL of TRIZOL was added to the pellet. Cells were transferred to a 2 mL microtube containing approximately 200 µL of sterile 0.02 mm glass beads and then disrupted by four cycles of 1 min at Mini-Beadbeater-96 (Biospec Products, USA), resting the tube on ice between cycles. Finally, RNA extraction followed the manufacturer's instructions by performing chloroform and ethanol washings. RNA integrity was evaluated via Agilent Bioanalyzer 2100 system (Agilent Technologies, USA), Nanodrop 1000 Spectrophotometer (Thermo Fisher Scientific, USA) and in 1% agarose gels.

3.5 RNA SEQUENCING AND DATA ANALYSIS

RNA-seq was performed by Centro de Genômica, the University of São Paulo on Illumina HiSeq 2500 system v4 using HiSeq SBS Kit v4, and 100bp (2x) paired-end reads. Libraries for RNA-Seq were prepared with TruSeq Stranded mRNA Sample

Prep LT Protocol (Illumina, USA) from RNA extractions of 21 independent samples. FastQC software was used to evaluate base quality distributions based on phred value (109). Raw reads were processed with Trimmomatic software (110), and once again analyzed for the quality of clean sequences on FastQC. Sequences were aligned using STAR (111). HTSeq-count version 0.9.1 tool (112) was used for counting the number of aligned sequences for each sample in each gene and estimate gene expression. Differentially expressed genes were detected by entering the count data into the R program (113) and using the DESeq2 package (114).

Genes were considered significantly differentially expressed with an adjusted *P*-values limit < 0.05 both for increasing and decreasing expression. The differentially expressed genes (DEG) overlap between conditions was assessed using Venn diagrams obtained with the Venn online platform[1]. The list of DEGs in the DESeq2 package was used for functional analysis to identify which genes and metabolic pathways are being activated or repressed in response to acetic acid, furaldehydes, and hydrolyzed. For this, the induced and the repressed genes were separated into different files. The individual lists were subjected to functional enrichment analysis using Fisher's Exact Test with a false discovery rate (FDR) < 0.05 in the GO_MWU tool[2]. This analysis assesses the significance of the representativeness of the GO (Gene Ontology) categories among DEG.

For the heatmap, the expression values of 630 genes differentially expressed in all conditions (adjusted *P*-value ≤ 0.05) were hierarchically clustered using MeV 4.9.0 program[3] with Pearson correlation metric and average linkage clustering. A distance threshold of 0.75 was used to split the gene tree into 7 clusters. Gene ontology annotation from each cluster was used as input to REVIGO (115) analysis to reduce redundancy and build the network. We used the GO terms database from *S. cerevisiae* and SimRel as the semantic similarity measure.

The transcriptome datasets generated during the current study are available in the NCBI with the accession number PRJNA666642.

[1] http://bioinformatics.psb.ugent.be/webtools/Venn/
[2] https://github.com/z0on/GO_MWU
[3] http://mev.tm4.org

3.6 QUANTIFICATION OF METABOLITES

Carbon sources (xylose and glucose) and extracellular metabolites xylitol, glycerol, acetate, HMF, and furfural concentrations were determined by High-Performance Liquid Chromatography (HPLC) (116) in samples withdrawn on different time points. Samples were centrifuged, and the supernatant was analyzed by HPLC (Acquity UPLC H Class, Waters, USA) equipped with a refractive index and a PDA detector. Metabolites were separated on an HPX-87 H column (Bio-Rad Laboratories, USA), using a 5 mM sulfuric acid mobile phase at a flow rate of 0.6 mL/min and temperature of 45 °C. Biomass was measured through OD600 using a spectrophotometer (SpectraMax M3, Molecular Devices, USA).

4 RESULTS AND DISCUSSION

4.1 EFFECT OF SUGARCANE BAGASSE HYDROLYSATE ON *P. pastoris* GROWTH

To evaluate the threshold tolerance of *P. pastoris* towards lignocellulosic hydrolysate, it was grown in microtiter plates containing YNB media with a final glucose concentration of 5.4 $g.L^{-1}$, and sugarcane hydrolysate addition from zero to 90% in 10% intervals (Fig. 6). The control (no lignocellulosic hydrolysate) started to grow after 2h of incubation and reached OD_{600} 0.43 after 26h of incubation. With the addition of hydrolysate, inhibition positively correlated with the increase in concentration. At 10% dilutions, the yeast grew slower than the control in the first 14h, when the profile became identical to the control, until a final OD_{600} 0.42. At 20% dilution, the lag phase was extended until 14h, when the yeast started to grow more rapidly. An apparent stationary phase was reached after 22h, with OD_{600} 0.27 until the end. From 30% hydrolysate and above no growth was observed (Fig. 6). However small absorbance variations in presence of hydrolysate from 50% to 90% were observed, which might be explained by the medium color being too dark and precipitation of hydrolysate particles.

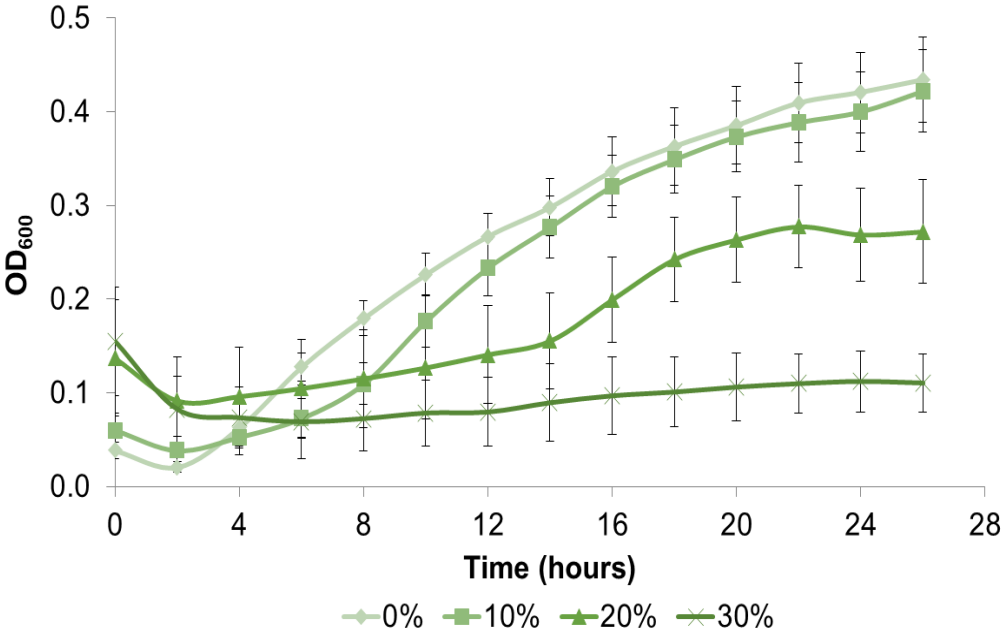

Figure 6 – *P. pastoris* growth in microtiter plates containing YNB supplemented with different dilutions of lignocellulosic hydrolysate and 5,4 g.L^{-1} glucose. Dilutions: zero to 30% lignocellulosic hydrolysate is shown. No growth was detected above 30%. The experiment was performed in technical triplicates. Source: Author's collection.

The concentrations chosen for further experiments were: 2 g.L^{-1} and 6 g.L^{-1} for acetic acid and 0.15 g.L^{-1} HMF and 0.9 g.L^{-1} of furfural (referred here as FH 1.5/0.25 g.L^{-1}) and 0.25 g.L^{-1} of HMF and 1.5 g.L^{-1} of furfural, (referred here as FH 0.9/0.15 g.L^{-1}). Lignocellulosic hydrolysate concentrations chosen were 10% and 30%. The concentrations of acetic acid chosen are similar to those found in 10% and 30% of the sugarcane bagasse hydrolysate used in this study. However, sugarcane bagasse hydrolysates have low amounts of furaldehydes in comparison to hydrolysates from other sources (70). To be able to observe relevant metabolic changes in presence of furaldehydes, the concentrations chosen for this inhibitor were similar to those found in 30% and 50% of the sugarcane bagasse hydrolysate used in this study. Similar inhibitor concentrations were observed in the range of inhibitors found in lignocellulosic hydrolysates from different sources after physicochemical pretreatment (40).

4.2 *P. pastoris* FERMENTATIVE PROFILE IN PRESENCE OF LIGNOCELLULOSE-DERIVED INHIBITORS

Based on the previous results (Fig. 6), the fermentative performance of *P. pastoris* was cultivated in YNB medium with the addition of acetic acid, furaldehydes, or lignocellulose hydrolysate, besides the control group with no inhibitors. Glucose and xylose concentrations were supplemented to 2% and 4%, respectively, in all cultivation conditions (Fig. 7).

The physiological response of the yeast varied drastically according to the evaluated condition and is summarized in figure 7a. In the control (media without inhibitors), 97% of the available glucose had been consumed after nine hours of incubation, and the yeast reached OD_{600} 18. Acetic acid did not extend the lag phase but reduced the yeast growth and sugar consumption rate in the first hours of cultivation. Indeed, the yeast consumed 81% and 60.6% of available glucose in the presence of 2 $g.L^{-1}$ and 6 $g.^{L}$-1 of acetic acid, respectively, compared to the control after 9h (Fig. 7b); however, the final yeast growth in 2 $g.L^{-1}$ of acetic acid was slightly higher than in control (OD_{600} = 22.8 ± 1 compared to 25.3 ± 2). In that case, the yeast was not only able to fully consume all the acetic acid present in the medium but also co-consume it along with glucose (Fig. 7b). This will be discussed in a further section.

An extended lag phase was observed in the presence of furaldehydes, with a reduction of 14% in yeast growth when compared to control (Fig. 7a). The prolonged lag phase correlates with reduced sugar consumption in both furaldehyde concentrations. In the first nine hours of cultivation, 61% and 47% of the available glucose was consumed by the yeast in presence of FH 0.9/0.15 and FH 1.5/0.25 respectively (Fig 7b). Furaldehydes are known to induce lag phase extension on other yeast, while they are converting it into less toxic compounds to then resume growth. Furfural, specifically, is more toxic than HMF when added in equimolar amounts in media, although it was more rapidly metabolized (39).

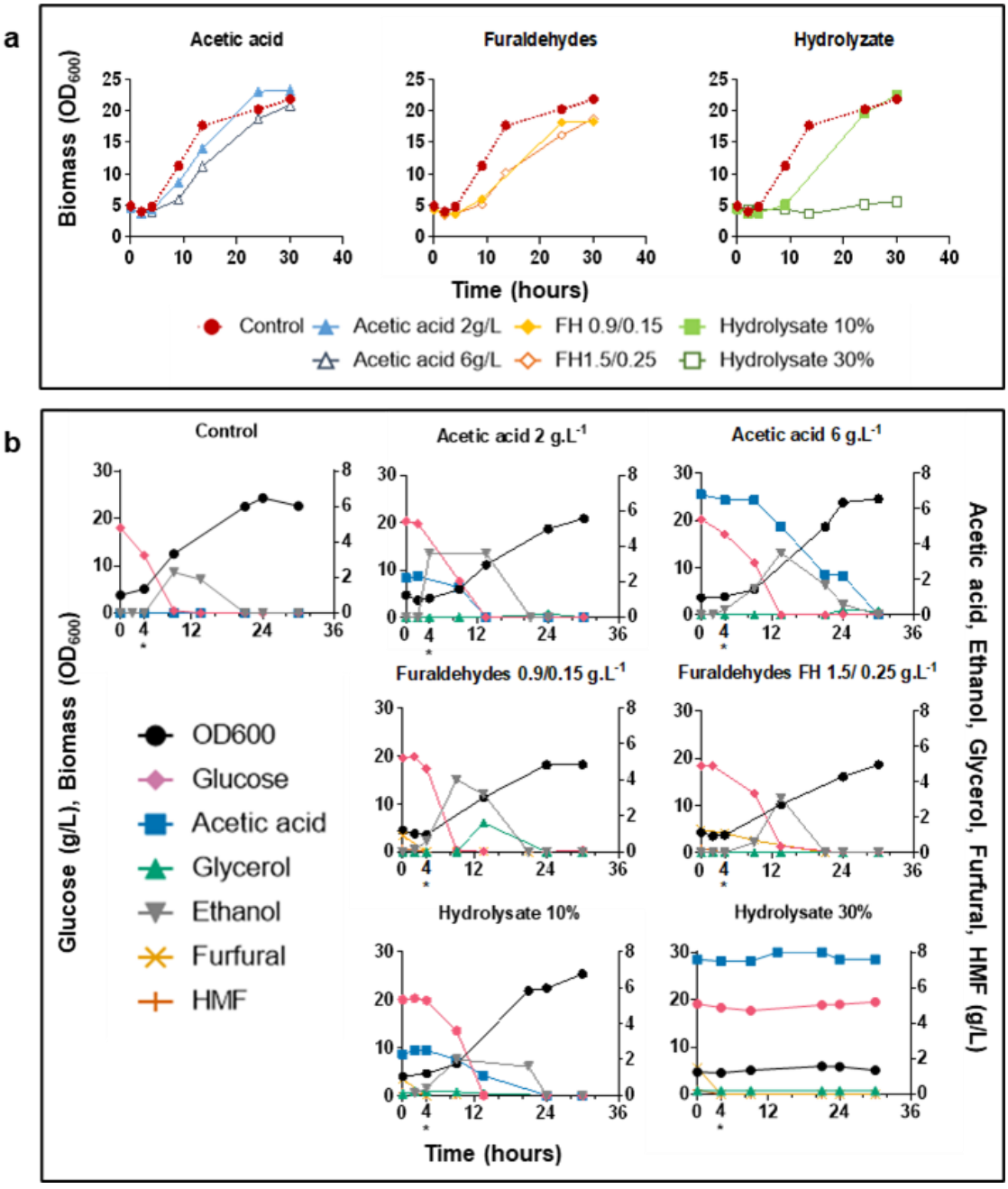

Figure 7 – *P. pastoris* growth profile in the presence of different inhibitors. a) growth curves of acetic acid (2 g.L-1 - closed triangle, and 6 g.L-1 open triangle), furaldehydes (0.15 g.L-1 HMF, 0.9 g.L-1 furfural - closed diamond and 0.25 g.L-1 HMF, 1.5 g.L-1 furfural - open diamond), and hydrolysate (10% - closed square, and hydrolysate 30% open square) against control (closed circle with dotted lines). b) substrate consumption and product formation in different conditions: control, acetic acid 2 g.L-1 , acetic acid 6 g.L-1 , FH 0.9/ 0.15 g.L-1 , FH 1.5/ 0.25 g.L-1 , hydrolysate 10%, hydrolysate 30%. Biomass (OD600, black circle), glucose (pink diamond), acetic acid (blue square), glycerol (green triangle), ethanol (gray upside-down triangle), furfural (yellow star), HMF (orange cross). Timepoint 4* highlights the timepoint where transcriptome samples were taken. Xylose concentration was constant through the cultivation. The experiments were performed in triplicate and the figure represents the profile of one biological replicate. Source: Author's collection.

The *P. pastoris* performance in the presence of acetic acid, furaldehydes, and sugarcane hydrolysate shown here demonstrates its relatively high tolerance to

lignocellulose-derived inhibitors, especially to acetic acid. Higher concentrations of furaldehydes (0.25 $g.L^{-1}$ HMF, 1.5 $g.L^{-1}$ furfural) or acetic acid 6 $g.L^{-1}$ delayed but did not impair *P. pastoris* growth. The total inhibition of yeast metabolism was only observed in the presence of hydrolysate 30%, which contains 6 $g.L^{-1}$ of acetic acid and FH 1.5/ 0.25 $g.L^{-1}$, besides other compounds. Yeast tolerance to the inhibitors is species and strain-specific (69,100,101). While some *S. cerevisiae* strains have shown sensibility to as few as 1 $g.L^{-1}$ of furaldehydes (39) and 4,8 $g.L^{-1}$ acetic acid (117), others have shown to be tolerant to concentrations as high as 10 $g.L^{-1}$ (118). As a reference, *O. polymorpha,* another methylotrophic yeast, could grow and produce xylitol in wheat straw hydrolysate containing up to 12.24 $g.L^{-1}$ of acetic acid and 4.17 $g.L^{-1}$ of total phenolics (68). Although a direct comparison between the yeast species is not possible due to the differences in experimental conditions, the results reported here demonstrated the *P. pastoris* can withstand lignocellulose-derived inhibitors even if inoculated at low cell density (DO_{600} 5).

4.3 TRANSCRIPTIONAL RESPONSE OF *P. pastoris* TOWARDS LIGNOCELLULOSE-DERIVED INHIBITORS

A genome-wide RNA-seq transcriptional profiling was used to understand the global cellular response of *P. pastoris* toward lignocellulose-derived inhibitors. For this, samples from the previous experiments (Fig. 7) were withdrawn after 4h of cultivation, RNA was extracted and sequenced. To identify differentially expressed genes (DEGs), the experimental data from cultivations in presence of inhibitors were normalized to the control condition. A total of 429,738 sequence reads were obtained after quality trimming. Samples were aligned to the *P. pastoris* (*K. phaffii)* str. WT (GenBank accession no. GCA_001708085) reference genome. Principal component analysis (PCA) based on expression patterns showed good reproducibility of the biological replicates and distinct isolation of hydrolysate 30% replicates from the other conditions (Fig. 8).

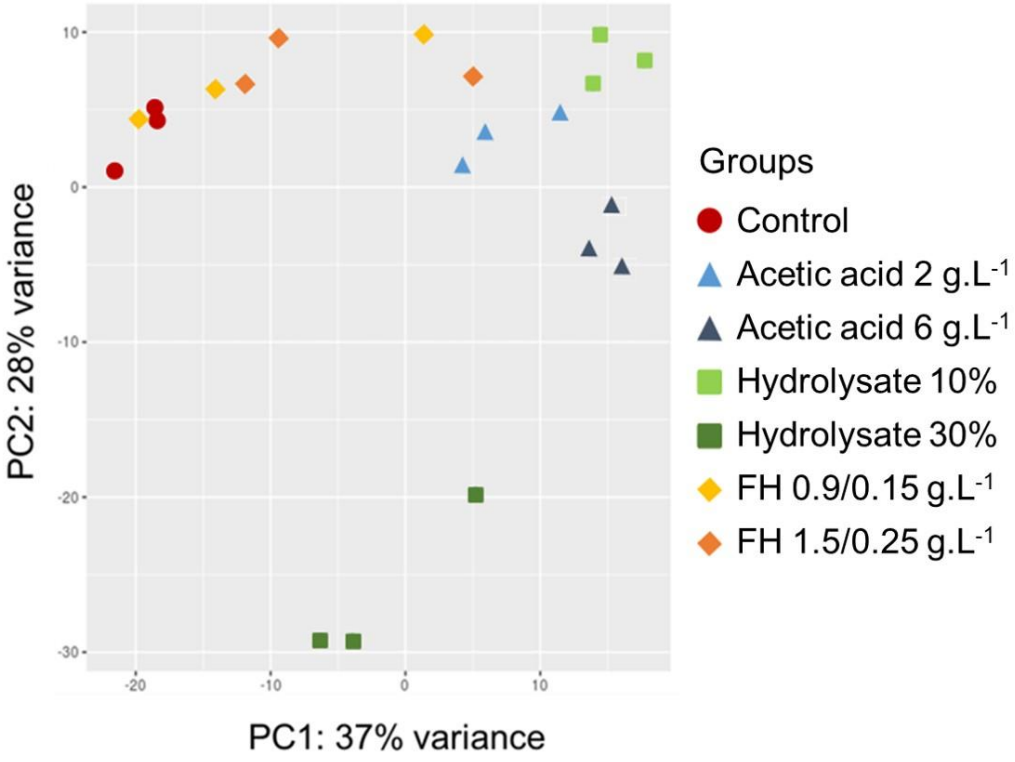

Figure 8 – Principal component analysis (PCA) of RNA-seq data of *P. pastoris* cultivated in different inhibitors. Dots represent samples and are colored according to the different conditions investigated: red circle: Control - minimal medium without inhibitor, light blue triangle: acetic acid 2 g.L^{-1}, dark blue triangle: acetic acid 6 g.L^{-1}, light green square: hydrolysate 10%, dark green square: hydrolysate 30%, yellow diamond: FH 0.9/ 0.15 g.L^{-1}, and orange diamond: FH 1.5/ 0.25 g.L^{-1}. Source: Author's collection.

The inhibitors incited a significant transcriptional response of *P. pastoris*. Out of 5040 genes found, a total of 3315 were differentially expressed (the adjusted *P*-value of <0.05 both for increasing and decreasing expression) (Fig. 9a). Most genes were differentially expressed in the presence of more than one inhibitor evaluated, with the bigger differences found between furaldehydes and acetic acid than hydrolysate to the two other conditions. From the total of DEGs, 234, 66, and 959 genes were exclusively differentially expressed in the presence of acetic acid, furaldehydes, or hydrolysate, respectively. Moreover, 630 genes are common to all three inhibitors (Fig. 9a). Acetic acid induced the differential expression of 2108 (64%) genes, sharing 1228 and 16 of them exclusively with hydrolysate and furaldehydes, respectively. In the presence of furaldehydes, the yeast showed the smallest amount of DEGs, summing up to 894 (27%) genes, whereas most DEGs were found in hydrolysate conditions with 2999 (90%). Out of all DEGs, 1194 had no annotation in the reference genome.

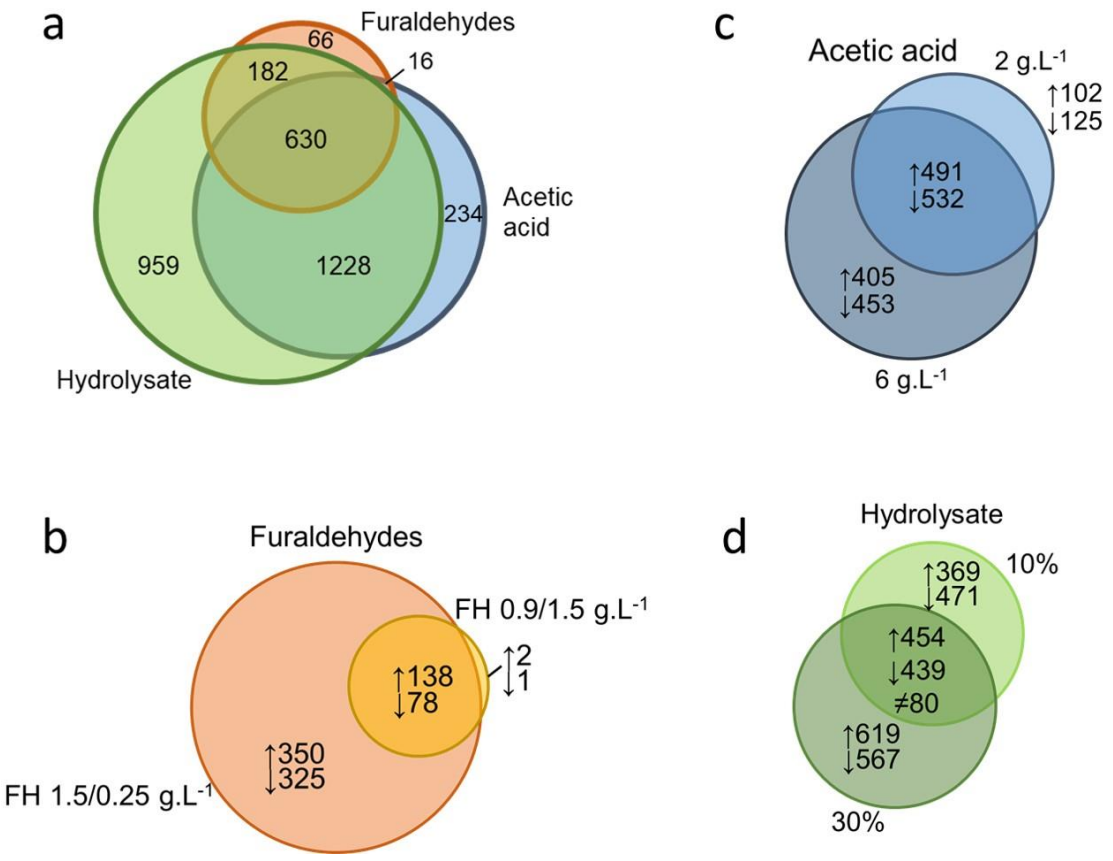

Figure 9 – Venn diagrams representing *P. pastoris* differentially expressed genes (DEGs) in response to lignocellulose-derived inhibitors. (a) three inhibitors, no concentration differentiation; (b) acetic acid 2 g.L-1 and 6 g.L-1 ; (c) furaldehydes FH 0.9/ 0.15 g.L-1 , and FH 1.5/ 0.25 g.L-1 ; (d) lignocellulosic hydrolysate 10% and 30%. Numbers account for DEGs that were differentially expressed in at least one of the two concentrations. The symbol ≠ stands for the 80 genes in which behavior changes depending on the hydrolysate concentration. Transcriptome samples were taken after 4 h of cultivation. Source: Author's collection.

The yeast transcriptional response to the inhibitors was observed to be dose-dependent, i.e., increased concentrations of acetic acid, furaldehydes, and hydrolysate lead to stronger inhibitory effects on yeast metabolism, increasing the time for the yeast to complete sugar consumption and grow. (Fig. 9b; c; d). In the presence of acetic acid, the yeast had 2108 DEGs, with 11% found exclusively at 2g. L^{-1} and 41% at 6g. L^{-1} of the acid, respectively (Fig. 9c). Similar responses were seen for furaldehydes and hydrolysate, where 0.3% (3 genes) and 28% of DEG were exclusively for the lower concentration of inhibitor and 75% and 40% for the highest concentrations, respectively (Fig. 9b and d). The conversion of furaldehydes happened within few hours of cultivation, and when the inhibitory effects were absent, the yeast exited the lag phase, started to consume sugars, and grow (Fig. 7A). The number of DEGs indicates that hydrolysate represented the biggest challenge to the yeast. (Fig. 9a). In

the total 2999 different genes were either up (1522) or downregulated (1557) in the presence of hydrolysate in at least one of the two concentrations. From 10% to 30%, a total of 80 genes changed their pattern of expression: 30 were from up to down-regulated and 50 from down to up-regulated (Fig. 9d). Furaldehydes induced the smallest response in terms of the number of DEGs, summing up to 891, followed by acetic acid, with 2108 (Fig. 9).

Dose-dependent responses have been shown for *S. cerevisiae* in the presence of the same inhibitors (119–123). In summary, for all conditions, a positive correlation was found between the increased concentration of the inhibitor, the physiological impairment of the yeast's growth, and the number of DEGs (Fig. 7 and 9). Similar observations were described for *S. cerevisiae* (121,124).

4.4 CENTRAL CARBON METABOLISM

The expression of the glycolysis pathway encoding genes was overall strongly down-regulated in presence of acetic acid and hydrolysate, except for the genes *FBA1-2* (fructose 1,6-bisphosphate aldolase) and *CDC19* (pyruvate kinase) that were overexpressed (Fig. 10). Furaldehydes did not increase or reduce the expression of most glycolysis encoding genes, which could be correlated with the less inhibitory profile seen during the fermentation (Fig. 7 and 10). On the other hand, genes encoding for glycerol metabolism enzymes were all inhibited, with expression levels of *GPD1* (glycerol-3-phosphate dehydrogenase) being downregulated in presence of acetic acid, furaldehydes, and hydrolysate.

The C2 metabolism, i.e., acetic acid and ethanol that integrate into the glycolysis pathway showed a mixed pattern of expression. In general, genes encoding enzymes involved in the production of acetate and ethanol did not show differential expression or were downregulated, such as *PDA1* (pyruvate dehydrogenase), *ADH2* (alcohol dehydrogenase), and *ALD5* (mitochondrial aldehyde dehydrogenase) (Fig. 10). Contrarily, the *ACS1* gene, which encodes an acetyl-CoA synthetase that can directly convert acetate to acetyl-CoA, and the tricarboxylic acid cycle (TCA) did not show significant differences in gene expression or were up-regulated. The most up-regulated genes from TCA were found when the yeast was cultivated in presence of acetic acid, which may be related to the consumption of acetic acid seen during the cultivations (Fig. 10).

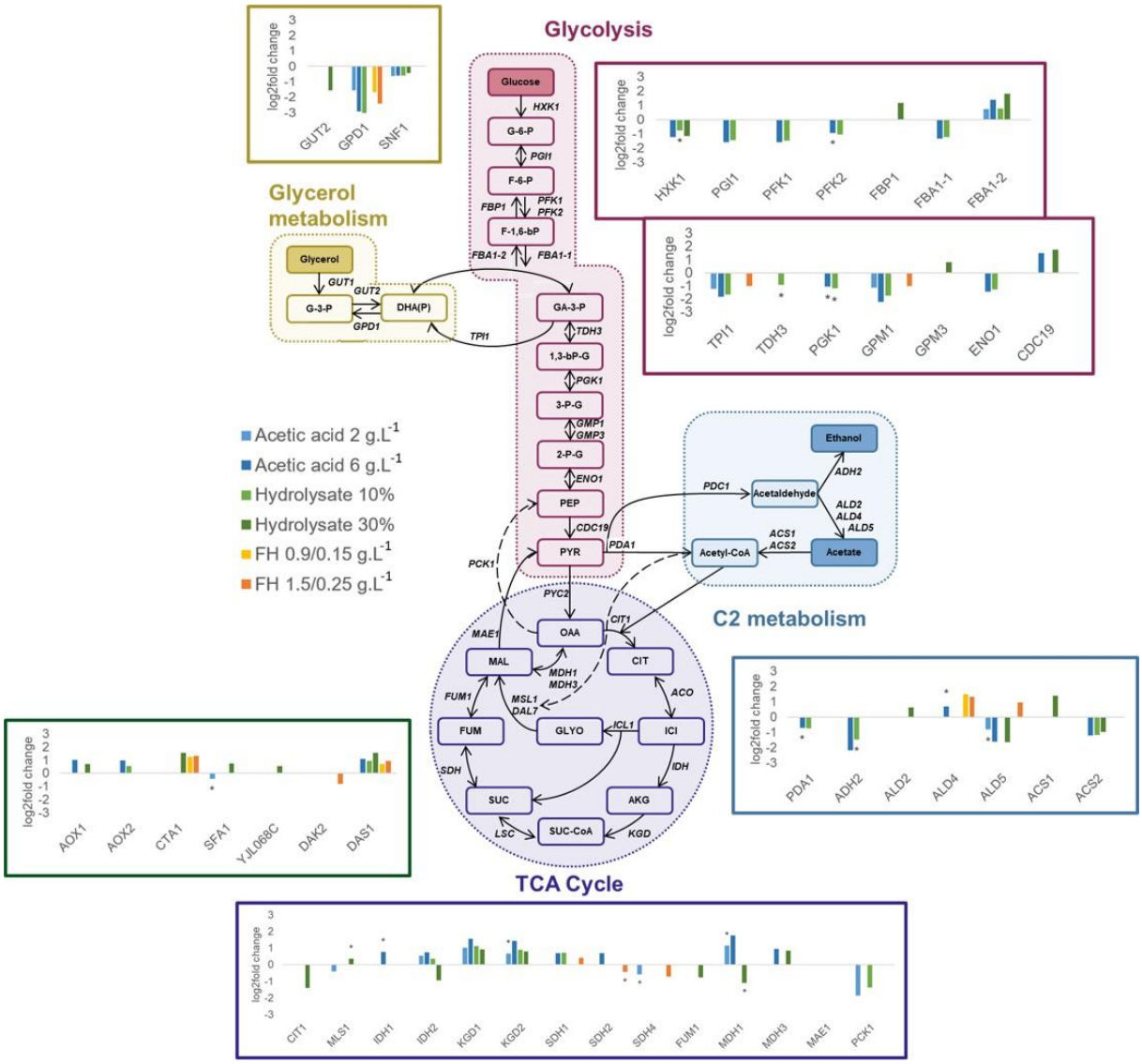

Figure 10 – Gene expression of central carbon metabolism pathways for *P. pastoris*. Possible carbon sources are presented in colored squares: glucose; glycerol, ethanol, acetate. Bar charts represent the transcriptional changes (log2 fold) of genes in acetic acid 2g.L-1 (light blue), 6g.L-1 (dark blue), FH 0.9/0.15g.L-1 (yellow), FH 1.5/0.25g.L-1 (orange), hydrolysate 10% (light green) and 30% (dark green) with P-values <0.05 or 0.1 (* on top of bar chart).
Metabolites: G-6-P: glucose 6-phosphate; F-6-P: fructose-6-phosphate, F-1,6-P: fructose 1,6-phosphate; G-3-P: glycerol 3-phosphate; GA-3-P: glyceraldehyde 3-phosphate; 1,3-bPG: 1,3-bisphosphoglycerate; 3-PG: 3-phosphoglycerate; 2-PG: 2-phosphoglycerate; PEP: phosphoenolpyruvate; PYR: pyruvate; DHA(P): dihydroxyacetone (phosphate); OAA: oxaloacetate; CIT: citrate; ICI: isocitrate; AKG: alpha-ketoglutarate; SUC: succinate; SUC-CoA: succinyl-Coenzyme A; FUM: fumarate; MAL: malate; GLYO: glyoxylate. Enzymes: HXK1: hexokinase; PGI1: phosphoglucose isomerase; PFK1/2: phosphofructokinase; FBP1: fructose-1,6-bisphosphatase; FBA1-1/1-2: fructose 1,6-bisphosphate aldolase; TPI1: triosephosphate isomerase; TDH3: glyceraldehyde-3-phosphate dehydrogenase; PGK1: 3-phosphoglycerate kinase; GPM1/3: phosphoglycerate mutase; ENO1: enolase I, phosphopyruvate hydratase; CDC19: pyruvate kinase, GUT1: glycerolkinase; GUT2: glycerol-3-phosphate dehydrogenase; GPD1: glycerol-3-phosphate dehydrogenase, SNF1: central kinase; PYC2: pyruvate carboxylase; CIT1: citrate synthase; ACO1/2: aconitase; ICL1: isocitrate lyase; DAL7: malate synthase; IDH1/2: isocitrate dehydrogenase; KGD1: alpha-ketoglutarate dehydrogenase complex; KGD2: dihydrolipoyl transsuccinylase; LSC1: succinyl-CoA ligase; SDH1/2/4: succinate dehydrogenase; FUM1: fumarase; MDH1: mitochondrial malate dehydrogenase; MDH3: malate dehydrogenase; MAE1: mitochondrial malic enzyme; PDC1 pyruvate decarboxylase; PDA1: pyruvate dehydrogenase (subunit from PDH complex); ALD2: cytoplasmic

aldehyde dehydrogenase; ALD4-1/4-2/5: mitochondrial aldehyde dehydrogenase; ADH2: alcohol dehydrogenase ACS1/2: acetyl-CoA synthetase; PCK1: phosphoenolpyruvate carboxykinase. Genes or conditions with P-values out of the threshold were not depicted. Source: Author's collection.

Pichia pastoris was able to co-consume glucose and acetic acid (Fig. 7b), which is not observed in most strains of *S. cerevisiae* (125). Those could be correlated with the reduced glucose consumption rate and toxicity of acetate (Fig. 7b). Acetate may lead to cytosol acidification by acetic acid dissociation in the cytosol, affecting cell metabolism and survival (Pampulha and Loureiro-Dias 1989; Sousa et al. 2012; Rego et al. 2014). The overexpression of genes responsible for the consumption and conversion of acetic acid may be a strategy to reduce its toxicity. This is corroborated by the transcriptional data that demonstrated that *P. pastoris* repressed glycolytic pathway and up-regulated C2/C3 metabolism in presence of inhibitors, especially in the presence of acetic acid and hydrolysate (Fig. 10).

These results are further supported by the recent results of Xu and coworkers (107) who have recently demonstrated that *P. pastoris* is capable of metabolizing acetate in the presence of glucose. The experimental data published by the authors do not explicitly demonstrate the co-consumption of glucose and acetate (as reported here); however, metabolite analyses suggest so. Also, contrary to *P. pastoris*, *S. cerevisiae* shows a Crabtree positive metabolism i.e., presents a fermentative metabolism even when it is cultivated in aerobiosis when glucose is present in high concentrations in the medium (Crabtree effect). Thus, *S. cerevisiae* shows a diauxic shift, where it switches from rapid fermentative growth once the preferred carbon source (glucose) has been exhausted to slower exponential growth by aerobic respiration using ethanol/acetate as carbon sources.

4.5 GENE ONTOLOGY ANALYSIS

The enrichment of gene ontology (GO) categories in response to acetic acid, furaldehydes, and hydrolysate was evaluated using differentially expressed genes for each inhibitor. Acetic acid resulted in the up-regulation of DEGs in the GO categories related to nucleic acid processing, in particular: RNA, methylation, and Rho protein signal transduction regulation (Supplemental Tables S1 and S2), and downregulation of oxi-reduction and macromolecules metabolic processes (Fig. 11a). From the eight genes present in the GO category methylation (*GCD10, GCD14, HSL7, MRM2, PPM1,*

PPM2, TGS1, PPR1), only *GCD10* and *MRM2* were not up-regulated in the presence of the other inhibitors evaluated (Supplemental Table S1). All Rho-related genes were up-regulated except in the hydrolysate 30% condition (Appendix B1 - Table S2). In the presence of furaldehydes, the oxi-reduction GO category was up-regulated (Fig. 11b), as well as the GO categories transmembrane transport, metal-ion, and iron-sulfur cluster assembly (Appendix B1 - Table S3 to S5). Protein-related processes, biosynthesis of small molecules (carboxylic acid), organophosphates, coenzymes, and phospholipids were expressively down-regulated (Fig. 11b). Finally, most identified GO terms found for lignocellulosic hydrolysate were down-regulated, being RNA processing, regulation of metabolic processes, and transmembrane transport the only up-regulated-related terms (Fig. 11c). A total of 25 DEGs within the transmembrane transport GO group were up-regulated in both hydrolysate conditions (Appendix B1 - Table S6), with 6 genes at least five times up-regulated in at least one of the two concentrations. (Table 3).

Acetic acid effects over nucleic acid synthesis and degradation (especially RNA) have been reported by Hasunuma and Kondo 2012 as a consequence of changes of intracellular pH homeostasis caused by undissociated weak acid diffusion into the cell (126). Cytoplasm acidification also affects lipid organization, membrane electrochemical gradient, and ATP availability, since it activates transmembrane proton pumps to regulate internal pH, as reviewed by (56). Overexpression of methyltransferases such as *PPR1* has been shown to improve *S. cerevisiae* growth and fermentation performance in the presence of acetic acid, presumably due to the decreased intracellular accumulation of reactive oxygen species (127). Since reactive oxygen species are also generated in the presence of furaldehydes (128), *PPR1* up-regulation in the presence of such compounds may also be advantageous. Interestingly, in this work, *PPR1* was up-regulated in all conditions evaluated, doubling its expression in acetic acid 6 $g.L^{-1}$. Another up-regulated group was the regulation of Rho protein signal transduction. Rho is a family of proteins which regulation affects numerous cell processes (129) and is essential for osmotic stress response (130) and low pH survival in yeast (131). However, further evaluation of the Rho role in lignocellulose-derived inhibitor tolerance must be performed.

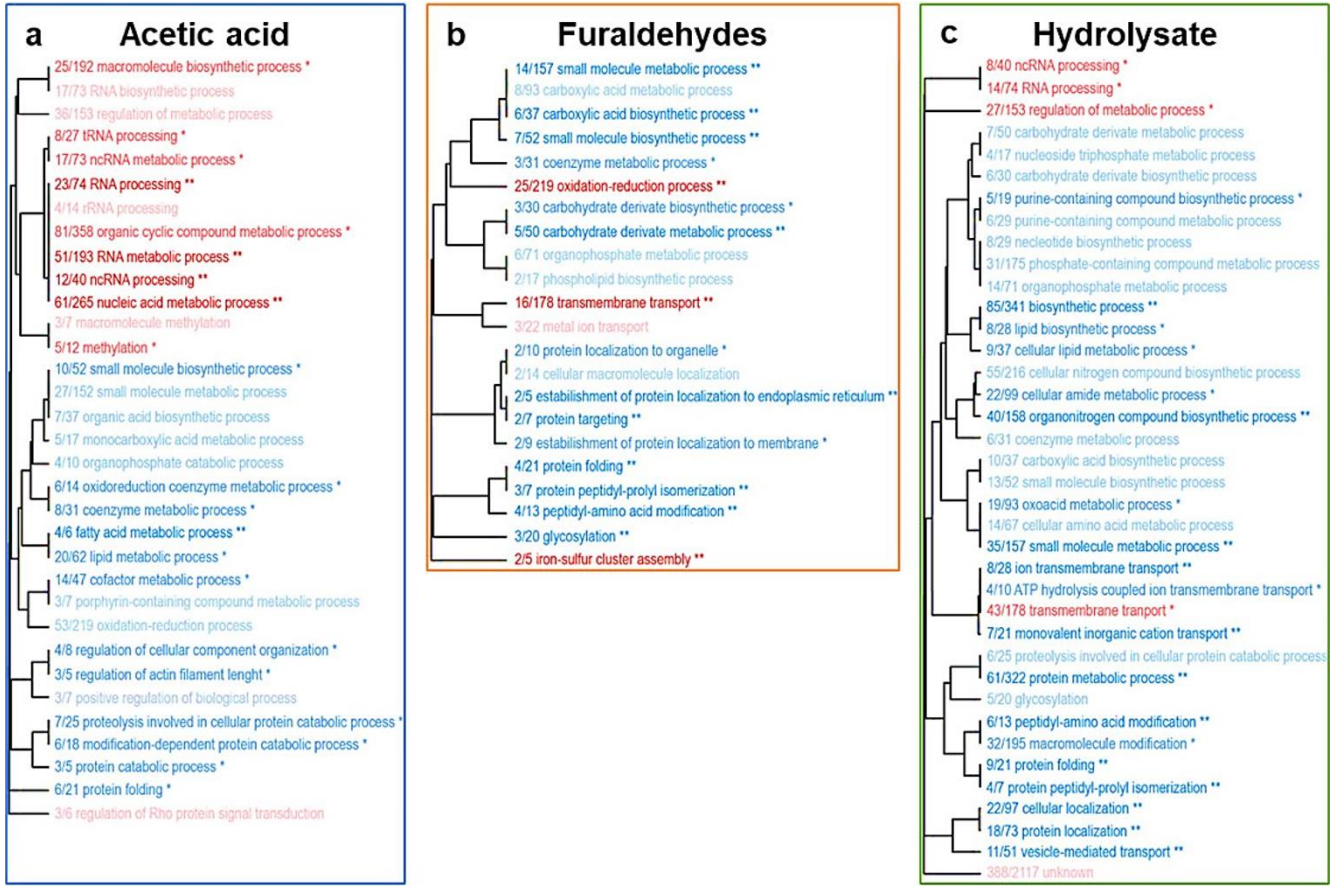

Figure 11 – Representativeness of GO categories for the DEGs of *P. pastoris* in presence of different inhibitors. The genes with differential expression in at least one concentration of acetic acid (a), furaldehydes (b), and hydrolysate (c) were used in the analysis. Up and down-regulated categories are shown in red and blue, respectively. P-values are equal to: (**) 0.01, (*) 0.05, (no symbol) 0.1. The numbers X/Y represent the number DEGs with that GO term found in the sample by the total number of genes with that GO term in the genome. Source: Author's collection.

These results indicate that *P. pastoris* response to furaldehydes is similar to previously reported for other yeast, like *S. cerevisiae.* The appearance of oxi-reduction GO term up-regulated in furaldehyde-containing conditions motivated a further investigation into this group. Oxi-reduction has been extensively described as relevant for yeast tolerance against furaldehydes (70,132). We observed that DEGs found in furaldehyde conditions (FH 0.9/ 0.15 g.L^{-1} and FH 1.5/ 0.25 g.L-1) related to oxidation-reduction process and transmembrane transport GO term category were induced. Among the 5 oxidoreductases with log 2-fold around 3 (*OYE3-2, QOR2, OYE3-1, NTL100*, and *NTL101*), only *QOR2* was not flavin mononucleotide (FMN)-dependent (Table 3). Genes *ZWF1* (glucose-6-phosphate dehydrogenase) and *ALD4* (mitochondrial aldehyde dehydrogenase) reported previously as important for furaldehydes tolerance, were also found overexpressed in *P. pastoris* (Table 3).

The increased expression of oxidoreductases in the presence of furaldehydes (Fig. 11) might be associated with the conversion of HMF and furfural to their less toxic forms, as reported previously for other yeasts (75,97,133,134). Among relevant oxidoreductase encoding genes found up-regulated in this work (Table 1), *ZWF1* (128) and *ALD4* (135) have been reported previously as capable of reducing HMF and furfural toxicity to the cell (136,137). Other oxidoreductases potentially involved in the detoxification of furaldehydes, but previously not shown, are *OYE3*, *QOR2*, and *NTL100* (Table 1). Another gene possibly related to furaldehydes tolerance is *YCT1* (Table 1), which encodes a cysteine transporter found to be up-regulated in the presence of many inhibitors in the yeast *Kluyveromyces marxianus* (138). Cysteine is related to the synthesis of glutathione, which is an important antioxidant molecule related to detoxification and oxidative stress response to HMF and furfural (139,140). Glutathione's importance in detoxification has also been related to the synthesis of sulfur amino acids and saving mechanisms in yeast (139), which may explain the iron-sulfur GO category up-regulation in presence of furaldehydes. Thus, *ZWF1, ALD4, OYE3, QOR2, NTL100 YCT1,* and *PPR1* are potential candidates for improving *P. pastoris* tolerance to the lignocellulose-derived inhibitors.

Table 3 – Selected differentially expressed genes (log 2-fold) of *P. pastoris* cultivated in presence of acetic acid, furaldehydes (HMF and furfural), and sugarcane biomass hydrolysate. In the case of P-value >0.05, the log2fold is not shown. Source: Author's collection.

			Log2fold						
Gene ID	**Gene Name**	**Description**	**Acetic acid 2 g.L^{-1}**	**Acetic acid 6 g.L^{-1}**	**FH 0.9/ 0.15 g.L^{-1}**	**FH 1.5/ 0.25 g.L^{-1}**	**Hydrolysate 10%**	**Hydrolysate 30%**	**GO term associated**
GQ6703442	*OYE3-2**	*Conserved NADPH oxidoreductase containing flavin mononucleotide (FMN), homologous to *OYE2P* with different ligand binding and catalytic properties, has potential roles in oxidative stress response and programmed cell death	-	-	3.12	3.47	3.30	-	Oxidation-reduction process
GQ6700918	*PPR1*	Pyrimidine pathway regulatory protein 1	0.53	1.03	-	0.40	0.82	0.28	Macromolecule methylation, macromolecule modification, methylation, ncRNA processing, organic cyclic compound metabolic process, RNA metabolic process, RNA modification, RNA processing, tRNA modification, tRNA processing
GQ6703443	*OYE3-1**	*Conserved NADPH oxidoreductase containing flavin mononucleotide (FMN), homologous to *OYE2P* with different ligand binding and catalytic properties, has potential roles in oxidative stress response and programmed cell death	-	-	2.72	3.09	3.12	-	Oxidation-reduction process
GQ6701720	*NTL101**	*Putative nitrilotriacetate monooxygenase family FMN-dependent oxidoreductase	-	-	2.45	2.94	2.20	-	Oxidation-reduction process
GQ6702862	*SOR1*	*Unique sorbitol dehydrogenase in *P. pastoris* whose promoter has an	2.26	3.77	-	1.32	3.01	2.36	Oxidation-reduction process

GQ6700554	*ALD4*	activity similar to *GAP* promoter using different carbon sources, expression in *S. cerevisiae* is induced in the presence of sorbitol or xylose Mitochondrial aldehyde dehydrogenase; required for growth on ethanol and conversion of acetaldehyde to acetate	-	-	1.51	1.32	-	-	Oxidation-reduction process
GQ6702661	-	-	2.74	2.12	0.81	0.91	2.48	2.63	Ion transport, cation transport, transmembrane transport, metal ion transport, localization
GQ6702266	*YCT1*	High-affinity cysteine transporter	-	1.35	4.43	4.81	2.87	-	Transmembrane transport, localization
GQ6703957	-	-	2.36	2.51	-	-	1.98	3.04	Transmembrane transport, localization
GQ6702465	*VBA1*	Vacuolar basic amino acid transporter 1	-	1.59	-	-	0.97	2.69	Transmembrane transport, localization
GQ6702095	*TPO3*	Polyamine transporter 3	2.53	1.81	-	-	1.98	2.63	Transmembrane transport, localization
GQ6703034	--	-	-	2.01	1.22	1.32	1.20	2.56	Transmembrane transport, localization
GQ6703337	**DUR3-2*	*Plasma membrane transporter for both urea and polyamines, expression in *S. cerevisiae* is highly sensitive to nitrogen catabolite repression and induced by allophanate, the last intermediate of the allantoin degradative pathway	-	2.34	1.32	4.92	-	1.14	Transmembrane transport, localization, localization,
GQ6705065	-	*Hypothetical protein not conserved	4.83	3.77	4.66	4.50	1.94	2.34	-
GQ6702974	-	*Hypothetical protein not conserved	2.81	2.71	2.90	3.90	-	1.54	-
GQ6700086	**RPH1*	*JmjC domain-containing histone demethylase	2.80	2.85	2.84	2.43	0.97	1.15	-
GQ6701722	**QOR2*	*Putative quinone oxidoreductase (NADPH:quinone reductase), similar to *Scheffersomyces stipitis QOR2*, and similarity to Zinc-binding dehydrogenases	-	-	2.86	-	2.73	3.22	Oxidation-reduction process,

GQ6701721	**NTL100*	*Putative nitrilotriacetate monooxygenase family FMN-dependent oxidoreductase	-	-	2.52	-	2.69	2.96	Oxidation-reduction process
GQ6701716	**CAO1*	*Copper amine oxidase similar to *Schizosaccharomyces pombe CAO1*	-	2.90	-	6.17	-	-	Amine metabolic process, oxidation-reduction process
GQ6702273	**DAL1*	*Allantoinase, expression in *S. cerevisiae* sensitive to nitrogen catabolite repression	2.36	3.73	2.59	2.78	-	-	Oxidation-reduction process
GQ6705251	-	*Hypothetical protein conserved (domain: *GAL4*)	2.75	3.12	2.81	2.56	-	-	Regulation of metabolic process, biological regulation
GQ6701500	**SOA1-6*	*Putative protein with similarity to allantoate permease, similar to the allantoate permease (*DAL5P*) subfamily of the major facilitator superfamily	-	-	2.18	-	2.51	3.08	Transmembrane transport, localization, localization,
GQ6703338	*AMD2*	-	-	2.65	1.89	4.47	-	1.66	-
GQ6701041	*CAR1*	-	-	2.36	-	3.85	-	1.89	-
GQ6703874	*NCS6*	-	2.79	2.79	2.80	1.53	-	1.54	-
GQ6704853	*PIC2*	Mitochondrial phosphate carrier protein 2 (Phosphate transport protein 2) (PTP 2) (Pi carrier isoform 2) (mPic 2)	2.75	3.38	2.88	2.67	-	-	-
GQ6702197	*PST2*	Protoplast secreted protein 2	-	-	4.56	1.27	-	0.90	-
GQ6703286	*STP3*	Zinc finger protein *STP3*	4.84	3.91	4.40	3.90	-	-	-
GQ6700345	*ZWF1*	Glucose-6-phosphate dehydrogenase (*G6PD*)			0.74	0.67			Single-organism carbohydrate metabolic process, small molecule metabolic process, oxidation-reduction process, hexose metabolic process,

*: information manually annotated.

To get a better insight into the patterns of response among the different inhibitors, a heat map was constructed with the 630 DEGs found for all three inhibitors (Fig. 9a) in at least one of the two concentrations evaluated (Fig. 12). The gene expression profiles from the cells grown in the presence of 2 $g.L^{-1}$ and 6 $g.L^{-1}$ acetic acid and hydrolysate 10% are more like to each other than to cells grown in hydrolysate 30%. This difference is consistent with the physiological (Fig. 7) and PCA analysis results (Fig. 8). As acetic acid is the most abundant inhibitor in the sugarcane bagasse used here, the proximity among acetic acid conditions and hydrolysate 10% indicates that a big part of yeast's response in this condition is related to the acid's presence.

Additionally, seven distinct clusters of differentially expressed genes and enriched GO annotations are found in the heat map (Fig. 12). Cluster 2 and most of 7 involve up-regulated categories mainly related to regulation of transcription from RNA pol II promoters, intracellular signal transduction, and nucleobase-containing compound metabolism. Most genes on cluster 7 were up-regulated, but a small part, especially in presence of acetic acid 6 g.L-1 was down-regulated. Cluster 4, 5, and 6 include down-regulated GO categories (except hydrolysate 30%) mostly related to transport, especially vesicle-mediated transport, regulation, initiation of transcription, and tRNA aminoacylation for protein translation. Clusters 1 and 3 showed both up and down-regulated categories, those related to nucleotides biosynthesis and oxidation-reduction processes.

The 10 genes most up-regulated in each condition, i.e., acetic acid, furaldehydes, and hydrolysate, in the two different concentrations, were identified (Supplemental Table S7). From those 60 genes, 26 were present in more than one condition (usually in the lower and higher concentration of the same inhibitor). Thus, a total of 34 unique genes were identified. Among those genes, 19 (*DUR3-2, GQ6705065, GQ6702974, RPH1, QOR2, OYE3-2, NTL100, CAO1, DAL1, GQ6705251, SOA1-6, AMD2, CAR1, NCS6, PIC2, PST2, SOR1, STP3, and YCT1*) were up-regulated in all tested conditions (Table 3).

As was previously discussed, transcriptome results indicate that *P. pastoris* response patterns to inhibitors are similar to what was previously reported for other yeasts. This is, therefore, also an opportunity to transpose previous strategies validated in other organisms to build more robust strains into *P. pastoris*. Investigation and validation of potential genes of interest previously undescribed for *P. pastoris* are also a possibility.

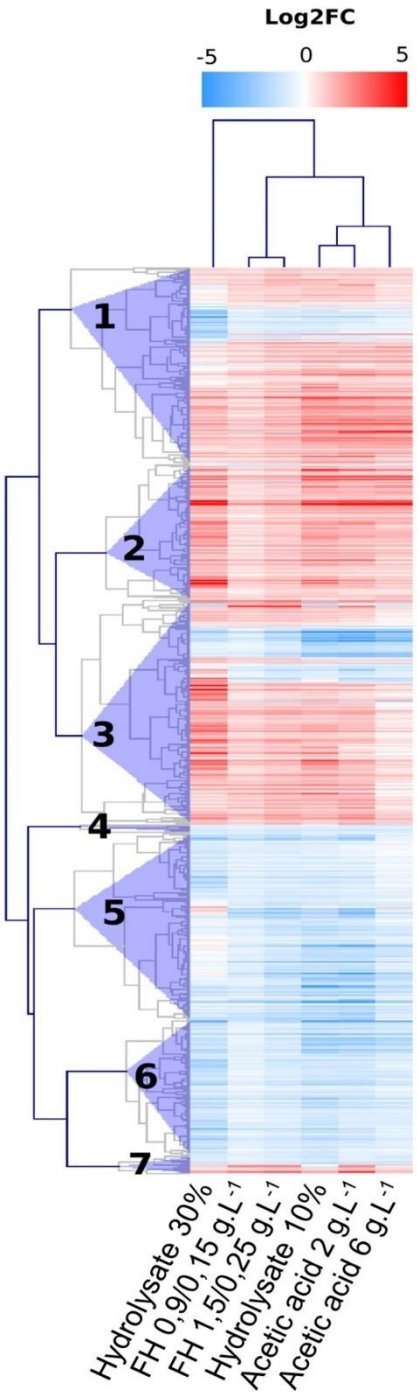

Figure 12 – Hierarchical clustering heat map of the 630 DEGs common in all categories. Changes in the expression are shown on a color scale, where red represents up-regulation and blue represents down-regulation. Each column relates to one inhibitory condition and each row represents one DEG. Genes were organized into seven clusters according to patterns of expression. Source: Author's collection.

5 CONCLUSIONS

Our study presents the a physiological and genome-wide transcriptome analysis of *P. pastoris* under the effect of major inhibitors found in lignocellulosic hydrolysates. The results reveal that acetic acid, furaldehydes, and sugarcane hydrolysate inhibit cell metabolism in a dose-dependent manner, and the yeast transcriptional response increases with increased concentrations of the inhibitors. Unlike other industrial yeast, such as *S. cerevisiae, a*cetic acid can be co-consumed by *P. pastoris* as an alternative carbon source, despite affecting yeast's growth rate. Additionally, gene clusters related to the response of *P. pastoris* to lignocellulose-derived inhibitors are described here, and candidate genes to improve yeast tolerance were identified.

Even though *P. pastoris* is a well-known and one of the favorite host organisms used as a tool in both academia and industry, little is known about its response to toxic compounds, and especially those present in lignocellulosic hydrolysate. Previous work combining lignocellulosic hydrolysate and *P. pastoris* seems to be focused on heterologously expressing enzymes of interest for lignocellulose detoxification (141–143). The results reported here have significant implications for better understanding and broadening biotechnological applications of *P. pastoris.* This could include, but not be limited to building more robust strains of *P. pastoris* for industrial applications, for example, resistance to low pH tolerances, (desirable to avoid contamination by other organisms), and use of acetic acid as carbon source, production of enzymes for in-loco detoxification of hydrolysate and others.

CHAPTER 2 – CHARACTERIZATION OF NEW *KOMAGATAELLA* STRAINS FOR TOLERANCE TO ACETIC ACID, FURALDEHYDES, AND XYLOSE CONSUMPTION

Pichia pastoris is one of the most widely used yeast in the pharmaceutical and biotechnological industries (4) and it is believed that more than five thousand different proteins have been produced in it (2). Having been discovered just under a century ago by Alexandre Guilliermond (1), the accumulated knowledge about this yeast is not as robust as other industrial species like *S. cerevisiae,* and it is still considered a "non-conventional" yeast (53,144). Technological advances in sequencing, genome editing techniques, characterization, and computational modeling have advanced our understanding of *P. pastoris* in recent decades, however, the available genetic diversity of this species is still very restricted. Nearly all research uses strain CBS7435 (NRRL Y-11430) and its derivatives (145) and if few natural isolates can be found in culture collections or have been thoroughly characterized (13).

Famous for growing at high cell densities using methanol, other sugars were investigated as possible carbon sources for *P. pastoris* like sorbitol, acetic acid, lactic acid, mannitol, fructose (21–25). However, despite the increasing interest in renewable sources and sustainable processes for applications in the bioeconomy, lignocellulose biomass has hardly been investigated as a plausible carbon source for *P. pastoris.* Although rare, previous growth tests for *P. pastoris* in lignocellulose-derived media were done, showing not all sugars are consumed by the yeast, which is also impaired by inhibitors present in the media (77). The current understanding is that *P. pastoris* is not capable of using xylose, a major component of lignocellulosic hydrolysate, as its main carbon source (21,146), unless genetically engineered towards it (26). Different groups have engineered *P. pastoris* for utilizing efficiently carbon sources derived from lignocellulose, like cellulose (27), and xylose, galactose (23,25).

One problem for lignocellulosic biomass utilization by yeast is the presence of inhibitory compounds. Acetic acid, HMF, and furfural are the most common inhibitors found in main lignocellulose-derived hydrolysates. They are formed during the pretreatment of biomass for the release of sugar monomers, as a result of hemicellulose deacetylation, hexoses, and pentose dehydration, respectively. In other yeasts, effects of lignocellulose-derived inhibitors include damages to the membrane, cell metabolism and macromolecules synthesis, growth inhibition, redox imbalance, and glycolysis enzyme impairment (59,71).

Given the huge biotechnological potential of lignocellulose biomass utilization, and the history of valuable industrial yeast of *P. pastoris*, in this chapter in total twenty-six strains of the *Komagataella* genus, twenty-five being different natural isolates, and

one commercial were characterized for their tolerance to acetic acid and furaldehydes, and their ability to consume xylose as the sole carbon source. Our results show for the that *P. pastoris* is, in fact, capable of metabolizing xylose into biomass, in contrast to what was previously thought (21,26,146). Fermentations and isotope labeling experiments show that xylose is being used by the yeast to produce biomass and xylitol. Preliminary metabolomics analysis shows the xylose is incorporated into the non-oxidative pentose pathway and glycolysis. However, all evaluated strains had slow growth on xylose, and optimization is necessary not only if any biotechnological application is aimed, but also to facilitate research.

One strategy proven to be a useful tool for metabolic engineering organisms towards desirable traits is adaptative laboratory evolution (ALE) (147). Typically, a series of cultures with a selective pressure agent is maintained for long periods at continuous dilution in a chemostat or serial transfer from culture flasks (80). ALE thrives from an unbiased selection of naturally occurring mutations and genetic variation for establishing a strain with desired characteristics (78,79). This strategy has been previously utilized to improve consumption of xylose in other yeast, like *S. cerevisiae* (83,148,149), or *Klebsiella oxytoca* (150), and *Lactobacillus pentosus* (151). Even though no major difference was observed between the original and adapted strains, one strain had an 85% improvement in growth after 158 days of adaptation.

This part of the study was performed in Dr. Diethard Mattanovich's lab (head of the Institute of Microbiology and Microbial Biotechnology) at the University of Natural Resources and Life Sciences, Vienna, Austria.

6 MATERIALS AND METHODS

6.1 YEAST STRAINS AND MAINTENANCE

All yeast strains used in this chapter are described in table 4. Stock cultures were grown in YPD (Yeast extract 1%, peptone 2%, and glucose 2%) medium, preserved in 30% glycerol, and maintained at -80ºC.

Table 4 – List of *Komagataella* strains used in Chapter 2, their codes, the origin of isolation, location, and collection source. Source: Author's collection.

	Strain		Code	Isolated from	Location	Source	Publication
1	*Komagataella*	*phaffii*	CBS 2612	*Quercus kelloggii*	California, USA	CBS-KNAW	(7)
2	*Komagataella*	*phaffii*	CBS 7435	Unknown	unknown	NRRL	(152)
3	*Komagataella*	*phaffii*	85- 263.1	*Quercus emoryi*	Tucson, Arizona, USA	UWOPS	Unpublished
4	*Komagataella*	*phaffii*	03- 318t1	*Quercus rubra*	Long Point, Ontario	UWOPS	Unpublished
5	*Komagataella*	*phaffii*	03- 328y3	*Quercus rubra*	Long Point, Ontario	UWOPS	Unpublished
6	*Komagataella*	*phaffii*	81- 86	*Quercus rubra*	unknown	UWOPS	Unpublished
7	*Komagataella*	*phaffii*	81- 18	Unknown	unknown	UWOPS	Unpublished
8	*Komagataella*	*phaffii*	82- 16	*Quercus rubra*	Pinery Ontario	UWOPS	Unpublished
9	*Komagataella*	*phaffii*	85- 348.1	*Quercus emoryi*	Tucson, Arizona, USA	UWOPS	Unpublished
10	*Komagataella*	*phaffii*	85- 926.1	*Quercus emoryi*	Tucson, Arizona, USA	UWOPS	Unpublished
11	*Komagataella*	*phaffii*	91- 119.3	*Quercus rubra*	Wilderness Area Pinery, Ontario	UWOPS	Unpublished
12	*Komagataella*	*phaffii*	91- 132.2	*Quercus rubra*	South of Road Backus Woods, Ontario	UWOPS	Unpublished
13	*Komagataella*	*pastoris*	DSMZ 70382 (CBS 704)	Chestnut tree	Lyon, France	DSMZ	(1)
14	*Komagataella*	*pastoris*	DSMZ 70877	Beech	Berlin, Germany	DSMZ	(6)
15	*Komagataella*	*pastoris*	CBS 9178	thick mass in the cavity of *Quercus sp.*	Pilis Mountain, Hungary	CBS-KNAW	(153)

16	*Komagataella*	*pastoris*	CBS 9173	exudate of *Acer platanoides*	Hungary	CBS-KNAW	(153)
17	*Komagataella*	*pastoris*	CBS 9185	exudate of *Fagus sylvatica*	Hungary	CBS-KNAW	(153)
18	*Komagataella*	*phaffii*	X33	-	-	Invitrogen®	-
19	*Komagataella*	*pseudopastoris*	CBS 9187	rotten wood (cavity of *Salix alba*)	Dorog, Hungary	CBS-KNAW	(153)
20	*Komagataella*	*pseudopastoris*	CBS 9189	rotten wood (*Salix alba*)	Budapest, Hungary	CBS-KNAW	(153)
21	*Komagataella*	*ulmi*	CBS 12361	elm tree exudate (*Ulmus americana*)	Peoria, Illinois, USA	CBS-KNAW	(154)
22	*Komagataella*	*ulmi*	91- 206.2	*Quercus rubra*	Long Point, Ontario	UWOPS	Unpublished
23	*Komagataella*	*ulmi*	03- 338t1	*Quercus rubra*	Long Point, Ontario	UWOPS	Unpublished
24	*Komagataella*	*populi*	CBS 12362	cottonwood tree exudate *(Populus deltoides*)	Peoria, Illinois, USA	CBS-KNAW	(154)
25	*Komagataella*	*kurtzmanii*	CBS 12817	fir flux	Catalina Mountains, Arizona, USA	CBS-KNAW	(155)
26	*Komagataella*	*mondaviorum*	CBS 15017	cottonwood tree exudate (*Populus deltoides*)	Davis, USA	CBS-KNAW	(12)
27	*Candida*	*lignohabitans*	CBS 10342	Gut of wood roaches	Border between Tennessee and North Carolina, USA		(156)

UWOPS: University of Western Ontario, Department of Biology Yeast Culture Collection

CBS-KNAW: Central Bureau of Fungal Cultures (Centraalbureau voor Schimmelcultures)

NRRL: The Northern Regional Research Laboratory (American Research Service culture collection - ARC)

DSMZ: German collection of microorganisms and cell cultures (Deutsche Sammlung von Mikroorganismen und Zellkulturen)

6.2 CULTURE MEDIA AND SOLUTIONS

6.2.1 PBS 10X (Stock)

KH_2PO_4 0.24%, Na_2HPO_4 *$2H_2O$ 1,8%, KCl 0.2%, NaCl 8.0%. The pH was adjusted to 7.4 and it was stored at room temperature. For use, the stock was diluted to 1X with sterile distilled water.

6.2.2 YNB

YNB (yeast nitrogen base, Sigma Aldrich Y0626) without amino acids (0.68% YNB, 2% ammonium sulfate), 4.10^{-5}% biotin, buffered to pH 5.5 with phthalate buffer (1.02% potassium hydrogen phthalate with 0.22% potassium hydroxide w/v).

6.3 SPOTTING ASSAY

Colonies grown in YPD medium (yeast extract 1%, peptone 2%, glucose 2%) from glycerol stocks were inoculated in 2 mL YPD in 24-deep well plates and incubated overnight at 28 ºC, 280 rpm in a rotary shaker. The next day, cells were washed twice and diluted in PBS to OD_{600} 0.3 into a 96-deep-well plate. Using a multichannel pipette, five series of 1:10 dilutions were done (up to $1:10^{-5}$ dilutions) on PBS (from row A to F, Fig. 13A). From the more diluted row (F) to least (A), 4ul of each was plated into the respective YNB agar plates supplemented with carbon sugar and/or inhibitors. Plates were then incubated at 30 º for 168 h. Pictures were taken every 24 hours on the scanner EPSON® perfection V750 PRO with the following settings: original type: film with area guidance, film type: positive film, resolution: 400 dpi, color 48 bits.

For inhibitor tests, YNB media was supplemented with 2% glucose (w/v). Concentrations of inhibitors in the final tests were: $2g.L^{-1}$ acetic acid, and 0.25 $g.L^{-1}$ HMF with 1.5 $g.L^{-1}$ furfural. Inhibitors were added to hot media after autoclavation. For xylose tests, YNB media was supplemented with 2% xylose (w/v).

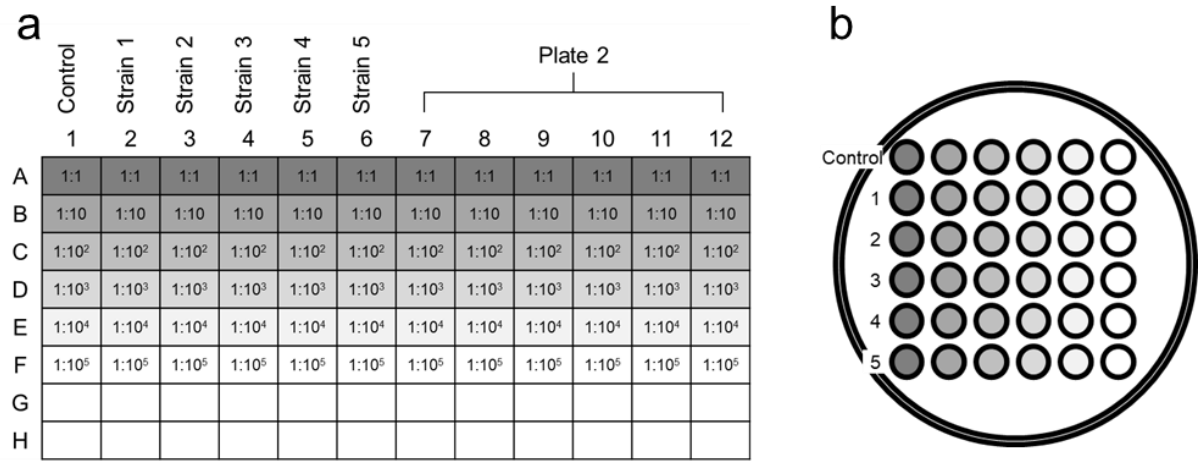

Figure 13 – Schematic representation of the 96-deep-well plate for dilutions (a) and agar plate (b) of the spotting assays. For each deep-well plate, two sets of strains (control plus five strains) were prepared. For each plate, lines represent the same inoculum in sequential 1:10 dilutions. Source: Author's collection.

6.4 FLASK GROWTH ASSAY

Colonies grown in YPD medium from glycerol stocks were inoculated in 50 mL YPG (yeast extract 1%, peptone 2%, glycerol 2%) and incubated overnight at 25 ºC and 180 rpm. The next day, cells were washed twice and inoculated to OD_{600} equals 12 on 25 mL of YNB supplemented with 2% xylose. The control media had no addition of xylose). Cultures were incubated at 25 ºC and 180 rpm for ten days (240 hours). A mixture of 50 U/mL penicillin and 50 µg/mL streptomycin were added as contamination control agents (Gibco) to each flask. Each culture was weighed before and after each sampling and the weight difference from evaporation was compensated with sterile water (w/v). Cell growth (OD_{600}) and metabolite profiles (HPLC analysis) were monitored, and dry cell weight was calculated from the pre-inoculum and final timepoints.

6.5 DRY CELL WEIGHT (DCW)

Cells from growth cultures were sampled, centrifuged and the supernatant removed. Pellets were washed twice with 0,1M HCl, resuspended in 1-3 mL distilled water, and transferred to pre-weighed glass tubes. Samples were dried at 100 ºC for 48 h. After cooling down, tubes were weighed on the same precision scale. Dry cell weight corresponds to the difference of tube weight with and without dried cells, divided

by the initial volume of cells that were sampled. Depending on the experiment volume, 1 ml, 5 ml, or 10 mL are taken for DCW.

6.6 ^{13}C XYLOSE CULTURE ASSAY

Colonies grown in YPD medium from glycerol stocks were inoculated in 100 mL YPG and incubated overnight at 25 ºC and 180 rpm. The next day, cells were washed twice with PBS, concentrated to OD_{600} 30, pre-inoculated into 30 mL of YNB supplemented with 2% xylose, and incubated overnight at 25ºC and 180rpm. Pre cultures were used to inoculate 10 mL YNB with 2% ^{12}C or 1-^{13}C D-xylose (Sigma-Aldrich) at a starting OD_{600} of 9. A mixture of 50 U/mL penicillin and 50 µg/mL streptomycin (Gibco) was added as a contamination control agent. Cells were grown at 25 ºC and 180 rpm for ten days and samples for OD600, cell dry mass, and HPLC analysis were taken at the beginning and end of the cultivation. All growth assays were performed in triplicates.

6.7 BIOMASS LABELING (ISOTOPE RATIO MASS SPECTROMETRY)

For ^{13}C content determination of enriched biomass, samples were taken at the end of cultivation, washed twice with 0,1M HCl, followed by once with reverse osmosis water. Volumes of cell culture corresponded to approximately 10mg of dry biomass. Cell pellets were stored at -20 ºC until sent for analysis. The ^{13}C /^{12}C content of yeast biomass was analyzed by Imprint Analytics GmbH (Austria) by elemental analysis with isotope ratio mass spectrometry (EA-IRMS).

6.8 ADAPTIVE LABORATORY EVOLUTION (ALE)

Adaptive laboratory evolution was done in two different phases, always in YNB media supplemented with 2% xylose. In the first phase (116 days), serial transfers were done when cultures reached OD_{600} of 4.0 to a new flask with 50 mL of media and an initial OD_{600} of 1.0. In the second phase (42 days) cells were sequentially transferred once a week to 25ml of media and an initial OD_{600} of 4.0. For every other transfer, cells were frozen in glycerol stocks (30%) and kept at -80ºC. Significant differences between original and ALE strains were done by the One-Way ANOVA test.

7 RESULTS AND DISCUSSION

7.1 TOLERANCE OF *Komagataella* STRAINS TO ACETIC ACID AND FURALDEHYDES

Spotting assays were used to identify *Komagataella* strains with potentially increased tolerance to inhibitors (acetic acid or furaldehydes) and xylose-consuming capacity. All tests were done in minimal media (YNB) in biological triplicates for the 26 *Komagataella* strains. Concentrations were chosen based on the results observed for the transcriptome (Chapter 1). For acetic acid, five concentrations were tested ranging from 2 $g.L^{-1}$ to 10 $g.L^{-1}$. however, no growth was observed in concentrations above 2 $g.L^{-1}$ of acetic acid even after 168 h of incubation. For furaldehydes, three concentrations were tested ranging from 0.25 $g.L^{-1}$ HMF and 1.5 $g.L^{-1}$ furfural to 0.55 $g.L^{-1}$ HMF and 2.9 $g.L^{-1}$, furfural. At the highest concentration, no growth was observed after 168 h of incubation. The inhibitors tests were performed in YNB supplemented with 2% glucose plates and the final concentrations of inhibitors chosen for further experiments were: 2 $g.L^{-1}$ of acetic acid; 0.25 $g.L^{-1}$ HMF, and 1.5 $g.L^{-1}$ furfural mixtures. The xylose tests were performed on YNB plates supplemented with 2% xylose at and were compared to glucose 2% plates. Two yeasts were chosen as control, with at least one added to all plates, making comparison among plates possible. The strain X33 was used as a *Komagataella* reference, and *C. lignohabitans* was used as a non-*Komagataella* and established xylose consuming reference (157).

The yeast strains grew consistently in the control (YNB supplemented with 2% glucose, no inhibitor) throughout all the experiments. To summarize overall observation among inhibitors, figure 14 shows the results of one of the replicates, for six strains (*K. phaffii* X33, *C. lignohabitans, K. phaffii CBS 2612, K. phaffii CBS 7435, K. phaffii 85-263.1 and K. phaffii 0.-318t1*), in three time points (48 h, 120 h, and 168 h). It was observed that *C. lignohabitans* has a similar profile to *Komagataella* strains in the presence of acetic acid but is more sensitive to furaldehyde.

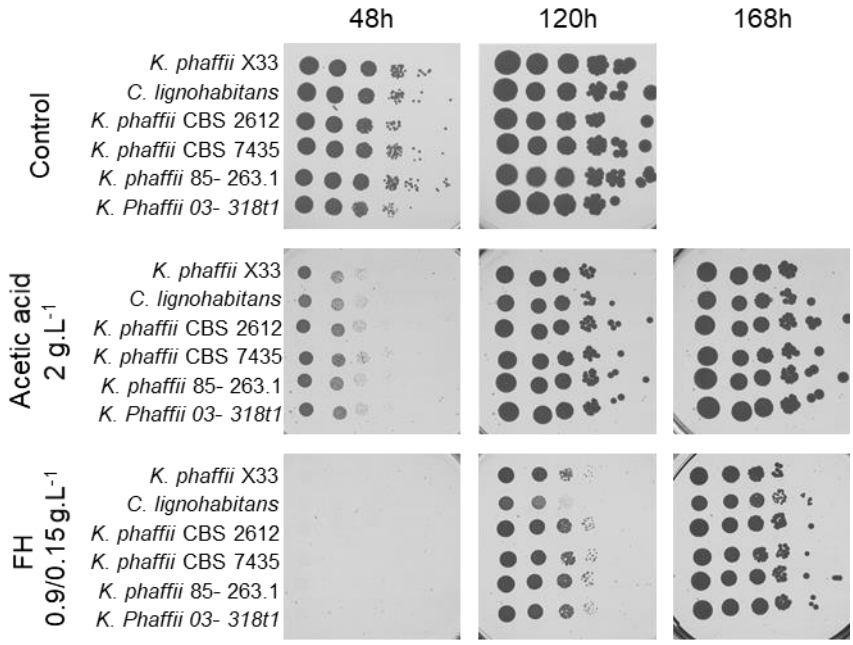

Figure 14 – Growth pattern of *Komagataella* strains in media with or without the addition of lignocellulose-derived inhibitors. Strains from top to bottom in each plate: *K. phaffii X33*, *K. phaffii CBS 2612*, *K. phaffii CBS 7435*, *K. phaffii 85-263.1,* and *K. phaffii 0.-318t1* and *C. lignohabitans* in YNB supplemented with 2% glucose plates without inhibitors (control), with the addition of acetic acid 2 $g.L^{-1}$, and furaldehydes (0.15 g.L-1 HMF and 0.9 g.L-1 of furfural) after 48 h, 120 h, or 168 h. The control plate was overgrown and discarded after 120 h. For each plate, columns represent the same inoculum in sequential 1:10 dilutions. Photos illustrate one from three biological replicates of each experiment. Source: Author's collection.

Cells from in control plates (no inhibitors) had grown more after 48 h, showing little difference among dilutions at 120 h, and being discarded. The presence of acetic acid delayed yeast's growth when compared to control, with the yeast growth pattern at 120 h in the presence of acetic acid being similar to that of the control at 48 h. Furaldehydes had the biggest impact on cell growth compared to other conditions. It showed no visible growth after 48 h, and at 120 h in the presence of furaldehydes, yeast growth had a pattern similar to 48 h in presence of acetic acid. By the end of the experiment, the growth observed in presence of furaldehydes was comparable to found in presence of acetic acid at 160h, or control after 120 h (Fig. 14).

Comparison between all the 26 *Komagataella* strains in acetic acid plates can be seen in figure 15. At 48 h, all strains had their growth initiated, with differences being more visible after the second dilution ($1:10^2$). In this replicate, some strains like *K. pastoris* stood out as a better grower, however, this is only one of the three replicates of the experiment. When functionally screening for better growers in the presence of

acetic acid, no strain consistently stood out throughout the triplicate. Variations are probably due to manipulation errors.

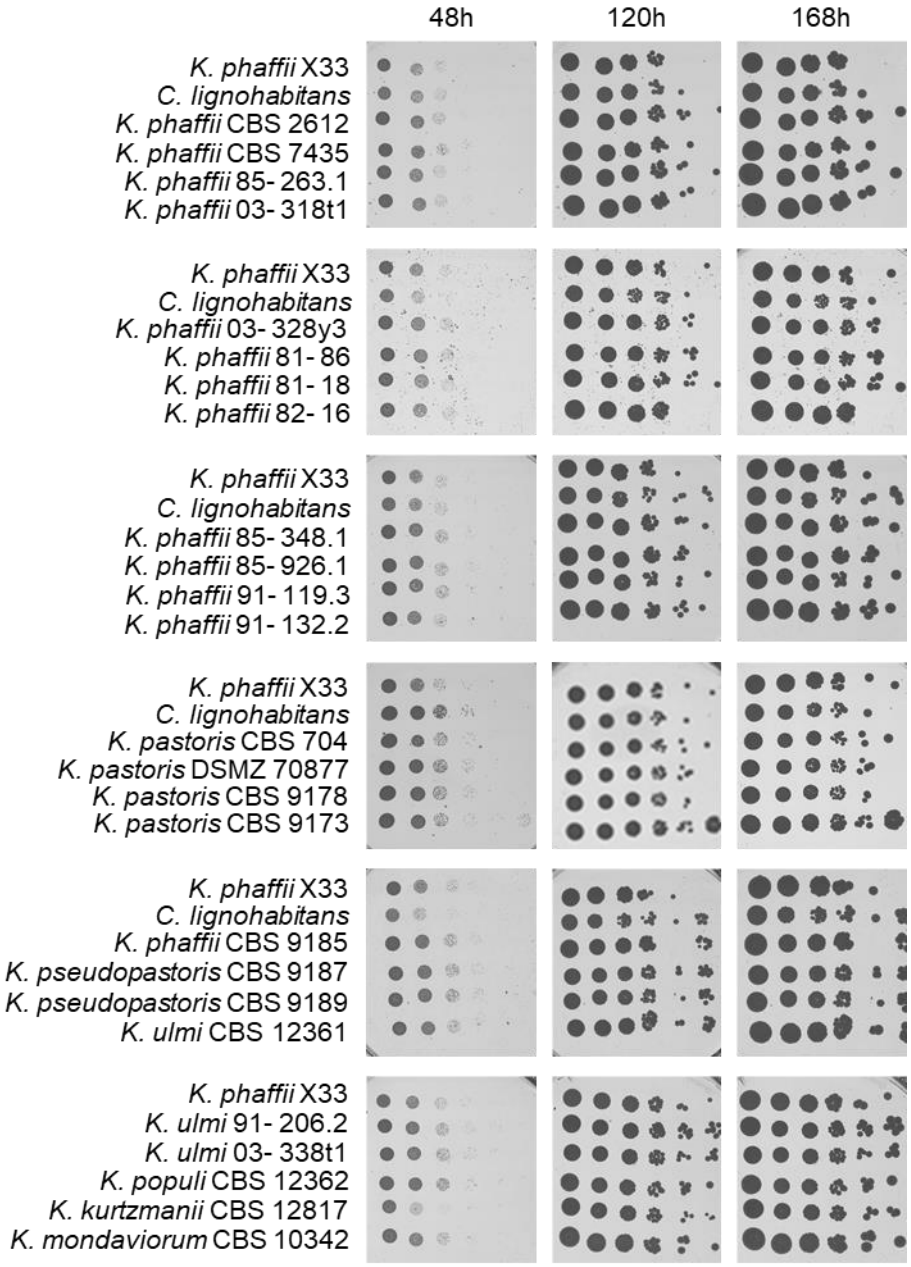

Figure 15 – Spotting assay of 26 strains of *Komagataella* strains in YNB supplemented with glucose-containing 2g.L^{-1} acetic acid in three different time points: 48 h, 120 h, and 168 h. For each plate, lines represent the same inoculum in sequential 1:10 dilutions. The experiments performed in biological triplicate show good reproducibility (i.e., similar performance of the strains), and the figure illustrates the results of one replicate. Source: Author's collection.

As designed for acetic acid, the spotting assays of the 26 yeast on plates supplemented with furaldehydes aimed to functionally screen for strains whose growth

stood out. An unexpected phenomenon, however, was observed in this test (Fig. 16). After 48 h, no growth could be observed for any of the 26 strains, but after 120 h, however, some plates had visible growth up to the fourth dilution (fourth column, $1:10^3$) while others had no visible growth (fourth plate). Comparing just the two controls in the final time point, for example, the three upper plates show a similar growth pattern, with at least one colony growing at the $1:10^4$ dilution (fifth column), while the fourth plate has nearly no growth visible after the second dilution (1:10). Inconsistent growth patterns are also present in the last two plates. The presence of the control strains *K. phaffii* X33 and *C. lignohabitans* in all of the plates (except *C. lignohabitans* in the bottom one) indicates that the differences are not related to strain's performances, but rather to experimental conditions employed.

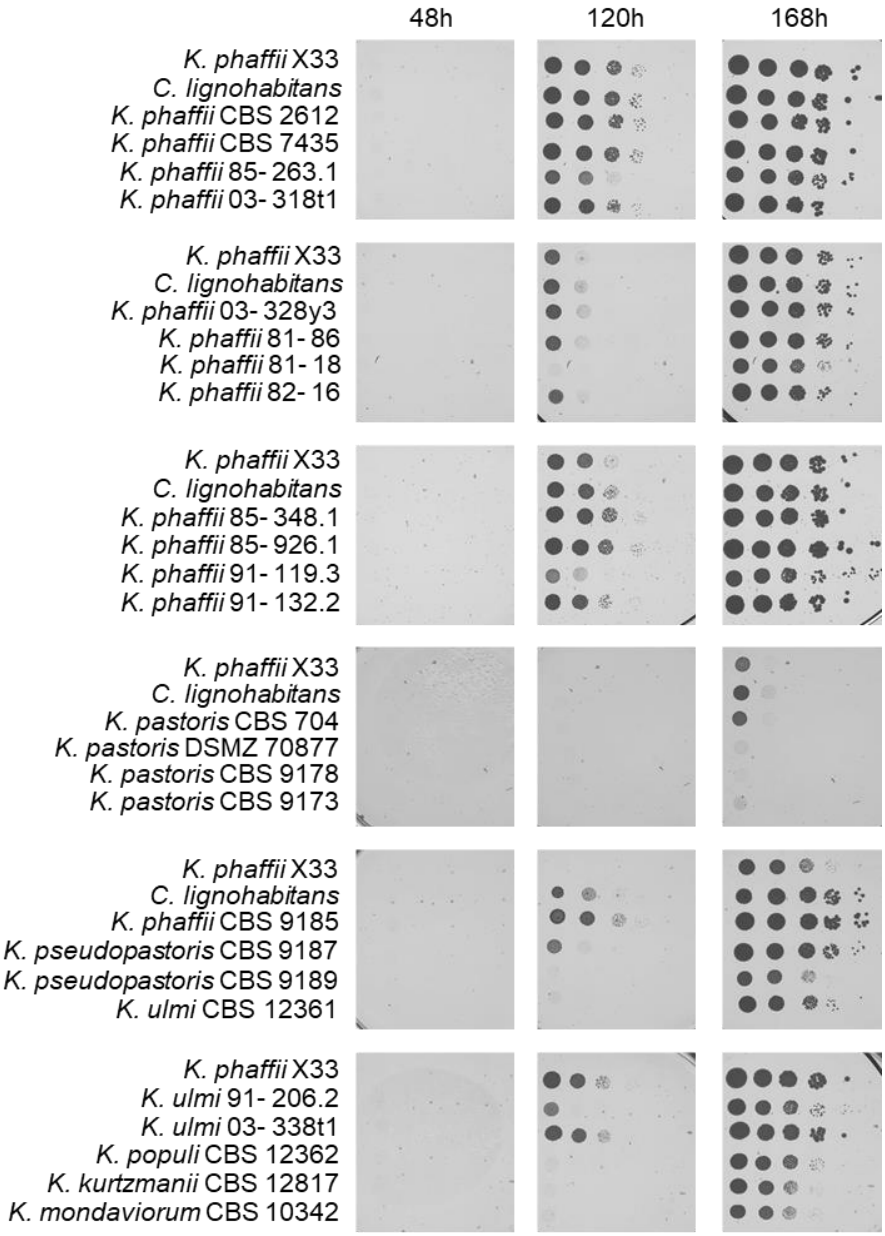

Figure 16 – Spotting assay of 26 strains of *Komagataella* in strains in media containing 0.15 g.L^{-1} HMF and 0.9 g.L^{-1} of furfural in three different time points: 48 h, 120 h, and 168 h. For each plate, lines represent the same inoculum in sequential 1:10 dilutions. The experiments performed in biological triplicate show good reproducibility (i.e., similar performance of the strains), and the figure illustrates the results of one replicate. Source: Author's collection.

To discard manipulation errors, the experiment was repeated two more times, however, every time aberrant behavior was observed. Since the furaldehyde tests were performed concomitantly to the acetic acid and xylose tests, from the same pre-inoculum and dilutions, and the media was made from the same stock solutions, the presence of furaldehydes in the media was determined as the only possible cause for the inconsistencies. One hypothesis investigated for the lack of reproducibility was that

the stacking order of the plates in the incubator was affecting the yeast's growth. A stacking test was done by plating the same strains (*K. phaffii* X33, *K. phaffii* 7435, *K. pastoris* CBS 704, K. populi CBS 12362*,* and *K. mondaviorum CBS 1517*) five times, randomly in YNB supplemented with 2% glucose with or without furaldehydes (0.15 $g.L^{-1}$ HMF and 0.9 $g.L^{-1}$ furfural). The stacking order was registered, and growth was monitored for 168 h (Fig. 17). The experiment was performed in biological triplicates.

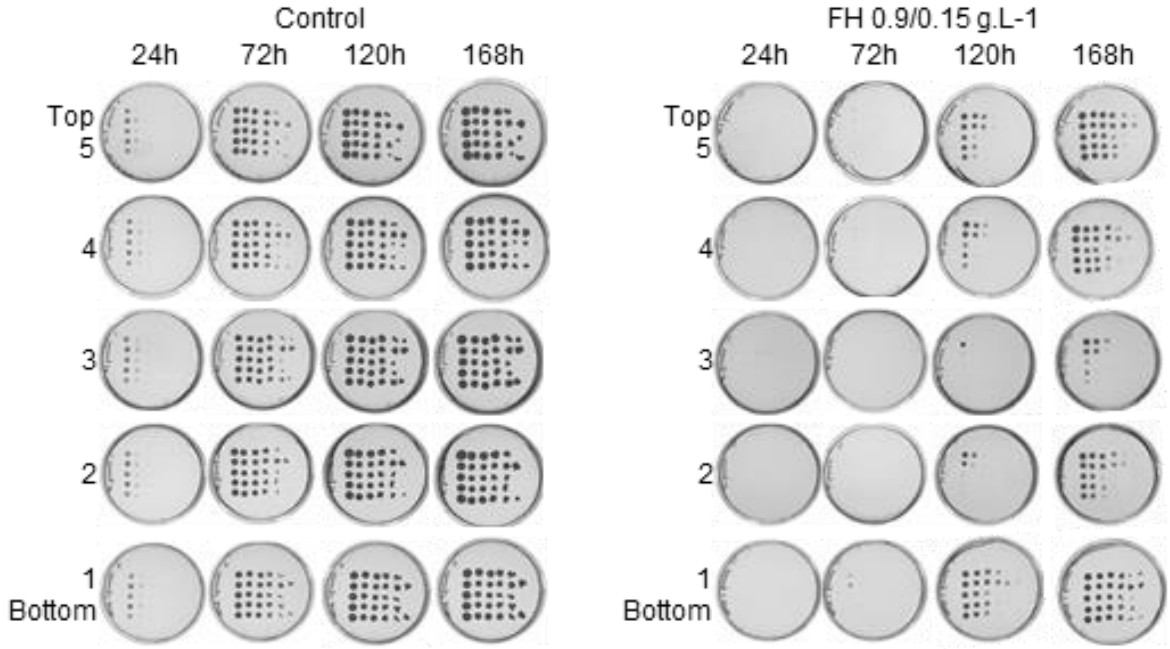

Figure 17 – Stacking experiment of *Komagataella* strains plated and stacked in the presence or absence of furaldehydes. Each plate contains the strains *K. phaffii X33*, *K. phaffii 7435*, *K. pastoris CBS 704*, *K. populi CBS 12362*, and *K. mondaviorum CBS 1517* plated from top to bottom. Strains were plated five times in either control (no inhibitor) or furaldehydes (FH 0.9/0.15$g.L^{1}$) and stacked (1 to 5 correspond to bottom to top plates). For each plate, rows represent the same inoculum in sequential 1:10 dilutions. The experiments performed in biological triplicate show good reproducibility (i.e., similar performance of the strains), and the figure illustrates the results of one replicate. Source: Author's collection.

The control plates had consistent and reproducible patterns for all the colonies at all of the positions of the stack. However, the results showed that depending on the stacking position, the same strains had different patterns of growth in media with furaldehydes (Fig. 17). No growth was visible after 24 h of incubation and after 72 h, only controls X33 and *C. lignohabitans* in the bottom plate had grown. The plate in the middle of the stack (#3) was more affected than the other four surrounding plates, with no visible growth until 72 h, only X33 without dilution at 120 h, and X33 and *C.*

lignohabitans growing at 1:10 dilution in the final timepoint. It is visible that the impairment is stronger from the middle of the stack outwards.

Contrary to the plates containing furaldehydes, the yeast in the control plate was not influenced by the stacking position, i.e., independent of the position, the yeast grows similarly in all cases (Fig. 17). This result may indicate that some factor is limiting growth in presence of furaldehyde in the “middle” plates. Since humidity and oxygenation should be homogeneous in the incubator, one possible hypothesis is the light degradation/oxidation of furaldehydes in the plates on the bottom and top of the stacking (158,159). Considering that the incubator had no internal light, and all plates would be exposed to similar luminosity daily during its scanning, it is unlikely that this is causing the observed effects.

Another explanation could be the temperature diffusion on stacking plates. Cell-growth troubleshooting manuals (160,161) suggest avoid stacking plates, for those in the bottom are closest to metal shelves and warm faster, and the middle ones would have more thermal stability. Mats Peterz in the 90s showed that the more plates there were in a stack, the longer was the time necessary to reach incubation temperature in the agar, and in piles from 3 to 9 plates, the middle one would always have a lower temperature than the ones in the edges (162). Higher temperatures could favor the conversion of furaldehydes since they are volatile (158,159). The underlying cause for growth inhibition when furaldehyde plates were stacked could not be determined in this work. Plate position in a stack is not an unprecedented factor in microorganism development. It has been shown for *Aspergillus* species that it formation of sexual structures varied according to stacking position (163).

Overall, the results for different *Komagataella* strains in the presence of inhibitors from lignocellulosic hydrolysate were not conclusive. Observations for acetic acid did not show any significant difference in growth between different *Komagataella* strains, but it should be considered that assays in solid media might not have enough sensibility, and not be ideal for this characterization. Furaldehyde resistance or sensibility was also not determined. However, inconsistencies throughout replicates lead to the investigation and finding that stacking plates could impact yeast growth patterns in the presence of furaldehydes. In future investigations, different assays should be used to better understand *Komagataella* growth in the presence of furaldehydes and different concentrations of acetic acid. Additionally, plate stacking should be avoided if possible, especially when working with furaldehydes.

7.2 XYLOSE METABOLISM BY *Komagataella* STRAINS

7.2.1 Spotting assays

Spotting assays with the 26 strains of *Komagataella* in plates containing YNB media supplemented with 2% xylose or glucose 2% were done in triplicates, monitored for 168 h, and pictures were taken every 48 h. Figure 18 shows the comparison between the growth of the 26 strains in glucose and xylose after 120 h of incubation. The strain *K. phaffii X33* was used as a reference in the first row of each plate.

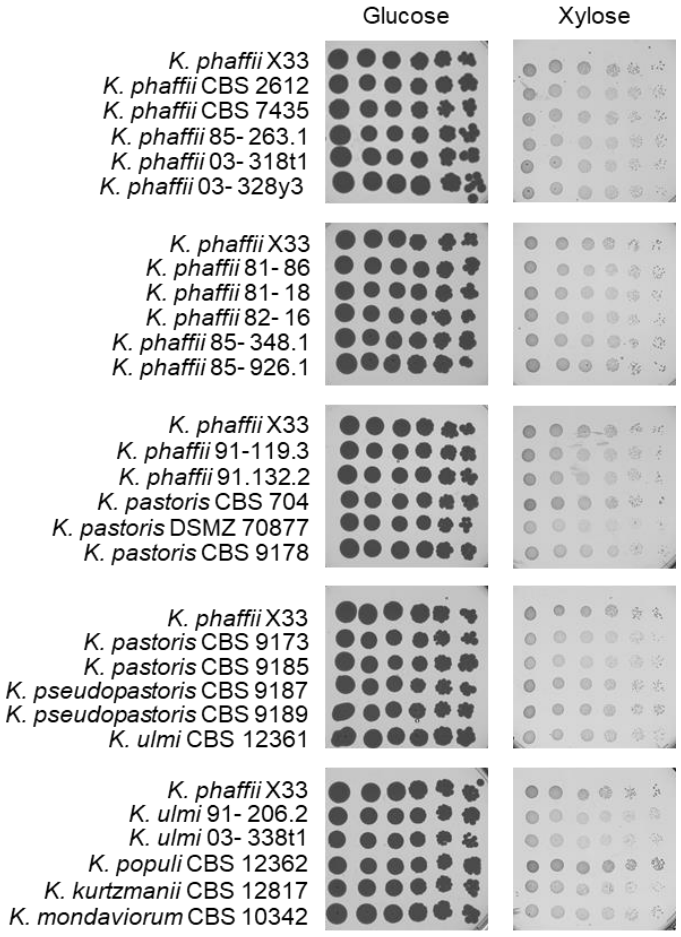

Figure 18 – Spotting assay of 26 *Komagataella* strains in YNB media with glucose and xylose after 120 h of incubation. The first line of each plate is the control *K. phaffii X33*. For each plate, lines represent the same inoculum in 10x serial dilutions. The experiments performed in biological triplicate show reproducibility (i.e., similar performance of the strains), and the figure illustrates the results of one replicate. Source: Author's collection.

All strains had similar patterns in glucose, being fully grown after 120 h, however, growth was significantly impaired when strains were grown on xylose (Fig.

18). Despite no strain having a remarkable growth compared to glucose, all *Komagataella* strains demonstrated the ability to grow on plates with xylose as the only carbon source. Since growth was very discrete, variations were also minimal, and the main parameter for defining the best growers was the final “density” of the spot by the end of the incubation. It was decided that three strains would be picked for future experiments. All *K. phaffii* strains had approximately the same pattern throughout the triplicate and the strain X33 was chosen as the commercial reference strain for further experiments. The other two chosen strains were natural isolates *K. pastoris CBS 704* and *K. populi CBS 12362* for having a final growth visually superior to other strains at the end of the assay in all replicates. To confirm this initial observation, and compare their growth side by side in the same plate, another spotting assay was done with the addition of an efficient xylose consumer strain *C. lignohabitans* (157). Selected strains were also plated alongside natural isolates *K. phaffii X33*, *K. phaffii* CBS2612, and *K. phaffii* CBS7435, and monitored for 168 h. (Fig. 19).

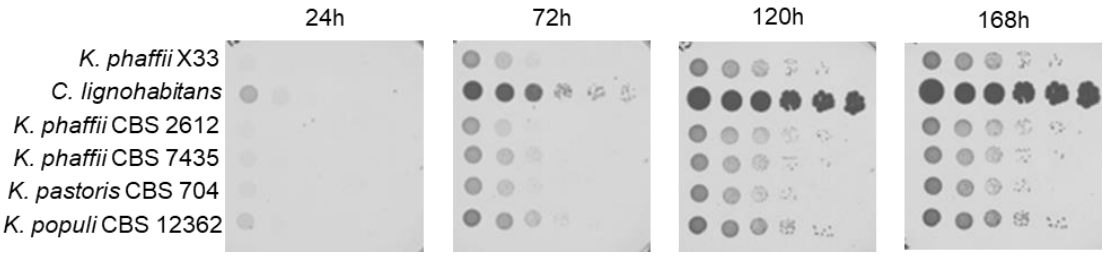

Figure 19 – Spotting assay of selected strains on xylose. From top to bottom, strains *K. phaffii X33*, *C. lignohabitans*, *K. phaffii CBS2612*, *K. phaffii CBS7435*, *K. pastoris,* and *K. populi* in different time points of growth. For each plate, lines represent the same inoculum in sequential 1:10 dilutions. Photos illustrate one from three biological replicates. Source: Author’s collection.

C. lignohabitans as expected showed growth faster than any other strain (Fig. 19). It showed high cell growth in 10^5 dilutions (last column) after 72 h, whereas all *Komagataella* strains, including the controls, have shown visible growth after 72 h of incubation in 10^2 dilutions (third column). Among *Komagataella* strains, and *K. populi* showed better growth, forming a colony in dilutions superior to $1x10^3$ (fourth column).

To this date, the literature indicates that *Komagataella* strains, specifically the old *P. pastoris* are not able to naturally utilize xylose as the sole carbon source (21,146). The construction of efficient xylose utilization by this yeast has been achieved however by the introduction of the xylose isomerase pathway (26). The growth visualized in the spotting assays is an indication of xylose metabolism in these yeasts.

Growth on plates might be due to some metabolic reserve from the cells. Further and detailed characterization of xylose metabolism on strains *K. phaffii X33*, *K. pastoris CBS 704*, and *K. populi CBS 12362* was done to support initial observations.

7.2.2 Characterization of selected *Komagataella* strains on xylose

The spotting assay results indicated that *Komagataella* growth on xylose was very slow, and maybe would be imperceptible if low cell density were used to start a culture. With that at sight, cultures for characterization of *Komagataella* strains growth in media with xylose 2% as the only carbon source were started with OD_{600} equals to 12 (circa 3.5 $g.L^{-1}$ cells) were done (Fig. 20).

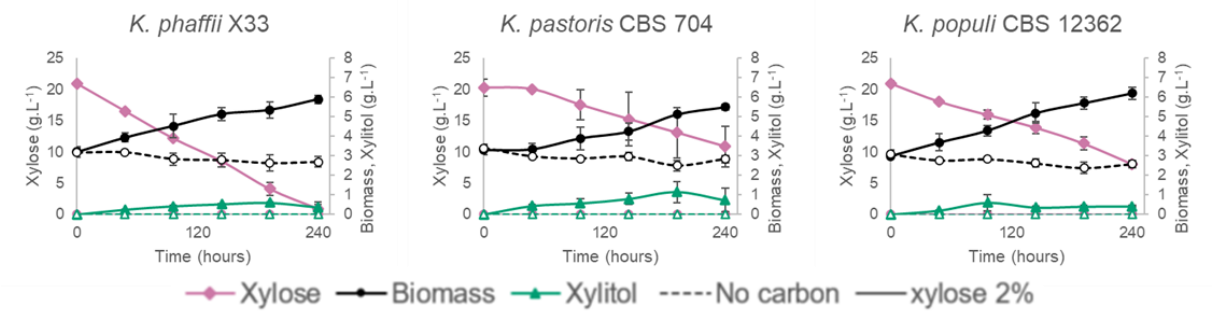

Figure 20 – Culture of strains *K. phaffii X33*, *K. pastoris*, and *K. populi* in YNB supplemented with 2% Xylose (continuous line) and YNB with no addition of carbon source (dotted line, -C). Data refers to triplicate average and standard deviation and are indicated in the picture. Source: Author's collection.

No cell growth or metabolite production was observed in the control conditions, ruling out that any visible growth or metabolite would come from any media component other than xylose. On average, *K. phaffii* X33 consumed 95.70% ± 3.2 of the available xylose after 10 days of incubation, nearly doubling (0.85 ± 0.08-fold) and produced a maximum of 0.60 ± 0.02 $g.L^{-1}$ of xylitol at 192h. *Komagataella pastoris* consumed 47.1%±13.2 of the available xylose and had the smallest growth in comparison to the other two (0.64±0.13 fold). The only strain that doubled the number of cells (1.04 ± 0.09-fold) was *K. populi,* consuming 61.4% of the available xylose and producing 0.42 ± 0.04 $g.L^{-1}$ of xylitol by the end of the fermentation. Xylitol production by the three strains is an indication that xylose is entering the oxi-reduction pathway. This is the xylose metabolic pathway most commonly found in yeasts (54). The liquid culture data supports the spotting assay observations indicating that *Komagataella* strains are

capable of consuming xylose. Growth is very slow and modest when compared to those reached in glucose, glycerol, or methanol(15), but the three *Komagataella* strains are using xylose from the media, growing and producing xylitol from it. Other metabolites like ethanol, acetic acid, and glycerol were not detected as fermentation products.

7.2.3 Confirmation of xylose metabolism by *K. phaffii* using labeled xylose.

Based on the results of the liquid culture assays, an isotope labeling experiment with 1-^{13}C xylose was designed to show the incorporation of xylose carbons into the cellular metabolites and yeast biomass. The *K. phaffii* X33 was chosen as an industrially relevant strain from the *K. phaffii* CBS 7435 background, and its connection to chapter 1. Precultures in YPG were grown overnight to obtain enough cells for the experiment. These were then cultured on YNB supplemented with 2% xylose for another day, to induce metabolism adaptation to xylose. Cultures were then grown in triplicates of YNB supplemented with either ^{12}C or 1-^{13}C xylose (2% w/v) for 10 days (Fig. 21).

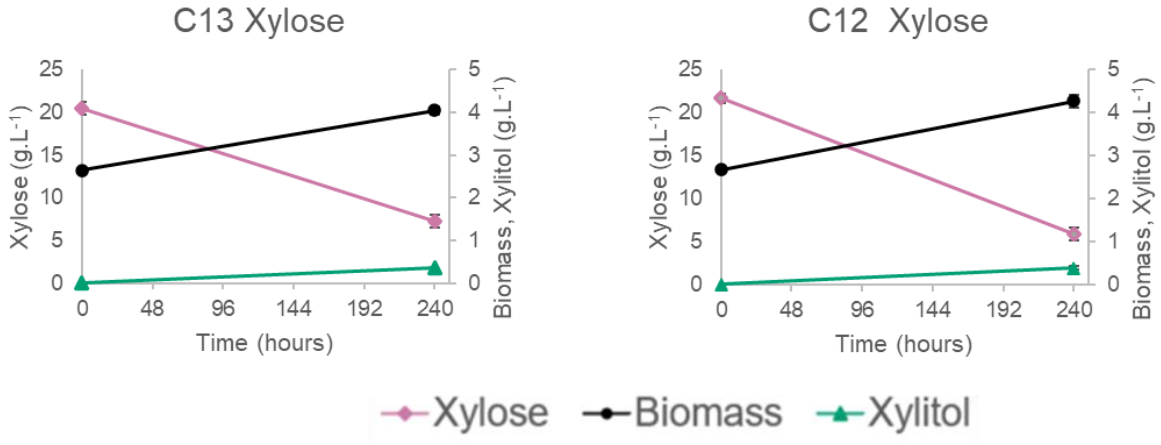

Figure 21 – *K. phaffii X33* cultivations in xylose isotopes ^{12}C and ^{13}C. Data refers to triplicate average and standard deviation and are indicated in the picture. Source: Author's collection.

The biological replicates and different isotope cultivations had similar growth patterns. After the ten days of cultivations, ^{13}C and ^{12}C consumed 66% and 70% of all available xylose, respectively. Cell growth and xylitol production were also very similar

for both conditions at 0.59 ± 0.03-fold and 0.37 ± 0.05 $g.L^{-1}$ for ^{12}C and 0.053 ± 0.06-fold and 0.38 ± 0.04 $g.L^{-1}$ for ^{13}C, respectively (Fig. 21).

The biomass ^{13}C content of the samples grown in the two types of xylose was measured by Elemental Analysis - Isotope Ratio Mass Spectrometry (EA-IRMS). In the non-labeled xylose, the ^{13}C represented an average of 1.1% of the total content, which is the exact estimated natural occurrence of this isotope (164). In the samples grown in medium with labeled xylose, an average total of 9% ^{13}C was found. Since the biomass increase in the cultivation was around 0.56-fold, this was slightly higher than the expected 5%.

The high percentage of ^{13}C found in yeast biomass compared to control samples confirms that the yeast is indeed incorporating the xylose from the media to its metabolism and using it to grow. The most common xylose metabolic pathway found in yeast is the oxi-reduction pathway, in which xylose is reduced to xylitol by xylose reductase, which is oxidized to xylulose by xylitol dehydrogenase. The latter is then phosphorylated to xylulose-5-phosphate by xylulokinase which enters the pentose phosphate pathway (54). The presence of xylitol as a product of fermentation is indicative that this is also the pathway found in *K. phaffii.* Further studies to better understand xylose metabolism in *K. phaffii* is necessary.

7.2.4 Adaptive laboratory evolution (ALE)

Although the capacity of *K. phaffii* strains to grow on xylose had been demonstrated, it was still very slow, and inefficient if compared to other carbon sources. To further improve growth and xylose utilization, the three selected strains *K. phaffii X33*, *K. pastoris CBS 704*, and *K. populi CBS 12362* were submitted to thirteen sequential transfers and more than 158 days of culture on minimal media supplemented with xylose. The comparison between original strains and the evolved population is shown in figure 22.

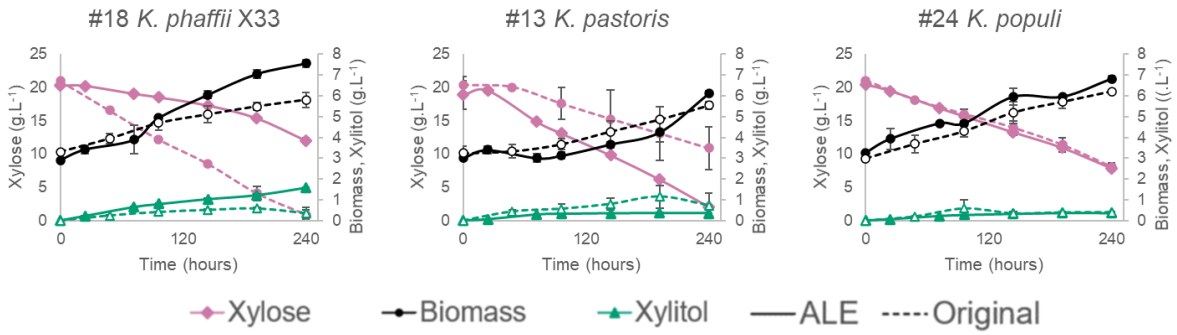

Figure 22 – Comparison of the general growth of strains before (dotted line) and after (continuous lines) undergoing adaptive laboratory evolution process. Data refers to triplicate average and standard deviation and are indicated in the picture. Source: Author's collection.

There were few differences between ALE strains performances on xylose. Adapted *K. phaffii X33* had an 85% improvement in final biomass generation almost the same amount of xylose consumed (88.6% ± 2.8) and produced xylitol (0.36 $g.L^{-1}$) after ten days. This could be the result of internal modulation and transcriptional changes for better metabolizing xylose. The gene-specific response of *K. phaffii* to different carbon sources (in this case, methanol, glycerol, or glucose) has been shown to happen mostly at its transcriptional level (165). Final glycerol production increased significantly in strain *K. pastoris CBS 704* from 0.26 ± 011 $g.L^{-1}$ to 1.59 ± 0.04 g. L^{-1}. The most unchanged strain was *K. populi CBS 12362* with all values remaining the same.

Because of the long time *Komagataella* strains need to grow in the presence of xylose, the number of generations under ALE selective pressure was limited and not enough to generate a major change in the xylose utilization profile. The strain that had the only significant improvement was *K. phaffii X33*.

8 CONCLUSIONS

P. pastoris is a highly studied organism, with great biotechnological potential (166), but most studies involving this yeast rely on two main strains and their derivates: *Komagataella pastoris* and *Komagataella phaffii* previously together under this name. (13). Biodiversity is one of the strategies that can be used to identify desirable industrial traits and expand the possibilities of use with this organism. Here, we present the characterization of 25 different natural and one commercial *Komagataella* isolates from six different species for LDI tolerance (acetic acid and furaldehyde), and xylose utilization.

Regarding the effect of the inhibitors, differences among the screened strains were not observed. In furaldehydes, however, a curious phenomenon was observed in which the stacking order of the plates had a direct effect on growth. The plates in the middle of the stack were the most affected by this. Our best hypothesis is that temperature changes between the stacked plates are involved in the volatilization of furaldehydes. The slower stabilization of temperature in the middle plates would be enough to maintain the cells exposed to furaldehydes' toxic effects for longer and having the biggest impairment in growth. Since we did not investigate it further, and currently do not have an explanation for it, we recommend avoiding stacking plates when working with furaldehydes. Close observation of other possible disturbances on microbial growth patterns because of substances added to media, or when stacking plates are also recommended.

The current understanding is that *P. pastoris* is not capable of using xylose, a major component of lignocellulosic hydrolysate, as its main carbon source (21,146), unless genetically engineered towards it (26). The characterization of the 26 strains on xylose revealed that actually, all screened strains were capable of slowly growing on xylose. Strains *K. phaffii* X33, *K. pastoris* CBS 704, and *K. populi* CBS 12362 were selected for further characterization experiments. This work shows for the that different strains of *Komagataella* are actually able to consume xylose as the only carbon source to grow and produce xylitol as a product. *Komagataella phaffii* X33 was able to consume up to 95.7% of the average 20 $g.L^{-1}$ of xylose available in the media, while *K. populi* CBS 12362 duplicated its biomass after ten days of incubation. All the strains produced xylitol as a fermentation product with *K. pastoris* reaching 1.59 $g.L^{-1}$, the

highest concentration among the three. Xylitol production is indicative that xylose undergoes an oxirreduction xylose catabolic pathway in these strains.

Isotope labeling experiments further confirmed that xylose was being metabolized by the yeast and incorporated into biomass since cells grown in ^{13}C-1 xylose had 9-times more ^{13}C content than the control. Preliminary metabolomic analysis shows the incorporation of the carbon isotope into the non-oxidative phase of the pentose phosphate pathway and glycolysis intermediates. Further research to unveil transporters, enzymes for xylose catabolism, and the specific transcriptional response of *Komagataella* strains to xylose remains to be investigated. To optimize *Komagataella* consumption and slow growth on xylose, the three selected strains were submitted to adaptative laboratory evolution (ALE) for 158 days. After adaptation, *K. phaffii* X33 had an 85% improvement in growth, despite the consumed xylose and xylitol production remaining the same. Other strain maintained their overall pattern from before adaptation. To our knowledge, this is the that *Komagataella* strains, specifically *K. phaffii* and *K. pastoris* are characterized for xylose consumption and metabolization. Metabolomic analysis of cells grown on either xylose or glucose is currently underway. Several further investigations regarding xylose incorporation and transporters, catabolic enzymes, transcriptional and translational responses will be necessary to fully comprehend *Komagataella*'s response to xylose as a carbon source.

CHAPTER 3 – OVEREXPRESSION OF THE GENE *HAA1* ENHANCES ACETIC ACID TOLERANCE IN *Pichia pastoris*

The use of lignocellulosic biomass hydrolysates requires microorganisms capable of metabolizing the sugars in presence of inhibitory compounds. Acetic acid can be considered one of the main inhibitors present in hydrolysates (59). Besides the possible utilization of lignocellulosic hydrolysate, tolerance to acetic acid is considered a desirable phenotype for microorganisms used in industrial biotechnology. The increase in overall tolerance to acetic acid could lead to further investigations towards building more robust strains with, for example, low pH tolerances, (avoids contamination by other organisms) and other biocatalysis products (20).

Acetic acid inhibits microbial growth due to the diffusion of uncoupled acid through the plasma membrane, causing cytoplasm acidification, metabolism inhibition, macromolecule degradation, and decreased synthesis rate and dissipation of membrane electrochemical gradient, as reviewed by (167). Genome-wide changes are required to overcome acetic acid toxic effects. Among these, cells respond by increasing saturated lipids from the plasmatic membrane and therefore reducing permeability (168). To compensate for cytoplasm acidification, plasma membrane ATPase removes excess protons, leading to ATP depletion and impacting cell growth (169).

The *S. cerevisiae* gene *HAA1* is a transcriptional factor involved in adaptation to weak acid stress (89,170,171). It was shown to be involved in circa 80% of acetic acid-responsive genes, and its deletion caused increased susceptibility of yeast to acetic acid, and its intracellular accumulation (170,172). The optimization either by metabolic engineering or adaptative laboratory evolution in this yeast was explored as a promising approach towards the development of strains for the industry with increased robustness against acetic acid (170,173). Overexpression of the *HAA1* gene in xylose-consuming strains improved yeast growth and sugar consumption abilities (174). In this chapter, the *S. cerevisiae HAA1* ortholog gene of *P. pastoris* was identified and overexpressed in this yeast. Then, the recombinant strain was physiologically characterized in the presence of acetic acid.

This part of the study was partially carried out in the context of my scholarship (Ernst Mach Worldwide 2019-2020), in collaboration with Dr. Diethard Mattanovich (head of the Institute of Microbiology and Microbial Biotechnology) at the University of Natural Resources and Life Sciences, Vienna, Austria.

8 MATERIALS AND METHODS

8.1 CELL STRAINS AND MAINTENANCE

8.1.1 *Escherichia coli*

Escherichia coli DH10B was used as a host for DNA manipulation. *E. coli* transformants were grown on LB (peptone 1.0%, yeast extract 0,5%, and NaCl 1% w/v, pH 7.2) supplemented with 50µg/ml geneticin (G418).

8.1.2 *Pichia pastoris X33* (Invitrogen)

The yeast *P. pastoris X33* (*Komagataella phaffii* - Invitrogen®) was used as the host for fermentations and DNA manipulation. Stock cultures of yeast were grown in YPD (yeast extract 1%, peptone 2%, glucose 2%) medium and were preserved in 30% glycerol and maintained at -80ºC. *P. pastoris* transformants were grown in selective media supplemented with geneticin (G418) 500µg/mL.

Whole-genome from *P. pastoris CBS 7435* was used as the template for *HAA1* amplification.

8.2 CULTURE MEDIA AND SOLUTIONS

All culture media utilized were diluted in distilled water and sterilized by humid heat at 120ºC for 15 min in the autoclave, except when specifically described.

8.2.1 PBS 10X (Stock)

KH_2PO_4 0.24%, Na_2HPO_4 *$2H_2O$ 1,8%, KCl 0.2%, NaCl 8.0%. The pH was adjusted to 7.4 and it was stored at room temperature. For use, the stock was diluted to 1X with sterile distilled water.

8.2.2 TAE Buffer 50x (Stock)

Tris Base 24.2%, glacial acetic acid 5.7%, 0.5 M EDTA pH 8 10%. The stock solution was stored at 4ºC. From stock, the buffer was diluted to 1X (concentration of use).

8.2.3 FB buffer (sterile filtrated)

KCl 0.7%, $CaCl_2$.2H2O 0.8%, 10% glycerol, 1.0% K-Acetate 1M pH 7.5. The buffer was sterile filtrated and stored at room temperature.

8.2.4 HEPES 1M pH 8,0 (Stock)

HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid) 23.8%. The pH was adjusted with NaOH to 8.0 and the solution was filter sterilized and stored at 4ºC.

8.2.5 DTT 1M (Stock)

DTT (DL-dithiothreitol, anhydrous) 15.4%. The solution was filter sterilized, kept in the dark (covered with aluminum foil), and at -20ºC.

8.2.6 Sorbitol 1M (Stock)

Sorbitol 18.2% was diluted in distilled water and stored at room temperature.

8.2.7 0,5 M EDTA

EDTA (disodium dihydrate) 18.6%, NaOH was added to adjust pH to 8.0.

8.2.8 YNB

YNB (yeast nitrogen base, Sigma Aldrich Y0626) without amino acids (0.68% YNB, 2% ammonium sulfate), 4.10^{-5}% biotin, buffered to pH 5.5 with phthalate buffer (1.02% potassium hydrogen phthalate with 0.22% potassium hydroxide w/v).

8.3 *HAA1* IDENTIFICATION

The putative *P. pastoris HAA1* identification was done by protein alignment of *S. cerevisiae HAA1* / YPR008W sequence to *P. pastoris CBS 7435.* The sequence with the higher similarity (53.61%) refers to the gene PP7435_Chr3-0991, (1,572 bp).

8.4 GOLDEN GATE ASSEMBLY (GGA)

Golden Gate Assembly reaction was performed according to (175) using NEB® Golden Gate Assembly Kit.

8.5 PCR – POLYMERASE CHAIN REACTION

For colony PCRs with yeast cells, a small amount picked with a 10 µl tip from isolated colonies selected from transformation were resuspended at 10µl 0,02 M NaOH (0.08% NaOH in distilled water) in a PCR reaction tube and boiled for 10 min at 99ºC. A total of 2 µl of the boiled content was used as the template for PCR. For bacteria, cells were picked from the colony and dissolved directly in the PCR mix. Both procedures were done using OneTaq® Quick-Load® DNA Polymerase from New Englands Biolabs, following the manufacturer´s protocol.

For average PCR, Q5® High-Fidelity DNA Polymerase from New Englands Biolabs, following the manufacturer´s protocol.

8.6 PRIMERS UTILIZED IN CHAPTER 3

The primers used are described in Table 5.

Table 5 – List of primers used in Chapter 3. Source: Author's collection.

Primer name	Sequence	Description
HAA1 A-FWD	Cg**GGTCTC**c C**ATG** GTGCTTATAAACGGCGT	Amplification of fragment A from *HAA1*
HAA1 A-REV	Cg**GGTCTC**C CtGT CTCCCCTTAGGCTTTA	
HAA1 B-FWD	Cg**GGTCTC**c ACaG TCTAGCCAATGTGCA	Amplification of fragment B from *HAA1*

HAA1 B-REV	Cg**GGTCTC**CCcTCTTACCGCAAGTACACTTAGC	
HAA1 C-FWD	Cg**GGTCTC**cGAgGACAAACACAAAGATGAAATA	Amplification of fragment C from *HAA1*
HAA1 C-REV	Cg**GGTCTC**CGaTCTCTTCGTTATCGCCCG	
HAA1 D-FWD	Cg**GGTCTC**cGAtCCCAAACACCTATCAGA	Amplification of fragment D from *HAA1*
HAA1 D-REV	Cg**GGTCTC**CAAGCTTAACGGTTCAATGGGAC	
J77	GAGGTATGTAGGCGGTGCTA	Confirmation of cassette assemble
J78	TTATATTTCTCTACAGGGGC	
B2_62	GTAAAACGACGGCCAGTT	Sequencing
B2_63	CAGGAAACAGCTATGAC	

Bold characters refer to *BsaI* restriction site. Underscore characters refer to start codon.

8.7 PLASMIDS ARE UTILIZED IN CHAPTER 3

The plasmids used are described in Table 6.

Table 6 – List of plasmids used in Chapter 3. Source: Author's collection.

Plasmid name	Elements	Selection Marker	Fusion site	Description
BB1_23	Empty vector	Kan/G418	2 and 3	Assemble of *HAA1*
BB1_12*GAP*	p*GAP*	Kan/G418	1 and 2	p*GAP* origin plasmid
BB1_12_p *RPP1b*	p *RPP1b*	Kan/G418	1 and 2	p *RPP1b* origin plasmid
BB1_34_RPS2tt	RPS2tt	Kan/G418	3 and 4	RPS2tt origin plasmid
BB3_14_empty	Empty vector	Kan/G418	1 and 4	Final backbone
p*GAP_HAA1*_ttRPS2	Final plasmid	Kan/G418	-	*HAA1* gene under the promotion of p*GAP* and terminator RPS2
p *RPP1b_HAA1*_ttRPS2	Final plasmid	Kan/G418	-	*HAA1* gene under the promotion of P *RPP1b* and terminator RPS2

8.8 CONSTRUCTION OF *HAA1* EXPRESSION VECTORS

Vectors containing the *HAA1* gene of *P. pastoris* were constructed using the GoldenPiCS modular cloning system (175) using golden gate assembly. The *HAA1* gene was cloned under the regulations of the constitutive promoters p*GAP* (glyceraldehyde-3-phosphate dehydrogenase promoter) or p *RPP1b* (ribosomal protein P1 beta) *RPP1B*. Both vectors contain the terminator RPS2tt. The integration site is *AOX1tt* and geneticin (G418) selective marker. The complete *HAA1* sequence can be found in Appendix B1.

8.9 AGAROSE GEL ELECTROPHORESIS

Integrity verifications and primary quantification of DNA were made by 1% agarose gels containing ethidium bromide 0.5 µg/ml. Samples were analyzed into Thermo Scientific electrophoresis with TAE 1X buffer and the results were visualized and photographed in the photo documentation system BioRad gel doc XR+. The DNA Ladder used as size and mass reference was the Quick-Load Purple 1kb DNA Ladder from New England BioLabs®.

8.10 CHEMICALLY COMPETENT *E. coli* CELL PREPARATION AND TRANSFORMATION.

E. coli DH10B colonies were inoculated in 5 mL LB media in 50 mL tubes and incubated overnight at 37ºC 180 rpm in a rotational shaker. The overnight culture was used to inoculate the main culture of 200 mL of LB medium to a final OD_{600} of 0.1m and incubated at the same conditions for approximately 2-3 hours until it reaches the final OD_{600} of 0.5. The culture was then transferred to four sterile 50 mL centrifuge tubes and left on ice for 10 minutes. Cells were centrifuged at 4ºC for 15 min and 1800x*g*. The supernatant was discarded, and cells were resuspended in 12.5 mL of prechilled FB buffer and pulled together in one single 50 mL tube. The cells were then left on ice for 40 min and centrifuged at 2,200 rpm for 15 min at 4ºC. The supernatant was once again removed and resuspended in a 15 mL FB buffer. In a pre-chilled 1.5 mL microcentrifuge tube, 200 µL aliquots were pipetted and stored at -80ºC.

For transformation, cells were thawed on ice and for each 50 µL of competent cells, up to 10 µL of DNA ware inserted. The solution was mixed by gently tapping and left on ice for thirty minutes. The mixture was heat-shocked in a 42ºC incubator for exactly thirty seconds and put back on the ice for five minutes. To the cells, 1 mL LB media without antibiotic was added, and it was incubated at 37ºC for between thirty minutes and one hour. Cells were then plated in selective media and incubated at 37ºC overnight.

8.11 ELECTRO-COMPETENT *P. pastoris* CELLS AND TRANSFORMATION

Pichia pastoris X33 colonies were inoculated in 5 mL YPD media in 50 mL tubes and incubated overnight at 30ºC and 180 rpm in a rotational shaker. The OD_{600} was measured and diluted to OD_{600} 0.1 in 200 mL of YPD into 2 mL flasks. When the OD_{600} reached between 1.0 and 1.5 (which should take between 4 to 5 hours) cells were centrifuged in four 50 mL flasks at 1500 x*g* for five minutes at 4ºC. The supernatant was discarded, and cells were resuspended in 100 mL of YPD supplemented with 20 mL of HEPES 1M pH 8.0 and 2.5 mL of DTT 1 M. Cells were incubated at 30º for 15 minutes and centrifuged for five minutes at 4ºC and 1500x*g.* After discarding the supernatant, cells were resuspended in 250 mL of ice-cold distilled water and centrifuged for five minutes at 1500x*g* at 4ºC. Cells were resuspended in 50 mL of ice-cold distilled water and once again centrifuged in the same conditions. Then, cells were resuspended in 10 mL of sorbitol 1M and centrifuged in the same conditions. The supernatant was discarded, and cells were then resuspended in 500 µL sorbitol 1 M and 60 µL of it were aliquoted in pre-chilled 1.5 mL microcentrifuge tubes and were stored at -80ºC.

For the transformation, cells of electro-competent *P. pastoris* were thawed on ice and then mixed gently with up to 25 µl of purified linearized DNA. As a negative control, the same average volume of water was added. The mixture was transferred into a chilled 2 mm electroporation cuvette and incubated on ice for ten minutes. The electroporation was done in the BioRad Gene Pulser™ using 2 kV voltage and 4ms time (25 µF and 200 Ω). Immediately after the shock, 1 mL of YPD media was added to the cuvettes, and its content was carefully transferred into a sterile 1.5 mL

microcentrifuge tube. Cells were regenerated for at least 1.5 hours at 30ºC and then plated into selective plates. Leftover cells were kept at 4ºC while plates were incubated at 30ºC for a minimum of 48 h.

8.12 SCREENING FOR ACETIC ACID-TOLERANT CLONES

The *P. pastoris X33* strains were inoculated in 24 deep-well plates containing 2 mL YPD (yeast extract 1%, peptone 2%, glucose 2%) and incubated at 25ºC and 280 rpm overnight. The following day, cells were washed twice with PBS 1X and resuspended in 2.5 mL sterile distilled water. Cells were transferred to a fresh 24-deep well plate to an initial OD_{600} of 0.3 in 2 mL. of assay media and incubated at 25ºC and 280 rpm for 48 h. Assay media was composed of YNB supplemented with 2% xylose, with either 2g.L-^{1} acetic acid, or no acid addition (control). Growth readings were performed by optical density at 600nm (OD_{600}) at Tecan Infinite 200 PRO (Life Sciences) with 200µL of the cultures harvested and diluted to densities within the equipment detection range with MilliQ water into 96-well ELISA plates.

8.13 SPOTTING ASSAY

The spotting assay was performed as described in Chapter two, section 6.3.

8.14 SEQUENCING

Chosen assembled plasmid was sequenced using primers B2_62 and B2_63 to confirm correct assembly by Microsynth AG (Switzerland).

8.15 INTRACELLULAR ACETIC ACID CONTENT ASSAY

The experiment was based on (89). The strains of *P. pastoris X33* were pre-inoculated in 5 mL YPD and grown overnight at 25ºC 180 rpm. On the next day, OD_{600} was measured and the necessary amount for an initial OD_{600} of 2 in 50 ml. Cells were then centrifuged and washed twice in PBS 1X. Cells were resuspended into the respective assay media. The OD_{600} was measured, and it was cultivated for two hours at 25ºC and 180 rpm. After two hours, OD_{600} was measured again, cells were

centrifuged ad 2,000 x*g* for 5 minutes at 4ºC, the supernatant was discarded. Cells were washed twice with ice-cold distilled water and re-suspended in 1 mL MilliQ water. Cells were transferred to 1.8mL cryo-tubes, with approximately 50 µl of glass beads 0.1mm. Cells were broken with one cycle of 40s at 5.5m/s at a FastPrep-24^{TM} bead beating grinder and lysis system. Cells were then centrifuged at 9000x*g* for five minutes. The supernatant was transferred to a fresh tube and boiled at 99ºC for 10 minutes. The content was centrifuged once again at 9000x*g* for five minutes and the supernatant was transferred to a fresh tube. The material was analyzed by HPLC. The experiment was carried out in biological triplicates.

Average YPD media with either: 2 g.L-$^{-1}$ or 5 g.L-$^{-1}$ acetic acid and final pH of 6.8 and 4.8, respectively. The control condition was composed of YPD media with no addition of acetic acid, and pH was adjusted with HCl. A one-way ANOVA statistical test was done to validate the significance of the values.

9 RESULTS AND DISCUSSION

9.1 *HAA1* OVEREXPRESSION

In *S. cerevisiae,* the gene *HAA1* has been proven to be an important transcriptional activator for response to acid stress. The transcription encodes a 694 aminoacid protein in the nucleus that regulates the transcription of several genes, many of which encode membrane proteins (170,171,173). Thus, aiming towards the creation of a strain of *P. pastoris* more resistant to acetic acid toxic effects, a putative ortholog of *S. cerevisiae*'s *HAA1* gene was identified and overexpressed in *P. pastoris.* For this, the aminoacids sequence of *S. cerevisiae* gene *HAA1* was aligned against the genome of *P. pastoris (K. phaffii CBS 7435).* The PP7435_Chr3-0991 gene had the best identity with 53.61% of similarity. The gene has since then been annotated as *HAA1* for the CBS 7435 strain and described as a transcriptional activator involved in response to weak acids, *TPO2, YRO2* transcriptional activation, and other genes encoding membrane stress proteins. *S. cerevisiae* contains a *HAA1* paralog gene; *CUP2,* identified as a Copper-binding transcription factor (176) that surged from the whole genome duplication; and relocalized from cytoplasm to nucleus. (177). The *HAA1* paralog gene *CUP2* was not identified in *P. pastoris.* In our transcriptome analysis (chapter 1) *p*-values of the analysis were above the defined threshold of 0.05, therefore *HAA1* expression levels could not be reliably analyzed. Since transcriptional factors have a fine regulation profile, it is possible that differences on expression levels are harder to be observed and more prone to error.

The *HAA1* gene was amplified from *P. pastoris CBS 7435* genome (Supplementary B2 – Complete *HAA1* sequence) from *P. pastoris CBS 7435* total genome. In four fragments (primer pairs *HAA1*-A, B, C, and D, Table 5, Fig. 23-1) that were later assembled into a backbone plasmid BB1_23 (165, Table 6) by golden gate assembly (GGA, Fig. 23-2). The fragmented amplification was necessary to remove the naturally occurring *BpiI* restriction site in the gene that could prevent the correct assemble in GGA reactions. For that, primer pairs A, B C, and D (Table 5) were designed with bases that would result in the formation of a *BpiI* restriction site being substituted for the coding of the same aminoacids as originally coded, with similar frequency, but without the formation of this restriction site (Appendix B3). At the flanking regions of the assembled *HAA1* gene, fusion sites [2] and [3] were added at

the beginning and end of the gene, respectively. These fusion sites are complementary to the end of the promoters ([2)] and the start of the terminator ([3]) chosen, and through GGA are constructed in the correct order into an empty vector (BB3_14_empty) with a selection marker for Kan/G418 and integration locus at the AOX gene (Fig. 23-3).

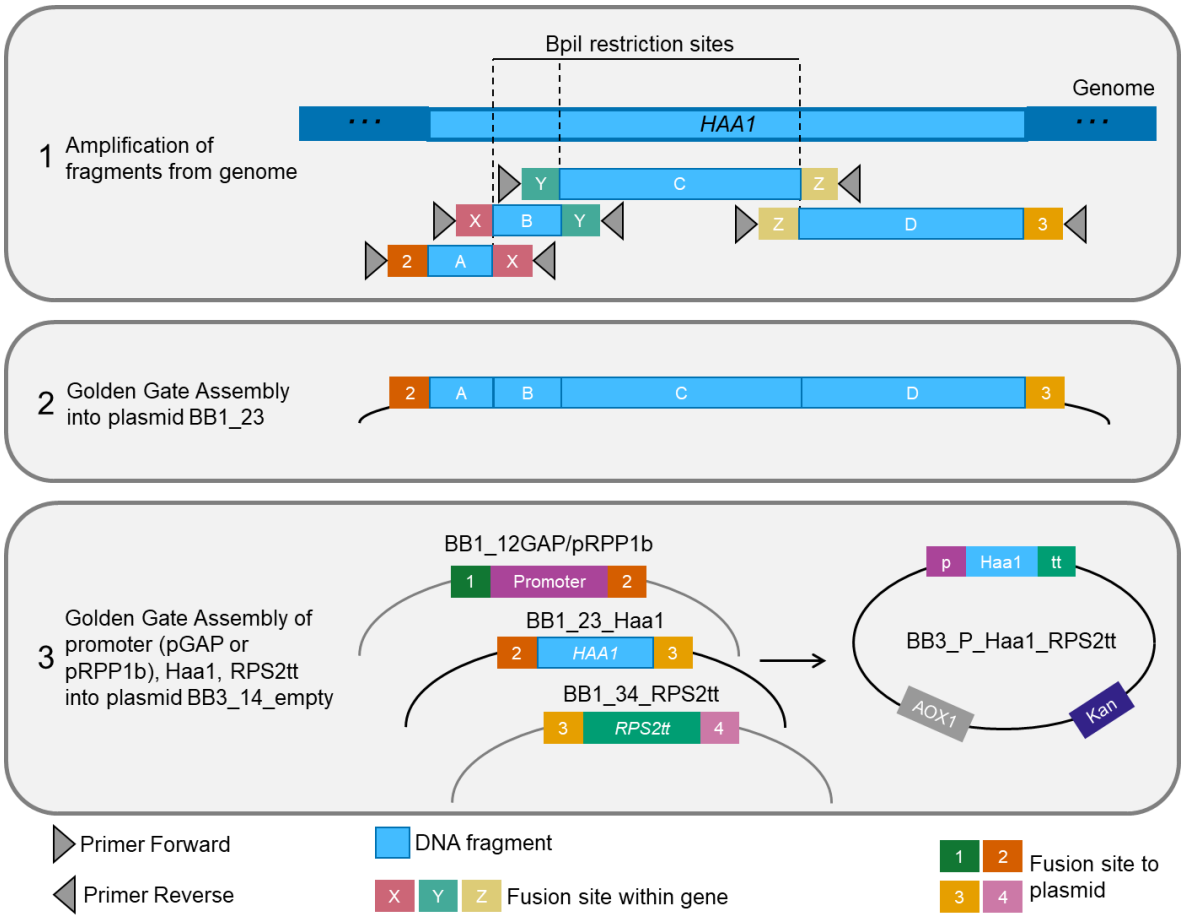

Figure 23 – Cloning strategy for *HAA1* overexpression. Source: Author's collection.

Plasmids containing the chosen promoters pP*GAP* and p *RPP1b*, and terminator p *RPP1b* were previously constructed in the correct wanted positions and were available at the Mattanovich-Gasser collection. Two different promoters were chosen to analyze the effects of the gene under different expression levels. The promoter p *RPP1b* has approximately 23% of the activity observed for the p*GAP* promoter in glucose and glycerol (175). Constructs were transformed in *E. coli* for plasmid multiplication and confirmed by colony PCR (Fig. 24).

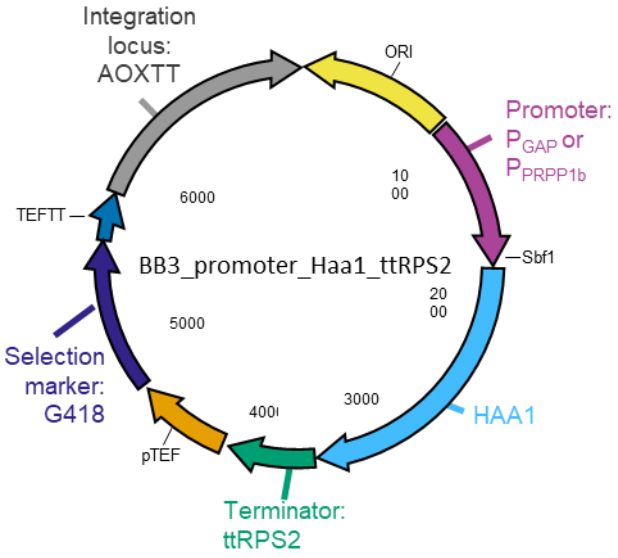

Figure 24 – Overview of plasmids for HAA1 expression in *P. pastoris* assembled. Plasmids were built using GoldenPiCS backbone 3 (BB3 2931bp) either p*GAP* (493bp) or p *RPP1b* (994bp) promoters, HAA1 (1567bp) and ttRPS2 (472bp). Source: Author's collection.

Cells were grown in liquid media with geneticin, plasmid DNA was extracted with Hi Yield® MINI Plasmid DNA Extraction Kit (Süd-Laborbedarf, Germany) and transformed into *P. pastoris X33*. Colonies in transformation plates were numbered and transferred to a fresh YPD+G418 plate. After growth, a colony was collected and prepared to be used as a PCR template for the determination of the correct integration of the *HAA1* overexpression cassette into the AOX gene using primers J77/J78 (Fig. 24, Table 5). Expected sizes for the PCR amplifications were: 572bp for empty plasmid, 3082bp for plasmid with p*GAP* promoter (p*GAP_HAA1*_ttRPS2, Table 6), and 3589bp for the plasmid with p *RPP1b* promoter (p *RPP1b_HAA1*_ttRPS2). An empty plasmid was used as the positive control (C+) and distilled water as the negative control (out of frame).

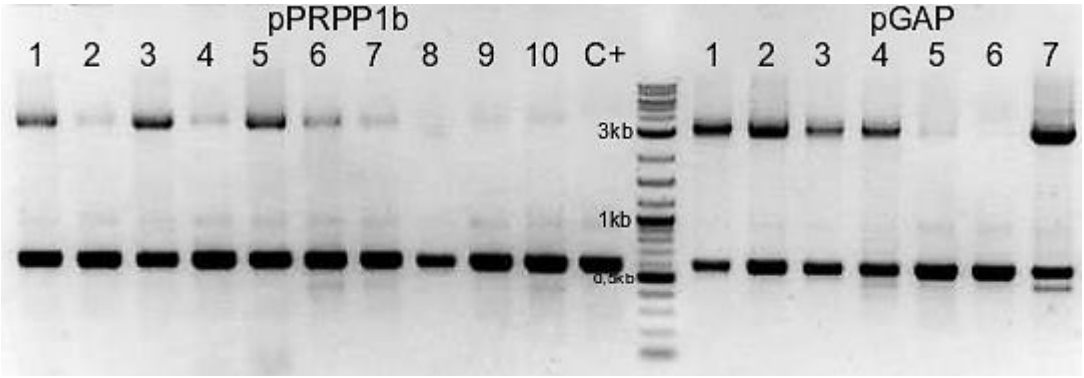

Figure 25 – Confirmation of the constructed systems for HAA1 overexpression. Electrophoretic analysis in 1% agarose gel. Molecular mass marker Quick-Load® Purple 1 kb Plus DNA Ladder (New England Biotechnology). PCR products: empty system C+ : 572bp, p*GAP_HAA1*_RPS2tt 3082bp, p *RPP1b_HAA1*_RPS2tt : 3589bp. Source: Author's collection.

A strong band around 600 bp that relates to the amplification of the empty plasmid was observed in all clones, while bands with the expected sizes were observed for both transformation cassettes. The presence of a band referring to an empty plasmid and bands of the correct expected genes in the same column suggests the amplification of promoter and terminator ends without the presence of the genes. PCR was repeated with different mixes, but the same result was obtained. Despite samples coming from isolated colonies and being plated in selective media, there is a possibility that some cells do not contain the *HAA1* gene. Colonies #1, #3, #5, #6, #7 for p *RPP1b-HAA1* and #1, #2, #3, #4, #6 for p*GAP-HAA1* were chosen and stored for future experiments.

9.2 SPOTTING ASSAYS

To functionally screen for increased tolerance to acetic acid, a spotting assay of selected strains was performed. Cells were plated on YPD pH 4.5 adjusted with HCl or YPD+acetic acid 2 $g.L^{-1}$ pH 4.5 and incubated for 72 h at 30◦C. This experiment was performed in biological triplicates (Fig. 26).

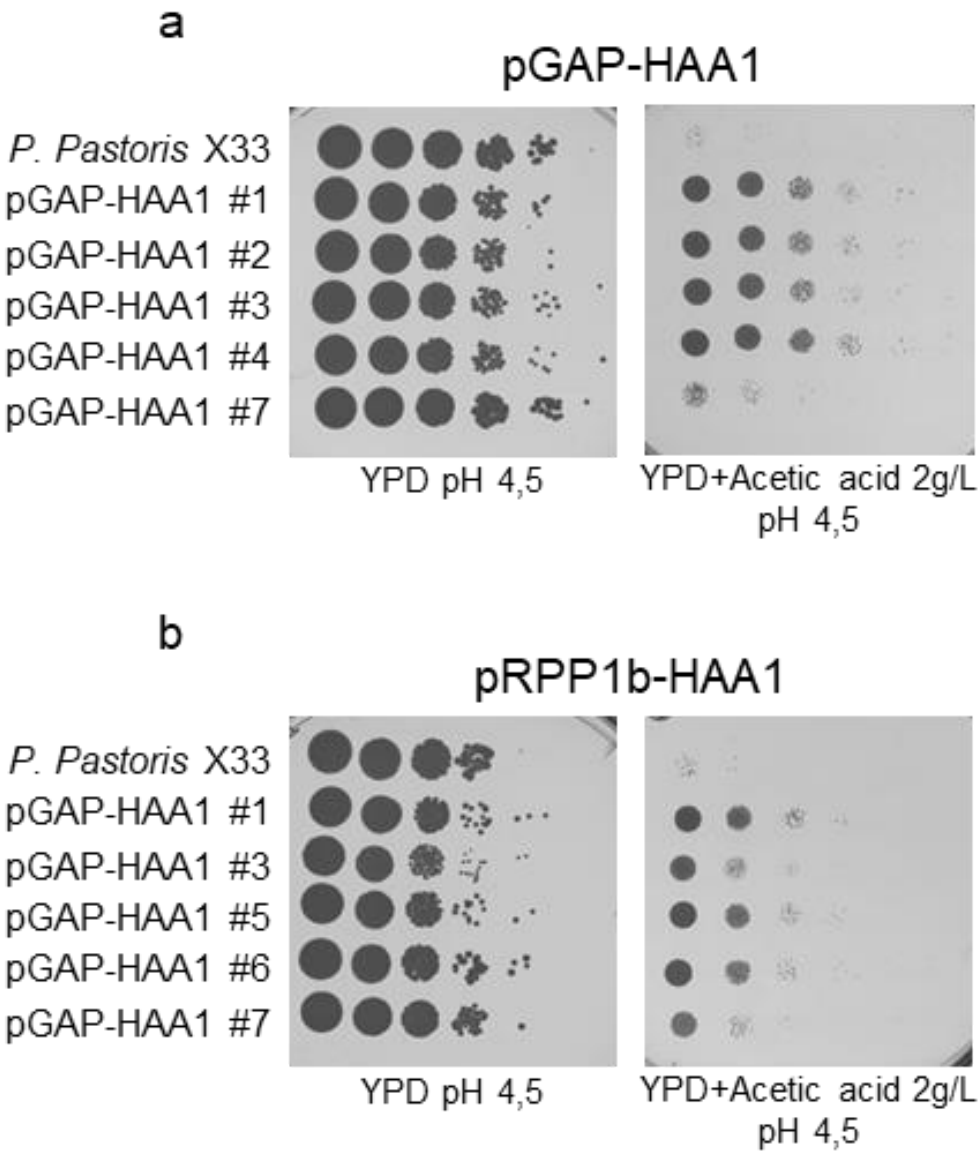

Figure – Spotting assay of *P. pastoris X33* clones overexpressing HAA1 under the promoter p*GAP* (A.) or p *RPP1b* (B.). Growth in YPD with or without acetic acid 2 $g.L^{-1}$ and pH 4.5 Normal X33 strain without the overexpression of *HAA1* was used as control. Photos of 48 h timepoint. Source: Author's collection.

Spotting assay shows that, compared to the control strain, most transformants have a better growth when acetic acid is added to the media. No relevant difference was observed in control media without acid. Clones with *HAA1* under the expression of p*GAP* have grown even better than p *RPP1b*, and both were better than the control strain. Curiously, clones #7 in both systems had near no growth compared to other strains in presence of acetic acid 2 $g.L^{-1}$, even though they grew normally in plates without acetic acid. When picked, it was observed that these colonies were smaller than the others. Since the presence of the gene was confirmed by PCR in both

colonies, the lack of growth could indicate mixed colonies or some problem in integration that was not identified by gel electrophoresis analysis. The colonies p*GAP-HAA1* #4 and p *RPP1b-HAA1* #5 were chosen as references for future experiments.

9.3 GROWTH ASSAYS IN LIQUID MEDIA

To confirm the apparent improvement conferred by the overexpression of *HAA1* observed in the spotting assays, cultivations of selected strains in liquid media were done. Colonies were grown in YNB media supplemented with 2 $g.L^{-1}$ of acetic acid was done in technical and biological replicates in 96-deep-well plates. Media YNB was used as a control condition and both conditions had pH 4,5 adjusted with HCl. The strain *P. pastoris X33* was used as a control strain.

The three strains grew identically in the control condition, with exponential phase starting after nine hours with an average of 42.4% of growth observed in their respective controls and reaching final growth after 24 h (Fig. 27). The strains overexpressing *HAA1* under p*GAP* and p *RPP1b* behaved alike strains in media without acetic acid, with growth starting after 9h. The only strain that showed an extended lag phase was X33, with 27.66% of its original growth at 9h, and 13% at 24 h, after which it started its exponential phase. The control X33 was also the last one to reach final growth above OD_{600} 20 and was, therefore, the most negatively affected by acetic acid. Strains overexpressing *HAA1* by the medium force promoter p *RPP1b* had intermediary effects between control and p*GAP* promoter strains. Under strong p*GAP* promotion, strains although significantly affected had the best growth of the three. Differences between strain growth in control and media containing acetic acid (AA) were most evident after 24 h, but strains reached final OD_{600} above 20 after 48 h of incubation.

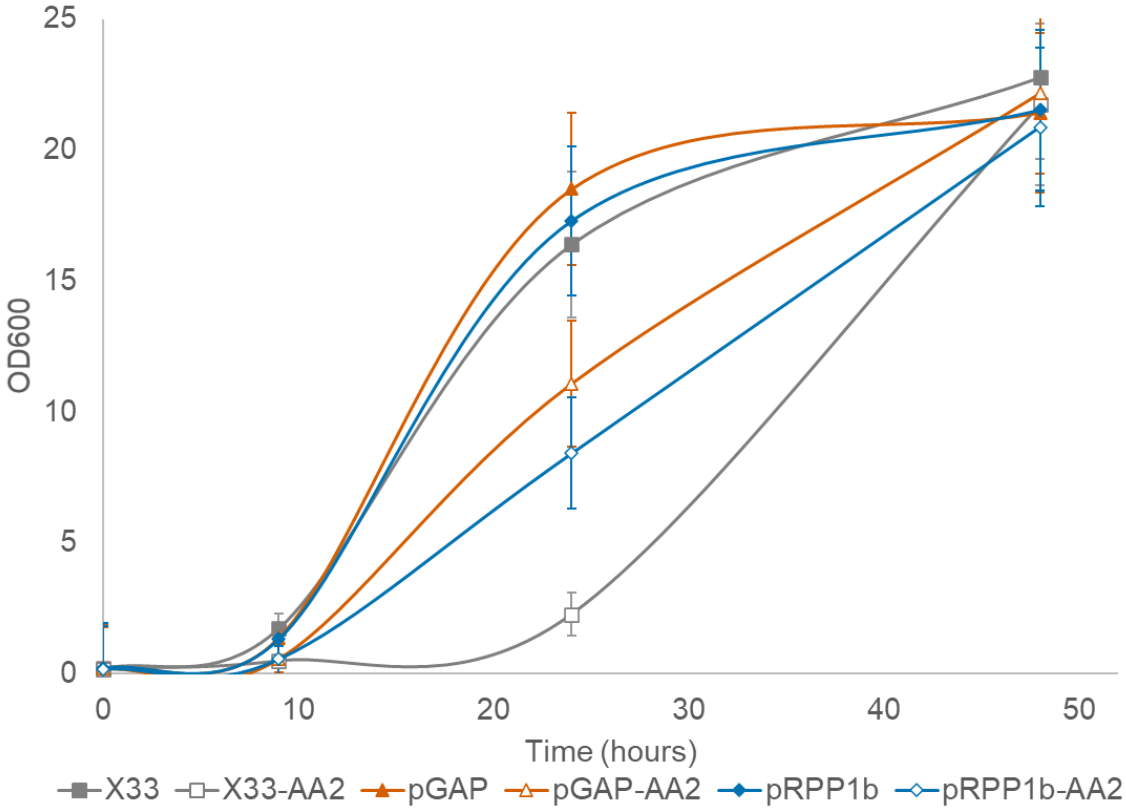

Figure 26 – Growth curves of strains overexpressing HAA1 under the promoters *pGAP (red triangle)* or p *RPP1b (blue diamond) in YNB media with or without supplementation of 2 g.L^{-1} of acetic acid (AA2)*. *P. pastoris X33* (gray square) was used as positive control strain. Cultures were made in 24-deep-well plates with 2 ml of YNB, and 200 µL of the culture was used in each timepoint for cell density determination. Data refers to triplicate average and standard deviation and are indicated in the picture. Source: Author's collection.

The tolerance to acetic acid observed in *P. pastoris* correlated with the promoter strength. This result is similar to reported for *S. cerevisiae* when *HAA1* native promoter was substituted for *TDH3,* the glyceraldehyde-3-phosphate dehydrogenase promoter, ortholog for p*GAP* in *P. pastoris* (178).

9.4 INTRACELLULAR ACETIC ACID ASSAY

The gene *HAA1* overexpression in *S. cerevisiae* has been correlated with the decrease in intracellular amounts of acetic acid (89). Following a protocol similar to the original experiment in *S. cerevisiae*, *P. pastoris* transformants overexpressing *HAA1* under the two different promoters were cultivated for two hours at 25ºC and 180 rpm

in YPD media containing either 2 g.L-$^{-1}$ or 5 g.L-$^{-1}$ acetic acid and final pH of 6.8 and 4.8, respectively. Control media pH was adjusted to the same values with HCl.

Intracellular contents of strains cultivated in media without acetic acid had similar concentrations of acetic acid close to 0.19 g/g in both pH 6.8 and 4.8 (Fig. 28). When added to the growth media, acetic acid concentration in the media can be directly correlated with its accumulation intracellularly. At 2 g.L$^{-1,}$ the average amount of acetic acid found inside the cells was 0.3 g.L^{-1}, and all three strains had statistically the same quantity. At 5 g.L^{-1} acetic acid in the media, it was 0.72 g.L^{-1} intracellularly, control strain was the one with the highest amounts of accumulated acid inside, while p *RPP1b-HAA1* and p*GAP-HAA1* amounts were statistically identical.

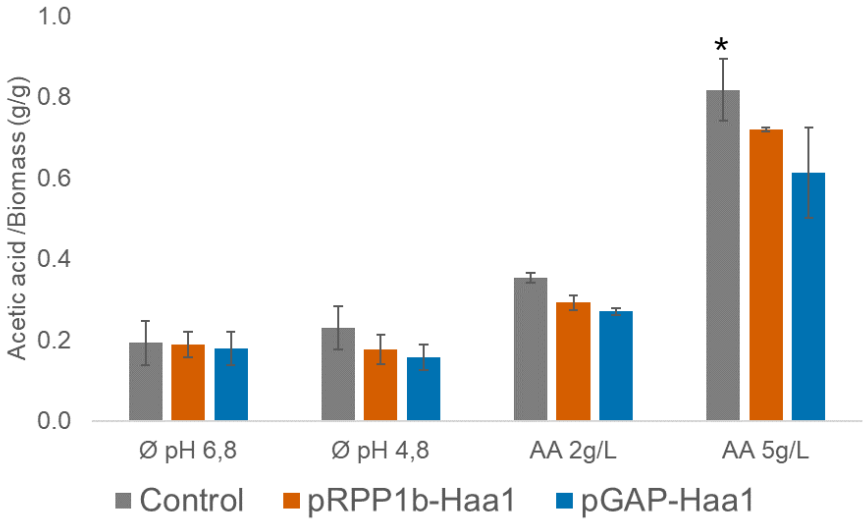

Figure 27 – Intracellular amounts of acetic acid on X33 (gray-left), p *RPP1b*-HAA1(orange-middle), and p*GAP*-HAA1(blue-right) strains. Strains were grown in media without the addition of acid (Ø) and with the addition of 2 g.L^{-1} (AA 2 g.L^{-1} pH 6.8) and 5 g.L^{-1} (AA 5 g.L^{-1} pH 4.8) of acetic acid. Quantifications were normalized per gram of cells. The experiments performed in biological triplicate and a statistically different sample from the group is identified with by asterisk (*). Source: Author's collection.

Mechanisms described for *S. cerevisiae* for decreasing acetic acid toxic effects include increasing saturated lipids from the plasma membrane and therefore reducing permeability, activating ATP-dependent proton pumps to externalize extra protons, among other changes. These adaptations control intracellular acidification by limiting diffusion of acid through the plasma membrane, and consequently, reducing intracellular amounts of acetic acid, as reviewed by (167). Reduction in intracellular amounts of acetic acid found in *P. pastoris* overexpressing *HAA1* shows that cell modifications to avoid the accumulation of the acid in the cytoplasm are probably

occurring. Transcriptome analysis of gene *HAA1* (chapter one) within the P-value < 0.05 was only observed for lignocellulose hydrolysate 30%, in which it was overexpressed. The gene *PMA1* which encodes for a plasma membrane H^+-ATPase (Pma1p) was shown to be involved positively express in response to acetic acid in *S. cerevisiae.* In chapter one, the *PMA1* gene was found to be increased 1.05-fold for acetic acid 2g.L^{-1}, 1.23-fold for acetic acid 2 g.$L^{-1,}$ and 0.94-fold for hydrolysate 30%. Data suggest that these responses are similar to those previously reported for other yeast, like *S. cerevisiae.*

10 CONCLUSIONS

One strategy to reduce acetic acid's effect on *Pichia*'s metabolism is the overexpression of transcriptional factor *HAA1*, previously described to confer tolerance to the referred acid in *S. cerevisiae.* Our experiments demonstrate that the tolerance is proportional to the promoter's strength, and a p*GAP* strong promoter can improve *Pichia*'s growth in the presence of acetic acid by 4.9-fold after 24 h of cultivation in deep-well plates. when compared to control. The impact of *HAA1* overexpression was also observed by the reduction of intracellular amounts of acetic acid found in the engineered strains.

The characterization of strains overexpressing *HAA1* in culture flasks in minimal media, and in lignocellulose hydrolysate are currently underway. Future experiments should include the investigation if these strains acquired tolerance uniquely to acetic acid or other organic or inorganic acids since it has been previously shown for *S. cerevisiae* that different strategies are highlighted for each type of organic acid (179). Other genes of potential interest for *P. pastoris* modification towards acetic acid tolerance are suggested in chapter one (Table 3).

11 FUTURE PROSPECTS

This work contributes to the understanding of *P. pastoris* responses to main components of lignocellulosic hydrolysate like inhibitors and xylose, The overall response of this yeast to acetic acid, furaldehydes, and sugarcane hydrolysate was outlined and improved. But there are yet, many questions to be answered. Several genes involved in the yeast response to lignocellulose derived inhibitors were identified and now their effect in improving robustness of strains need to be validated. Similarly, the gene *HAA1* gene from *P. pastoris* was identified and used to increase the yeast tolerance towards acetic acid. Thoroughly analyses of effects of HAA1 on yeast metabolism should be carried out. Additionally, the inhibitory effect of phenolic compounds on yeast metabolism was not investigated and is an important factor to be considered. The applications of these findings, nevertheless, don't have to be restricted to lignocellulosic hydrolysate utilization. Strain robustness is overall a valuable tool in biotechnological processes.

It was also demonstrated here that *P. pastoris* is capable of metabolizing xylose as the sole carbon source in minimal media. Complementary experiments are currently underway to provide a better understanding of this process. Transcriptional levels of the putative xylose pathway genes in the presence of glucose or xylose as carbon source are being analyzed by quantitative PCR. Additionally, the incorporation of ^{13}C from xylose into pentose-phosphate pathway metabolites is being investigated. Finally, optimization (through genetic engineering and/or adaptive laboratory evolution) can enable *P. pastoris* to use xylose efficiently and be applied in several biotechnological processes.

Different growth patterns on xylose support the idea that microbial diversity is a great source of desirable traits in yeasts of biotechnological interest. The seven species currently found in the *Komagataella* genus were all isolated in either Europe or the United States of America. Search for new *Komagataella* strains in different continents and climates, (like Brazil) have a huge potential of finding new species, strains, and desirable traits. When screening different species for tolerance to inhibitors, however, liquid assays are probably more sensitive and accurate than spotting assays, especially for furaldehydes.

Lastly, we hope to continue our work with *P. pastoris* and lignocellulosic hydrolysate components, in particular on xylose utilization. Thus, we hope to contribute

to the development of new biotechnological processes, and optimization of current ones, towards the construction of more sustainable and ecological industrial processes.

REFERENCES[4]

1. Guilliermond A. *Zygosaccharomyces Pastori*, nouvelle espèce de levures à copulation hétérogamique. In: mycologique de France. S, editor. Bulletin trimestriel de la Société mycologique de France [Internet]. Paris: (Épinal); 1920. p. 203–11. Available from: http://www.biodiversitylibrary.org/item/106463

2. Schwarzhans JP, Luttermann T, Geier M, Kalinowski J, Friehs K. Towards systems metabolic engineering in *Pichia pastoris*. Biotechnol Adv [Internet]. 2017;35(6):681–710. Available from: http://dx.doi.org/10.1016/j.biotechadv.2017.07.009

3. Peña DA, Gasser B, Zanghellini J, Steiger MG, Mattanovich D. Metabolic engineering of *Pichia pastoris*. Metab Eng [Internet]. 2018 Nov;50(April):2–15. Available from: https://doi.org/10.1016/j.ymben.2018.04.017

4. Bernauer L, Radkohl A, Lehmayer LGK, Emmerstorfer-Augustin A. *Komagataella phaffii* as Emerging Model Organism in Fundamental Research. Vol. 11, Frontiers in Microbiology. 2021. p. 1–16.

5. Phaff HJ, Miller MW, Shifrine M. The taxonomy of yeasts isolated from *Drosophila* in the Yosemite region of California. Antonie Van Leeuwenhoek. 1956;22(1):145–61.

6. Yamada Y, Matsuda M, Maeda K, Mikata K. The Phylogenetic Relationships of Methanol-assimilating Yeasts Based on the Partial Sequences of 18S and 26S Ribosomal RNAs: The Proposal of *Komagataella* Gen. Nov. (Sacchar omy cet aceae). Biosci Biotechnol Biochem. 1995;59(3):439–44.

7. Kurtzman CP. Description of *Komagataella phaffii* sp. nov. and the transfer of *Pichia pseudopastoris* to the methylotrophic yeast genus *Komagataella*. Int J Syst Evol Microbiol. 2005;55(2):973–6.

8. De Schutter K, Lin YC, Tiels P, Van Hecke A, Glinka S, Weber-Lehmann J, et al. Genome sequence of the recombinant protein production host *Pichia pastoris*. Nat Biotechnol. 2009;27(6):561–6.

9. Mattanovich D, Graf A, Stadlmann J, Dragosits M, Redl A, Maurer M, et al.

[4] According to the Vancouver style of referencing.

Genome, secretome and glucose transport highlight unique features of the protein production host *Pichia pastoris*. Microb Cell Fact. 2009;8:1–13.

10. Zahrl RJ, Peña DA, Mattanovich D, Gasser B. Systems biotechnology for protein production in *Pichia pastoris*. FEMS Yeast Res. 2017;17(7):1–15.

11. Caspeta L, Shoaie S, Agren R, Nookaew I, Nielsen J. Genome-scale metabolic reconstructions of *Pichia stipitis* and *Pichia pastoris* and in silico evaluation of their potentials. BMC Syst Biol [Internet]. 2012;6(1):24. Available from: http://www.biomedcentral.com/1752-0509/6/24

12. Naumov GI, Naumova ES, Boundy-Mills KL. Description of *Komagataella mondaviorum* sp. nov., a new sibling species of *Komagataella (Pichia) pastoris*. Antonie van Leeuwenhoek, Int J Gen Mol Microbiol [Internet]. 2018;111(7):1197–207. Available from: https://doi.org/10.1007/s10482-018-1028-6

13. Braun-Galleani S, Dias JA, Coughlan AY, Ryan AP, Byrne KP, Wolfe KH. Genomic diversity and meiotic recombination among isolates of the biotech yeast *Komagataella phaffii* (*Pichia pastoris*). Microb Cell Fact [Internet]. 2019;18(1):1–13. Available from: https://doi.org/10.1186/s12934-019-1260-4

14. Ogata K, Nishikawa H, Ohsugi M. A Yeast Capable of Utilizing Methanol. Agric Biol Chem [Internet]. 1969 Oct 9;33(10):1519–20. Available from: https://academic.oup.com/bbb/article/33/10/1519-1520/5977590

15. Higgins DR, Cregg JM. Pichia Protocols [Internet]. 1st editio. Higgins DR, Creggx JM, editors. Vol. 103. Humana Press; 1998. 284 p. Available from: http://books.google.com/books?id=Ku2wPAAACAAJ

16. Wegner GH. Emerging applications of the methylotrophic yeasts. FEMS Microbiol Lett. 1990;87(3–4):279–83.

17. Cregg JM, Madden KR, Barringer KJ, Thill GP, Stillman CA. Functional characterization of the two alcohol oxidase genes from the yeast *Pichia pastoris*. Mol Cell Biol. 1989;9(3):1316–23.

18. Hamilton SR, Gerngross TU. Glycosylation engineering in yeast: the advent of fully humanized yeast. Curr Opin Biotechnol. 2007;18(5):387–92.

19. Gasser B, Mattanovich D. Antibody production with yeasts and filamentous

fungi: On the road to large scale? Biotechnol Lett. 2007;29(2):201–12.

20. Paes BG, Almeida JRJR. Genetic improvement of microorganisms for applications in biorefineries. Chem Biol Technol Agric [Internet]. 2014 Dec 29;1(1):21. Available from: http://www.chembioagro.com/content/1/1/21

21. Inan M, Meagher MM, Inan M, Meagher MM. Non-Repressing Carbon Sources for Alcohol Oxidase (AOXI) Promoter of *Pichia pastoris*. Pap Biochem Eng. 2001;92(November):585–9.

22. Xie J, Zhou Q, Du P, Gan R, Ye Q. Use of different carbon sources in cultivation of recombinant *Pichia pastoris* for angiostatin production. Enzyme Microb Technol. 2005;36(2–3):210–6.

23. Riley R, Haridas S, Wolfe KH, Lopes MR, Hittinger CT, Göker M, et al. Comparative genomics of biotechnologically important yeasts. Proc Natl Acad Sci U S A. 2016;113(35):9882–7.

24. Zepeda AB, Pessoa A, Farías JG. Carbon metabolism influenced for promoters and temperature used in the heterologous protein production using *Pichia pastoris* yeast. Brazilian J Microbiol [Internet]. 2018;49(December):119–27. Available from: https://doi.org/10.1016/j.bjm.2018.03.010

25. Potvin G, Zhang Z, Defela A, Lam H. Screening of Alternative Carbon Sources for Recombinant Protein Production in *Pichia pastoris*. Int J Chem React Eng [Internet]. 2016;14(1):251–7. Available from: https://www.degruyter.com/document/doi/10.1515/ijcre-2015-0092/html

26. Li P, Sun H, Chen Z, Li Y, Zhu T. Construction of efficient xylose utilizing *Pichia pastoris* for industrial enzyme production. Microb Cell Fact. 2015;14(1):1–10.

27. Kickenweiz T, Glieder A, Wu JC. Construction of a cellulose-metabolizing *Komagataella phaffii* (*Pichia pastoris*) by co-expressing glucanases and β-glucosidase. Appl Microbiol Biotechnol. 2018 Feb;102(3):1297–306.

28. Gassler T, Sauer M, Gasser B, Egermeier M, Troyer C, Causon T, et al. The industrial yeast *Pichia pastoris* is converted from a heterotroph into an autotroph capable of growth on CO2. Nat Biotechnol [Internet]. 2020 Feb 16;38(2):210–6. Available from: http://dx.doi.org/10.1038/s41587-019-0363-0

29. Anwar Z, Gulfraz M, Irshad M. Agro-industrial lignocellulosic biomass a key to

unlock the future bio-energy: A brief review. J Radiat Res Appl Sci [Internet]. 2014;7(2):163–73. Available from: http://linkinghub.elsevier.com/retrieve/pii/S1687850714000119

30. Karimi K, Taherzadeh MJ. A critical review of analytical methods in pretreatment of lignocelluloses: Composition, imaging, and crystallinity. Bioresour Technol [Internet]. 2016;200:1008–18. Available from: http://dx.doi.org/10.1016/j.biortech.2015.11.022

31. Betts WB, Dart RK, Ball AS, Pedlar SL. Biosynthesis and Structure of Lignocellulose. 1991;139–55.

32. Hernández-Beltrán JU, Hernández-De Lira IO, Cruz-Santos MM, Saucedo-Luevanos A, Hernández-Terán F, Balagurusamy N. Insight into pretreatment methods of lignocellulosic biomass to increase biogas yield: Current state, challenges, and opportunities. Appl Sci. 2019;9(18).

33. Rubin EM. Genomics of cellulosic biofuels. Nature. 2008;454(7206):841–5. License Number: 5079491355522.

34. Jönsson LJ, Martín C. Pretreatment of lignocellulose: Formation of inhibitory by-products and strategies for minimizing their effects. Bioresour Technol. 2016;199:103–12.

35. Rezende CA, De Lima M, Maziero P, Deazevedo E, Garcia W, Polikarpov I. Chemical and morphological characterization of sugarcane bagasse submitted to a delignification process for enhanced enzymatic digestibility. Biotechnol Biofuels [Internet]. 2011;4(1):54. Available from: http://www.biotechnologyforbiofuels.com/content/4/1/54

36. Østby H, Hansen LD, Horn SJ, Eijsink VGH, Várnai A. Enzymatic processing of lignocellulosic biomass: principles, recent advances and perspectives [Internet]. Vol. 47, Journal of Industrial Microbiology and Biotechnology. Springer International Publishing; 2020. 623–657 p. Available from: https://doi.org/10.1007/s10295-020-02301-8

37. Passos D de F, Pereira N, Castro AM de. A comparative review of recent advances in cellulases production by *Aspergillus*, *Penicillium* and *Trichoderma* strains and their use for lignocellulose deconstruction. Curr Opin Green Sustain

Chem [Internet]. 2018;14:60–6. Available from: https://doi.org/10.1016/j.cogsc.2018.06.003

38. Ferdeş M, Dincă MN, Moiceanu G, Zabava BŞ, Paraschiv G. Microorganisms and enzymes used in the biological pretreatment of the substrate to enhance biogas production: A review. Sustain. 2020;12(17).

39. Almeida JR, Modig T, Petersson A, Hähn-Hägerdal B, Lidén G, Gorwa-Grauslund MF. Increased tolerance and conversion of inhibitors in lignocellulosic hydrolysates by *Saccharomyces cerevisiae*. J Chem Technol Biotechnol [Internet]. 2007 Apr [cited 2014 Jun 23];82(4):340–9. Available from: http://onlinelibrary.wiley.com/doi/10.1002/jctb.1676/full. License Number: 5079500936159.

40. Kim D. Physico-chemical conversion of lignocellulose: Inhibitor effects and detoxification strategies: A mini review. Molecules. 2018;23(2).

41. Kim D, Ximenes EA, Nichols NN, Cao G, Frazer SE, Ladisch MR. Maleic acid treatment of biologically detoxified corn stover liquor. Bioresour Technol [Internet]. 2016;216:437–45. Available from: http://dx.doi.org/10.1016/j.biortech.2016.05.086

42. Cao G, Ximenes E, Nichols NN, Frazer SE, Kim D, Cotta MA, et al. Bioabatement with hemicellulase supplementation to reduce enzymatic hydrolysis inhibitors. Bioresour Technol [Internet]. 2015;190:412–5. Available from: http://dx.doi.org/10.1016/j.biortech.2015.04.064

43. Ballesteros I, Negro MJ, Oliva JM, Cabañas A, Manzanares P, Ballesteros M. Ethanol production from steam-explosion pretreated wheat straw. In: Applied Biochemistry and Biotechnology. 2006. p. 496–508.

44. Linde M, Jakobsson EL, Galbe M, Zacchi G. Steam pretreatment of dilute H2SO4-impregnated wheat straw and SSF with low yeast and enzyme loadings for bioethanol production. Biomass and Bioenergy. 2008;32(4):326–32.

45. Ximenes E, Kim Y, Ladisch MR. Biological Conversion of Plants to Fuels and Chemicals and the Effects of Inhibitors. In: Aqueous Pretreatment of Plant Biomass for Biological and Chemical Conversion to Fuels and Chemicals.

2013. p. 39–60.

46. Kim Y, Ximenes E, Mosier NS, Ladisch MR. Soluble inhibitors/deactivators of cellulase enzymes from lignocellulosic biomass. Enzyme Microb Technol [Internet]. 2011;48(4–5):408–15. Available from: http://dx.doi.org/10.1016/j.enzmictec.2011.01.007

47. Martín C, Klinke HB, Thomsen AB. Wet oxidation as a pretreatment method for enhancing the enzymatic convertibility of sugarcane bagasse. Enzyme Microb Technol. 2007;40(3):426–32.

48. Michelin M, Ximenes E, de Lourdes Teixeira de Moraes Polizeli M, Ladisch MR. Effect of phenolic compounds from pretreated sugarcane bagasse on cellulolytic and hemicellulolytic activities. Bioresour Technol [Internet]. 2016;199:275–8. Available from: http://dx.doi.org/10.1016/j.biortech.2015.08.120

49. Cara C, Ruiz E, Ballesteros M, Manzanares P, Negro MJ, Castro E. Production of fuel ethanol from steam-explosion pretreated olive tree pruning. Fuel. 2008;87(6):692–700.

50. Ferreira-Leitão V, Perrone CC, Rodrigues J, Franke APM, Macrelli S, Zacchi G. An approach to the utilisation of CO2 as impregnating agent in steam pretreatment of sugar cane bagasse and leaves for ethanol production. Biotechnol Biofuels [Internet]. 2010 Jan;3(1):7. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2861027&tool=pmcentrez&rendertype=abstract

51. Moysés DN, Reis VCB, de Almeida JRM, de Moraes LMP, Torres FAG. Xylose fermentation by *Saccharomyces cerevisiae*: Challenges and prospects. Int J Mol Sci. 2016;17(3):1–18.

52. Hahn-Hägerdal B, Galbe M, Gorwa-Grauslund MF, Lidén G, Zacchi G. Bio-ethanol - the fuel of tomorrow from the residues of today. Trends Biotechnol. 2006;24(12):549–56.

53. Rebello S, Abraham A, Madhavan A, Sindhu R, Binod P, Bahuleyan AK, et al. Non-conventional yeast cell factories for sustainable bioprocesses. FEMS Microbiol Lett. 2018;365(April):1–10.

54. Zhao Z, Xian M, Liu M, Zhao G. Biochemical routes for uptake and conversion of xylose by microorganisms. Biotechnol Biofuels [Internet]. 2020;13(1):1–12. Available from: https://doi.org/10.1186/s13068-020-1662-x

55. Hahn-Hägerdal B, Karhumaa K, Fonseca C, Spencer-Martins I, Gorwa-Grauslund MF. Towards industrial pentose-fermenting yeast strains. Vol. 74, Applied Microbiology and Biotechnology. 2007. p. 937–53.

56. Hasunuma T, Kondo A. Development of yeast cell factories for consolidated bioprocessing of lignocellulose to bioethanol through cell surface engineering. Biotechnol Adv. 2012;30(6):1207–18. License number: 5079501489953.

57. Almeida JRM, Runquist D, Sànchez i Nogué V, Lidén G, Gorwa-Grauslund MF. Stress-related challenges in pentose fermentation to ethanol by the yeast *Saccharomyces cerevisiae*. Biotechnol J [Internet]. 2011 Mar [cited 2013 Oct 25];6(3):286–99. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21305697

58. Jönsson LJ, Alriksson BB, Nilvebrant N-OO. Bioconversion of lignocellulose: Inhibitors and detoxification. Biotechnol Biofuels. 2013;6(1):1–10.

59. Vanmarcke G, Demeke MM, Moreno MRF, Thevelein JM. Biotechnology for Biofuels Identification of the major fermentation inhibitors of recombinant 2G yeasts in diverse lignocellulose hydrolysates. Biotechnol Biofuels [Internet]. 2021;1–15. Available from: https://doi.org/10.1186/s13068-021-01935-9

60. Rumbold K, van Buijsen HJJ, Overkamp KM, van Groenestijn JW, Punt PJ, Werf MJVD. Microbial production host selection for converting second-generation feedstocks into bioproducts. Microb Cell Fact. 2009;8:1–11.

61. Yang Y, Hu M, Tang Y, Geng B, Qiu M, He Q, et al. Progress and perspective on lignocellulosic hydrolysate inhibitor tolerance improvement in *Zymomonas mobilis*. Bioresour Bioprocess [Internet]. 2018;5(1). Available from: https://doi.org/10.1186/s40643-018-0193-9

62. Zha Y, Hossain AH, Tobola F, Sedee N, Havekes M, Punt PJ. *Pichia anomala* 29X: A resistant strain for lignocellulosic biomass hydrolysate fermentation. FEMS Yeast Res. 2013;13(7):609–17.

63. Martín C, Jönsson LJ. Comparison of the resistance of industrial and laboratory

strains of *Saccharomyces* and *Zygosaccharomyces* to lignocellulose-derived fermentation inhibitors. Enzyme Microb Technol. 2003;32(3–4):386–95.

64. Hou X, Yao S. Improved inhibitor tolerance in xylose-fermenting yeast *Spathaspora passalidarum* by mutagenesis and protoplast fusion. Appl Microbiol Biotechnol [Internet]. 2012 Mar [cited 2014 Mar 20];93(6):2591–601. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22116630

65. Moreno AD, Carbone A, Pavone R, Olsson L, Geijer C. Evolutionary engineered *Candida intermedia* exhibits improved xylose utilization and robustness to lignocellulose-derived inhibitors and ethanol. Appl Microbiol Biotechnol. 2019;103(3):1405–16.

66. Cottier F, Tan ASM, Chen J, Lum J, Zolezzi F, Poidinger M, et al. The transcriptional stress response of *Candida albicans* to weak organic acids. G3 GENES, GENOMES, Genet [Internet]. 2015;5(4):497–505. Available from: http://g3journal.org/lookup/doi/10.1534/g3.114.015941

67. Delgenes JP, Moletta R, Navarro JM. Effects of lignocellulose degradation products on ethanol fermentations of glucose and xylose by *Saccharomyces cerevisiae*, *Zymomonas mobilis*, *Pichia stipitis*, and *Candida shehatae*. Enzyme Microb Technol. 1996;19(3):220–5.

68. Yamakawa CK, Kastell L, Mahler MR, Martinez JL, Mussatto SI. Exploiting new biorefinery models using non-conventional yeasts and their implications for sustainability. Bioresour Technol [Internet]. 2020;309:123374. Available from: https://doi.org/10.1016/j.biortech.2020.123374

69. Soares CEVF, Bergmann JC, Almeida JR. Variable and dose-dependent response of *Saccharomyces* and non-*Saccharomyces* yeasts toward lignocellulosic hydrolysate inhibitors. 2020;1–22.

70. Almeida JRM, Bertilsson M, Gorwa-Grauslund MF, Gorsich S, Lidén G. Metabolic effects of furaldehydes and impacts on biotechnological processes. Appl Microbiol Biotechnol [Internet]. 2009 Mar [cited 2013 Oct 23];82(4):625–38. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19184597

71. Palmqvist E, Hahn-Hägerdal B. Fermentation of lignocellulosic hydrolysates. II: Inhibitors and mechanisms of inhibition. Bioresour Technol. 2000

Aug;74(1):25–33.

72. Adeboye PT, Bettiga M, Aldaeus F, Larsson PT, Olsson L. Catabolism of coniferyl aldehyde, ferulic acid and p-coumaric acid by *Saccharomyces cerevisiae* yields less toxic products. Microb Cell Fact. 2015;14(1):1–14.

73. Mira NP, Palma M, Guerreiro JF, Sá-Correia I. Genome-wide identification of *Saccharomyces cerevisiae* genes required for tolerance to acetic acid. Microb Cell Fact. 2010;9:1–13.

74. Brandt BA, Jansen T, Görgens JF, van Zyl WH. Overcoming lignocellulose-derived microbial inhibitors: advancing the *Saccharomyces cerevisiae* resistance toolbox. Biofuels, Bioprod Biorefining. 2019;13(6):1520–36.

75. Petersson A, Almeida JRM, Modig T, Karhumaa K, Hahn-Hägerdal B, Gorwa-Grauslund MF, et al. A 5-hydroxymethyl furfural reducing enzyme encoded by the *Saccharomyces cerevisiae* ADH6 gene conveys HMF tolerance. Yeast. 2006;23(6):455–64.

76. Radecka D, Mukherjee V, Mateo RQ, Stojiljkovic M, Foulquié-Moreno MR, Thevelein JM. Looking beyond *Saccharomyces*: The potential of non-conventional yeast species for desirable traits in bioethanol fermentation. FEMS Yeast Res. 2015;15(6):1–13.

77. Cavka A, Jornsson LJ. Comparison of the growth of filamentous fungi and yeasts in lignocellulose-derived media. Biocatal Agric Biotechnol [Internet]. 2014;3(4):197–204. Available from: http://dx.doi.org/10.1016/j.bcab.2014.04.003

78. Portnoy VA, Bezdan D, Zengler K. Adaptive laboratory evolution — harnessing the power of biology for metabolic engineering. Curr Opin Biotechnol [Internet]. 2011;22(4):590–4. Available from: http://dx.doi.org/10.1016/j.copbio.2011.03.007

79. Dragosits M, Mattanovich D. Adaptive laboratory evolution – principles and applications for biotechnology. Microb Cell Fact [Internet]. 2013;12(1):64. Available from: http://microbialcellfactories.biomedcentral.com/articles/10

80. Tong Si, Jiazhang Lian and HZ 7. Strain Development by Whole-Cell Directed Evolution. In: Alcalde M, editor. Directed Enzyme Evolution: Advances and

Applications [Internet]. 1st ed. Springer International Publishing; 2017. p. 173–200. Available from: https://www.springer.com/gp/book/9783319504117

81. Heer D, Sauer U. Identification of furfural as a key toxin in lignocellulosic hydrolysates and evolution of a tolerant yeast strain. Microb Biotechnol. 2008;1(6):497–506.

82. Hughes SR, Gibbons WR, Bang SS, Pinkelman R, BischoV KM, Slininger PJ, et al. Random UV-C mutagenesis of *Scheffersomyces* (formerly *Pichia*) *stipitis* NRRL Y-7124 to improve anaerobic growth on lignocellulosic sugars. J Ind Microbiol Biotechnol. 2012;39(1):163–73.

83. Koppram R, Albers E, Olsson L. Evolutionary engineering strategies to enhance tolerance of xylose utilizing recombinant yeast to inhibitors derived from spruce biomass. Biotechnol Biofuels. 2012;5:32.

84. Wang S, Sun X, Yuan Q. Strategies for enhancing microbial tolerance to inhibitors for biofuel production: A review. Bioresour Technol [Internet]. 2018;258:302–9. Available from: https://doi.org/10.1016/j.biortech.2018.03.064

85. Ruegg TL, Kim EM, Simmons BA, Keasling JD, Singer SW, Soon Lee T, et al. An auto-inducible mechanism for ionic liquid resistance in microbial biofuel production. Nat Commun. 2014;5:1–7.

86. Frederix M, Hütter K, Leu J, Batth TS, Turner WJ, Rüegg TL, et al. Development of a native *Escherichia coli* induction system for ionic liquid tolerance. PLoS One. 2014;9(7):1–10.

87. Tomas CA, Welker NE, Papoutsakis ET. Overexpression of groESL in *Clostridium acetobutylicum* results in increased solvent production and tolerance, prolonged metabolism, and changes in the cell's transcriptional program. Appl Environ Microbiol. 2003;69(8):4951–65.

88. Wang PM, Zheng DQ, Chi XQ, Li O, Qian CD, Liu TZ, et al. Relationship of trehalose accumulation with ethanol fermentation in industrial *Saccharomyces cerevisiae* yeast strains. Bioresour Technol. 2014;152:371–6.

89. Tanaka K, Ishii Y, Ogawa J, Shima J. Enhancement of Acetic Acid Tolerance in *Saccharomyces cerevisiae* by Overexpression of the HAA1 Gene , Encoding a Transcriptional. 2012;78(22):8161–3.

90. Bui LM, Lee JY, Geraldi A, Rahman Z, Lee JH, Kim SC. Improved n-butanol tolerance in *Escherichia coli* by controlling membrane related functions. J Biotechnol. 2015;204:33–44.

91. Wang J, Zhang B, Zhang J, Wang H, Zhao M, Wang N, et al. Enhanced succinic acid production and magnesium utilization by overexpression of magnesium transporter mgtA in *Escherichia coli* mutant. Bioresour Technol. 2014;170:125–31.

92. Qureshi AS, Zhang J, Bao J. High ethanol fermentation performance of the dry dilute acid pretreated corn stover by an evolutionarily adapted *Saccharomyces cerevisiae* strain. Bioresour Technol. 2015;189:399–404.

93. Mohamed ET, Wang S, Lennen RM, Herrgård MJ, Simmons BA, Singer SW, et al. Generation of a platform strain for ionic liquid tolerance using adaptive laboratory evolution. Microb Cell Fact. 2017;16(1):1–15.

94. Qi F, Kitahara Y, Wang Z, Zhao X, Du W, Liu D. Novel mutant strains of *Rhodosporidium toruloides* by plasma mutagenesis approach and their tolerance for inhibitors in lignocellulosic hydrolyzate. J Chem Technol Biotechnol. 2014;89(5):735–42.

95. Sridhar M, Kiran Sree N, Venkateswar Rao L. Effect of UV radiation on thermotolerance, ethanol tolerance and osmotolerance of *Saccharomyces cerevisiae* VS1 and VS3 strains. Bioresour Technol. 2002;83(3):199–202.

96. Gasser B, Mattanovich D. A yeast for all seasons – Is *Pichia* pastoris a suitable chassis organism for future bioproduction? FEMS Microbiol Lett. 2018;365(17):1–4.

97. Liu ZL, Slininger PJ, Dien BS, Berhow MA, Kurtzman CP, Gorsich SW. Adaptive response of yeasts to furfural and 5-hydroxymethylfurfural and new chemical evidence for HMF conversion to 2,5-bis-hydroxymethylfuran. J Ind Microbiol Biotechnol [Internet]. 2004 Sep [cited 2014 Mar 20];31(8):345–52. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15338422

98. Skerker JM, Leon D, Price MN, Mar JS, Tarjan DR, Wetmore KM, et al. Dissecting a complex chemical stress: chemogenomic profiling of plant hydrolysates. Mol Syst Biol [Internet]. 2013 Jan 18;9(1):674. Available from:

https://onlinelibrary.wiley.com/doi/abs/10.1038/msb.2013.30

99. Modig T, Lidén G, Taherzadeh MJ. Inhibition effects of furfural on alcohol dehydrogenase, aldehyde dehydrogenase and pyruvate dehydrogenase. Biochem J. 2002;363(3):769–76.

100. Sitepu I, Selby T, Lin T, Zhu S, Boundy-mills K, Avenue OS, et al. Carbon source utilization and inhibitor tolerance of 45 oleaginous yeast species. 2015;41(7):1061–70.

101. Modig T, Almeida JRM, Gorwa-Grauslund MF, Lidén G. Variability of the response of *Saccharomyces cerevisiae* strains to lignocellulose hydrolysate. Biotechnol Bioeng. 2008;100(3):423–9.

102. Nocon J, Steiger MG, Pfeffer M, Sohn SB, Kim TY, Maurer M, et al. Model based engineering of *Pichia pastoris* central metabolism enhances recombinant protein production. Metab Eng. 2014;24(August):129–38.

103. Siripong W, Wolf P, Kusumoputri TP, Downes JJ, Kocharin K, Tanapongpipat S, et al. Metabolic engineering of *Pichia pastoris* for production of isobutanol and isobutyl acetate. Biotechnol Biofuels [Internet]. 2018;11(1):1–16. Available from: https://doi.org/10.1186/s13068-017-1003-x

104. Melo NTM, Mulder KCL, Nicola AM, Carvalho LS, Menino GS, Mulinari E, et al. Effect of pyruvate decarboxylase knockout on product distribution using *Pichia pastoris* (*Komagataella phaffii*) engineered for lactic acid production. Bioengineering [Internet]. 2018;5(1):17. Available from: http://www.mdpi.com/2306-5354/5/1/17

105. Vogl T, Hartner FS, Glieder A. New opportunities by synthetic biology for biopharmaceutical production in *Pichia pastoris*. Curr Opin Biotechnol [Internet]. 2013;24(6):1094–101. Available from: http://dx.doi.org/10.1016/j.copbio.2013.02.024

106. Almeida JRM, Rodrigues CM, Justen F, Ramos TGS. *Komagataella phaffii* recombinante produtora de ácido xilônico e método para a produção de ácido xilônico a partir de xilose. INPI - BR102018001359. 2018;

107. Xu Q, Bai C, Liu Y, Song L, Tian L, Yan Y, et al. Modulation of acetate utilization in *Komagataella phaffii* by metabolic engineering of tolerance and

metabolism. Biotechnol Biofuels [Internet]. 2019;12(1):1–14. Available from: https://doi.org/10.1186/s13068-019-1404-0

108. Morais Junior WG, Pacheco TF, Trichez D, Almeida JRM, Gonçalves SB. Xylitol production on sugarcane biomass hydrolysate by newly identified *Candida tropicalis* JA2 strain. Yeast. 2019;36(5):349–61.

109. Andrews S, Krueger F, Segonds-Pichon A, Biggins L, Christel K, Wingett S. FastQC: a quality control tool for high throughput sequence data [Internet]. Babraham Institute. 2012. Available from: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/

110. Bolger AM, Lohse M, Usadel B. Trimmomatic: A flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30(15):2114–20.

111. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: Ultrafast universal RNA-seq aligner. Bioinformatics. 2013;29(1):15–21.

112. Anders S, Pyl PT, Huber W. HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics. 2015 Jan;31(2):166–9.

113. R Foundation for Statistical Computing RCT. R: A language and environment for statistical computing. 2016.

114. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(12):1–21.

115. Supek F, Bošnjak M, Škunca N, Šmuc T. REVIGO Summarizes and visualizes long lists of gene ontology terms. Gibas C, editor. PLoS One. 2011 Jul;6(7).

116. Veras HCT, Parachin NS, Almeida JRM. Comparative assessment of fermentative capacity of different xylose-consuming yeasts. Microb Cell Fact. 2017;16(1):1–8.

117. Ludovico P, Sousa MJ, Silva MT, Leão C, Côrte-Real M. *Saccharomyces cerevisiae* commits to a programmed cell death process in response to acetic acid. Microbiology [Internet]. 2001 Sep 1;147(9):2409–15. Available from: https://www.microbiologyresearch.org/content/journal/micro/10.1099/00221287-147-9-2409

118. Stratford M, Nebe-von-Caron G, Steels H, Novodvorska M, Ueckert J, Archer

DB. Weak-acid preservatives: pH and proton movements in the yeast *Saccharomyces cerevisiae*. Int J Food Microbiol [Internet]. 2013 Feb;161(3):164–71. Available from: http://dx.doi.org/10.1016/j.ijfoodmicro.2012.12.013

119. Pampulha ME, Loureiro-Dias MC. Activity of glycolytic enzymes of *Saccharomyces cerevisiae* in the presence of acetic acid. Appl Microbiol Biotechnol. 1990;34(3):375–80.

120. Palmqvist E, Almeida JS, Hahn-Hägerdal B. Influence of furfural on anaerobic glycolytic kinetics of *Saccharomyces cerevisiae* in batch culture. Biotechnol Bioeng. 1999;62(4):447–54.

121. Dong Y, Hu J, Fan L, Chen Q. RNA-Seq-based transcriptomic and metabolomic analysis reveal stress responses and programmed cell death induced by acetic acid in *Saccharomyces cerevisiae*. Sci Rep [Internet]. 2017;7(January):1–16. Available from: http://dx.doi.org/10.1038/srep42659

122. Liu ZL, Slininger PJ, Gorsich SW. Enhanced biotransformation of furfural and hydroxymethylfurfural by newly developed ethanologenic yeast strains. Appl Biochem Biotechnol [Internet]. 2005;121(1–3):0451–60. Available from: http://link.springer.com/10.1385/ABAB:121:1-3:0451

123. Larsson S, Palmqvist E, Hahn-Hägerdal B, Tengborg C, Stenberg K, Zacchi G, et al. The generation of fermentation inhibitors during dilute acid hydrolysis of softwood. Enzyme Microb Technol. 1999;24(3–4):151–9.

124. Li J, Zhu K, Zhao H. Transcriptome analysis reveals the protection mechanism of proanthocyanidins for *Saccharomyces cerevisiae* during wine fermentation. Sci Rep [Internet]. 2020;10(1):1–12. Available from: http://dx.doi.org/10.1038/s41598-020-63631-2

125. Sousa MJ, Ludovico P, Rodrigues F, Leão C, Côrte-Real M. Stress and cell death in yeast induced by acetic acid. In: Cell Metabolism - Cell Homeostasis and Stress Response. 2012. p. 73–100.

126. Madshus IH. Regulation of intracellular pH in eukaryotic cells. Biochem J. 1988;250(1):1–8.

127. Zhang MM, Zhao XQ, Cheng C, Bai FW. Improved growth and ethanol

fermentation of *Saccharomyces cerevisiae* in the presence of acetic acid by overexpression of SET5 and PPR1. Biotechnol J. 2015;10(12):1903–11.

128. Gorsich SW, Dien BS, Nichols NN, Slininger PJ, Liu ZL, Skory CD. Tolerance to furfural-induced stress is associated with pentose phosphate pathway genes *ZWF1*, *GND1*, *RPE1*, and *TKL1* in *Saccharomyces cerevisiae*. Appl Microbiol Biotechnol. 2006;71(3):339–49.

129. Etienne-Manneville S, Hall A. Rho GTPases in cell biology. Nature. 2002;420(6916):629–35.

130. Annan RB, Wu C, Waller DD, Whiteway M, Thomas DY. Rho5p is involved in mediating the osmotic stress response in *Saccharomyces cerevisiae*, and its activity is regulated via Msi1p and Npr1p by phosphorylation and ubiquitination. Eukaryot Cell. 2008;7(9):1441–9.

131. Fletcher E, Feizi A, Kim SS, Siewers V, Nielsen J. RNA-seq analysis of *Pichia anomala* reveals important mechanisms required for survival at low pH. Microb Cell Fact. 2015;14(1):1–11.

132. Ask M, Bettiga M, Duraiswamy VR, Olsson L. Pulsed addition of HMF and furfural to batch-grown xylose-utilizing *Saccharomyces cerevisiae* results in different physiological responses in glucose and xylose consumption phase. Biotechnol Biofuels. 2013;6(1):1–15.

133. Almeida JR, Modig T, Röder A, Lidén G, Gorwa-Grauslund M-F. *Pichia stipitis* xylose reductase helps detoxifying lignocellulosic hydrolysate by reducing 5-hydroxymethyl-furfural (HMF). Biotechnol Biofuels [Internet]. 2008 Jan [cited 2013 Nov 16];1(1):12. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2464581&tool=pmcentrez&rendertype=abstract

134. Horváth IS, Taherzadeh MJ, Niklasson C, Lidén G. Effects of furfural on anaerobic continuous cultivation of *Saccharomyces cerevisiae*. Biotechnol Bioeng. 2001;75(5):540–9.

135. Liu ZL. Molecular mechanisms of yeast tolerance and in situ detoxification of lignocellulose hydrolysates. Appl Microbiol Biotechnol. 2011;90(3):809–25.

136. Ma M, Liu ZL. Comparative transcriptome profiling analyses during the lag

phase uncover *YAP1, PDR1, PDR3, RPN4*, and *HSF1* as key regulatory genes in genomic adaptation to the lignocellulose derived inhibitor HMF for *Saccharomyces cerevisiae*. BMC Genomics. 2010;11(1):12–4.

137. Heer D, Heine D, Sauer U. Resistance of *Saccharomyces cerevisiae* to high concentrations of furfural is based on NADPH-dependent reduction by at least two oxireductases. Appl Environ Microbiol. 2009;75(24):7631–8.

138. Wang D, Wu D, Yang X, Hong J. Transcriptomic analysis of thermotolerant yeast: *Kluyveromyces marxianus* in multiple inhibitors tolerance. RSC Adv. 2018;8(26):14177–92.

139. Fauchon M, Lagniel G, Aude JC, Lombardia L, Soularue P, Petat C, et al. Sulfur sparing in the yeast proteome in response to sulfur demand. Mol Cell. 2002;9(4):713–23.

140. Ask M, Bettiga M, Mapelli V, Olsson L. The influence of HMF and furfural on redox-balance and energy-state of xylose-utilizing *Saccharomyces cerevisiae*. Biotechnol Biofuels. 2013;6(1):22.

141. Chahed H, Boumaiza M, Ezzine A, Marzouki MN. Heterologous expression and biochemical characterization of a novel thermostable *Sclerotinia sclerotiorum* GH45 endoglucanase in *Pichia pastoris*. Int J Biol Macromol [Internet]. 2018;106:629–35. Available from: https://doi.org/10.1016/j.ijbiomac.2017.08.062

142. Gündüz Ergün B, Çalık P. Lignocellulose degrading extremozymes produced by *Pichia pastoris*: Current status and future prospects. Bioprocess Biosyst Eng. 2016;39(1):1–36.

143. Bernardi AV, Yonamine DK, Uyemura SA, Dinamarco TM. A thermostable *Aspergillus fumigatus* gh7 endoglucanase over-expressed in *Pichia pastoris* stimulates lignocellulosic biomass hydrolysis. Int J Mol Sci. 2019;20(9).

144. Schwartz C, Wheeldon I, Ann-kathrin L. Genome and metabolic engineering in non-conventional yeasts : Current advances and applications. 2017;2:198–207.

145. Ahmad M, Hirz M, Pichler H, Schwab H. Protein expression in *Pichia pastoris*: Recent achievements and perspectives for heterologous protein production. Appl Microbiol Biotechnol. 2014;98(12):5301–17.

146. Sohn SB, Graf AB, Kim TY, Gasser B, Maurer M, Ferrer P, et al. Genome-scale metabolic model of methylotrophic yeast *Pichia pastoris* and its use for in silico analysis of heterologous protein production. Biotechnol J. 2010;5(7):705–15.

147. Conrad TM, Lewis NE, Palsson BO. Microbial laboratory evolution in the era of genome-scale science. Vol. 7, Molecular Systems Biology. 2011.

148. Demeke MM, Dietz H, Li Y, Foulquié-moreno MR, Mutturi S, Deprez S, et al. Development of a D-xylose fermenting and inhibitor tolerant industrial *Saccharomyces cerevisiae* strain with high performance in lignocellulose hydrolysates using metabolic and evolutionary engineering. 2013;1–24.

149. Promdonkoy P, Mhuantong W, Champreda V, Tanapongpipat S, Runguphan W. Improvement in d-xylose utilization and isobutanol production in *S. cerevisiae* by adaptive laboratory evolution and rational engineering. J Ind Microbiol Biotechnol. 2020;47(6–7):497–510.

150. Cha JW, Jang SH, Kim YJ, Chang YK, Jeong KJ. Engineering of *Klebsiella oxytoca* for production of 2,3-butanediol using mixed sugars derived from lignocellulosic hydrolysates. GCB Bioenergy. 2020;12(4):275–86.

151. Cubas-Cano E, González-Fernández C, Tomás-Pejó E. Evolutionary engineering of *Lactobacillus pentosus* improves lactic acid productivity from xylose-rich media at low pH. Bioresour Technol [Internet]. 2019;288(March):121540. Available from: https://doi.org/10.1016/j.biortech.2019.121540

152. Cregg JM, Barringer KJ, Hessler AY, Madden KR. *Pichia pastoris* as a host system for transformations. Mol Cell Biol [Internet]. 1985;5(12):3376–85. Available from: http://mcb.asm.org/lookup/doi/10.1128/MCB.5.12.3376

153. Dlauchy D, Tornai-Lehoczki J, Fülöp L, Péter G. *Pichia* (*Komagataella*) *pseudopastoris* sp. nov., a new yeast species from Hungary. Antonie van Leeuwenhoek, Int J Gen Mol Microbiol. 2003;83(4):327–32.

154. Kurtzman CP. *Komagataella populi* sp. nov. and *Komagataella ulmi* sp. nov., two new methanol assimilating yeasts from exudates of deciduous trees. Antonie van Leeuwenhoek, Int J Gen Mol Microbiol. 2012;101(4):859–68.

155. Naumov GI, Naumova ES, Tyurin O V., Kozlov DG. *Komagataella kurtzmanii* sp. nov., a new sibling species of *Komagataella* (*Pichia*) *pastoris* based on multigene sequence analysis. Antonie van Leeuwenhoek, Int J Gen Mol Microbiol. 2013;104(3):339–47.

156. Blackwell M, Urbina H, Frank R. *Scheffersomyces cryptocercus*: A new xylose-fermenting yeast associated with the gut of wood roaches and new combinations in the *Sugiyamaella* yeast clade. Mycologia. 2013;105(3):650–60.

157. Bellasio M, Peymann A, Steiger MG, Valli M, Sipiczki M, Sauer M, et al. Complete genome sequence and transcriptome regulation of the pentose utilizing yeast *Sugiyamaella lignohabitans*. FEMS Yeast Res. 2016;16(4):1–14.

158. National Center for Biotechnology Information. PubChem Compound Summary for CID 7362, Furfural [Internet]. National Library of Medicine. 2021 [cited 2021 Apr 30]. Available from: https://pubchem.ncbi.nlm.nih.gov/compound/furfural

159. National Center for Biotechnology Information. PubChem Compound Summary for CID 237332, 5-Hydroxymethylfurfural [Internet]. National Library of Medicine. 2021 [cited 2021 Apr 30]. Available from: https://pubchem.ncbi.nlm.nih.gov/compound/5-hydroxymethylfurfural

160. Merck. Common Cell Culture Problems : Poor Cell Growth Cell growth : an overview Common Causes of and Solutions to Poor Cell Growth [Internet]. 2021. p. 2–5. Available from: https://www.sigmaaldrich.com/technical-documents/protocols/biology/cell-culture/poor-cell-growth.html

161. Ryan JA, Ph D. Corning Guide for Identifying and Correcting Common Cell Growth Problems Technical Bulletin [Internet]. Available from: https://beta-static.fishersci.com/content/dam/fishersci/en_US/documents/programs/scientific/technical-documents/technical-bulletins/corning-guide-identifying-cell-growth-problems-technical-bulletin.pdf

162. Peterz MEG. Temperature in agar plates and its influence on the results of quantitative microbiological food analyses. 1991;14:59–66.

163. Ashour AO. Control of Sexual Reproduction in *Aspergillus* Species [Internet]. University of Nottingham; 2014. Available from:

http://eprints.nottingham.ac.uk/13987/1/ADEL_THESIS_%28FINAL_VERSION%29.pdf

164. Lide DR. CRC Handbook ofChemistry and Physics. CRC Handbook of Chemistry and Physics. 2005. 1089–1342 p.

165. Prielhofer R, Cartwright SP, Graf AB, Valli M, Bill RM, Mattanovich D, et al. Pichia pastoris regulates its gene-specific response to different carbon sources at the transcriptional, rather than the translational, level. BMC Genomics [Internet]. 2015 Mar 11;16(1):167. Available from: http://www.ncbi.nlm.nih.gov/pubmed/25887254

166. Mattanovich D, Sauer M, Gasser B. Yeast biotechnology: Teaching the old dog new tricks. Microb Cell Fact. 2014;13(1):1–5.

167. Mira NP, Teixeira MC, Sá-Correia I. Adaptive response and tolerance to weak acids in *Saccharomyces cerevisiae*: A genome-wide view. Omi A J Integr Biol. 2010;14(5):525–40.

168. Lindberg L, Santos AX, Riezman H, Olsson L, Bettiga M. Lipidomic Profiling of *Saccharomyces cerevisiae* and *Zygosaccharomyces bailii* Reveals Critical Changes in Lipid Composition in Response to Acetic Acid Stress. Oresic M, editor. PLoS One [Internet]. 2013 Sep 4;8(9):e73936. Available from: https://dx.plos.org/10.1371/journal.pone.0073936

169. Carmelo V, Santos H, Sá-Correia I. Effect of extracellular acidification on the activity of plasma membrane ATPase and on the cytosolic and vacuolar pH of *Saccharomyces cerevisiae*. Biochim Biophys Acta - Biomembr. 1997;1325(1):63–70.

170. Fernandes AR, Mira NP, Vargas RC, Canelhas I, Sá-Correia I. *Saccharomyces cerevisiae* adaptation to weak acids involves the transcription factor Haa1p and Haa1p-regulated genes. Biochem Biophys Res Commun. 2005;337(1):95–103.

171. Keller G, Ray E, Brown PO, Winge DR. Haa1 , a Protein Homologous to the Copper-regulated Transcription Factor Ace1 , Is a Novel Transcriptional Activator *. 2001;276(42):38697–702.

172. Mira NP, Becker JD, Sá-Correia I. Genomic expression program involving the Haa1p-regulon in *Saccharomyces cerevisiae* response to acetic acid. Omi A J

Integr Biol. 2010;14(5):587–601.

173. Swinnen S, Henriques SF, Shrestha R, Ho PW, Correia IS, Nevoigt E. Improvement of yeast tolerance to acetic acid through Haa1 transcription factor engineering : towards the underlying mechanisms. Microb Cell Fact. 2017;1–15.

174. Cunha JT, Costa CE, Ferraz L, Romaní A, Johansson B, Sá-correia I. HAA1 and PRS3 overexpression boosts yeast tolerance towards acetic acid improving xylose or glucose consumption : unravelling the underlying mechanisms. 2018;

175. Prielhofer R, Barrero JJ, Steuer S, Gassler T, Zahrl R, Baumann K, et al. GoldenPiCS: A Golden Gate-derived modular cloning system for applied synthetic biology in the yeast *Pichia pastoris*. BMC Syst Biol. 2017;11(1):1–14.

176. Buchman C, Skroch P, Welch J, Fogel S, Karin M. The CUP2 gene product, regulator of yeast metallothionein expression, is a copper-activated DNA-binding protein. Mol Cell Biol [Internet]. 1989;9(9):4091–5. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=362479&tool=pmcentrez&rendertype=abstract

177. Byrne KP, Wolfe KH. The Yeast Gene Order Browser: Combining curated homology and syntenic context reveals gene fate in polyploid species. Genome Res. 2005;15(10):1456–61.

178. Inaba T, Watanabe D, Yoshiyama Y, Tanaka K, Ogawa J, Takagi H, et al. An organic acid-tolerant HAA1-overexpression mutant of an industrial bioethanol strain of *Saccharomyces cerevisiae* and its application to the production of bioethanol from sugarcane molasses. AMB Express. 2013;3:1–7.

179. Fletcher E, Feizi A, Bisschops MMM, Hallström BM. Evolutionary engineering reveals divergent paths when yeast is adapted to di ff erent acidic environments. Metab Eng [Internet]. 2017;39(September 2016):19–28. Available from: http://dx.doi.org/10.1016/j.ymben.2016.10.010

APPENDIX

Appendix A – Published papers

Paes, B.G., Steindorff, A.S., Formighieri, E.F. *et al.* Physiological characterization and transcriptome analysis of *Pichia pastoris* reveals its response to lignocellulose-derived inhibitors. *AMB Expr* **11,** 2 (2021). https://doi.org/10.1186/s13568-020-01170-9

Paes *et al. AMB Expr* (2021) 11:2
https://doi.org/10.1186/s13568-020-01170-9

AMB Express

ORIGINAL ARTICLE Open Access

Physiological characterization and transcriptome analysis of *Pichia pastoris* reveals its response to lignocellulose-derived inhibitors

Barbara G. Paes[1,2], Andrei Stecca Steindorff[1], Eduardo F. Formighieri[1], Ildinete Silva Pereira[1,2] and João Ricardo M. Almeida[1,3*]

Abstract

The negative effects of lignocellulose-derived inhibitors such as acetic acid and furaldehydes on microbial metabolism constitute a significant drawback to the usage of biomass feedstocks for the production of fuels and chemicals. The yeast *Pichia pastoris* has shown a great biotechnological potential for producing heterologous proteins and renewable chemicals. Despite its relevance, the performance of *P. pastoris* in presence of lignocellulose-derived inhibitors remains unclear. In this work, our results show for the first time the dose-dependent response of *P. pastoris* to acetic acid, furaldehydes (HMF and furfural), and sugarcane biomass hydrolysate, both at physiological and transcriptional levels. The yeast was able to grow in synthetic media with up to 6 $g.L^{-1}$ acetic acid, 1.75 $g.L^{-1}$ furaldehydes or hydrolysate diluted to 10% (v/v). However, its metabolism was completely hindered in presence of hydrolysate diluted to 30% (v/v). Additionally, the yeast was capable to co-consume acetic acid and glucose. At the transcriptional level, *P. pastoris* response to lignocellulose-derived inhibitors relays on the up-regulation of genes related to transmembrane transport, oxidoreductase activities, RNA processing, and the repression of pathways related to biosynthetic processes and central carbon metabolism. These results demonstrate a polygenetic response that involves detoxification activities, and maintenance of energy and cellular homeostasis. In this context, *ALD4, OYE3, QOR2, NTL100, YCT1,* and *PPR1* were identified as target genes to improve *P. pastoris* tolerance. Altogether, this work provides valuable insights into the *P. pastoris* stress tolerance, which can be useful to expand its use in different bioprocesses.

Keywords: *Komagataella phaffii*, *Pichia pastoris*, Inhibitors, Lignocellulosic hydrolysate, Acetic acid, Furaldehydes

Introduction

Lignocellulosic biomass is an abundant raw material that can be converted by physicochemical and microbial processes into different products, such as biofuels, building-block chemicals, and high added-value chemicals (Anwar et al. 2014; Paes and Almeida 2014). Before microbial fermentation, the biomass needs to undergo pretreatment and hydrolysis to release the monosaccharides present in the biomass. During pretreatment, compounds that inhibit microbial metabolism are also released or formed during dehydration of pentoses and hexoses, hemicellulose deacetylation, or lignin breakdown (Almeida et al. 2007; Jönsson and Martín 2016). These inhibitors can be classified into three main groups: furaldehydes, such as 2-furaldeyde (furfural) and 5-hydroxymethyl-2-furaldehyde (HMF), weak acids (acetic acid, formic acid, and levulinic acid), and phenolic compounds (vanillin, syringaldehyde, coniferyl aldehyde,

*Correspondence: joao.almeida@embrapa.br
[1] Laboratory of Genetics and Biotechnology, Embrapa Agroenergia, Parque Estação Biológica, PqEB – W3 Norte Final s/no, Brasília, DF 70.770-901, Brazil
Full list of author information is available at the end of the article

Appendix B – Supplementary material

B1 - Supplementary Tables

Supplementary tables can be downloaded here.

B2 - *Pichia pastoris* (*Komagataella phaffii*) *HAA1* complete sequence

ATGGTGCTTATAAACGGCGTCAAATACGCCTGTGAACGATGTATTCGTGGACAT
AGGGTGTCTGCTTGTACTCATACTGATCAACCGTTAGTAATGATAAAGCCTAAGG
GGAGACCGTCTAGCCAATGTGCATTTTGCCGTGAACAAAGGAAGTTGCGGAAC
CATCATGCTAAGTGTACTTGCGGTAAGAAGACAAACACAAAGATGAAATACCAT
GAAAATGGTTGTGCCTGTTCCAATGATGTAGGCACTGTTAAACTTGAGGGCCAC
TGTACCTGTTGTCACCCAAAGAAGAAAAAGTCACCACCGGAAAAAGAAAAATTA
GATGAGCAGGATAAGACTAAAAACCCTGGCCACAAACTGAAAACTGACACAAAG
AATGGAGACGATTTGTTCCTTGAACAGCAACCCATTAATTCCGAGTTAGTTTCTT
GGGATATGGGACTCTCGCCCTCCGTCTCAACATCTATAATGGAATATCCGTCGT
CAACTAATTTATTAGCCCAGGTCGAGAAAAGCATGAGTTCTGGATCTGCAGTATC
CGGGCCTCGTCAAGTTGGAATGGAGCAGTTGGATAACTTTCATAGAGTCGGGG
ACGATACGAAACATGCCGGGGAGATTTCAGTTCCTTTGACCGAGTATACCCAAC
CACTACTCTCCATGGACCAGAACCTTAATCATCTACTAAGAGATAGGAAGCTGG
ATGATGCATTGTCGTCGGCTTCTCGTGAAGCTTCCAGAAGTGCCTCTCGAGCCT
CTAGAAGGGCGGGCGATAACGAAGAGACCCCAAACACCTATCAGAATATTCCGT
ATCCCTTTTCCTCAGGAGGATTGTTGGAAATGCTTAACGAGGAAAAATCCTACAA
GTCGGAGATACATTCATCCCGTACGGCAAGTAAACAGCAACAATATCCTCTTCA
GCCACAATCCACTTGGAACAGTGGAGGTAATACGAGTACGAACACAAACGGAG
CTGCTTCGTTTGGTCCCGAGGAAGTTTTTCCTCTGTATCCTTTAATTGGCCCTGG
TGGTACCCCCGATTATGAACCTTACACATCCCCTCCCTCTAACTTACAAGCGAAC
TTAACTGGTTCGAGTCATATTAGCTCAAAATCCGGGACCTACTCCAGCCATGCA
CCACACCATCATGGACATCATCACCATCACCATTTGATGCATCAGTTTACACCGT
ACTCAGCTCCAAAACCTGCTAATAGTATCCATTCCAGCTCCAGCTCCGTTAGTAT
TGGAGCCCAAGTTCCTGGTTCTGTATCTCCAAGTCTTTCCTCTCAACGTTCGTTC
CGCAGCATGGACAGTACGGATTCCTTGGTAGGGTTCCAGCTGAACAATACAAAA
ACGAACGCTACTCTAACAAGGCTTGACGAAGAATTTGGTGATATGCCGATGATG
AAACCGGTAGCTACAGTTGTCATGGATGACATCCTTGCTGGACCTTCTAATATAG
GTGAAAATGGAAATCTACCAGGTAATAATGAAACTGGAACTTTGTCAGAGGAGC
AGCCAGATGATTTCTTCTTTGGCCTAGATATGGGGGTCCCATTGAACCGTTAA

B3 – *Pichia pastoris* (*Komagataella phaffii*) mutated *HAA1* sequence

Mutation points are in lowercase, highlighted in red.

ATGGTGCTTATAAACGGCGTCAAATACGCCTGTGAACGATGTATTCGTGGACAT
AGGGTGTCTGCTTGTACTCATACTGATCAACCGTTAGTAATGATAAAGCCTAAGG
GGAGACaGGGAGACCcGC**g**GGTCTC**c**ACaGTCTAGCCAATGTGCATTTTGCCGT
GAACAAAGGAAGTTGCGGAACCATCATGCTAAGTGTACTTGCGGTAAGAgGGGA
GACCcGCgGGTCTCcGA**g**GACAAACACAAAGATGAAATACCATGAAAATGGTTGT
GCCTGTTCCAATGATGTAGGCACTGTTAAACTTGAGGGCCACTGTACCTGTTGT
CACCCAAAGAAGAAAAAGTCACCACCGGAAAAAGAAAAATTAGATGAGCAGGAT
AAGACTAAAAACCCTGGCCACAAACTGAAAACTGACACAAAGAATGGAGACGAT
TTGTTCCTTGAACAGCAACCCATTAATTCCGAGTTAGTTTCTTGGGATATGGGAC
TCTCGCCCTCCGTCTCAACATCTATAATGGAATATCCGTCGTCAACTAATTTATT
AGCCCAGGTCGAGAAAAGCATGAGTTCTGGATCTGCAGTATCCGGGCCTCGTC
AAGTTGGAATGGAGCAGTTGGATAACTTTCATAGAGTCGGGGACGATACGAAAC
ATGCCGGGGAGATTTCAGTTCCTTTGACCGAGTATACCCAACCACTACTCTCCA
TGGACCAGAACCTTAATCATCTACTAAGAGATAGGAAGCTGGATGATGCATTGT
CGTCGGCTTCTCGTGAAGCTTCCAGAAGTGCCTCTCGAGCCTCTAGAAGGGCG
GGCGATAACGAAGAGAtCGGAGACCcGCgGGTCTCcGA**t**CCCAAACACCTATCAG
AATATTCCGTATCCCTTTTCCTCAGGAGGATTGTTGGAAATGCTTAACGAGGAAA
AATCCTACAAGTCGGAGATACATTCATCCCGTACGGCAAGTAAACAGCAACAAT
ATCCTCTTCAGCCACAATCCACTTGGAACAGTGGAGGTAATACGAGTACGAACA
CAAACGGAGCTGCTTCGTTTGGTCCCGAGGAAGTTTTTCCTCTGTATCCTTTAAT
TGGCCCTGGTGGTACCCCCGATTATGAACCTTACACATCCCCTCCCTCTAACTT
ACAAGCGAACTTAACTGGTTCGAGTCATATTAGCTCAAAATCCGGGACCTACTC
CAGCCATGCACCACACCATCATGGACATCATCACCATCACCATTTGATGCATCA
GTTTACACCGTACTCAGCTCCAAAACCTGCTAATAGTATCCATTCCAGCTCCAGC
TCCGTTAGTATTGGAGCCCAAGTTCCTGGTTCTGTATCTCCAAGTCTTTCCTCTC
AACGTTCGTTCCGCAGCATGGACAGTACGGATTCCTTGGTAGGGTTCCAGCTGA
ACAATACAAAAACGAACGCTACTCTAACAAGGCTTGACGAAGAATTTGGTGATAT
GCCGATGATGAAACCGGTAGCTACAGTTGTCATGGATGACATCCTTGCTGGACC
TTCTAATATAGGTGAAAATGGAAATCTACCAGGTAATAATGAAACTGGAACTTTG
TCAGAGGAGCAGCCAGATGATTTCTTCTTTGGCCTAGATATGGGGGTCCCATTG
AACCGTTA