

Engineering *Kluyveromyces marxianus* as an Acid-Tolerant Platform for Sustainable
Biomanufacturing

Mackenzie Thornbury

A Thesis

In the Department

of

Biology

Presented in Partial Fulfillment of the Requirements

For the Degree of

Doctor of Philosophy (Biology)

At Concordia University

Montreal, Quebec, Canada

October 2025

© Mackenzie Thornbury, 2025

**CONCORDIA UNIVERSITY
SCHOOL OF GRADUATE STUDIES**

This is to certify that the thesis prepared

By: Mackenzie Thornbury

Entitled: Engineering *Kluyveromyces marxianus* as an Acid-Tolerant Platform for Sustainable
Biomanufacturing

and submitted in partial fulfillment of the requirements for the degree of

Doctor Of Philosophy Biology

complies with the regulations of the University and meets the accepted standards with respect to
originality and quality.

Signed by the final examining committee:

_____ Chair

Dr. David Walsh

_____ External Examiner

Dr. David Stuart

_____ External to the Program

Dr. Paul Joyce

_____ Examiner

Dr. Christopher Brett

_____ Examiner

Dr. Isabelle Benoit Gelber

_____ Thesis Supervisor

Dr. Vincent Martin

Approved by _____

Dr. Robert Weladji, Graduate Program Director

2025-10-31

Date of Defense

Dr. Pascale Sicotte, Dean, Faculty of Arts and Science

Abstract

Engineering *Kluyveromyces marxianus* as an Acid-Tolerant Platform for Sustainable Biomanufacturing

Mackenzie Thornbury, Ph.D

Concordia University, 2025

The yeast *Kluyveromyces marxianus* is an attractive candidate for industrial bioproduction due to its rapid growth rate, thermotolerance, and ability to utilize a wide range of substrates, including lactose from dairy byproducts. However, its use in metabolic engineering has been limited by underdeveloped genetic tools. This thesis addresses this gap by developing these tools with a focus on enabling fumaric acid biosynthesis as a model pathway for organic acid biomanufacturing.

To build a foundation for engineering, we developed and characterized a standardized promoter library compatible with the Yeast Toolkit system and optimized CRISPR-based editing methods, allowing precise and efficient control of gene expression and genome modification. These tools enabled the development of CRISPR activation, interference and deletion system in *K. marxianus* which facilitated a genome-wide study of fumaric acid tolerance, where a pooled CRISPR screen identified new contributors to acid stress resistance. In particular, the DHA1-family transporters *QDR2* and *QDR3* were found to improve growth in fumaric acid, underscoring the role of transport in tolerance mechanisms.

Metabolic engineering strategies were then applied to increase fumaric acid production. Deletion of *FUM1* enabled fumarate accumulation via the TCA cycle, while a modified reductive TCA pathway from *Rhizopus oryzae* further improved yields. When combined with *QDR2* and

QDR3 co-overexpression, fumaric acid titers doubled relative to the base strain. Additional interventions, including heterologous glutamate dehydrogenase expression, provided modest improvements in redox balance and by-product reduction, though acetate accumulation persisted as a major challenge.

Overall, this work establishes enabling genetic engineering tools for *K. marxianus*, expands knowledge of its acid stress responses, and demonstrates rational design strategies for fumaric acid biosynthesis. While acetic acid overflow remains a barrier to high titers, these advances strengthen the case for *K. marxianus* as a versatile platform for bioproduction and open new opportunities for valorizing dairy byproducts.

Acknowledgments

I want to thank my supervisor Dr. Vincent Martin for his guidance and support throughout this degree. I am extremely grateful that you took me on as a doctoral student and gave me the structure to learn the foundations of metabolic engineering and synthetic biology while also allowing me the independence to pursue my own interests. I'd also like to thank Dr. Malcolm Whiteway who was a co-supervisor in practice, but not in name. Thank you for all your feedback, troubleshooting help and science history chats.

I'd also like to acknowledge all the help and feedback my committee gave me throughout the years. Thank you to Dr. Isabelle Benoit Gelber and Dr. David Kwan for your invaluable feedback. Also thank you to Dr. Christopher Brett for subbing into my committee for my defense, I really appreciate that.

Thank you to the entire Martin lab for making the lab a fun and productive place to be. Huge thank you to the KM team: Jaya, Lauren, Iain, Sydney, Adrien, James, Joseph and Raha for being great friends but also great scientists to share in wading through the beginning of working with a non-conventional organism. Special thanks to Bjorn, Meghan, and Mo who despite not working on the same project gave hours of their time to teach me wet lab, dry lab, data visualization, and writing techniques. I treasure your friendship and am a better scientist by a mile because of you :).

For their technical help, I'd like to thank Marcos di Falco, Lauren Narcross, Nick Gold, and Chris Dykstra for teaching and guidance with the HPLC. I'd like to thank Nick Gold, James Bagley, and Flavia Araujo from the Concordia Foundry for help with preparing the promoter library, and analysis of the CRISPR-AID genome-wide screen results. Dr. Aida Tafirishi was

critically helpful in getting CHOPCHOP running on the command line for the CRISPR-AID sgRNA and my supportive and brilliant partner Rory Coles helped with all my python trouble shooting. I'd also like to thank my extremely talented and patient sister Olivia for editing this thesis.

Thank you to the funding sources that supported my studies: Génome Québec, Genome Canada, Agropur, the FRQNT, the SynBioApps program, NSERC Vanier Canada Graduate Scholarship, and Concordia University.

Contribution of authors

Chapter 1: Introduction

Vincent Martin supervised the project. Mackenzie Thornbury wrote and researched the literature review.

Chapter 2: Domestication of *Kluyveromyces marxianus*

Vincent Martin and Malcolm Whiteway supervised the project and edited the manuscript. Mackenzie Thornbury was the primary writer of the manuscript, Adrien Knoops wrote the results section for Figures 2.4 and 2.5, James Dhaliwal wrote the results section and prepared Figure S2.6, and Joseph Christan Utomo wrote the results section for Figure 2.6. Jaya Joshi collected the data that was used in Figure 2.1. Iain Summerby-Murray prepared samples for PacBio sequencing and annotated the genome sequence. Mackenzie Thornbury designed and built the promoter library. She and Sydney Johnson collected data for Figure 2.3 and Sydney Johnson analyzed the data. Adrien Knoops created the three inducible promoters, characterized them and analyzed the data in Figure 2.4. Mackenzie Thornbury designed the guides used in Figure 2.5 and James Dhaliwal and Adrien Knoops designed the experiment, collected the data, and analyzed the results. Lauren Narcross and Joseph Christian Utomo ran the fermentations of Y-1190 and analyzed the data (Figure 2.6). Mackenzie Thornbury performed experiments and data analysis, and prepared figures.

Chapter 3: Tri-function CRISPR system to enhance fumaric acid tolerance

Vincent Martin and Malcolm Whiteway supervised the project and edited the manuscript. Mackenzie Thornbury wrote the manuscript. Adrien Knoops created the inducible promoters that were used in the CRISPR-AID strain, as well as the design of the guide plasmids for combinatorial

CRISPR system. Mackenzie Thornbury supervised Raghad Abushain and together they generated CRISPR-AID guide library. Raghad Abushahin and Mackenzie Thornbury tested and sampled the CRISPR-AID library (Figure S3.2, Figure 3.3). Raha Parvizi Omran and Lalit Kumar performed the fed-batch bioreactor experiment (Figure 3.6) and Mackenzie Thornbury performed all other experiments. Mackenzie Thornbury performed the data analysis and prepared all figures.

Chapter 4: Metabolic engineering of primary metabolism for fumaric acid production

Vincent Martin supervised the project, and Mackenzie Thornbury performed the design of experiments, the metabolite analysis, the growth assays, data analysis, figure preparation and most of the plasmid and strain construction. Raha Parvizi Omran created the rTCA^{mod} plasmid (ROK022).

Table of contents

Abstract	iii
Acknowledgments	v
Contribution of authors.....	vii
List of figures.....	xii
List of tables.....	xiii
List of supplementary figures	xiv
List of supplementary tables	xv
List of abbreviations	xvi
Chapter 1: Introduction to the thesis	1
1.1. Thesis overview	1
1.2. Organic acids as platform chemicals.....	3
<i>1.2.1. Industrial relevance of organic acids</i>	<i>3</i>
<i>1.2.2. Fumaric acid.....</i>	<i>4</i>
<i>1.2.3. Metabolic pathways to fumaric acid production</i>	<i>6</i>
1.3. Microbial hosts for organic acid production	8
<i>1.3.1. Model hosts and their limitations</i>	<i>8</i>
<i>1.3.2. Fumaric acid production by non-conventional organisms.....</i>	<i>11</i>
<i>1.3.3. Lactose as a feedstock.....</i>	<i>13</i>
1.4. Kluyveromyces marxianus as a synthetic biology chassis.....	14
<i>1.4.1. Biology and physiology of K. marxianus</i>	<i>14</i>
<i>1.4.2. Industrial relevance</i>	<i>16</i>
<i>1.4.3. Advances in genetic engineering tools.....</i>	<i>17</i>
Chapter 2: Domestication of Kluyveromyces marxianus.....	19
2.1. Abstract.....	19
2.2. Introduction.....	20
2.3. Materials and methods	22
<i>2.3.1. Kluyveromyces strain selection and media optimization.....</i>	<i>22</i>
<i>2.3.2. Transformation</i>	<i>24</i>
<i>2.3.3. Genome sequencing and bioinformatics.....</i>	<i>24</i>
<i>2.3.4. Promoter set characterization</i>	<i>26</i>
<i>2.3.5. Genome editing using CRISPR-Cas9.....</i>	<i>28</i>
<i>2.3.6. Investigation of strain haploidization</i>	<i>29</i>
<i>2.3.7. Cultivation of K. marxianus in a bioreactor.....</i>	<i>31</i>

2.4. Results	32
2.4.1. <i>K. marxianus</i> strain selection	32
2.4.2. Transformation efficiency	33
2.4.3. Genome sequence analysis	33
2.4.4. Analysis of lactose permease variation.....	35
2.4.5. Promoter set characterization	36
2.4.6. CRISPR genome editing optimization.....	39
2.4.7. Testing of diploid stability	42
2.4.8. Optimization of <i>K. marxianus</i> fermentation in permeate.....	42
2.5. Discussion	45
2.6. Supplementary figures and tables	50
Chapter 3: Tri-functional CRISPR system to enhance fumaric acid tolerance.....	68
3.1. Abstract.....	68
3.2. Introduction.....	69
3.3. Materials and methods	71
3.3.1. Plasmids, strains, and growth media	71
3.3.2. <i>K. marxianus</i> transformations and CRISPR-Cas9 Editing.....	71
3.3.3. Establishing CRISPR-AID in <i>K. marxianus</i>	72
3.3.4. Genome-wide CRISPR-AID sgRNA library design and construction	73
3.3.5. CRISPR-AID sgRNA enrichment analysis	77
3.3.6. Validation of CRISPR-AID hits	78
3.3.7. Growth rate analysis.....	78
3.3.8. Metabolite analysis	79
3.3.9. Fed-batch cultivation	80
3.3.10. Protein alignments	81
3.4. Results	81
3.4.1. Creation of a fumaric acid producing <i>K. marxianus</i>	81
3.4.2. Validation of CRISPR-AID in <i>K. marxianus</i>	82
3.4.3. CRISPR-AID library screen.....	85
3.4.4. Validation of hits from CRISPR-AID enrichment.....	87
3.4.5. Fumaric acid production	90
3.5. Discussion.....	94
3.6. Supplementary figures and tables	101
Chapter 4: Metabolic engineering of primary metabolism for fumaric acid production..	121

4.1. Abstract.....	121
4.2. Introduction.....	122
4.3. Materials and methods	126
4.3.1. <i>Strains, media, and plasmids</i>	<i>126</i>
4.3.2. <i>Heterologous genes.....</i>	<i>127</i>
4.3.3. <i>K. marxianus transformation.....</i>	<i>127</i>
4.3.4. <i>Growth assays.....</i>	<i>128</i>
4.3.5. <i>Metabolite analysis</i>	<i>129</i>
4.3.6. <i>RT-qPCR.....</i>	<i>129</i>
4.4. Results	130
4.4.1. <i>Deletion of putative carboxylic acid importers increase fumaric acid titers</i>	<i>130</i>
4.4.2. <i>Engineering the TCA cycle for enhanced fumaric acid production</i>	<i>133</i>
4.4.3. <i>Redox balancing fumaric acid production strains to optimize titers</i>	<i>139</i>
4.4.4. <i>Reducing carbon loss by redirecting flux around the pyruvate node</i>	<i>144</i>
4.5. Discussion.....	146
4.6. Supplemental tables and figures.....	150
Chapter 5: Conclusion and Perspectives	155
5.1. Preamble	155
5.2. <i>K. marxianus</i> as a biomanufacturing chassis.....	156
5.2.1. <i>Industrial relevance</i>	<i>156</i>
5.2.2. <i>Dual nature of diploidy in <i>K. marxianus</i> Y-1190.....</i>	<i>158</i>
5.3. CRISPR-AID as a functional genomics tool.....	159
5.3.1. <i>Establishing CRISPR-AID in <i>K. marxianus</i>.....</i>	<i>159</i>
5.3.2. <i>Expanding CRISPR-AID utility beyond growth-coupled screens in <i>K. marxianus</i></i>	<i>161</i>
5.4. Metabolic engineering insights and future directions	162
5.4.1. <i>Decreasing acetic acid through metabolic engineering</i>	<i>162</i>
5.5. Conclusion	164
Bibliography	166

List of figures

Figure 1.1. Metabolic pathways involved in fumaric acid production.	7
Figure 1.2. Common strategies for fumaric acid production	9
Figure 2.1. Growth rate comparison of aerobic growth of 19 <i>K. marxianus</i> strains	33
Figure 2.2. Alignment of Lac12 amino acid sequences across <i>K. marxianus</i> strains.....	36
Figure 2.3. Characterization of promoter strength across media.	38
Figure 2.4. Characterization of aTc-inducible promoters.	39
Figure 2.5. CRISPR-Cas9 genome-editing efficiencies in <i>K. marxianus</i>	41
Figure 2.6. Fermentation Profile of <i>K. marxianus</i> Y-1190.....	44
Figure 2.7. Promoter comparison to Lang <i>et al.</i> 2020.	47
Figure 3.1. Effect of <i>FUM1</i> deletion on fumaric acid production and tolerance.....	82
Figure 3.2. CRISPR-AID functionality in <i>K. marxianus</i>	83
Figure 3.3. Library workflow and summary statistics.	86
Figure 3.4. Validation of CRISPR-AID library hits in Y-1190 Δ <i>FUM1</i> strain.....	89
Figure 3.5. Effect of <i>QDR2</i> and <i>QDR3</i> on fumaric acid production in Y-1190 Δ <i>FUM1</i>	91
Figure 3.6. Growth and metabolite profiles of <i>K. marxianus</i> Y-1190.....	93
Figure 3.7. Alignment of <i>K. marxianus</i> Qdr2p and Qdr3p.....	98
Figure 4.1. Strategies for engineering central metabolism for fumaric acid production.	124
Figure 4.2. PPP engineering strategy for enhancing NADH pools.....	125
Figure 4.3. Elucidation of putative carboxylic transports in fumaric acid import and production.	132
Figure 4.4. <i>R. oryzae</i> rTCA expression on fumaric acid production.	134
Figure 4.5. Effect of modified rTCA on fumaric acid production in <i>QDR2</i> overexpression strain.	136
Figure 4.6. Metabolite and growth profiles for rTCA ^{mod} in MTK130hom background.....	138
Figure 4.7. Effect of NADH recycling on fumaric acid production.	141
Figure 4.8. Exploring ScGDH enzyme expression in enhancing fumaric acid and reducing by- products.....	143
Figure 4.9. Effect of Overexpression of <i>KmPCK</i> and <i>ScPCK</i> on fumaric acid titer.	145

List of tables

Table 1.1. Comparison of model organism engineering for fumaric acid production.....	10
Table 2.1. Dairy permeate dry composition analyzed by Inductively Coupled Plasma Spectroscopy conducted by Mérieux Nutrisciences, Dorval, Québec	23
Table 2.2. <i>K. marxianus</i> Y-1190 summary assembly	34
Table 3.1. Summary Coverage of the <i>K. marxianus</i> CRISPR-AID Library.....	74
Table 3.2. Significantly upregulated sgRNAs from CRISPR-AID Screen	87

List of supplementary figures

Figure S2.1. Growth rates of 19 strains of <i>K. marxianus</i>	50
Figure S2.2. Transformation efficiency of our top growing strains of <i>K. marxianus</i> using a modified Gietz transformation method.	51
Figure S2.3. k-mer frequency of PacBio raw reads. k-mer coverage of Y-1190 using Jellyfish and Genome Scope	51
Figure S2.4. Relative promoter expression in <i>K. marxianus</i>	52
Figure S2.5. CRISPR-Cas9 genome editing efficiencies using one gRNA (sequence 2) for genomic deletion and integration via HDR.	53
Figure S2.6. Ploidy and mating type analysis.	54
Figure S2.7. Growth curves of Y-1190 and homologous diploid JD044	55
Figure S2.8. Primary and haplotig sequence comparison for two example genes	55
Figure S3.1. Transformation efficiency with CRISPR-AID system.	101
Figure S3.2. Testing fumaric acid concentrations to be used in CRISPR-AID enrichment experiments	101
Figure S3.3. Fumaric acid titers in SC-Lactose and 0.5x Permeate with QDR2 and QDR3 overexpressed.	102
Figure S3.4. Production of acetic acid by QDR2 and QDR3 overexpression strains.	102
Figure S3.5 Fed-batch fermentation profiles of <i>K. marxianus</i> Y-1190.	103

List of supplementary tables

Table S2.1. Genome Contig Size.....	56
Table S2.2. List of promoters	56
Table S2.3. List of yeast strains used in this study	57
Table S2.4. List of primers and synthetic sequences	58
Table S2.5. List of plasmids used in this study.....	63
Table S3.1. List of plasmids used in this study.....	104
Table S3.2. List of oligonucleotides used in this study	108
Table S3.3. List of yeast strains used in this study	112
Table S3.4. sgRNA scoring rubric.....	113
Table S3.5. List of sgRNA used in the CRISPR-AID genome-wide guide library.....	114
Table S3.6. Genome-wide CRISPR-AID library coverage	114
Table S3.7. Library sequencing read coverage per condition.....	115
Table S3.8. Significantly enriched sgRNA from the genome-wide CRISPR-AID screen in fumaric acid.	115
Table S4.1. List of strains used in this study	150
Table S4.2. List of plasmids used in this study.....	151
Table S4.3. List of oligonucleotides and sgRNA used in this study.....	153

List of abbreviations

α KG: α -ketoglutarate
 α KGD: α -ketoglutarate dehydrogenase
ACH1: acetyl-CoA hydrolase
ADP: adenosine diphosphate
ATP: adenosine triphosphate
ATP7: ATP synthase, subunit d
AMP: adenosine monophosphate
ARS: Autonomously replicating sequences
CRISPR: Clustered regularly interspaced short palindromic repeats
DCW: Dry cell weight
DNA: Deoxyribonucleic acid
FBA: Flux balance analysis
GDH: Glutamate dehydrogenase
GRAS: Generally recognized as safe
LSC: Ligase of Succinyl-CoA
OAA: Oxaloacetic acid
PCK: Phosphoenolpyruvate carboxykinase (gene)
PCR: Polymerase chain reaction
PEP: Phosphoenolpyruvate
PEPCK: Phosphoenolpyruvate carboxykinase (enzyme)
PPP: Pentose phosphate pathway
QDR: Quinidine resistance
QPS: Qualified presumption of safety
RNA: ribonucleic acid
rTCA: Reductive tricarboxylic acid
SAM: S-adenosyl methionine
SDH: Succinate dehydrogenase
SNP: Single nucleotide polymorphisms

TCA: Tricarboxylic acid

1. Chapter 1: Introduction to the thesis

1.1. Thesis overview

This thesis explores the genetic engineering of the non-conventional yeast *K. marxianus* to enable its use as a robust chassis for bioproduction of the organic acid fumaric acid as a target for valorization from the dairy waste product, milk permeate. My first objective is to expand the genetic engineering toolkit for *K. marxianus*, including standardizing and characterizing a modular promoter library compatible with the Yeast Toolkit (MoClo) system and enhancing CRISPR-based genetic engineering tools. My second objective is to leverage these tools to enhance acid tolerance through a genome-wide CRISPR guide enrichment screen in fumaric acid, with the aim of identifying fumaric acid-specific stress responses that contribute to the acid tolerance of *K. marxianus*. Once a fumaric acid tolerant strain is built, my third objective is to enhance fumaric acid yields through rational metabolic engineering.

Chapter 1 provides the background necessary to understand the context and motivation for this work. It introduces organic acids as platform chemicals, with a particular focus on fumaric acid and its production in microbial hosts. The chapter also reviews the biology, physiology, and industrial relevance of *K. marxianus*, highlighting its potential as a non-conventional yeast chassis. Finally, it surveys the current genetic tools available for *K. marxianus*, establishing the foundation for the engineering strategies presented in later chapters.

Chapter 2 focuses on the development and characterization of synthetic biology tools in the dairy isolate *K. marxianus* Y-1190. This includes the construction of a modular promoter library optimized for gene expression during growth on lactose. In parallel, a high-efficiency

CRISPR-Cas9 system is established, employing a dual-guide strategy to enhance gene disruption and integration, an essential platform for downstream metabolic engineering. Additionally, the formulation of milk permeate media and bioreactor cultivation parameters are optimized to support robust growth of *K. marxianus*. Together, these resources provide a foundational toolkit for precise gene regulation and metabolic control in this emerging non-conventional yeast chassis.

Chapter 3 builds on this toolkit to present a genome-wide CRISPR-AID screen aimed at enhancing tolerance to fumaric acid stress. CRISPR activation, interference, and deletion are first validated in *K. marxianus*, marking the first report of CRISPR-mediated gene activation in this species. Genome-wide sgRNA libraries were constructed for each perturbation and screened in the presence of 25–42 mM fumaric acid to enrich for guides conferring improved fumaric acid tolerance. Among the top ten significant hits, three were validated in a CRISPR-Cas free background: activation of *QDR2* and *QDR3*, and deletion of *ATP7*. Overexpression of *QDR2* and *QDR3* in combination led to a 7.2-fold increase in fumaric acid yield. This chapter reveals novel gene targets that enhance tolerance to fumaric acid and demonstrates the utility of genome-wide functional screens in accelerating strain improvement.

Chapter 4 employs rational metabolic engineering to redirect carbon flux and enhance fumaric acid production. By identifying a fumaric acid importer, introducing a heterologous rTCA pathway, implementing NADH balancing strategies, and leveraging insights from the CRISPR-AID enrichment study, we achieved further improvements in fumaric acid titers. Together, these efforts illustrate a multi-layered and integrated approach to strain development.

Overall, this work establishes *K. marxianus* as a versatile platform for fumaric acid production by combining tool development, functional genomics, and rational metabolic engineering. Foundational resources, including a publicly available annotated genome of *K.*

marxianus Y-1190, a modular promoter library, a CRISPR-Cas9 system and optimized growth conditions, enabled precise genetic control. Genome-wide CRISPR-AID screening identified key genes enhancing both fumaric acid tolerance and production, while targeted pathway engineering, including a heterologous rTCA pathway, NADH balancing, and a fumaric acid importer, further improved titers. Together, these efforts create a multi-layered strategy that begins to unlock the biotechnological potential of *K. marxianus* for sustainable biomanufacturing of fumaric acid.

1.2. Organic acids as platform chemicals

1.2.1. Industrial relevance of organic acids

Organic acids hold substantial industrial relevance due to their diverse applications across food, pharmaceutical, and polymer sectors. Microbial fermentation and biomanufacturing have become leading methods for producing key organic acids sustainably and at scale. For example, *Aspergillus niger* is widely employed to produce citric acid, a critical acidulant, flavouring agent, and preservative in the food and beverage industries^{1,2}. Another organic acid, lactic acid, is produced by *Lactobacillus* species and recombinant yeast^{3,4}. It is primarily used as a food preservative and pH regulator but also increasingly as a monomer for manufacturing polylactic acid, a biodegradable plastic used in packaging, agricultural films, and biomedical devices⁵⁻⁷.

Microbial manufacturing of organic acids initially relied on naturally occurring strains selected for their innate capacity to overproduce the acid of interest, such as *A. niger* for citric acid and *Lactobacillus* species for lactic acid described above. Over time, advances in metabolic engineering and synthetic biology have made it possible to reprogram a wider range of microbes to achieve higher yields, improved tolerance to acid accumulation, and utilization of diverse feedstocks. For example, while succinic acid was historically produced through petrochemical routes, recombinant strains of yeast such as *Saccharomyces cerevisiae*⁸, and *Candida krusei*⁹, as

well as the recombinant bacterium *Basfia succiniciproducens*¹⁰, have been developed to convert renewable sugars into succinic acid efficiently, enabling commercial biobased production¹¹. This transition from relying on native producers to designing custom microbial cell factories has broadened the economic and environmental feasibility of organic acid biomanufacturing.

Looking ahead, we aim to expand the biomanufacturing of organic acids by developing efficient bioprocesses for fumaric acid production, which remains largely petrochemical today¹². Although advances in strain engineering have enabled promising yields for several acids, many challenges remain unresolved: product toxicity, low titers, and downstream recovery will require continued innovation to realize sustainable, large-scale bioproduction.

1.2.2. *Fumaric acid*

Fumaric acid is a four-carbon dicarboxylic acid that is an intermediate in the tricarboxylic acid (TCA) cycle. Beyond its central metabolic role, fumaric acid has attracted attention for its industrial relevance and potential as a renewable bioproduct. In an effort to focus on economic growth towards biomanufacturing, the U.S. Department of Energy published the report "Top Value Added Chemicals from Biomass" in 2004¹³. The report identified 12 chemicals that were considered valuable not only at their base level, but also by the chemical's ability to produce families of similar derivatives. Fumaric acid met these criteria and was identified as one of these top value added chemicals¹³. Its many applications in resin production, food and beverage additives and as a plasticizer reflect its value as a versatile platform chemical.

Fumaric acid is used primarily in resins (56 %), food and beverage additives (22 %), and plasticizers (5 %)¹⁴. While the majority of fumaric acid goes towards the making of paper resins, it is a key monomer in unsaturated polyester resins and alkyd resins used to manufacture fiberglass-

reinforced plastics, coatings, and composite materials for automotive, construction, marine, and electric vehicle applications¹⁴. In resin polymerization, maleic anhydride is often used as a cheaper option, however fumaric acid is gaining popularity due to its non-toxic nature and its ability to form resins with greater polymer hardness^{12,14}.

In the food and beverage industry, fumaric acid is widely used as both an acidulant and a preservative¹⁴. Compared to citric acid, one of the most common acidulants, fumaric acid is approximately 1.5 times more acidic, delivering greater sourness and flavor intensity per unit of weight¹⁵. Its strong acidifying properties not only enhance tartness but also help stabilize flavour profiles and inhibit microbial growth. For these reasons, it is commonly incorporated into beverages, baked goods, confectioneries, and gelatin-based desserts to improve taste, extend shelf life, and maintain product quality^{14,15}.

Outside of food and resins, fumaric acid has also found utility in medicine. Derivatives of fumaric acid, specifically esters, are used to treat the immune-mediated diseases multiple sclerosis and psoriasis¹². The U.S. Food and Drug Administration (FDA) has approved fumaric acid derivatives to treat relapsing multiple sclerosis, although the mechanism of action of these esters is still unclear. While the FDA has not yet approved fumarate esters for the treatment of psoriasis, they are authorized for this use in Germany and have been occasionally prescribed off-label in the United States^{16,17}. In the case of psoriasis, the fumarate esters are thought to act by switching the immune profiles of T helper cells from the more pro-inflammatory Th1 response to the less inflammatory Th2 response¹⁸.

Fumaric acid is an attractive target for biomanufacturing due to its wide range of industrial applications and potential for sustainable production. The U.S. Department of Energy has

identified key barriers to bio-based fumaric acid production, including the need to improve microbial catalysts to increase yield and productivity while reducing unwanted coproducts¹³. Developing efficient microbial platforms for fumaric acid production could enable cost-effective and scalable bioprocesses, further driving its adoption as a sustainable chemical building block. The following section reviews current microbial technologies and engineering strategies aimed at overcoming these limitations.

1.2.3. *Metabolic pathways to fumaric acid production*

Fumaric acid is a central intermediate of cellular metabolism, connecting energy generation with biosynthetic pathways. In *S. cerevisiae* and *Escherichia coli*, it can be synthesized through several distinct routes that intersect primary carbon and nitrogen metabolism (Fig. 1.1). As an intermediate of the TCA cycle, fumaric acid can be created by both oxidative and reductive pathways. Through the oxidative route acetyl-CoA enters the TCA cycle by condensing with oxaloacetate to form citrate, which then undergoes a series of oxidative reactions that ultimately yield fumaric acid through the oxidation of succinate¹⁹ (Fig 1.1, blue). In the reductive TCA (rTCA) pathway, oxaloacetate is reduced to malate and then dehydrated to fumarate²⁰ (Fig 1.1, red). The glyoxylate shunt is a modified branch of the TCA cycle that bypasses the decarboxylation steps, allowing cells to conserve carbon during growth on two-carbon substrates. In this pathway, isocitrate is cleaved into succinate and glyoxylate, which then combines with acetyl-CoA to form malate, linking directly to fumaric acid production^{21,22} (Fig 1.1, green). While the TCA cycles are primary methods of fumaric acid generation in the cell, two other minor pathways exist for fumaric acid generation that are both involved in nitrogen metabolism (Fig 1.1, purple). In the urea cycle, argininosuccinate is cleaved into arginine and fumarate²³, while in the purine cycle, adenylosuccinate is converted into AMP and fumarate²⁴.

A study by Zelle *et al.* compared four different routes to malate production: the rTCA, the oxidative TCA cycle, the glyoxylate shunt, and a modified glyoxylate shunt in which malate cannot be further oxidized to oxaloacetate. Among these pathways, the rTCA demonstrated the highest molar yield from glucose, producing up to 2 mol of malate per mol of glucose. This high yield is attributed in part to the net fixation of CO₂ in the rTCA, in contrast to the oxidative TCA cycle, which loses carbon as CO₂ during decarboxylation steps²⁵. Because malate is only one step away from fumaric acid in the TCA, these findings highlight the potential of the rTCA for fumaric acid production. Building on this understanding, metabolic engineering studies have explored and combined these pathways to optimize fumarate production in both yeast and bacterial hosts.

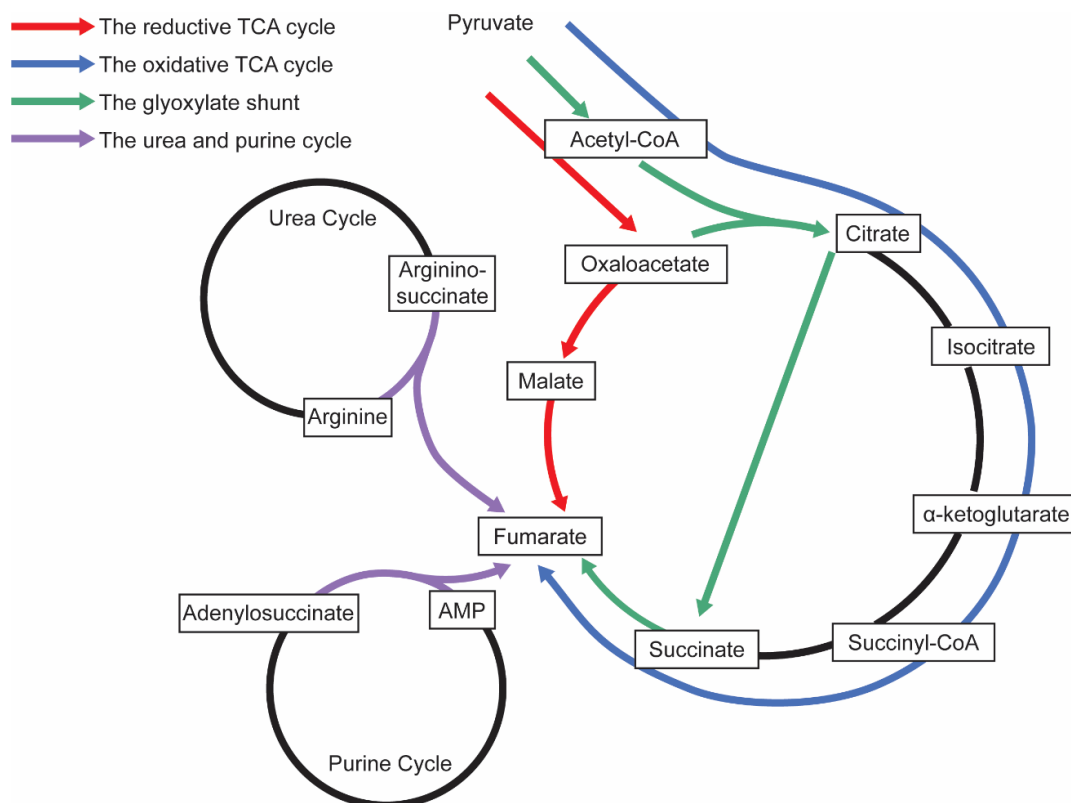


Figure 1.1 Metabolic pathways involved in fumaric acid production. AMP; adenosine monophosphate. Adapted from Chen *et al.*, *Met Eng*, 2016, <https://doi.org/10.1016/j.ymben.2015.07.007>, licensed under a [Creative Commons Attribution 4.0 International License](#).

1.3. Microbial hosts for organic acid production

1.3.1. Model hosts and their limitations

A wide range of strategies were developed to engineer microbial hosts for fumaric acid production, as summarized in Table 1.1. While *E. coli* and *S. cerevisiae* have served as the most extensively studied model organisms, both exhibit inherent limitations that can constrain large-scale, economically viable biomanufacturing processes. *E. coli* offers exceptional genetic tractability but demands extensive pathway rewiring to circumvent tight regulatory controls and minimize by-product formation, while *S. cerevisiae* and *E. coli* are both limited by feedstock options. Additionally, industrial *E. coli* strains are vulnerable to bacteriophage infection, which can rapidly spread through bioreactors and disrupt large-scale fermentation processes²⁶. These challenges underscore the need to consider non-conventional microbial hosts that naturally combine metabolic flexibility, tolerance to low pH, and diverse substrate utilization.

A clear distinction among the reported strategies is the degree of metabolic intervention required in *E. coli* to accumulate significant quantities of fumaric acid. High-titer strains, such as those developed by Song *et al.*^{27,28}, and Chen *et al.*²⁹, necessitated multiple deletions of central metabolic genes including *fumABC*, *ptsG*, *iclR*, and *arcA*, along with overexpression of carboxylation enzymes and transporters to rewire carbon flux^{27–29} (Fig. 1.2 A). Even with these modifications, process success hinged on tight control of pH and oxygenation in fed-batch systems to counteract acid accumulation and overflow metabolism. For example, Song *et al.* 2015 achieved 6.24 g L⁻¹ in a shake flask but were able to attain 35.1 g L⁻¹ in a fed-batch bioreactor with pH and oxygen control²⁸.

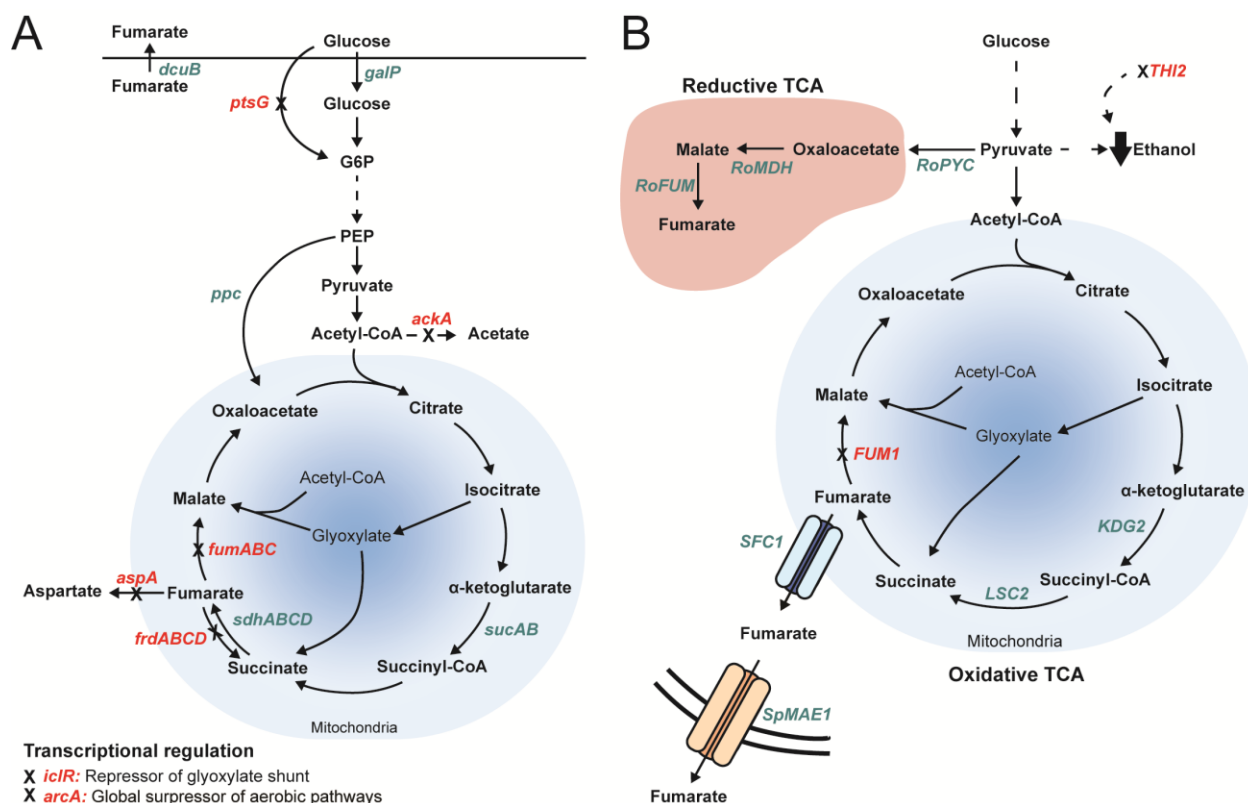


Figure 1.2 Common strategies for fumaric acid production in A. *E. coli* and B. *S. cerevisiae*. Red text represents deletion, teal text represents overexpression. Panel A is adapted from Song *et al.*, *Biotechnol Bioeng*, 2013, <https://doi.org/10.1002/bit.24868>, licensed under a [Creative Commons Attribution License 3.0](#).

In *S. cerevisiae*, a different strategy has proven effective: introducing *Rhizopus oryzae* genes encoding pyruvate carboxylase, malate dehydrogenase, and fumarase to establish an efficient rTCA branch (Fig. 1.2 B). This approach has allowed engineered yeast strains to achieve fumaric acid titers exceeding 30 g L^{-1} ³⁰ without as many deletions of endogenous pathways as seen in *E. coli*. However, even in yeast, achieving these yields still requires high-glucose feedstocks, and the organism's moderate acid tolerance can restrict productivity as titers rise. These considerations point to the broader potential of non-conventional organisms that naturally thrive in low-pH environments, grow rapidly on inexpensive or mixed carbon sources, and possess inherent metabolic pathways better-suited to organic acid production; organisms such as

acidophilic yeasts, filamentous fungi and certain bacterial strains. Harnessing these traits could reduce the burden of extensive genetic modification and alleviate some of the process limitations that have characterized work in conventional hosts. As the field advances, expanding the toolkit to include diverse, underexplored microbes may prove critical for developing more robust and cost-effective fumaric acid bioprocesses.

Table 1.1. Comparison of model organism engineering for fumaric acid production

Organism	Titer	Gene Deletions	Gene Overexpressions	Carbon Source	Vessel	Citation
<i>E. coli</i>	1.32 g L ⁻¹ (11.38 mmol L ⁻¹)	<i>iclR</i> , <i>fumABC</i> , <i>frdABCD</i> , <i>arcA</i>	<i>carAB</i> , <i>argI</i> , <i>rocF</i>	Glucose	Shake flask	³¹
<i>E. coli</i>	86.3 % per mole of consumed glucose	<i>ackA</i> , <i>pta</i> , <i>poxB</i> , <i>ldhA</i> , <i>adhE</i> , <i>ptsG</i> ,	<i>BsupycA</i> <i>Mtukgd</i> , <i>glk</i> , <i>galP</i>	Glucose	750 mL vented shake-flasks	³²
<i>E. coli</i>	1.53 g g ⁻¹ DCW	<i>fumABC</i> , <i>frdABCD</i> , <i>iclR</i> , <i>arcA</i> , <i>ptsG</i> , <i>mdh</i>	<i>glk</i> , <i>galp</i> , <i>ppc</i> , <i>acs</i>	Glucose	500 mL shake flask	³³
<i>E. coli</i>	9.42 g L ⁻¹	<i>iclR</i> , <i>fumABC</i> , <i>frdABCD</i> , <i>arcA</i> ,	<i>carAB</i> , <i>argI</i> , <i>rocF</i> , <i>dcuB</i>	Glucose	5 L bioreactor, fed-batch	³⁴
<i>E. coli</i>	22.4 g L ⁻¹	<i>ldhA</i> , <i>pflB</i> , <i>poxB</i> , <i>pta</i> , <i>ackA</i> , <i>frdBC</i> , <i>fumABC</i> ,	<i>Aflpyc</i> , <i>cs</i> , <i>acn</i> , <i>icl</i> , <i>sdhABCD</i>	Glucose	5 L bioreactor, fed-batch	²⁹
<i>E. coli</i>	28.2 g L ⁻¹	<i>iclR</i> , <i>fumABC</i> , <i>arcA</i> ,	<i>ppc</i> , <i>galP</i>	Glucose	6.6 L fermentor, fed-batch	²⁷

		<i>ptsG</i> , <i>aspA</i> , <i>lacI</i>				
<i>E. coli</i>	35.1 g L ⁻¹	<i>iclR</i> , <i>fumABC</i> , <i>ptsG</i> , <i>aspA</i> , <i>lacI</i>	<i>ppc</i> , <i>gltA</i>	Glucose, supplemented with aspartic acid	6.6 L fermentor, fed-batch	²⁸
<i>E. coli</i> E2	41.5 g L ⁻¹	<i>fumABC</i> , <i>aspA</i>	<i>ppc</i> , <i>sucAB</i>	Glycerol and corn steep liquor	1.3 L reactor, fed-batch	³⁵
<i>S. cerevisiae</i>	1.675 g L ⁻¹	<i>FUM1</i>	<i>RoPYC</i> , <i>SFC1</i>	Glucose	250 mL shake flask	³⁶
<i>S. cerevisiae</i>	3.18 g L ⁻¹		<i>RoMDH</i> , <i>PYC2</i>	Glucose	250 mL shake flask	³⁷
<i>S. cerevisiae</i>	5.64 g L ⁻¹	<i>FUM1</i> , <i>THI2</i>	<i>RoPYC</i> , <i>RoFUM</i> , <i>RoMDH</i>	Glucose	250 mL shake flask	³⁸
<i>S. cerevisiae</i>	33.13 g L ⁻¹	<i>FUM1</i> , <i>THI2</i>	<i>RoPYC</i> , <i>RoMDH</i> - <i>RoFUM1</i> ^{P160A} , <i>KGD2</i> - <i>SUCLG2</i> , <i>SpMAE1</i> , <i>SFC1</i>	Glucose	250 mL shake flask	³⁰

1.3.2. Fumaric acid production by non-conventional organisms

As the genetic engineering tools for microbes expands beyond traditional model organisms, the use of non-conventional microbes in biomanufacturing has gained significant traction. A central question in the field now concerns strategy: is it more efficient to engineer a model organism to exhibit desired traits, or to start with a non-conventional organism that already possesses those traits and optimize it for production? Several non-model yeasts have emerged as promising hosts due to their inherent metabolic capabilities. For instance, *Kluyveromyces marxianus* has been explored for bioethanol, aroma compounds, and biosurfactant production³⁹;

Komagataella phaffii (formerly *Pichia pastoris*) is widely used for recombinant protein and enzyme production⁴⁰; *Yarrowia lipolytica* serves as a robust chassis for producing lipids, carotenoids, and biosurfactants⁴¹; and Cargill uses a strain of low pH yeast, *Issatchenkia orientalis*, for lactic acid production⁴².

Some work has started to explore fumaric acid in non-conventional yeast. Liuming Liu's group has expanded to *Candida glabrata* (previously *Torulopsis glabrata*) using similar approaches described in their *S. cerevisiae* papers where they achieved fumaric acid titers of 9.2 to 21.6 g L⁻¹ ⁴³⁻⁴⁵. *C. glabrata* is a known industrial producer of pyruvate and was chosen for fumaric acid production due to its high pyruvate production, acid tolerance, and tolerance to high glucose⁴⁶. By recombinant expression of the rTCA pathway from *R. oryzae*, they were able to achieve relatively high titers of 21.6 g L⁻¹ ⁴⁴.

A study into C4-dicarboxylic acid production used *K. phaffii* to produce fumaric, malic, and succinic acid. By overexpressing pyruvate carboxylase and malate dehydrogenase, malic acid increased to titers of 42.28 g L⁻¹ ⁴⁷. Titer increases were more modest for succinic and fumaric acid reaching 9.42 g L⁻¹ and 0.42 g L⁻¹, respectively; however, these results demonstrate that *K. phaffii* is a valid chassis for organic acid production⁴⁷.

High organic acid tolerance has also been observed in the *Pichia* genus. Pyne *et al.* screened 108 strains of yeast from 66 distinct species in the presence of citric or adipic acid and found that of the top 15 most acid tolerant strains, 47% were from the *Pichia* genus⁴⁸. One of these strains from the *Pichia* genus, *P. kudriavzevii*, had been previously identified as a robust host for organic acid production^{49,50}. They chose *P. occidentalis* for further exploration of *cis,cis*-muconic acid production, the precursor to adipic acid, and achieved the highest titers by bioproduction to date, 1.72 x higher than production in *S. cerevisiae*⁴⁸. The success of *P. occidentalis* in surpassing

S. cerevisiae for *cis*, *cis*-muconic acid production underscores how host selection can dramatically impact titer performance. As metabolic engineering continues to push the boundaries of microbial production, the strategic choice of organism will be key to unlocking scalable and economically viable processes.

These examples underscore the growing recognition that non-conventional yeasts can offer unique physiological traits, such as acid tolerance, high precursor flux or robust growth, that make them attractive platforms for organic acid production. However, another equally critical factor in selecting a microbial host is its ability to metabolize alternative or low-cost feedstocks. The economic and environmental sustainability of bioproduction hinges not only on product yield but also on the carbon source used to reach that yield⁵¹. Many industrial waste streams, particularly in the dairy, food, agricultural and forestry sectors, contain non-hexose sugars or complex substrates that *S. cerevisiae* cannot utilize efficiently^{52–54}. This opens the door for yeasts with specialized metabolic capabilities, such as *K. marxianus*, a thermotolerant species capable of rapidly metabolizing lactose, a major component of dairy waste⁵⁵. By coupling the inherent advantages of non-conventional hosts with access to underutilized feedstocks, *K. marxianus* emerges as a compelling candidate for sustainable bioproduction.

1.3.3. Lactose as a feedstock

Although produced in large quantities during dairy processing, lactose remains one of the least valorized components of the dairy value chain. Cream, protein concentrate, skim milk, and cheese, reliably generate profit; however, the lactose-rich by-product milk permeate is generally sold into low-value markets in North America. While permeate is sometimes used to create a low-cost sweetener, growing rates of lactose intolerance and rising demand for lactose-free products

may shrink that market⁵⁶. For decades, developing commercially viable and sustainable uses for the lactose has remained an unresolved industry challenge.

In Canada, the annual production of ultrafiltered whey and milk permeate powder is an estimated 30–50 million kilograms annually, while U.S. estimates reach 500–600 million kg (Agropur, personal communication, January 2025). Within the context of a circular economy where economic growth is decoupled from resource depletion, lactose represents a renewable carbon source with considerable potential. However, many metabolic engineering efforts continue to focus on hosts like *S. cerevisiae*, *K. phaffi*, *P. occidentalis*, and *Y. lipolytica*, which are unable to metabolize lactose⁵⁷. Even when *K. marxianus*, a naturally lactose-consuming yeast is used, strains with limited lactose metabolism are often chosen^{58,59}. Optimizing *K. marxianus* for industrial lactose conversion offers an opportunity to simultaneously resolve a persistent processing bottleneck and enable more sustainable biomanufacturing strategies.

1.4. *Kluyveromyces marxianus* as a synthetic biology chassis

1.4.1. Biology and physiology of K. marxianus

K. marxianus is a non-conventional yeast species that has emerged as a promising host for industrial biotechnology due to its unique physiological traits⁶⁰. As a facultative aerobe and member of the *Saccharomycetaceae* family, *K. marxianus* is phylogenetically close to *S. cerevisiae* yet displays several advantageous features that distinguish it from the model yeast. Its rapid growth rate⁶¹, broad pH tolerance⁶², thermotolerance⁶³, and ability to metabolize a wide range of sugars^{57,64} make it well-suited to industrial applications. Most notably, *K. marxianus* possesses a native ability to metabolize lactose⁶⁵, setting it apart from many other yeast species and positioning it as a natural candidate for valorizing dairy by-products.

K. marxianus has eight chromosomes and is homothallic, able to exist as haploid or diploid with the ability to switch between mating types^{66–68}. This switching occurs via a two-transposase system first identified in a related yeast *Kluyveromyces lactis* and later confirmed in *K. marxianus*^{69–72}. Despite this capacity to switch mating types, *K. marxianus* strains tend to exist as haploids, diploids, and in some cases triploids, rather than a mixed population^{66,69,73}. A comparative analysis of 14 *K. marxianus* strains found that all diploid and triploid isolates were derived from dairy environments, while haploid strains were isolated from more diverse, non-dairy niches⁶⁶. Notably, the diploid strains exhibited high levels of heterozygosity, although this was unevenly distributed across the genome: 25–51 % of each diploid genome showed regions of loss of heterozygosity, a pattern not observed in triploid strains. This heterozygosity is thought to contribute to the stress tolerance of these strains in dairy environments and may contribute to the robustness and adaptability of these same strains in industrial contexts⁶⁶.

The ability of *K. marxianus* to utilize lactose is enabled by a suite of endogenous genes, including the lactose permease *LAC12* and the β -galactosidase *LAC4*, which allow the organism to import and hydrolyze lactose into glucose and galactose⁶⁵. Once imported and cleaved, glucose and galactose are channeled into central carbon metabolism through glycolysis and the Leloir pathway, respectively⁷⁴. Beyond lactose, *K. marxianus* exhibits broad substrate versatility, capable of metabolizing a range of sugars including glucose, fructose, galactose, sucrose, raffinose and inulin, as well as more complex substrates such as lignocellulosic hydrolysates and plant-based feedstocks^{57,75}. This flexibility opens the door for cost-effective fermentation processes based on agro-industrial residues.

In contrast to *S. cerevisiae*, which ferments glucose to ethanol under high sugar and aerobic conditions (Crabtree effect)⁷⁶, *K. marxianus* preferentially directs carbon flux toward respiration

rather than fermentation when oxygen is available⁷⁷. As a result, it produces biomass and metabolic intermediates more efficiently, with reduced ethanol by-product formation. This trait is particularly advantageous for industrial production processes that seek to maximize yield of non-ethanol products, including organic acids, aroma compounds, and bio-based chemicals⁷⁸.

1.4.2. Industrial relevance

K. marxianus is a non-conventional yeast species that holds GRAS (Generally Recognized as Safe) status in the United States⁷⁹ and has been granted QPS (Qualified Presumption of Safety) status in the European Union⁸⁰. Its established safety profile has facilitated its adoption in a range of commercial applications, particularly within the food, feed, and industrial biotechnology sectors. As mentioned above, certain strains can reach growth rates as high as 0.60 h⁻¹, metabolize a wide range of carbon sources, tolerate temperatures upwards of 50 °C, and exhibit strong acid tolerance⁶⁷. Having a high growth rate is particularly desirable in biomanufacturing when the product is the microorganism itself (e.g., baker's yeast)⁸¹ or when product formation is coupled to growth (e.g. protein production, central metabolism)⁸². The feedstock usage versatility for *K. marxianus* opens opportunities to reduce raw material cost by utilizing low-cost or waste-derived substrates, such as lactose in ultrafiltered milk permeate. Glucose as a substrate is a significant cost for bioprocessing; a techno-economic analysis of biodiesel production estimated that glucose accounted for 79 % of raw materials cost, amounting to ~\$17 million annually in 2014⁸³.

Thermotolerance in *K. marxianus* also provides important industrial advantages. Higher-temperature fermentation reduces cooling costs by minimizing water and electricity demand, while also lowering contamination risk^{84,85}. Additionally, processes involving complex carbon sources like lignocellulosic biomass requires a saccharification step, during which enzymes function optimally at elevated temperatures. Thermotolerant strains can perform simultaneous

saccharification and fermentation (SSF), reducing enzyme usage, operational time, and overall cost^{86,87}.

Due to these advantages, strains of *K. marxianus* are employed in the production of enzymes such as inulinase⁸⁸, β -galactosidase⁶⁷, and pectinase^{89–91}, as well as value-added products like ethanol⁹². While it's been used to produce bioethanol for fuel applications for many years, it has more recently been used to create ethanol spirits for consumption; companies like Vodka in Almonte, Ontario, have leveraged lactose metabolism of *K. marxianus* to turn milk permeate into a vodka-like spirit⁹³. These traits, combined with its regulatory approval and history of safe use, position *K. marxianus* as a versatile and increasingly valuable platform organism for sustainable biomanufacturing.

1.4.3. Advances in genetic engineering tools

K. marxianus is an emerging non-conventional yeast with a broad range of industrially relevant traits. Despite its relatively recent rise in prominence, a variety of genetic tools and transformation techniques have already been established. *K. marxianus* is readily transformable with exogenous DNA^{94,95}, and several autonomously replicating sequences (ARS) have been identified⁹⁶, including one that functions across multiple budding yeast species⁹⁷. In addition to utilizing strong *S. cerevisiae* promoters⁹⁸, *K. marxianus*-specific promoter libraries have been developed to enable more precise control of gene expression^{99,100}. A comprehensive genetic toolkit has also been published, which characterizes a broader set of endogenous promoters and introduces defined terminators, ARS elements, centromeric sequences, and validated chromosomal safe-harbor sites for stable genomic integration¹⁰¹.

Beyond plasmid-based systems, several approaches have been developed for genome-scale engineering in *K. marxianus*. These include Cre-Lox recombination¹⁰², genome shuffling¹⁰³,

random integration via non-homologous end joining (NHEJ)¹⁰⁴, and a modular recombination-based method called Promoter-based Gene Assembly and Simultaneous Overexpression (PGASO)¹⁰⁵. With the advent of CRISPR-Cas9, multiple high-efficiency gene editing strategies have been adapted for *K. marxianus*, enabling both gene deletions and targeted insertions^{72,106–109}.

CRISPR-based tools have already been applied to engineer *K. marxianus* strains for a variety of industrial applications. These include enhancing the production of valuable compounds such as 2-phenylethanol (2-PE)^{59,110} and ethyl acetate¹¹¹, improve glycerol utilization¹¹² and fatty acid biosynthesis⁷², enable antibody production¹¹³ and humanization¹¹⁴, and to support starch metabolism¹¹⁵. More recently, the toolkit has been expanded to incorporate CRISPR interference (CRISPRi), which has been used to boost ethyl acetate production¹¹⁶ and has been deployed in a genome-wide screen to improve monoterpene biosynthesis¹¹⁷. While these advances have firmly established *K. marxianus* as a genetically tractable and industrially relevant chassis, ongoing efforts to improve editing efficiency, expand regulatory part libraries, and standardize toolkits across diverse strains will further unlock its full biotechnological potential.

2. Chapter 2: Domestication of *Kluyveromyces marxianus*

Reprinted (adapted) with permission from:

Thornbury M, Knoops A, Summerby-Murray I, Dhaliwal J, Johnson S, Utomo JC, Joshi J, Narcross L, Remondetto G, Pouliot M, Whiteway M, Martin VJJ. Sequencing of a Dairy Isolate Unlocks *Kluyveromyces marxianus* as a Host for Lactose Valorization. ACS Synth Biol. 2025 Jul 8. doi: 10.1021/acssynbio.5c00096. Epub ahead of print. PMID: 40629255. Copyright 2025 American Chemical Society.

2.1. Abstract

The use of genetically modified non-conventional yeast provides significant potential for the bioeconomy by diversifying the tools available for the development of sustainable and novel products. In this study, we sequenced and annotated the genome of *Kluyveromyces marxianus* Y-1190 to establish it as a platform for lactose valorization. The strain was chosen for rapid growth on lactose-rich dairy permeate, high transformation efficiency, and ease of culturing in bioreactors. Genomic sequencing revealed that *K. marxianus* Y-1190 possesses single nucleotide polymorphisms associated with efficient lactose metabolism. The strain is diploid with notable genomic heterogeneity, which appears to be critical for its robust growth and acid tolerance. To further exploit this platform strain, we developed protocols for gene and chromosome manipulation using CRISPR editing, constructed and validated a series of promoters compatible with MoClo vectors, and designed synthetically inducible promoters for *K. marxianus*. These tools enable precise control over gene expression, allowing for the tailored optimization of metabolic pathways and production processes. The synthetic promoters provide flexibility for dynamic expression tuning, while the CRISPR-based editing protocols facilitate targeted genetic modifications with high efficiency. Together, these advancements significantly enhance the genetic toolbox for *K. marxianus*, positioning it as a versatile platform for industrial biotechnology. These tools open new opportunities for the sustainable production of biobased

chemicals, fuels, and high-value products, leveraging lactose-rich feedstocks to contribute to a circular economy.

2.2. Introduction

The search for efficient and cost-effective hosts for industrial applications is an ongoing process in biotechnology. *Kluyveromyces marxianus*, a non-conventional yeast species, has gained considerable attention in recent years due to specific properties that make it a suitable candidate for a variety of industrial applications⁶⁰. *K. marxianus* is known for its high growth rate⁶¹, together with broad pH tolerance⁶², heat tolerance⁶³, and ability to metabolize a wide range of sugars^{57,64}. These characteristics make it an ideal candidate for use in industrial processes such as bioethanol production^{58,118,119}, the production of value-added products such as enzymes^{120–122}, and bioactive compounds^{123,124}. *K. marxianus* strains have been isolated from multiple sources including dairy products, fermented foods, and soil⁵⁷. Several strains of *K. marxianus* have been studied for their potential use in industrial applications. Among the commonly studied strains for industrial applications are NBRC1777 and DMKU3–1042, isolated from soil in Thailand and Japan, respectively, as well as CBS 6556, which was isolated from the fermented drink pozol^{125–127}. Although these strains can grow on lactose as a carbon source, they are not from the dairy lineage of *K. marxianus*⁶⁶ and may not be suited for applications in the dairy industry.

Despite being produced in great quantities as a byproduct of dairy processing, lactose is the weak link of the modern dairy industry. All other outputs of dairy processing, cream, skim milk, cheese and protein concentrates, sell into markets that are consistently profitable. In contrast, the lactose byproduct of dairy processing, referred to as permeate, is a low value commodity that in North America is sold into highly volatile markets. In many cases the available price for lactose does not justify the cost of concentrating and drying this byproduct for sale as animal feed,

consequently, it must be disposed of at an environmental and economic cost. Milk and whey permeates are often used as cheap options for food sweeteners, however, with the rate of lactose intolerance and the popularity of lactose-free products, that market for permeate may be shrinking⁵⁶. Developing markets for the lactose fraction of nonfat solids is a longstanding issue that the dairy industry has been grappling with for years.

Advances in genetic engineering and the development of molecular tools have paved the way for the manipulation of *K. marxianus* to improve its industrial potential. To date, two promoter libraries^{99,100}, a genetic toolkit¹⁰¹, and standard protocols for CRISPR have been developed for *K. marxianus*^{72,106,107}. While these tools have been instrumental in broadening the field to work with *K. marxianus*, few are modularized in the yeast toolkit (YTK)¹²⁸ standard and the majority work with nondairy isolates that can respire but not ferment lactose, rather than dairy isolates that can do both⁵⁵.

In the current study, we aimed to domesticate a strain of *K. marxianus* capable of transforming the low-value lactose found in permeate into higher-value specialty chemicals. We screened 19 strains of *K. marxianus* available from public collections and identified six candidates with robust growth on both minimal lactose medium and lactose-rich permeate. We further screened the 19 strains on permeate with fumaric acid added to pH 2.2, since for bioprocessing, limiting the addition of a base during fermentation can streamline the process and cut down costs¹²⁹. Many molecules of commercial interest are acidic, such as the nine organic acids that have been designated by the United States Department of Energy as chemical building blocks that can be produced from biomass^{1,13}. Aside from its broad substrate utilization and high acid tolerance, *K. marxianus* is thermotolerant, growing at temperatures up to 52 °C⁶⁷. It is also one of the fastest-growing eukaryotic organisms reported, with maximal growth rates ($\mu_{\max}$) at 0.6–0.8 h⁻¹

^{161,67}. These traits are particularly advantageous for industrial applications: thermotolerance can reduce cooling costs during fermentation, while a high growth rate enables shorter production cycles, thereby improving overall process efficiency. Having a strain that can be leveraged to metabolize lactose and be organic acid-tolerant would be an asset to the circular economy. We tested the plasmid transformation efficiency of the characterized strains from these screens, identifying three strains that grew well in all conditions and had high transformation efficiencies.

We have generated the first publicly available annotated genome of the lactose-using *K. marxianus* strain Y-1190. We have further modularized and characterized 23 promoters sourced from a promoter library¹⁰⁰ and the *K. marxianus* toolkit¹⁰¹ for activity on lactose, glucose, and supplemented permeate in both strain Y-1190 and a soil strain CBS 6556. Additionally, we took three of the strongest promoters and engineered them to be tetracycline inducible. We further developed an improved CRISPR-Cas9 system for genomic deletions and integrations using a combination of a double-guide array and an NHEJ inhibitor and finally, we optimized the fermentation strategy for Y-1190 in a bioreactor to ensure its industrially relevant potential as a production strain.

2.3. Materials and methods

2.3.1. Kluyveromyces strain selection and media optimization

Nineteen strains of *K. marxianus*, and one control strain of *Saccharomyces cerevisiae* (CEN.PK), were revived from frozen stocks and cultured on YPD (10 g L⁻¹ Bacto Yeast Extract, 20 g L⁻¹ Bacto peptone, and 20 g L⁻¹ glucose) agar plates at 30 °C. Three independent colonies of each strain were selected to inoculate 800 µL of YPD medium in a 96-deep-well plate. Once the cultures reached visible turbidity, all the strains were back-diluted into 800 µL of YPD in a 96-deep-well plate and incubated for 8 h to reach an OD₆₀₀ of 1.8. After an 8 h incubation, a diluted

intermediate plate was created to achieve an OD₆₀₀ of 0.9. A 10 µL inoculum from the diluted plate was used to inoculate 170 µL of the relevant growth medium. The final volume of each well was 180 µL, and the initial OD₆₀₀ of all strains was 0.05. Growth rates ($\mu_{\max}$) were calculated in R using the growth rates package¹³⁰.

To test the growth of *K. marxianus* strains on various sugars, strains were grown in YNB-Glucose (YNBG, 20 g L⁻¹), YNB-Lactose (YNBL, 20 g L⁻¹), and in supplemented permeate, a filtered permeate from industrial dairy processing waste (Agropur Cooperative, Table 2.1) supplemented with vitamins and microelements (8.36 g L⁻¹ urea, 2.07 g L⁻¹ KH₂PO₄, 1.15 g L⁻¹ MgSO₄·7H₂O, 1 g L⁻¹ NiSO₄·6H₂O, 5 mL L⁻¹ vitamin stock, and 5 mL L⁻¹ microelements). The composition of the vitamin and microelement stock was described previously¹³¹. For the acid tolerance test, 5 g L⁻¹ of fumaric acid was added to the supplemented permeate, and 12 N HCL was added dropwise to the final pH of 2.2.

Table 2.1. Dairy permeate dry composition analyzed by Inductively Coupled Plasma Spectroscopy conducted by Mérieux Nutrisciences, Dorval, Québec

Components	Value (± St. Dev)
Total Solid (%)	5.93 ± 0.32
Ash (%)	0.47 ± 0.027
Total Protein (%)	0.36 ± 0.018
Protein ATC 12% (%)	0.14 ± 0.009
True Protein (%)	0.22 ± 0.023
Fat Mojonnier (%)	0.03 ± 0.008

Lactose (%)	4.94 ± 0.061
Calcium (mg Kg ⁻¹)	245.8 ± 13.646
Phosphorus (mg Kg ⁻¹)	334.6 ± 12.462
Potassium (mg Kg ⁻¹)	1394 ± 28.81
Sodium (mg Kg ⁻¹)	339.6 ± 6.804
Magnesium (mg Kg ⁻¹)	71.62 ± 2.191

2.3.2. Transformation

K. marxianus was transformed using the Gietz method with changes to heat shock temperature and time¹³². Briefly, cells from three OD₆₀₀ units of log phase cultures were harvested by centrifugation (1860 g) for 3 min, washed with dH₂O, and transformed in a total volume of 150 µL using the concentrations as described previously¹³². The transformations were incubated at room temperature for 20 min prior to heat shock at 47 °C for 15 min. Transformations were recovered overnight in YPD at 30 °C without shaking and subsequently back diluted 1:5 into YPD and the appropriate antibiotic for 24–48 h for optional outgrowth, before plating on selection media.

2.3.3. Genome sequencing and bioinformatics

2.3.3.1. Yeast strains and library preparation

K. marxianus strain Y-1190 was subjected to whole genome sequencing. Y-1190 was cultured overnight in 5 mL YPD at 30 °C and 300 rpm. Cells were harvested by centrifugation and DNA was purified using the MasterPure Yeast DNA Purification Kit (Lucigen) following the

manufacturer's instructions for cells harvested from liquid cultures except that DNA was eluted in nuclease-free H₂O rather than TE buffer to avoid EDTA contamination. DNA quantity was measured using a Qubit 4 fluorometric assay (Invitrogen) and was 560 ng μL^{-1} . All yeast strains used in this study are listed in Table S2.3.

2.3.3.2. *Long-read sequencing, genome assembly, and annotation*

Genomic DNA was sequenced (HistoGenetics Inc., Ossining NY) using the PacBio SMRT Sequel II platform following library prep using the SMRTbell HiFi kit (Pacific Biosciences) which generated 7.31 Gb reads for *K. marxianus* Y-1190. Demultiplexing and de novo genome assembly was achieved using the SMRTlink de novo assembly pipeline¹³³.

Low complexity and highly repetitive regions were masked using RepeatMasker¹³⁴. AUGUSTUS v3.4.0¹³⁵ was used for ab initio gene prediction. To assign identity, predicted genes were queried by BLASTX 2.10.0+ or DIAMOND¹³⁶ against the nonredundant database v5 or v2023-09-21, respectively, using default parameters in Omicsbox with the taxonomy filter set to 4910 (*Kluyveromyces*). Predicted genes were aligned with BLAST against *Saccharomycetales* to assign identity, annotation performed using InterProScan along with GO mapping and annotation. Predicted proteins were queried against the AybraH yeast protein database, however, since *K. marxianus* is not present in this database, NCBI BLAST/diamond returned better matches. Annotated genome files are available at NCBI (BioProject: PRJNA957743, Accession: JBJBSF000000000). BUSCO (v5.4.5) completeness scores were used to assess genome completeness using the following parameters: Lineage *Saccharomycetes*, Mode Transcriptome, and e-value $<10 \times 10^{-3}$ along with the OrthoDB v10 database.

2.3.4. Promoter set characterization

2.3.4.1. Promoter selection

Two *K. marxianus* promoter parts from Rajkumar *et al.* were obtained from Addgene¹⁰¹. Twenty-one promoters from Lang *et al.* were a generous gift from Dr. Ian Wheeldon (University of California Riverside). These promoters were not in a modular cloning format and were cloned into the YTK MoClo format¹²⁸, including removing unwanted type IIS restriction sites, before building and testing (pMTKM09-pMTKM29). All primers used in this study can be found in Table S2.4.

The three anhydrotetracycline (aTc) inducible promoters were designed by replacing endogenous sequences with two bacterial *tetO* modules (TCCCTATCAGTGATAGAGA) separated by two nucleotides (AT) 10 bp after the identified TATA box of the strongly expressed promoters *pTEF3*, *pPGK1*, and *pTDH3* from *K. marxianus*. Those engineered promoters were amplified from *K. marxianus* Y-1190 genomic DNA and any BsaI or Esp3I sites present within their sequences were removed using truncated primers. Esp3I sequences were added with primer overhangs at 3' and 5' ends and the resulting amplicons were assembled with GoldenGate cloning into level 0 plasmids from the modular cloning suite (pMTKM30-pMTKM32)¹²⁸. The engineered promoters were cloned upstream of mNeonGreen and the *K. marxianus* *PDC1t* terminator following the MoClo syntax. In parallel, a *K. lactis* codon-optimized version of the bacterial repressor gene *tetR* fused to the Nuclear Localization Signal SV40 was ordered as a synthetic gene (Twist Bioscience) and cloned into a level 0 MoClo entry plasmid. The repressor was next cloned between the *pSSA3* promoter and the *tINU1* terminator from *K. marxianus* following the MoClo syntax. Finally, the two transcriptional units were fused together within a plasmid containing the 5' and 3' homology arms for integration into the *ARO1* safe harbor site, as previously described¹⁰¹.

Those plasmids were digested with NotI-HF and used as donor DNA together with a pCas plasmid containing two guides targeting the *ARO1* locus for genomic integration (see section genome engineering for details). Finally, the resulting strains were cured of their pCas plasmid by restreaking transformants on non-selective YPD plates for multiple passages, followed by verification of plasmid loss through the inability to grow on selective medium. Strains were inoculated overnight in YPD prior to their dilution at an OD₆₀₀ of 0.1 in Synthetic Complete (SC) medium and grown in 96-deep-well plates for 6 h until OD₆₀₀ of ~ 1 and fluorescence was measured using a plate reader (Infinite200, Tecan).

2.3.4.2. Promoter reporter plasmid construction

Promoter activity reporter plasmids consisted of the promoters listed in Table S2.2, 5' to the sfGFP from the Yeast ToolKit paper, and the KmtPGK terminator from the *K. marxianus* biological parts paper^{101,128}. These plasmids were created using Golden Gate¹³⁷. Briefly, 40 fmol of each insert was combined with 20 fmol of backbone plasmid (YTK083) along with 0.5 µL BsaI-HF (NEB), 0.5 µL T4 ligase, 1.5 µL 10x T4 ligase buffer, and water to a 15 µL total volume. Assembly was carried out by digestion at 37 °C for 5 min followed by ligation at 16 °C for 5 min for 30 cycles. A final 37 °C incubation was done for 60 min digestion followed by an inactivation step of 80 °C for 15 min. The resulting plasmids were transformed into chemically competent *E. coli* DH5α cells and stored in glycerol at –80 °C for later use. Plasmids were sequenced by nanopore sequencing (Plasmidsaurus). Once confirmed by sequencing these 23 plasmids were transformed into *K. marxianus* strains Y-1190 and CBS 6556 for promoter strength testing. All plasmids used and created in this study can be found in Table S2.5.

2.3.4.3. Promoter strength testing

Four independent colonies from each promoter plasmid were selected and grown overnight in 700 μ L of YPD with 100 μ g mL⁻¹ nourseothricin selection. Cells from overnight cultures were washed twice in water and an intermediate 96-well plate was created at an OD₆₀₀ of 1 in water. This plate was used to seed our experiment in 96-deep-well plates at an OD₆₀₀ of 0.1 in either YNBL, YNBG, and 0.5x supplemented permeate. These plates were grown at 30 °C with shaking at 300 rpm for 6 h and 24 h at which point fluorescence and OD₆₀₀ were measured using a plate reader (Infinite200, Tecan). The interquartile range (IQR) was calculated for each promoter and data points that fell outside 1.5*IQR on either side were removed. Figures were created in R.

2.3.5. Genome editing using CRISPR-Cas9

A series of SpCas9/gRNA vectors originally published by Ryan *et al.*, 2014¹³⁸, and later simplified by Bean 2022¹³⁹, were modified for optimal functionality in *K. marxianus* by replacing the 2 μ sequence with a panARS(opt) replication origin⁹⁷. The resulting plasmids display a BsaI GFP-dropout for easy gRNA cloning via GoldenGate and are available on Addgene (pMTKM01-pMTKM08, see Table S2.5). To design the different gRNAs used in this study, we used CCTop¹⁴⁰, and selected guides displaying at least a CRISPRater score¹⁴¹ of 0.60. GoldenGate cloning was performed as previously described to introduce one guide into the vector¹³⁹. To express two gRNA from the same transcriptional unit within one vector, we used the tRNAgly self-splicing module as previously reported¹⁴². We used the already published pCD4 plasmid¹⁴³ containing a tRNAgly together with a SpCas9 sgRNA scaffold as a template for amplification with primers containing the two sgRNAs followed by two BsaI sites as overhangs. The amplicon generated was next purified and 80 fmol was added to a GoldenGate reaction containing 20 fmol of vector as described in the previous section.

The donor DNA used for the optimization of the protocol for CRISPR-Cas9 genome editing with HDR was amplified by PCR from the pMT187 and pMT188 plasmids. Briefly, *pPDC1*, mNeonGreen, *tPDC*, and the 3' and 5' regions of the *URA3* genes were amplified by PCR with overlapping overhangs and cloned into the pJET plasmid (Thermo Fisher Scientific) according to the manufacturer's instructions. The plasmid was next used as a template to generate donor DNA with different recombination arm sizes with different sets of primers. The Δ DNL4 strain was generated following a protocol previously published¹⁰¹.

For the deletion/integration test of the *URA3* locus, *K. marxianus* was transformed as described in the previous section with 500 ng of Cas9 plasmid vector. For integrations, 500 ng of Cas9 plasmid vector and 1–2 μ g of donor DNA were provided. After recovery overnight in 500 μ L of YPD, cells were plated on YPD with 100 μ L mL⁻¹ nourseothricin selection for the Cas9 plasmid. For outgrowth, 100 μ L of the recovery culture was inoculated into 400 μ L of YPD with selection for the Cas9 plasmid and grown for 72 h in 96-deep-well plates. For NHEJ chemical inhibition, NU7026 (Cayman Chemicals) was added at a final concentration of 25 μ g mL⁻¹ at every stage of the transformation until plating. After 48 h of growth on selective media at 30 °C, at least 12 colonies per technical replicate were picked with a colony picker (Qpix400, MolecularDevice) and inoculated in 96-well microplates in SC medium supplemented or not with 5-FOA at a final concentration of 1 g L⁻¹. Cells were grown at 200 rpm for 16 h at 30 °C, and OD₆₀₀ was measured together with green fluorescence to evaluate *URA3* inactivation and mNeonGreen integration, respectively.

2.3.6. Investigation of strain haploidization

To test the diploid-haploid transition in *K. marxianus* strain Y-1190, cells were sporulated as described by Wu *et al.*,¹⁰³ with minor modifications. An overnight culture grown at 30 °C in

YPD was used to inoculate 5 mL of YPA (20 g L⁻¹ potassium acetate, 20 g L⁻¹ bactopectone, 10 g L⁻¹ yeast extract) at an OD₆₀₀ of 0.2. The cells were grown for 3–5 h to a final OD₆₀₀ of 0.8–1.0, washed once with sterile dH₂O (5,000g, 5 min), and suspended in 5 mL of 2% potassium acetate. Following 16 h of incubation at 30 °C with shaking, approximately 50% of the cells had formed asci. Cells from 1 mL of the sporulated culture were washed twice with dH₂O and suspended in a final volume of 500 µL. To eliminate nonsporulated cells, cell walls were digested by the addition of 25 µL of 5 unit µL⁻¹ 100-T Zymolyase (Zymo Research, E1004) with 5 µL of β-mercaptoethanol. The mixture was incubated at 30 °C for 30 min without shaking followed by the addition of 200 µL of 1.5% NP-40 and incubation at 30 °C for 30 min. Asci were dispersed by incubation in a sonicating water bath for 30 s intervals (10–15 repetitions). Once sufficient dispersal was achieved, the mixture was washed twice with dH₂O, serially diluted to 10⁻⁵, and 100 µL was plated on YPD. Colonies were then screened by colony-PCR both to identify potential haploids and to determine their mating type.

To trap potential haploids in a mating competent form, the transposases *α3* and *Kat1* were targeted by CRISPR-NHEJ using the two-gRNA strategy described above. Guides were designed such that successful inactivation caused the removal of a 1532 bp (*Kat1*) or 652 bp (*α3*) sequence, resulting in a size difference that can be detected by gel electrophoresis after PCR amplification of the gene region. Strain Y-1190 was transformed with the construct targeting *Kat1* (pLN026), followed by plasmid curing and subsequent transformation with the construct targeting *α3* (pLN028). The resulting strain (JD032) had a truncation in both *α3* and *Kat1* as verified by PCR and by sequencing.

Propidium iodide staining followed by flow cytometry¹⁴⁴ was used to determine the ploidy of potential haploid strains recovered from sporulation. Samples were analyzed using an Accuri

C6 flow cytometer with autosampler. Known *S. cerevisiae* diploid and haploid strains, as well as the *K. marxianus* Y-1190 parental diploid strain were used as controls for comparison with candidate haploids.

2.3.7. Cultivation of *K. marxianus* in a bioreactor

Controlled fermentations were conducted in 3 L BioBundle bioreactors (Applikon). The fermentation temperature was maintained at 30 °C. Batch fermentations were performed without controlling the pH, while fed-batch fermentation pH was controlled at pH 4.0 by automated titration with 4 N NaOH. Dissolved oxygen was maintained at 35% air saturation with an aeration rate of 1 L min⁻¹. Concentrations of O₂ and CO₂ in the off-gas were analyzed using a Tandem Multiplex Gas Analyzer (Magellan BioTech). Bioreactor inoculum was grown in a 500 mL shake flask containing 50 mL of batch medium at 30 °C overnight. Cells were washed and suspended in 0.9% NaCl before inoculation in 1 L of batch medium with starting OD₆₀₀ of ~ 0.3. As a batch medium, filtered permeate with or without supplementation was used as described above. For fed-batch fermentation, the bioreactor was operated in batch mode until lactose exhaustion, followed by the fed-batch phase with the automated addition of a feeding medium. For the feeding medium, three times ultrafiltration-concentrated permeate (Agropur Cooperative) with supplementation (23.07 g urea, 6.19 g KH₂PO₄, 3.45 g MgSO₄·7H₂O, 3 g NiSO₄·6H₂O, 22.5 mL vitamin stock, and 22.5 mL trace element stock per liter) was used. Nickel was added to promote urea utilization as mentioned by Dubencovs *et al.*¹⁴⁵. Samples for OD measurements and HPLC analysis were collected every hour for 8 h per day with the two-day total fed-batch running time. Dry cell weight (g L⁻¹) was calculated using a gravimetrically determined conversion factor of 0.22 g L⁻¹ per OD₆₀₀. Bioreactor data were collected and analyzed using BioXpert.V2 (Applikon).

The theoretical biomass yield on lactose was calculated based on the black box bioenergetics mathematical model of the microbial growth¹⁴⁶, which also provides the data for Gibbs free energy formation of lactose. The assumption made for this calculation was that the biomass of *K. marxianus* has the formula of $\text{CH}_{1.63}\text{O}_{0.54}\text{N}_{0.16}$ as mentioned by Krzystek and Ledakowicz¹⁴⁷. Observed biomass yield on lactose was calculated using the $Y = X/S$ formula. Lactose, glycerol, and ethanol concentrations were analyzed from 10 μL of fermentation samples diluted 1:10 in dH_2O using HPLC (Agilent 1290 Infinity II) equipped with an Aminex Fast Acid Analysis column (100×7.8 mm, Bio-Rad) and a 10 mM H_2SO_4 mobile phase at a flow rate of 0.8 mL min^{-1} . Lactose, glycerol, and ethanol were detected using RI, while fumaric acid was detected using a UV detector at 210 nm. Chromatograms were collected and analyzed using OpenLab CDS version 2.7 (Agilent Technologies).

2.4. Results

2.4.1. *K. marxianus* strain selection

To identify a strain of *K. marxianus* that not only grows well on lactose, but also grows well on permeate, a substrate that differs in composition from commonly used laboratory media, we compared the growth rate of 19 *K. marxianus* strains grown in supplemented permeate with those grown in YNB lactose (YNBL) medium (Fig. 2.1 A and B). These 19 strains were selected from public databases based on their isolation from dairy or acidic environments (Table S2.3). Additionally, we included CBS 6556, a strain isolated from Pozol, due to its established use in the *K. marxianus* research community. We identified six strains that grew at a rate above 0.4 h^{-1} in both media: Y-2415 (also known as CBS 397), Y-1190, Y-12992, Y-12990, Y-1163, and L388 (Fig. 2.1 A).

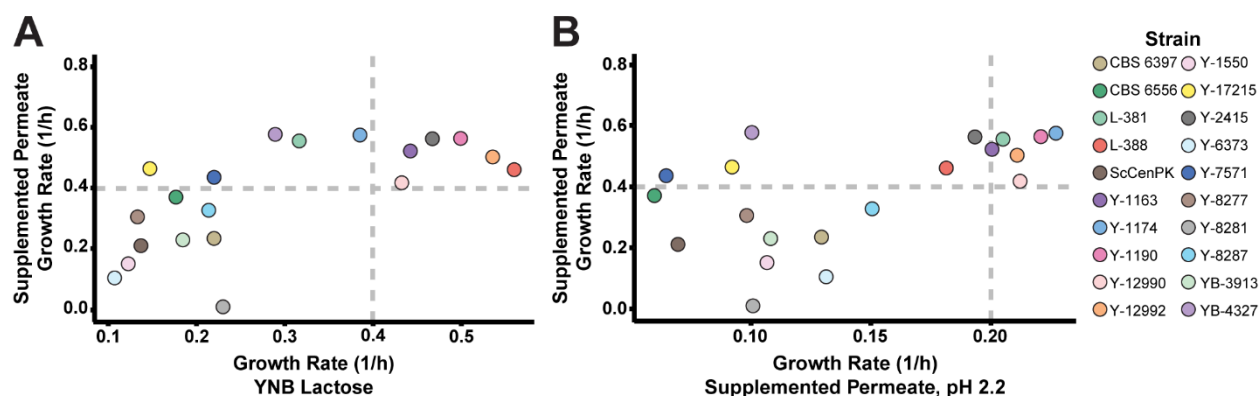


Figure 2.1 Growth rate comparison of aerobic growth of 19 *K. marxianus* strains on A. YNBL compared with supplemented permeate and B. Supplemented permeate compared with supplemented permeate at pH 2.2. Each data point is the average of three independent replicates. The complete data set can be found in Figure S2.1.

2.4.2. Transformation efficiency

To identify the most genetically tractable isolates, we tested the DNA transformation efficiency of the six strains selected for robust growth on permeate and low pH (Fig. 2.1 B). The strains were transformed with a modified Gietz procedure using an empty vector plasmid containing a spacer and nourseothricin resistance¹³². From this data, three strains, Y-1190, Y-12990, and Y-12992, had comparable transformation efficiencies of 1.14×10^4 , 1.44×10^4 , and 2.38×10^4 cfu μg^{-1} DNA respectively (Fig. S2.2), and surpassed our growth rate thresholds in all three media conditions (Fig. 2.1). We selected strain Y-1190 based on its growth in acid and lactose, its reliable transformation efficiency as well as its previous use in a CRISPR-Cas9 study that highlighted its genetic tractability through high transformation efficiency, amenability to CRISPR-Cas9-mediated genome editing, and propensity to exist in the haploid state⁷².

2.4.3. Genome sequence analysis

We created a publicly available annotated genome for *K. marxianus* strain Y-1190 to facilitate the use of this yeast in strain engineering projects. Strain Y-1190 was sequenced by

PacBio technology, which generated two assembly files: one primary assembly and one secondary assembly containing haplotigs, establishing the diploid nature of the strain. The primary assembly resulted in 10 contigs and the secondary assemblies resulted in 66 haplotigs (Table 2.2). To confirm the near-chromosome level of our assemblies, each contig was queried using NCBI BLAST while restricted to sequences from *K. marxianus* (taxid 4911). Primary contigs were found to be close to chromosome length and total assembly size matched closely to the *K. marxianus* reference strain DMKU3-1042; 1.1×10^7 for Y-1190 compared to 1.09×10^7 in DMKU3-1042 (Table S2.1).

Table 2.2. *K. marxianus* Y-1190 summary assembly

	Primary Contigs	Haplotigs
Sequences	10	66
GC%	40.1 %	40.6 %
N50	1,590,052	935,761
Assembly Length	10,951,246	9,309,754
Gene Annotations	4853	4108

Before gene prediction, low complexity and highly repetitive regions were masked using RepeatMasker (see Methods). Repetitive sequences made up 245885/10951256 bp (2.25%) while low complexity sequences made up 39868/10951256 bp (0.36%). Using AUGUSTUS, gene

prediction resulted in the identification of 4853 putative genes for Y-1190 with an average gene length of 1549 nt. Gene prediction for the haplotig data identified 4108 genes. The primary assembly was classified as 98.08% complete single-copy, 0.09% as complete duplicated, 0.28% as fragmented, and 1.54% as missing. In the haplotig assembly 76.28% were classified as complete single-copy, 1.59% as complete duplicate, 0.89% as fragmented, and 21.24% as missing. When predicting genes by BUSCO, missing genes could mean that there was no significant hit in the database for that gene. This is highly possible since there are only two *K. marxianus* genomes in the BUSCO database, both non-dairy strains¹⁴⁸.

2.4.4. Analysis of lactose permease variation

Previous work has identified three groups of *K. marxianus* haplotypes, A, B and C, which differ on their ability to metabolize lactose⁵⁵. While all three haplotypes can metabolize lactose, only haplotype B can both ferment and respire lactose. This is thought to be due to polymorphisms in the lactose permease *Lac12-1*, which increased lactose transport 10-fold in dairy isolate CBS 397 (also known as Y-2415) compared to non-dairy isolate CBS 6556⁶⁵. With full access to the Y-1190 genome, we conducted a multisequence alignment of *LAC12-1* genes from Y-1190, dairy isolate CBS 397, and non-dairy isolates CBS 6556, NBRC1777, and DMKU3-1042 (Fig. 2.2). Our analysis revealed that the *LAC12-1* genes of Y-1190 and CBS 397 exhibit 19 identified polymorphisms, aligning with those observed in the Varela *et al.* study, and identifies the *LAC12-1* gene in Y-1190 as encoding the variant predicted to have higher activity (Fig. 2.2).

to the modular cloning standard described for *Saccharomyces cerevisiae*¹²⁸ and tested them alongside two other promoters from Rajkumar *et al.*¹⁰¹ for a total of 23 promoters (Table S2.2).

Our modularized promoter library was cloned upstream of sfGFP and was transformed into two strains of *K. marxianus*: the nondairy strain CBS 6556 and the dairy strain Y-1190. The library was tested in YNBL, YNB glucose (YNBG) and supplemented permeate (Fig. 2.3, Fig. S2.4). To ensure the design of the constructs would work both in lactose and permeate, we compared the promoter strengths of those two conditions (Fig. 2.3 A). Generally, the relationship is linear, with promoters tending to drive overall higher expression in supplemented permeate than YNBL, apart from *pTDH3* which performs 4.3-fold better in YNBL than supplemented permeate. Similarly, the relationship between promoter expression in YNBL and YNBG is linear with the only exception being *pLAC4* which is only expressed in YNBL, as expected (Fig. 2.3 B). After characterizing the promoters in strain Y-1190 across different media, we compared the promoter expression profiles between Y-1190 and the soil-derived strain CBS 6556 in both YNBL (Fig. 2.3 C) and YNBG (Fig. 2.3 D). There were several promoters that fell outside a strictly linear relationship. As an example, *pINU1* (light orange) consistently had higher expression in Y-1190 while *pRPE1* (light brown) consistently had higher expression in CBS 6556 regardless of the carbon source (Fig. 2.3 CD). Both *pPDC1* (medium blue) and *pNC1* (medium green) were consistently high for both strains in all conditions, although with differing magnitudes of expression (Fig. 2.3 CD). All expression data for strains CBS 6556 and Y-1190 in YNBL, YNBG, and supplemented permeate at 6 h and 24 h can be found in the supplemental data (Fig. S2.4).

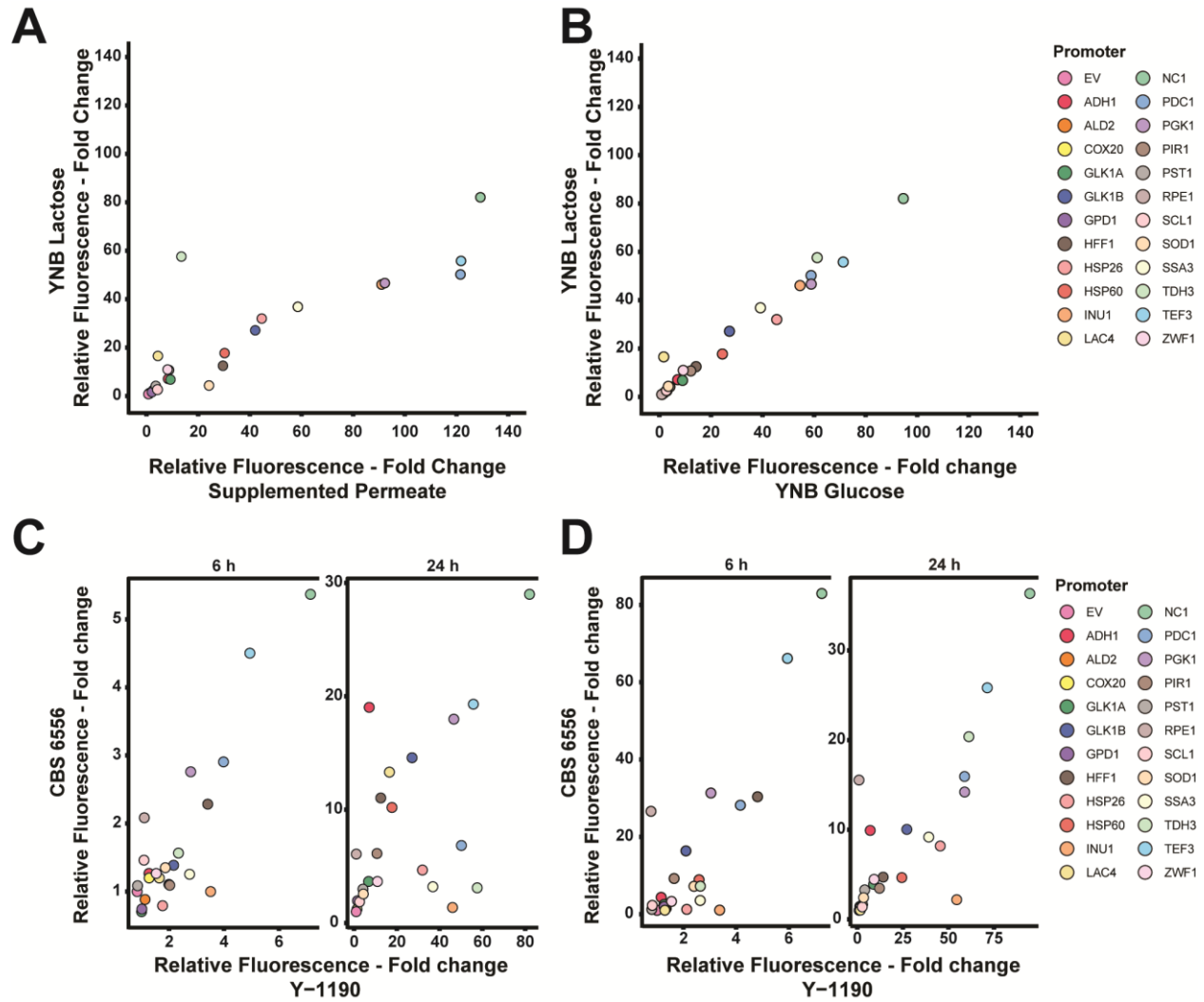


Figure 2.3 Characterization of promoter strength across media. Comparison of relative expression of GFP in Y-1190 in **A**. YNBL and supplemented permeate **B**. YNBL and YNBG. Comparison of relative expression of GFP in CBS 6556 and Y-1190 grown in **C**. YNBL at 6 and 24 h **D**. YNBG at 6 and 24 h. Each data point is the average of four biological replicates, the IQR method was used to remove outliers. The individual replicates and their standard deviation are reported in Figure S2.4.

In parallel, we developed three different anhydrotetracycline (aTc) inducible promoters as tools for genetic expression modulation in *K. marxianus*. We introduced two bacterial *tetO* modules into the strongly expressed promoters *pTEF3*, *pPGK1*, and *pTDH3* from *K. marxianus*, and fused those promoters to a mNeonGreen reporter system¹⁴⁹. This transcriptional unit was used in tandem with a codon-optimized bacterial TetR protein containing a nuclear localization site

(NLS) under the control of the strong promoter *pSSA3*. Those constructs were integrated in the genome of *K. marxianus* at the previously published *ARO1* safe harbor site¹⁰¹ using the optimized integration protocol presented in the next section, and fluorescence levels upon aTc addition were measured (Fig. 2.4 AB). While the *pTEF3_{tetO2}* and *pPGK1_{tetO2}* engineered promoters displayed high induction (19- and 14-fold increase) upon aTc addition, together with tight control in the absence of the inducer, *pTDH3_{tetO2}* displayed higher expression in the absence of the inducer and an overall lower level of induction (8-fold increase) due to the leaky expression.

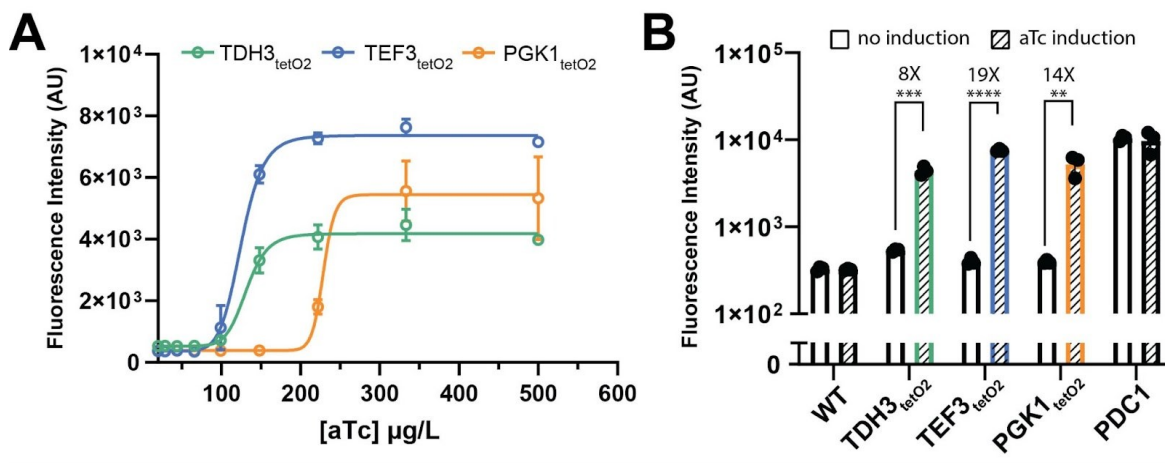


Figure 2.4 Characterization of aTc-inducible promoters. A. Titration of promoter activity upon various aTc concentrations. B. Fold change upon maximal induction.

2.4.6. CRISPR genome editing optimization

Robust genomic engineering of *K. marxianus* Y-1190 is crucial for its use as a platform for biomanufacturing diverse molecules. As endogenous gene deletions and genomic integrations are critical for this purpose, we optimized a CRISPR-based integration/deletion protocol for efficient genome editing. We first adapted a GFP-dropout pCas9 plasmid from the *S. cerevisiae* Mylo Toolkit¹³⁹ by swapping the 2 µ replication origin with the yeast broad-species pan(OPT)ARS⁹⁷. The resulting plasmid (pMTKM01) was used to build a single guide plasmid targeting the endogenous *URA3* gene from *K. marxianus* Y-1190. We tested whether the transformation of this

plasmid was sufficient to integrate a linear donor DNA containing 900 bp of 5' and 3' *URA3* homology arms flanking a strongly expressed mNeonGreen. Although this strategy was able to inactivate the *URA3* gene, we were unable to integrate the mNeonGreen cassette (Fig. 2.5 A). NHEJ prevalent DNA repair was reported by several studies to interfere with HDR-based genetic integration in different *K. marxianus* strains^{101,104,108}. To reduce activity of the NHEJ pathway in Y-1190 to favor HDR-based integrations, we generated a NHEJ ligase mutant (Δ DNL4)¹⁵⁰, and tested the use of a NHEJ chemical inhibitor (NU7026)^{151,152}. In parallel, we tested whether prolonged liquid growth with selective pressure for the pCas plasmid after transformation (outgrowth) would favor transformants with modified Cas9 target sites via HDR, but none of those strategies led to a successful HDR integration or *URA3* inactivation (Fig. 2.5 A). We next tested a two gRNA approach to simultaneously target different areas of the CDS, a strategy commonly used in mammalian cell culture transformations^{153,154}. This double cut approach, in combination with the NHEJ inhibition strategies, facilitated the mNeonGreen donor DNA integration at the *URA3* locus (Fig. 2.5 A). The most successful protocol, with 100% of the tested colonies displaying genomic integration, used the two-guide strategy in combination with NHEJ inactivation via genetic deletion, followed by outgrowth. Finally, to verify that the increase in efficiency was due to the use of two guides and was not gRNA sequence specific, we performed the one-guide experiment again with the second gRNA targeting *URA3* (Fig. S2.5). Similarly to the first single gRNA, we were not able to generate any integrations, confirming the importance of the two-guide strategy.

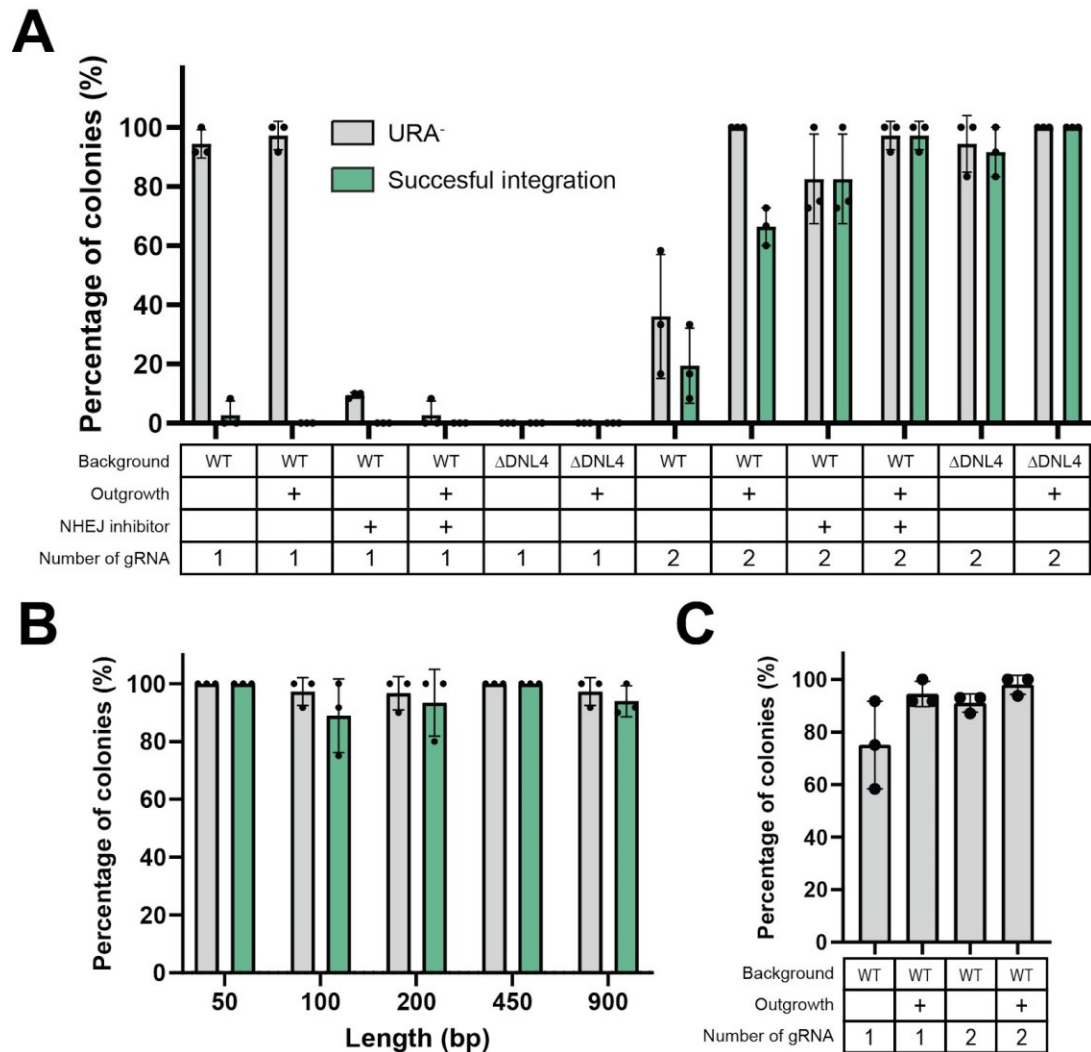


Figure 2.5 CRISPR-Cas9 genome-editing efficiencies in *K. marxianus*. **A.** Optimization of genomic deletion and integration via HDR. **B.** Homologous arm size effect on genomic deletion and integration efficiency via HDR. **C.** Optimization of genomic deletion via NHEJ. Grey bars represent URA³- colonies as defined by growth in 5-FOA supplemented media. Green bars in A and B represent mNeonGreen integration as defined by measureable fluorescence. Each dot represents a biological replicate, and the error bars represent the standard deviation.

We then used our optimized method with outgrowth and *DNL4* deletion to identify the minimal homology arm size necessary for genomic integration. We amplified the mNeonGreen cassette with different sets of primers to generate 900 bp, 450 bp, 200 bp, 100 bp, and 50 bp *URA3* homology arms and tested their efficiencies (Fig. 2.5 B). High integration rates were measured for

homology arms size as low as 50 bp, consistent with previous reports in other *K. marxianus* strains¹⁰⁷. Finally, we investigated if gene deletions could be generated only by NHEJ, without providing donor DNA. We tested the effect of the outgrowth and the double cut strategies on this approach. Although a single cut was enough to generate some *URA3* inactivation, 100% inactivation could be reached either using the double cut or the outgrowth strategies (Fig. 2.5 C).

2.4.7. Testing of diploid stability

Generation of a stable haploid of *K. marxianus* Y-1190 would enable a more streamlined CRISPR protocol as well as the ability to use mating-based assays and tools for genetic manipulation. *K. marxianus* is characterized as homothallic, able to exist as haploid or diploid with the ability to spontaneously switch between mating types^{68,69}. This switching occurs via a two-transposase system first identified in *K. lactis* and later confirmed in *K. marxianus*^{69–72}. To stabilize mating competent haploid strains, we used CRISPR-Cas9 to delete the two transposases, $\alpha 3$ and *Kat1* (Fig. S2.6). We then sporulated the Y-1190 Δ *Kat1* Δ $\alpha 3$ strain and used two separate methods, classical tetrad dissection and NP-40 based spore retrieval¹⁰³ to generate haploids. Over 100 tetrads were dissected, and hundreds of cells went through the NP-40 based protocol, yet we retrieved only one haploid, JD044, which subsequently autodiploidized (Fig. S2.6). Such autodiploidization has also been noted for *C. albicans* cells after selection for the haploid state¹⁵⁵. In all media tested diploidized JD044 had a lower growth rate than the wild type Y-1190, thus reducing its utility as a platform for genetic engineering (Fig. S2.7). For this reason, we continued our work with the wild-type diploid Y-1190.

2.4.8. Optimization of *K. marxianus* fermentation in permeate

After identifying a dairy strain of *K. marxianus* highly suitable for genetic manipulation, we evaluated its growth performance in a controlled fermentation environment using small scale

bioreactors and permeate. *K. marxianus* Y-1190 cells were first grown on permeate without pH control for 48 h to observe the intrinsic biomass yield and pH decline while using this substrate (Fig. 2.6 A). At the end of the batch fermentation, strain Y-1190 had a biomass yield of 0.15 g g⁻¹ lactose, consuming all lactose by approximately 30 h. The calculation of theoretical biomass yield of *K. marxianus* on lactose using a black box bioenergetics mathematical model¹⁴⁶ results in theoretical biomass of 0.323 g g⁻¹ lactose. Therefore, the biomass yield under our condition was only half of the theoretical yield, suggesting limitation of nutrients in the media.

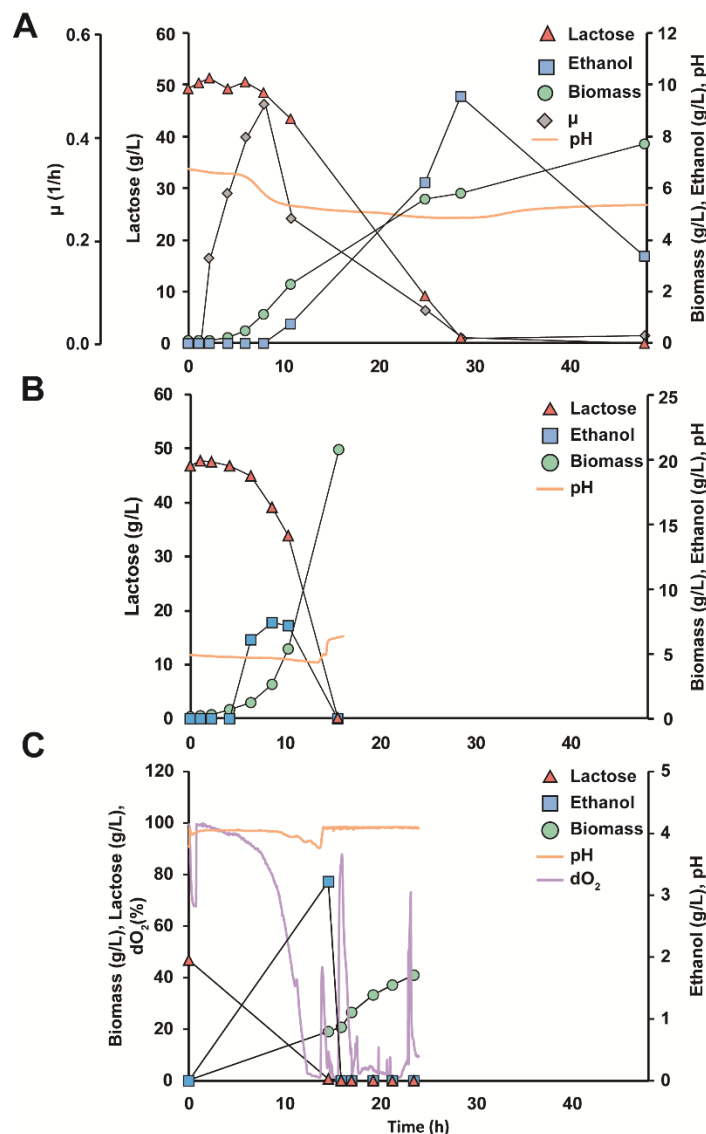


Figure 2.6 Fermentation Profile of *K. marxianus* Y-1190. A. Batch fermentation in non-supplemented permeate. B. Batch fermentation in supplemented permeate. C. Fed-batch fermentation in supplemented permeate with pH maintained at 4.

To optimize media composition and increase biomass yields we adjusted the composition of the permeate. Based on composition measurement, the total nitrogen from protein in the permeate is only 0.001 g L^{-1} (Table 2.1). The optimum mole of nitrogen per mole of carbon for *K. marxianus* growth is estimated to be 0.07^{145} . With around 50 g L^{-1} lactose in the permeate (Table 2.1), the predicted minimum nitrogen needed is 1.69 g L^{-1} , indicating a lack of nitrogen in the permeate. Similarly, some minerals and vitamins, such as potassium, phosphorus, sulfur, magnesium, vitamins, and microelements, were also lacking in permeate when compared to the required amounts for the yeast fermentation process¹⁵⁶. Thus, to increase *K. marxianus* biomass yield, we added urea as a nitrogen source and the lacking nutrients into the permeate¹⁴⁵. The batch fermentation of *K. marxianus* Y-1190 using supplemented permeate in a controlled fermentor without pH control showed a maximum uptake rate of lactose of 6.50 g L^{-1} lactose per hour (Fig. 2.6 B), with no lactose detected after 15 h. Under these conditions, a biomass yield of 0.43 g g^{-1} lactose was obtained.

Having established the necessary nutrient supplementation for efficient *K. marxianus* fermentation in permeate, we began optimizing the fed-batch phase for biomass yield by balancing the amount of supplied lactose with the oxygen supply of the fermentor. Oxygen limitation was observed during the batch phase, where ethanol was accumulated (Fig. 2.6 B). An increase in pH can occur when the rate of ammonium production from urea metabolism is faster than the assimilation of ammonium into biomass^{157,158}. This rise in pH has been associated with lactose accumulation^{157,159}, so we adjusted the pH to 4 to optimize the consumption of lactose. Under these

conditions we characterized the biomass yield in fed-batch mode using supplemented permeate (Fig. 2.6 C). No residual lactose was observed and the biomass yield of *K. marxianus* Y-1190 reached 0.51 g g⁻¹ lactose. While there is a discrepancy between theoretical yield (0.323 g g⁻¹ lactose) and our observed yield, our results are consistent with previous reports, which have documented biomass yields ranging from 0.37 g g⁻¹ lactose to 0.45 g g⁻¹ lactose¹⁶⁰. This result indicates that a high biomass yield of *K. marxianus* strain Y-1190 can be obtained in supplemented permeate, underlining the potential of this strain being used for the valorization of industrial dairy waste.

2.5. Discussion

There is a growing body of literature for *K. marxianus* that highlights its inherent capacity for industrial fermentation⁶⁰, lignocellulosic biomass degradation⁵⁸, and protein expression^{121,161}. New and improved techniques to increase the genetic tractability of this organism have been published^{72,103,107,162}, but few studies have harnessed the strain's inherent ability to metabolize lactose, even with the incredible potential of lactose as a feedstock for valorization. In Canada alone, dairy companies produce an estimated 30–50 million kg of lactose-rich ultrafiltered whey and milk permeate powder annually, while in the U.S., estimates reach 500–600 million kg (Agropur, personal communication, January 2025). Lactose, often considered a low-value byproduct, could play a crucial role in the circular economy - a framework aimed at decoupling economic growth from resource consumption by regenerating and recycling materials¹⁶³. Despite the accessibility of this otherwise low value carbon source, many metabolic engineering studies often focus on yeasts such as *S. cerevisiae*, *Pichia pastoris*, *Pichia occidentalis*, and *Yarrowia lipolytica* that cannot metabolize lactose⁵⁷. Those studies that do use *K. marxianus* often use non-dairy derived strains with less efficient lactose metabolizing properties^{58,65,101}. Here we identify *K.*

marxianus Y-1190 as having the combined characteristics of robust growth on lactose-containing permeate and tolerance to acid, making an excellent basis for an industrially relevant strain (Fig. 2.1).

To begin our domestication process, we sequenced the genome of *K. marxianus* Y-1190. Our genomic sequencing and observation of genome copy number established the diploid nature of this organism (Fig. S2.6). Further, the haplotigs generated by PacBio sequencing point to a diploid organism with considerable heterozygosity. We observed that the primary assembly contained 4,853 predicted genes, while the haplotig assembly contained 4,108 predicted genes, representing 85% of the predicted genes in the primary assembly. This proportion aligns with the total assembly sizes: the primary assembly spans 1.1×10^7 bp, whereas the haplotigs collectively cover 8.2×10^6 bp, representing 85% of the primary assembly's size across 66 fragments (Table 2.2). The changes observed at the DNA level also translate to amino acid changes (Fig. S2.8) suggesting there could be functional differences between the different alleles. The homothallic nature of *K. marxianus* means that it can be *MATa*, *MAT α* , diploid, or a mixed population^{57,68}. The preferred ploidy of *K. marxianus* is thought to be either haploid or diploid dependent on the strain, however, the mechanism of this preference is currently unknown. Previously published work that sequenced *K. marxianus* strain SP1 (NRRL Y50883) using PacBio saw similar high levels of heterozygosity¹⁶⁴. It is possible that this heterozygosity made generating a stable haploid nearly impossible. Due to the heterologous nature of the genome, the single-nucleotide polymorphisms across alleles may play a role in strain fitness and viability. We were able to obtain one viable clone, JD044, however, we observed that this was not a stable haploid, and though lacking the mating machinery, the genome duplicated to a homozygous diploid state (Fig. S2.6). Losing heterozygosity also cost Y-1190 its high growth rate suggesting that heterozygosity is important

for fitness (Fig. S2.7). This was an unexpected outcome as previously published work described Y-1190 as preferring the haploid state, and after removing the mating type-switching transposases, a stable haploid with mating capability was identified⁷². However, ploidy variation in *K. marxianus* has been observed previously. The ploidy of the commonly used strain CBS 6556 has been debated⁶⁷ and short-read sequencing has made it difficult to answer these questions.

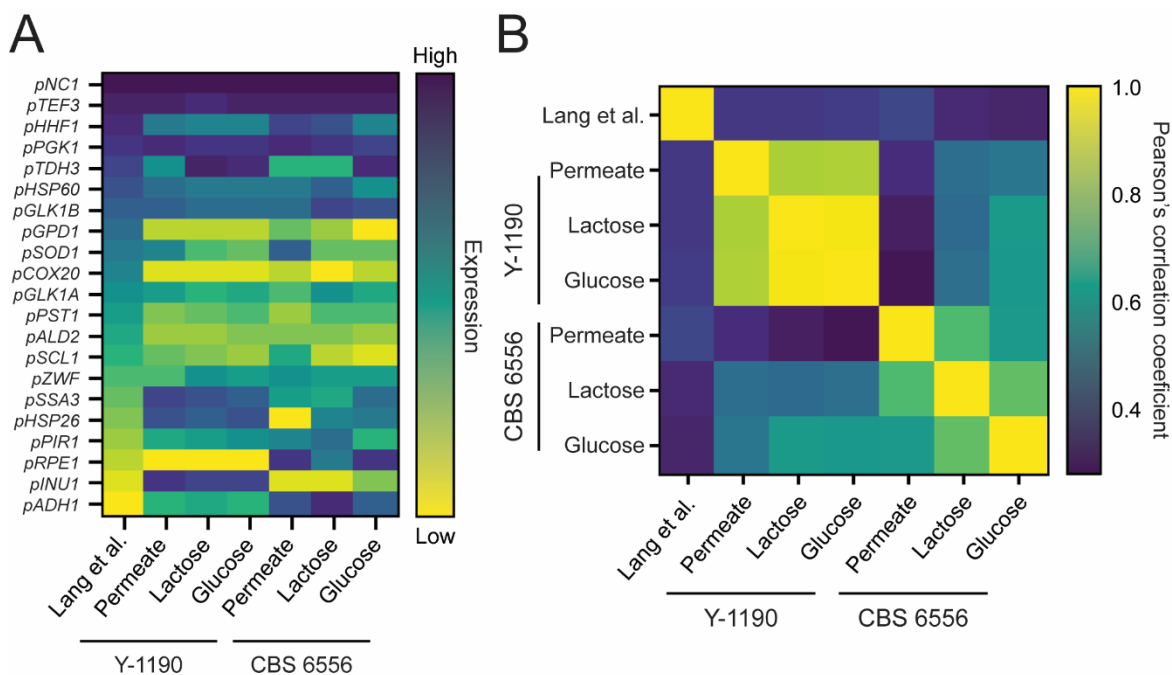


Figure 2.7. Promoter comparison to Lang *et al.* 2020. **A.** Common promoters used in both studies were ranked, with dark blue representing the highest expression and yellow representing the lowest. **B.** The ranked lists were compared by calculating Pearson's correlation coefficient, shown in a heatmap. The correlation values range from $r = 1.0$ (yellow) to $r = 0.4$ (dark blue).

To enhance metabolic engineering tools in *K. marxianus*, we characterized 23 promoters from strains Y-1190 and CBS 6556 across three media types: supplemented permeate, YNBL, and YNBG. Twenty-one of the 23 promoters were sourced from Lang *et al.*¹⁰⁰ allowing a comparison of promoter strength across strains of *K. marxianus* (Fig. 2.7 A). Notably, three high-expressing promoters (*pNC1*, *pTEF3*, *pPGK1*), one mid-expressing promoter (*pGLK1B*), and one low-

expressing promoter (*pRPE1*) exhibited consistent rankings (Fig. 2.7). Overall, the Pearson's correlation coefficients between our findings and Lang *et al.* were weakly positive.

Rajkumar *et al.* tested 16 promoters from CBS 6556, sharing only five with our study, complicating direct comparisons. Our study also overlaps with only four promoters from the *S. cerevisiae* toolkit, yet *pTDH3* and our highest expressor, *pNC1* (homologous to *S. cerevisiae* *pCCW12*), are noteworthy for maintaining high expression both in *S. cerevisiae* and *K. marxianus*¹²⁸.

Promoter strength comparisons within this study reveal greater similarity across strains than media types (Fig. 2.7 B). For example, Y-1190 promoter strengths in supplemented permeate and YNBG were more similar than between Y-1190 and CBS 6556 in YNBG, suggesting strain-specific regulatory differences. Variations in transcription factors or regulatory proteins could explain these discrepancies, as Y-1190 is diploid while CBS 6556 is haploid, potentially affecting expression through gene dosage or chromatin structure^{165,166}. This is reinforced by a previous study that showed promoters maintain their relative activity across different media types¹⁶⁷.

pINUI was identified as a differentially expressed promoter; although typically low in the absence of inulin, it is highly expressed in Y-1190 (Fig. 2.3, Fig. 2.7 A). This suggests regulatory differences or sequence variations affecting activator binding. Overall, the relationship between promoter expression in CBS 6556 and Y-1190 appears linear across media, except in supplemented permeate due to CBS 6556's poor growth (Fig. 2.7 B, Fig. S2.4).

Several CRISPR-based genome-editing systems have been developed for *K. marxianus* in recent years, including Cas9-based approaches that enhance HDR by inactivating NHEJ pathways^{72,101,106,108,109,115}. While these systems have demonstrated successful genome

modification, they often rely on permanent genetic alterations that may introduce long-term instability^{168,169}. To address this, we explored alternative strategies that maintain wild-type backgrounds. Using a two-guide approach to target a single locus, we achieved 100% deletion efficiency and 71% integration efficiency in a wild-type strain (Fig. 2.5 A), representing a significant improvement over previously reported 18% integration efficiency under similar conditions¹⁰⁸.

We further combined this method with NHEJ suppression strategies, through deletion of *DNL4* or chemical inhibition of NHEJ, resulting in increased integration efficiencies of 100% and 98%, respectively. Notably, in wild-type backgrounds where NHEJ remains active, the two-guide system offers the additional advantage of enabling efficient screening for gene inactivation. By targeting a gene with two guides, NHEJ removes the intervening sequence, allowing rapid identification of deletion events by PCR without the need for donor DNA. These results, in combination with our sequencing data, promoter library characterization, inducible promoters, and fermentation optimization have expanded the genetic toolkit for *K. marxianus*, increasing the potential for metabolic engineering applications and enabling more precise control over gene expression for improved biotechnological outcomes.

2.6. Supplementary figures and tables

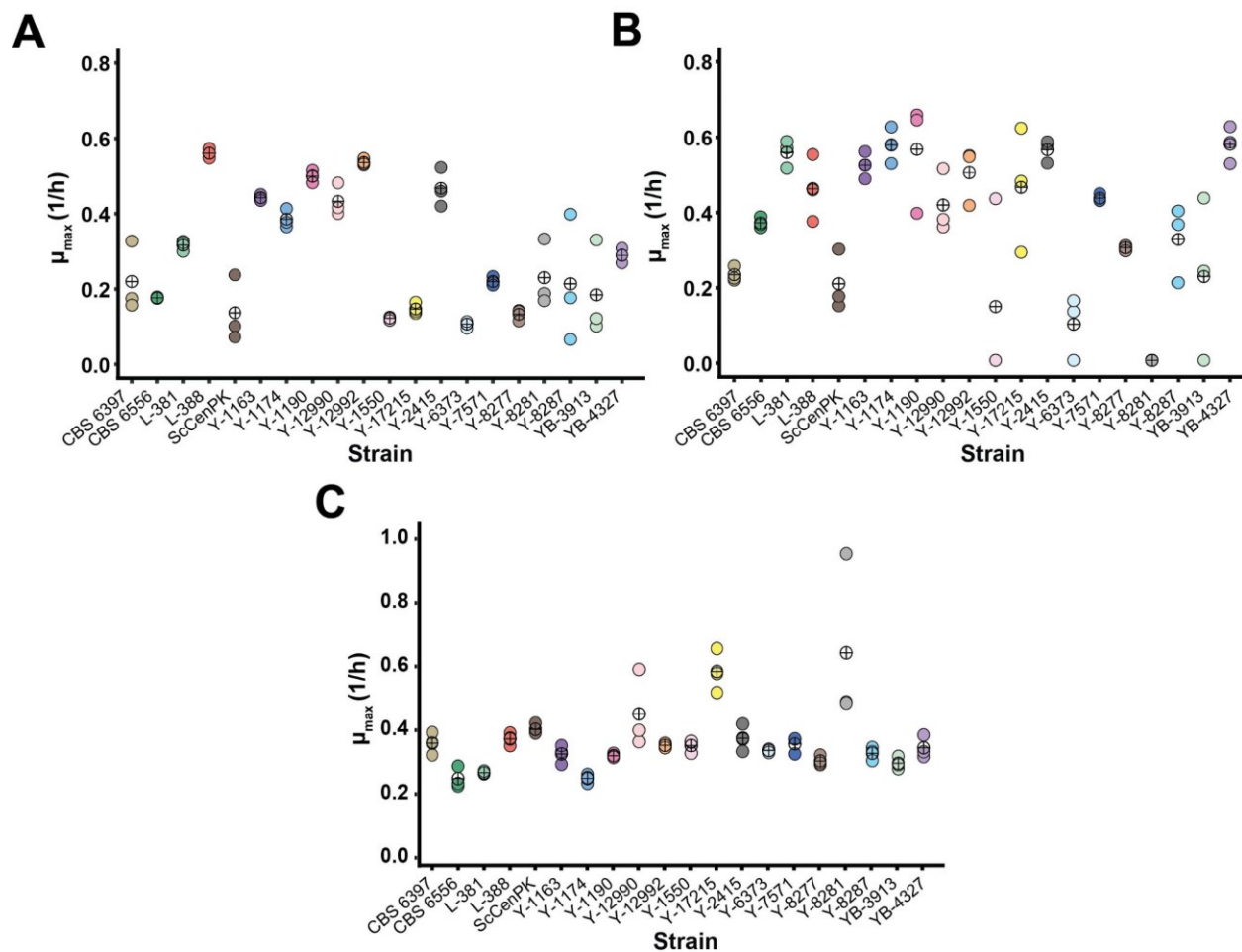


Figure S2.1. Growth rates of 19 strains of *K. marxianus* on A. YNBL, B. Supplemented permeate and C. YNBG. Each coloured dot represents a biological replicate, the + symbol represents the average of the three replicates.

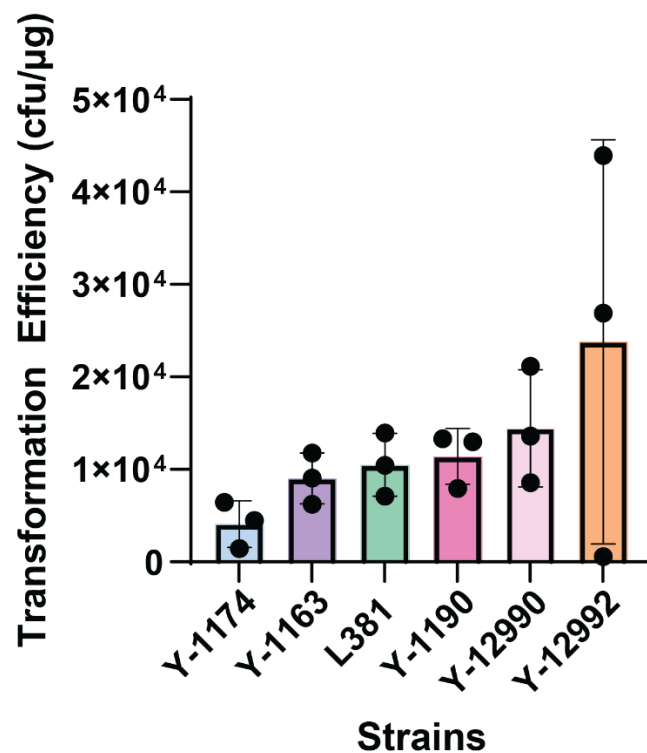


Figure S2.2. Transformation efficiency of our top growing strains of *K. marxianus* using a modified Gietz transformation method.

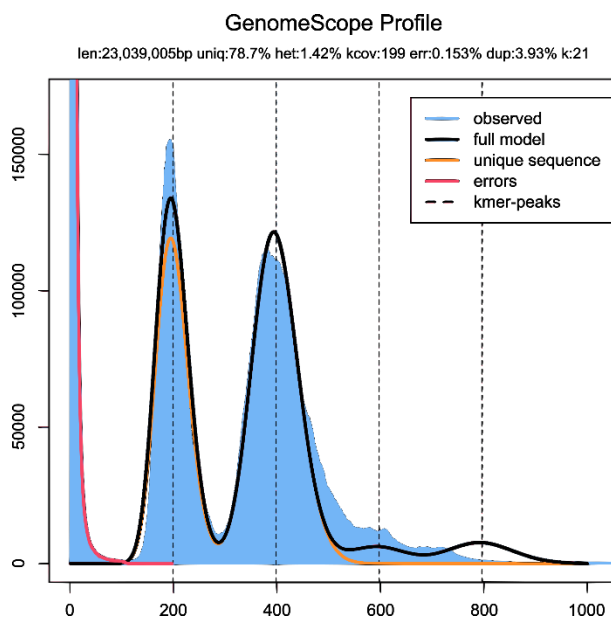


Figure S2.3. k-mer frequency of PacBio raw reads. k-mer coverage of Y-1190 using Jellyfish and Genome Scope

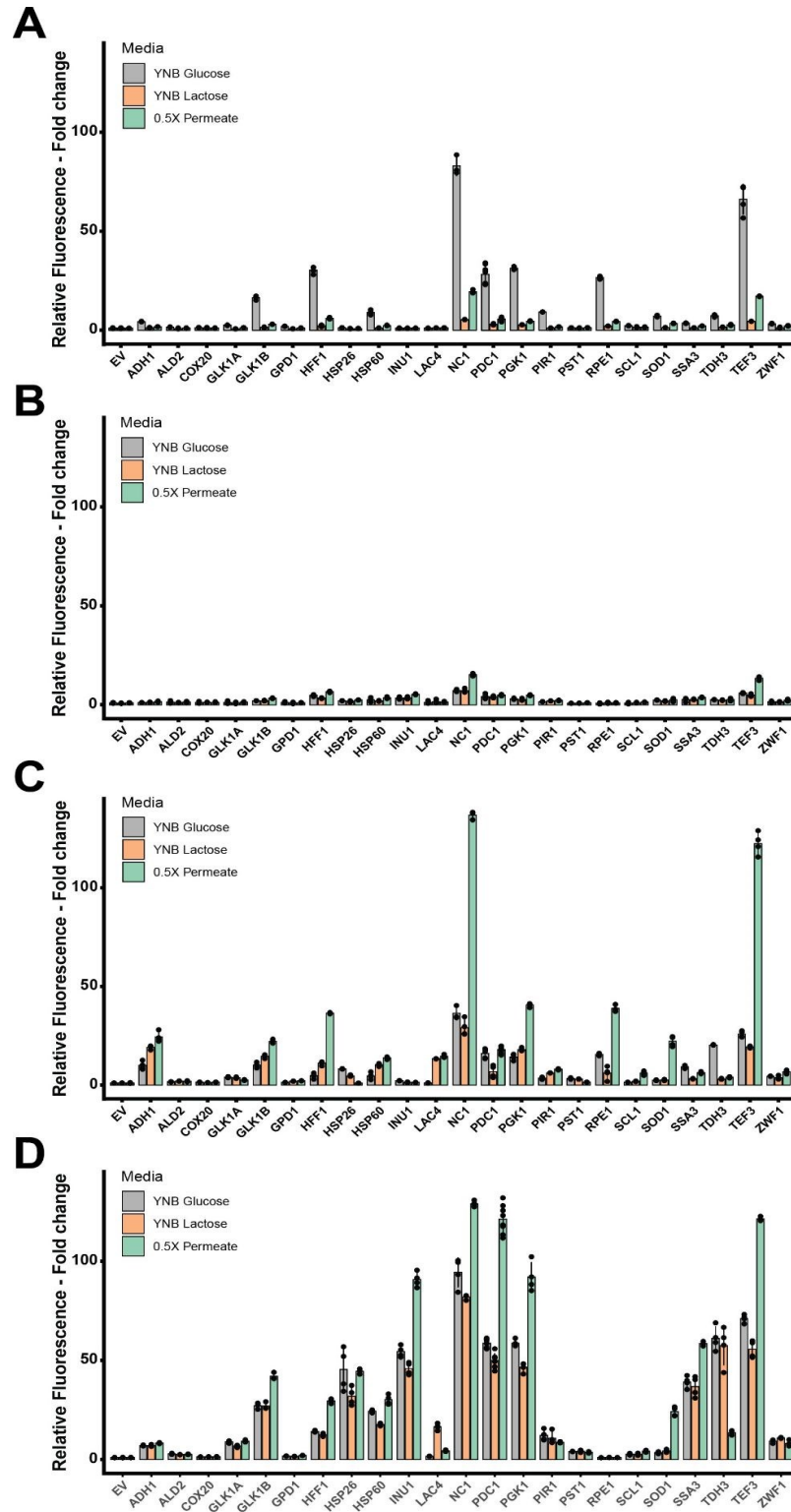


Figure S2.4. Relative promoter expression in *K. marxianus* A. CBS 6556 at 6 h, B. Y-1190 at 6 h, C. CBS 6556 at 24 h, D. Y-1190 at 24 h, YNBL, and supplemented permeate. Each promoter

was tested in 4 biological replicates, the IQR method was used to remove outliers, and the error bars represent the standard deviation.

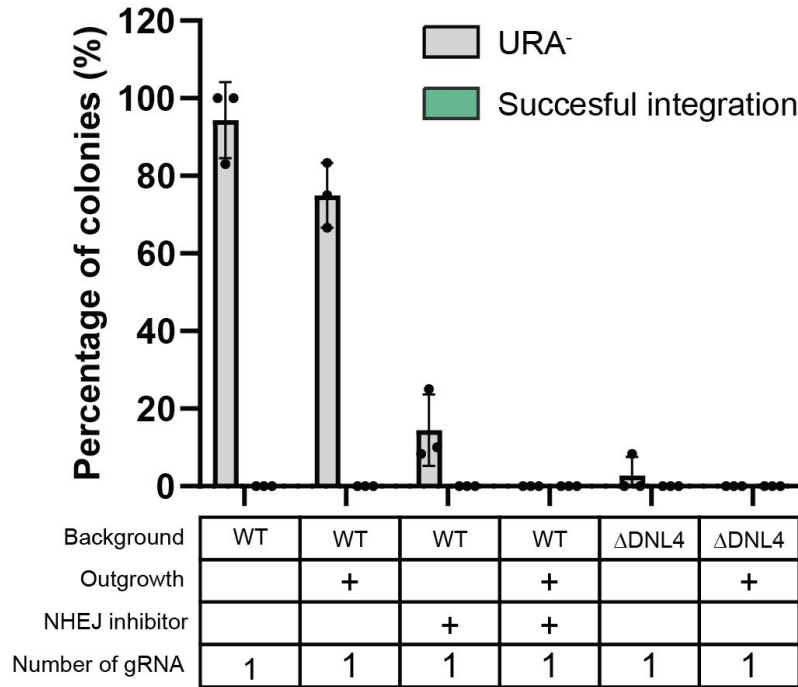


Figure S2.5. CRISPR-Cas9 genome editing efficiencies using one gRNA (sequence 2) for genomic deletion and integration via HDR. Each dot represents a biological replicate, and the error bars represent the standard deviation.

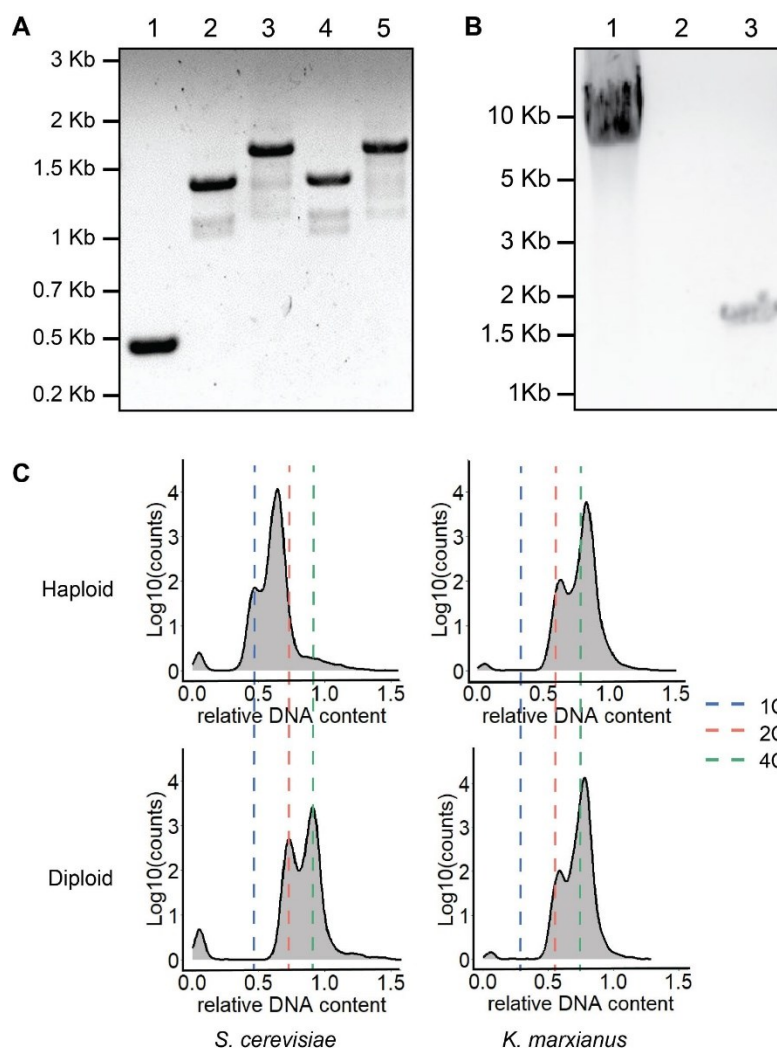


Figure S2.6. Ploidy and mating type analysis **A.** Detection of disrupted *Kat1* and $\alpha 3$ in *K. marxianus* by PCR and agarose gel electrophoresis. Lane 1: *Kat1* amplification by primers MT169/MT172 (2kb in WT, 489bp when truncated). Lane 2: $\alpha 3$ amplification from the *MAT* locus by primers JD108/JD110 (2.1kb in WT, 1.4kb when truncated). Lane 3: $\alpha 3$ amplification from the *MAT* locus by primers JD088/JD109 (2.4kb in WT, 1.7kb when truncated). Lane 4: $\alpha 3$ amplification from the *HML* locus by primers JD108/JD109 (2.1kb in WT, 1.4kb when truncated). Lane 5: $\alpha 3$ amplification from the *HML* locus by primers JD088/JD109 (2.3kb in WT, 1.6kb when truncated). **B.** Determination of mating type following sporulation of *K. marxianus* Y-1190 ($\Delta Kat1:\Delta \alpha 3$). Lane 1: Amplification of *MAT* locus using primers JD103/JD106 (6.8kb if *MAT* α , 3.8kb if *MAT* α). Lane 2: Amplification of *MAT* locus using primers MT240/MT241 (1.1kb if *MAT* α). Lane 3: Amplification of *MAT* locus using primers MT240/MT242 (1.5kb if *MAT* α). **C.** Determination of ploidy following sporulation of *K. marxianus* Y-1190 ($\Delta Kat1:\Delta \alpha 3$). Propidium iodide staining followed by flow cytometry was performed to determine the ploidy of a known *S. cerevisiae* diploid (bottom left), haploid (top left), *K. marxianus* diploid (Y-1190 $\Delta Kat1:\Delta \alpha 3$, bottom right), and the sporulated *MAT* α “haploid” (top right).

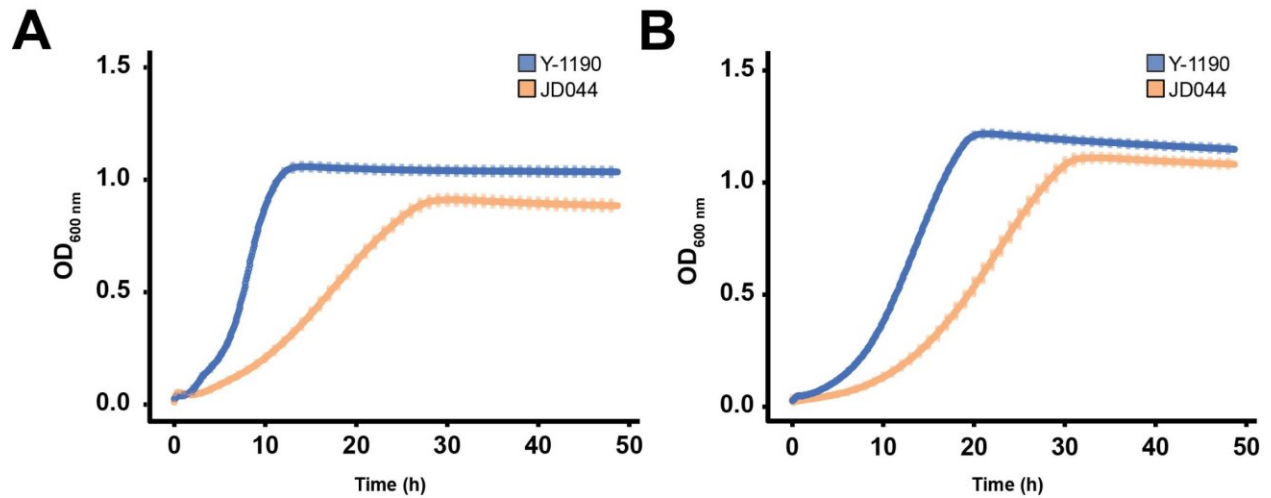


Figure S2.7. Growth curves of Y-1190 and homologous diploid JD044 in A. YNBL and B. Supplemented permeate. Growth was measured using OD₆₀₀ every 5 min for 48 hr. The data points are an average of three biological replicates and the error bars represent standard deviation.

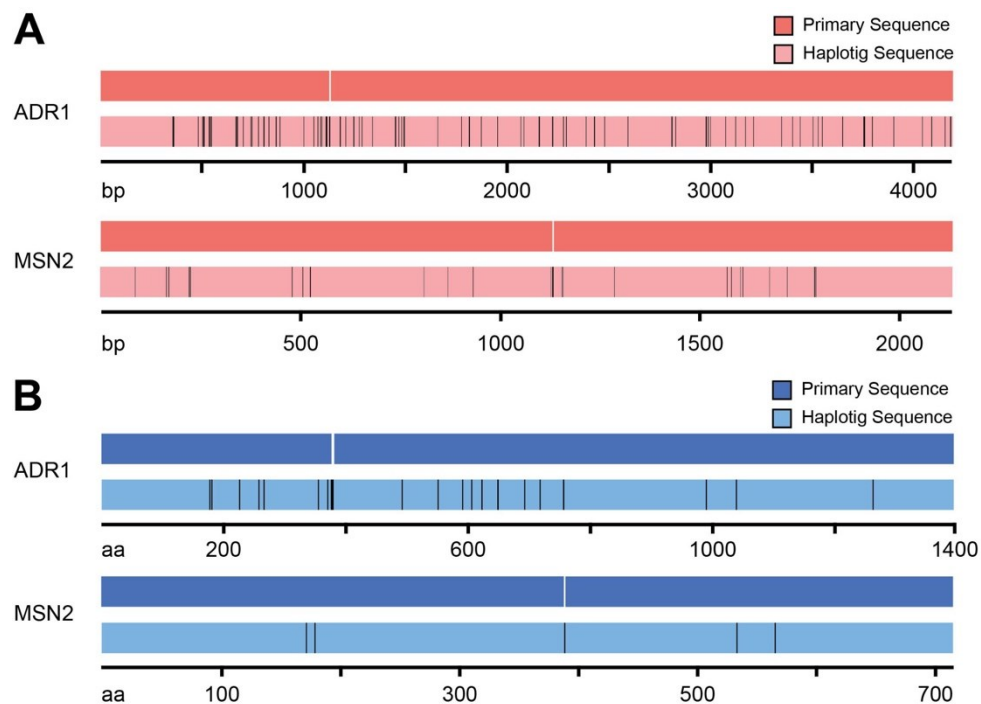


Figure S2.8. Primary and haplotig sequence comparison for two example genes A. Nucleotide sequences B. Amino acid sequences. Black lines represent changes in the A) nucleotide sequence and B) amino acid sequence.

Table S2.1. Genome Contig Size

Strain	DMKU-1042		Y1190	
	Name	Size (bp)	Name	Size (bp)
	chr7	9.63E+05	0F	1.74E+06
	chr1	1.75E+06	1F	1.74E+06
	chr2	1.71E+06	2F	1.71E+06
	chr3	1.59E+06	3F	1.59E+06
	chr4	1.42E+06	4F	1.45E+06
	chr6	1.20E+06	5F	1.23E+06
	chr8	9.40E+05	7F	9.80E+05
	chr5	1.35E+06	14F	2.36E+05
			11F	2.27E+05
	mt	4.63E+04	32	4.09E+04
Total		1.10E+07		1.10E+07

Table S2.2. List of promoters

Name	Source	Citation
<i>ADH1pr</i>	<i>K. marxianus</i> CBS 6556	100
<i>ALD4pr</i>	<i>K. marxianus</i> CBS 6556	100
<i>COX20pr</i>	<i>K. marxianus</i> CBS 6556	100
<i>GLK1Apr</i>	<i>K. marxianus</i> CBS 6556	100
<i>GLK1Bpr</i>	<i>K. marxianus</i> CBS 6556	100
<i>GPD1pr</i>	<i>K. marxianus</i> CBS 6556	100
<i>HFF1pr</i>	<i>K. marxianus</i> CBS 6556	100
<i>HSP26pr</i>	<i>K. marxianus</i> CBS 6556	100
<i>HSP60pr</i>	<i>K. marxianus</i> CBS 6556	100
<i>INU1pr</i>	<i>K. marxianus</i> CBS 6556	100
<i>NC1pr</i>	<i>K. marxianus</i> CBS 6556	100
<i>PGK1pr</i>	<i>K. marxianus</i> CBS 6556	100
<i>PDC1pr</i>	<i>K. marxianus</i> CBS 6556	101
<i>PIR1pr</i>	<i>K. marxianus</i> CBS 6556	100
<i>PST1pr</i>	<i>K. marxianus</i> CBS 6556	100
<i>RPE1pr</i>	<i>K. marxianus</i> CBS 6556	100
<i>SCL1pr</i>	<i>K. marxianus</i> CBS 6556	100
<i>SOD1pr</i>	<i>K. marxianus</i> CBS 6556	100

<i>SSA3pr</i>	<i>K. marxianus</i> CBS 6556	¹⁰⁰
<i>TDH3pr</i>	<i>K. marxianus</i> CBS 6556	¹⁰⁰
<i>TEF3pr</i>	<i>K. marxianus</i> CBS 6556	¹⁰⁰
<i>ZWF1pr</i>	<i>K. marxianus</i> CBS 6556	¹⁰⁰
<i>LAC4pr</i>	<i>K. marxianus</i> CBS 6556	¹⁰¹
<i>TDH3_{tetO2}</i>	<i>K. marxianus</i> Y-1190	This study
<i>TEF3_{tetO2}</i>	<i>K. marxianus</i> Y-1190	This study
<i>PGK1_{tetO2}</i>	<i>K. marxianus</i> Y-1190	This study

Table S2.3. List of yeast strains used in this study

Species	Strain Name	Genetic modification	Aliases	Isolated From	Source
<i>K. marxianus</i>	Y-12990		Lachance UWO-PS SM 3-4	Cheese	NRRL
<i>K. marxianus</i>	Y-6373			Citrus peel juice	NRRL
<i>K. marxianus</i>	Y-8281		CBS 608, NCYC 143	Unknown (NRRL)	NRRL
<i>K. marxianus</i>	CBS 6556		Y-7571, NCYC 2597, ATCC 26548, KCTC 17555	Pozol	Westerdijk Institute
<i>K. marxianus</i>	Y-1174		ATCC 64884, ATCC 8619, Laffer 98	Whey	NRRL
<i>K. marxianus</i>	L-388				Lallemand
<i>K. marxianus</i>	CBS 6397				Westerdijk Institute
<i>K. marxianus</i>	L-381				Lallemand
<i>K. marxianus</i>	Y-1163		ATCC 8608, DSM 5420, KY 5463, LAFFER 85, VTT C-75024	Dairy cream	NRRL
<i>K. marxianus</i>	Y-8277		ATCC 22295, ATCC 56495, BKM Y-1517, CBS 4857, CCY 43-3-1, UCD-FST 71-14	Cow's back	NRRL
<i>K. marxianus</i>	Y-12992		Lachance UWO-PS SM 16-10	Cheese	NRRL
<i>K. marxianus</i>	Y-17215		CBS 6923, CSIR Y-808	Garden soil	NRRL
<i>K. marxianus</i>	YB-3913		CBS 3002, Dietrichson V-435	Expectorate	NRRL
<i>K. marxianus</i>	YB-4327		ATCC 16045, ATCC 56494, BKM Y-1242, BKM Y-1494, CBS 2762, NCYC 970, Santa Maria Y2, UCD-FST 71-13, CRRC 21623, VTT C-81109	Yogurt	NRRL
<i>K. marxianus</i>	Y-1190		ATC 8635, ATCC 28910, BI CZAS 326/5, BKM Y-2012, CCY 51-1-1, DEM 1128, DSM 5421, Henneberg 236, IBFM Y-	Yogurt	NRRL

			613, Laffer 132, VTT C-81110		
<i>K. marxianus</i>	Y-1550		CBS 396, ATCC 10606, IFO 0260, CCRC 21722	AIR	NRRL
<i>K. marxianus</i>	Y-2415		CBS 397, ATCC 46537, ATCC 56752, NCYC 851, CCRC 21477, DBVPG 6164, DSM 5422, IFO 1735, TAGUCHI KCM-003, TAGUCHI KCM-0230, UCD-FST 71-58	Yogurt	NRRL
<i>K. marxianus</i>	Y-8287		ATCC 22296, ATCC 56501, BKM Y-1545, CBS 5671, CCY 43-5-1, UCD-FST 71-15, van der Walt K.210	Bantu beer, sorghum	NRRL
<i>K. marxianus</i>	Y-7571		CBS 6556, NCYC 2597, ATCC 26548, KCTC 17555	Pozol	NRRL
<i>S. cerevisiae</i>	CEN.PK 113_13D	Δ URA3 mata			
<i>K. marxianus</i> Y-1190	AK081	Δ DNL4			This work
<i>K. marxianus</i> Y-1190	AK086	ARO1::SSA3p-TetR-NLS-INU1t-TDH3-tetO2p-mNeonGreen-PDC1t			This work
<i>K. marxianus</i> Y-1190	AK087	ARO1::SSA3p-TetR-NLS-INU1t-TEF3-tetO2p-mNeonGreen-PDC1t			This work
<i>K. marxianus</i> Y-1190	AK088	ARO1::SSA3p-TetR-NLS-INU1t-PGK1-tetO2p-mNeonGreen-PDC1t			This work
<i>K. marxianus</i> Y-1190	JD032	Δ Kat1 $\Delta\alpha$ 3			This work
<i>K. marxianus</i> Y-1190	JD044	Δ Kat1 $\Delta\alpha$ 4 MAT α			This work

Table S2.4. List of primers and synthetic sequences

Application	Oligo ID	Nucleotide sequence (5' – 3')
Amplification of URA3 donor DNA with 900bp homology arms	MT293	GCAAGGGAATTTGCCACC
Amplification of URA3 donor DNA with 900bp homology arms	MT298	TGGTATGCTTCTTGTGCATTC
Amplification of URA3 donor DNA with 450bp homology arms	MT299	CAAGGATGCTCATCACAATACG
Amplification of URA3 donor DNA with 450bp homology arms	MT300	AGTATAGACGCATGAAGTCCTTC
Amplification of URA3 donor DNA with 200bp homology arms	AK772	TGATCGTTTTTAACATTTTCAAA GG

Amplification of URA3 donor DNA with 200bp homology arms	AK773	AAATTTACATAATGTACTTCAC C
Amplification of URA3 donor DNA with 100bp homology arms	AK774	AAATACTACCGTAAAAAGTGTAAGTTTC
Amplification of URA3 donor DNA with 100bp homology arms	AK775	CCATTACACTCTAACATCTCTGTATAG
Amplification of URA3 donor DNA with 50bp homology arms	AK776	AACAGCAAGAGGAAGTATCATC
Amplification of URA3 donor DNA with 50bp homology arms	AK777	TAACTTCTGATCCGAGTACACTC
Hybridization of donor DNA for DNL4 deletion	AK702	TGTCAACTTCTAACAATGAGTGT TTATTGGTTCATGAGAAAAAGAA CAAAAAGAATCACACAATCTTT GTTGGAGCTAGTTGGAGCTAGAA GAGGGTTT
Hybridization of donor DNA for DNL4 deletion	MT481	GTGTCCGAGTTCTTCGGTAATCTG AGATAATGACACAATGCTTTGAT TAAAGTGTTCCTTTTATATTATA AACCCTCTTCTAGCTCCAA
Amplification diagnostic for DNL4 deletion	MT482	ATAAATCCATTCTAAGGCAACT
Amplification diagnostic for DNL4 deletion	MT483	TGCGGAAACCTTCTGTATATAA
Hybridization for sgRNA cloning into pUCC001 targeting DNL4	MT152	CGTCAACCATTAAACTCTGGAG A
Hybridization for sgRNA cloning into pUCC001 targeting DNL4	MT126	AAACTCTCCAGAGTTTAAATGGT T
Hybridization for one sgRNA cloning to generate pAK214	AK1144	GAAGTAGGTCTCGGACTTTGAAG ACACATGTAGATATCTGTTTGGAGACCGTAGCC
Hybridization for one sgRNA cloning to generate pAK214	AK1145	GGCTACGGTCTCCAAACAGATAT CTACATGTGTCTTCAAAGTCCGAGACCTAGTTC
Amplification of tRNAgly and SpCas9 scaffold for URA3 double sgRNA cloning into pJD005	JD118	AAAAGGTCTCAGACTGAAGACAC ATGTAGATATCTgttttagagctagaatag caagttaaataag
Amplification of tRNAgly and SpCas9 scaffold for URA3 double sgRNA cloning into pJD005	JD119	TTTTGGTCTCTAAACGGCGTCATG ATCAACCAATCtgcgcaagcccggaatc
Hybridization for one sgRNA cloning to generate pAK219	AK1166	GACTTTGATTGGTTGATCATGAC GCC
Hybridization for one sgRNA cloning to generate pAK219	AK1167	AAACGGCGTCATGATCAACCAAT CAA
Amplification of the synthetic gene TetR-NLS for pAK014 cloning	AK874	GCATCGTCTCATCGGTCTCATATG TCCAGGCTAGATAAATCTAAAG
Amplification of the synthetic gene TetR-NLS for pAK014 cloning	AK875	ATGCCGTCTCAGGTCTCAGGATC TACACCTTCCTTTCTTCTTC
Amplification of the 5' side of the TDH3 promoter for pAK015 cloning	AK876	GCATCGTCTCATCGGTCTCAAAC GCACTATATCAGGCCTCCAC
Amplification of the 5' side of the TDH3 promoter for pAK015 cloning	AK877	GGCTACCGTCTCATATCTCTATCA CTGATAGGGAAAAATGTGGTCTTT ATATATGGCAG
Hybridization of the 3' side of the TDH3 promoter for pAK015 cloning	AK878	GGCTACCGTCTCAGATATCCCTA TCAGTGATAGAGAAGTCTATAAT AATTATATCTCTTCCTTGTTCCTC CCAAACCAAGTCTT
Hybridization of the 3' side of the TDH3 promoter for pAK015 cloning	AK879	ATGCCGTCTCAGGTCTCACATAT GTGAATGTGTAAAAAGTGTGTGTG TACTGTTGTTAAATCTAAAAGAC TTGGTTTGGGAGGAA

Amplification of the 5' side of the TEF3 promoter for pAK016 cloning	AK880	ATGTCGTCTCATCGGTCTCAAAC GAAGCAAAGAAGTAACAGCAGG
Amplification of the 5' side of the TEF3 promoter for pAK016 cloning	AK881	GGCTACCGTCTCATATCTCTATCA CTGATAGGGATTTCAATCTCCTTT TATATTCAAAAAAATTAC
Hybridization of the 3' side of the TEF3 promoter for pAK016 cloning	AK882	GGCTACCGTCTCAGATATCCCTA TCAGTGATAGAGATTGATTTGTA GTTTATTTACTAGTATTCTCATAG TTCTAAC
Hybridization of the 3' side of the TEF3 promoter for pAK016 cloning	AK883	ATGCCGTCTCAGGTCTCACATAC TTTAATGTTACTTCTCTTGAGTT AGAACTATGAGAACTAGTAAA TAAAC
Amplification of the 5' side of the GLK1 promoter for pAK017 cloning	AK884	GCATCGTCTCATCGGTCTCAAAC GTTACCCTCACTCTTTCACATTAC
Amplification of the 5' side of the GLK1 promoter for pAK017 cloning	AK885	GGCTACCGTCTCATATCTCTATCA CTGATAGGGAGGGACACCTATAT TTATACAATTTTTCAC
Hybridization of the 3' side of the GLK1 promoter for pAK017 cloning	AK886	GGCTACCGTCTCAGATATCCCTA TCAGTGATAGAGAAGACTTGATA AGTATTTCCCTAGTAATACCCCA CCAACCATCGGTTGATACATTTT AATAC
Hybridization of the 3' side of the GLK1 promoter for pAK017 cloning	AK887	ATGCCGTCTCAGGTCTCACATATT TTGTATCTTTATATAGGTAGTGTG TATTAAAATGTATCAACCGATGG TTG
Primer outside of homology region of ARO1 for diagnostic PCR	MT487	CTTTACGTGATCGAGACCG
Primer outside of homology region of ARO1 for diagnostic PCR	MT488	TCTAGCCATTTGATATCAACAAA C
Bacterial TetR synthetic gene K. lactis codon optimized and fused to a NLS sequence	TetR-SV40	ATGTCCAGGCTAGATAAAATCTAA AGTCATTAACCTCGGCCTTAGAAC TACTAAATGAAGTAGGTATCGAA GGCTTGACTACAAGAAAATTGGC CCAGAAGTTAGGTGTAGAGCAGC CCACTCTATATTGGCACGTTAAA AATAAAAGGGCCTTGCTTGACGC CTTGGCTATCGAGATGCTAGATC GTCATCACACACATTTCTGCCCAT TAGAAGGTGAAAGTTGGCAGGAT TTCTTAAGAAATAATGCAAAATC GTTTCAAGATGCGCTCTATTATCCCA TAGAGATGGGGCAAAGGTCCATT TGGGCACGAGGCCGACGGAGAA ACAGTACGAGACTTTAGAGAACC AGTTGGCCTTTTTGTGTCAACAA GGATTTAGTTTGGAAAATGCCTT GTATGCCTTATCGGCAGTCGGAC ACTTTACGCTAGGTTGCGTATTA GAAGACCAAGAACACCAGGTTGC AAAGGAGGAAAGGGAAACTCCC ACAACCGACTCAATGCCTCCACT TTTGAGACAGGCCATAGAATTGT TCGACCATCAGGGTGCAGAGCCA GCTTTTTTATTCGGCTTGGAATTG ATTATATGCGGATTGGAAAAACA ATTGAAATGTGAATCTGGGTCCG GAGGACCGAAGAAGAAAAGGAA GGTGTAG

Amplification for the generation of ARO1 double guide in the making of pMT193	MT489	AAAAGGTCTCAGACTTTAACTG GCTTCAAATCATAAgttttagagctagaaa tagcaagttaaaataag
Amplification for the generation of ARO1 double guide in the making of pMT193	MT491	TACGGGTCTCTAAACCCCTTAAT GTTTACGCCAAATtgcgcaagcccggaatc
Amplification of panARS(OPT) for gibson reaction to produce plasmids pMT306-pMT312	MT675	GATCCTTTGATCTTTTCTACACTA GTtcaacatcttggataatatcagaatg
Amplification of panARS(OPT) for gibson reaction to produce plasmids pMT306-pMT312	MT676	GTTGAAGAAAGAGTATACTACAT AACACATATACtagtctgattatgattgac g
Amplification of ADHp into golden gate level 0	MT001	GCATCGTCTCAAACGCATCATC CATCCACACACG
Amplification of ADHp into golden gate level 0	MT002	ATGCCGTCTCACATATGTGTTG TGTATATGATTTGTTTTTAAAT G
Amplification HHF1 into golden gate level 0	MT003	GCATCGTCTCAAACGGTAAGTAA TGGAGGATTGATATGTTT
Amplification HFF1 into golden gate level 0	MT004	ATGCCGTCTCACATAGTTTATTTA TTGATTGCTGTTTATTGC
Diagnostic primer for ADHp, use with MT006	MT005	CTTCGAGCGCTGAAGTATC
diagnostic primer for promoters targeting CAM	MT006	ccctaaagggtttattgagaatatg
diagnostic primer for HHF1, use with MT006	MT007	GTTCCATGGGCCTGAAC
Get ALD2p into golden gate level 0	MT012	GCATCGTCTCAAACGCTCTAAAA CCCATGTCCG
Get ALD2p into golden gate level 0	MT013	ATGCCGTCTCACATAATTTCTACT TTGTATACGGCTG
ALD2 Insert check with MT006	MT014	AACCTAGTAGGAGTGGTACTC
Get GLK1Ap into golden gate level 0	MT015	GCATCGTCTCAAACGCCTTCACA AAAGAAAAGTATTTTG
Get GLK1Ap into golden gate level 0	MT016	ATGCCGTCTCACATAATTTAAGTT TTATGTTACTAGTTTCCG
GLK1A insert check with MT006	MT017	CACTTAATATGGGAACCATTTCC
Get GLK1Bp into golden gate level 0	MT018	GCATCGTCTCAAACGTTGCTAGC CGTTTTTCC
Get GLK1Bp into golden gate level 0	MT019	ATGCCGTCTCACATATTTTGTAA GTGTGTGTGTTG
GLK1B insert check with MT006	MT020	GGAATGGAAATGGAAATGGG
Get GPD1p into golden gate level 0	MT021	GCATCGTCTCAAACGAAAGGGAC AATCGGAGAATAG
Get GPD1p into golden gate level 0	MT022	ATGCCGTCTCACATAATCTGAAA ACTGATTCTCCAC
GPD1 insert check with MT006	MT023	CTAGTCCGATTCTCTCTCC
Get HSP26p into golden gate level 0	MT027	GCATCGTCTCAAACGCGCGCATG ACAAGGG
Get HSP26p into golden gate level 0	MT028	ATGCCGTCTCACATACTCGTAAT CGCTTTTGTCTTAG
HSP06 insert check with MT006	MT029	TACGAGAGAGTTCTTTGCATG
Get PGKp into golden gate level 0	MT030	GCATCGTCTCAAACGTTACCCTC ACTCTTTCACATTAC
Get PGKp into golden gate level 0	MT031	ATGCCGTCTCACATATTTTGTATC TTTATATAGGTAGTGTG
PGK1 insert check with MT006	MT032	CATTGTGAATGGTATTGTAGC

Get PST1p into golden gate level 0	MT033	GCATCGTCTCAAACGCATTTTCAG AGTTCTTTTGTC
Get PST1p into golden gate level 0	MT034	ATGCCGTCTCACATAATCTGTTA ATATGAACAGGTC
PST insert check with MT006	MT035	ACAGGCGACAAATGACAG
Get SCL1p into golden gate level 0	MT036	GCATCGTCTCAAACGCAGGCTTG CGCTTCTC
Get SCL1p into golden gate level 0	MT037	ATGCCGTCTCACATATCTGTTCAC ACTTGATAACG
SCL1 insert check with MT006	MT038	GTTACACCTCTAATCTACCACC
Get SSA3p into golden gate level 0	MT039	GCATCGTCTCAAACGCTTGCCAC TATCAAGTGAG
Get SSA3p into golden gate level 0	MT040	ATGCCGTCTCACATAAATTCTTCT TCTTTTTAGCTCTG
SSA3 insert check with MT006	MT041	TAAGTAGCGCAGAAGACTG
Get TDH3p into golden gate level 0	MT042	GCATCGTCTCAAACGCACTATAT CAGGCCTCCAC
Get TDH3p into golden gate level 0	MT043	ATGCCGTCTCACATATGTGAATG TGAAAAGTGTGTG
TDH3 insert check with MT006	MT044	CAACCAACAACACCCCTTC
Get TEF3p into golden gate level 0	MT045	GCATCGTCTCAAACGAACACCGA TGAAGCAAAGAAG
Get TEF3p into golden gate level 0	MT046	ATGCCGTCTCACATACTTTAATGT TACTTCTCTTGGAGTTAG
TEF3 insert check with MT006	MT047	ATGCGACTGGTTCTGAG
Get RPE1p into golden gate level 0	MT053	GCATCGTCTCAAACGTGGGCAAC TTCTCAAAC
Get RPE1p into golden gate level 0	MT054	ATGCCGTCTCACATACACAGCTA GAAACTGG
RPE1 remove internal bsmbI site (use bsai)	MT055	GCATGGTCTCATCTGTACTAAGC ACCCCC
RPE1 remove internal bsmbI site (use bsai)	MT056	ATGCGGTCTCACAGACGTGTTTG GATTATAGGG
RPE1 Insert check with MT006	MT057	CAGGAAGTAGCAGCATGTAC
Get ZWFp into GG level 0	MT062	GCATCGTCTCAAACGCACCCGCC TAAAAAAAAAAGG
Get ZWFp into GG level 0	MT063	ATGCCGTCTCACATAGGGCTTGT TCAAATGTTTCAG
ZWFp diagnostic with MT006	MT064	TGCAGCAGCTTAGTGTG
Get HSP60 into GG level 0	MT065	GCATCGTCTCAAACGGAAGATAC AGAGAAACGACC
Get HSP60 into GG level 0	MT066	ATGCCGTCTCACATACCCGGGTG TATTGTATTG
HSP60 diagnostic with MT006	MT067	TTGCAGCTTTCCCAGC
Get PIR1 into GG level 0	MT068	GCATCGTCTCAAACGGCCTGGGT TTTCAAATTGAC
Get PIR1 into GG level 0	MT069	ATGCCGTCTCACATAGTGTATAA ATTGAAATGTTTGGATTG
PIR1 diagnostic with MT006	MT070	GGCAAGAGAGTACAGGTAC
Get SOD1 into GG level 0	MT071	GCATCGTCTCAAACGCAATACCA CCCGCCAG
Get SOD1 into GG level 0	MT072	ATGCCGTCTCACATAGGGGATTA ATTAATTGTATGTGTTATG
SOD1 diagnostic with MT006	MT073	CTTACTAGTCAGCTAGCTAGG

Get INU1 into GG level 0	MT074	GCATCGTCTCAAACGATACCACC CGCCATGG
Get INU1 into GG level 0	MT075	ATGCCGTCTCACATAGGATCTAA CAAAAAAAAAAATTAAATGTGTC
INU1 diagnostic with MT006	MT076	CTGCTCCAAAATCCATCTTC
Amplification of tRNAgly and SpCas9 scaffold for Kat1 double sgRNA cloning into pLN012 to generate pLN026	LN1176	GGTCTCACGTCTGAAGAGTGATTC ACAGTTTCGTTTTAGAGCTAGAA ATAGCAAGTTAAAA
Amplification of tRNAgly and SpCas9 scaffold for Kat1 double sgRNA cloning into pLN012 to generate pLN026	LN1177	GGTCTCAaaacCATTGTAGTAAGA GATACTTGC GCAAGCCCCGGAATC
Amplification of tRNAgly and SpCas9 scaffold for $\alpha 3$ double sgRNA cloning into pLN012 to generate pLN028	LN1198	GGTCTCACGTCAATTGGCTTTTGA AAAACGGGTTTTAGAGCTAGAAA TAGCAAGTTAAAA
Amplification of tRNAgly and SpCas9 scaffold for $\alpha 3$ double sgRNA cloning into pLN012 to generate pLN028	LN1199	ggtctcaaaacAATATATGCAGAGTTT GTCATGCGCAAGCCCCGGAATC
Diagnostic primer to detect truncation of Kat1	MT169	GATCTGTCAAACGTGCAC
Diagnostic primer to detect truncation of Kat2	MT172	TGTACTAGCATTGTATGATTTCAC
Diagnostic primer to detect truncation of $\alpha 3$	JD088	CCTTTTCCGAAGAAGTTAGC
Diagnostic primer to detect truncation of $\alpha 3$	JD108	CATTGTGATTCCGACACC
Diagnostic primer to detect truncation of $\alpha 3$ in HML locus	JD109	CATTGTGGCAAGAAGAGC
Diagnostic primer to detect truncation of $\alpha 3$ in MAT locus	JD110	GATATGAGTCCTTATGCATTGC
Diagnostic primer to determine mating type	JD103	AACAAACGAATGCTTTGG
Diagnostic primer to determine mating type	JD106	AGCGTCTGTCTCAAATATTCC
Diagnostic primer to determine mating type	MT240	TGCAACCAACCAATCCCTTCCAA ATTCT
Diagnostic primer to determine mating type	MT241	TCTTCCTTGAACCCGAAGCAAAA GATC
Diagnostic primer to determine mating type	MT242	AACTTCAATCCCCGACCCACCGC AGTC

Table S2.5. List of plasmids used in this study

Plasmid ID	Characteristics	Selection	Reference	Addgene ID
pBBK05	pCAS-Tyr-[gRNA: BsaI GFP dropout]	NatR, AmpR	¹³⁹	178989
pMTKM01	pCAS-Tyr-[gRNA: BsaI GFP dropout] with pan(OPT)ARS replication origin	NatR, AmpR	This work	233473
pMTKM02	pCAS-Tyr-[gRNA: BsaI GFP dropout] with pan(OPT)ARS replication origin	AmpR, splitNatR	This work	233474
pMTKM03	pCAS-Tyr-[gRNA: BsaI GFP dropout] with pan(OPT)ARS replication origin	AmpR, G418	This work	233475

pMTKM04	pCAS-Tyr-[gRNA: BsaI GFP dropout] with pan(OPT)ARS replication origin	AmpR, splitG418	This work	233476
pMTKM05	pCAS-Tyr-[gRNA: BsaI GFP dropout] with pan(OPT)ARS replication origin	AmpR, HygR	This work	233477
pMTKM06	pCAS-Tyr-[gRNA: BsaI GFP dropout] with pan(OPT)ARS replication origin	AmpR, splitHygR	This work	233478
pMTKM07	pCAS-Tyr-[gRNA: BsaI GFP dropout] with pan(OPT)ARS replication origin	AmpR, -URA	This work	233479
pMTKM08	pCAS-Tyr-[gRNA: BsaI GFP dropout] with pan(OPT)ARS replication origin	AmpR, split-URA	This work	233480
pMTKM09	Level 0 plasmid MoClo with KM promoter PST	CamR	This work	233481
pMTKM10	Level 0 plasmid MoClo with KM promoter GLK1A	CamR	This work	233482
pMTKM11	Level 0 plasmid MoClo with KM promoter GLK1B	CamR	This work	233483
pMTKM12	Level 0 plasmid MoClo with KM promoter SCL1	CamR	This work	233484
pMTKM13	Level 0 plasmid MoClo with KM promoter ZWF	CamR	This work	233485
pMTKM14	Level 0 plasmid MoClo with KM promoter RPE1	CamR	This work	233486
pMTKM15	Level 0 plasmid MoClo with KM promoter GPD1	CamR	This work	233487
pMTKM16	Level 0 plasmid MoClo with KM promoter ALD2	CamR	This work	233488
pMTKM17	Level 0 plasmid MoClo with KM promoter HSP26	CamR	This work	233489
pMTKM18	Level 0 plasmid MoClo with KM promoter HSP60	CamR	This work	233490
pMTKM19	Level 0 plasmid MoClo with KM promoter PIR1	CamR	This work	233491
pMTKM20	Level 0 plasmid MoClo with KM promoter SSA3	CamR	This work	233492
pMTKM21	Level 0 plasmid MoClo with KM promoter SOD1	CamR	This work	233493
pMTKM22	Level 0 plasmid MoClo with KM promoter INU1	CamR	This work	233494
pMTKM23	Level 0 plasmid MoClo with KM promoter ADH1	CamR	This work	233495
pMTKM24	Level 0 plasmid MoClo with KM promoter PGK	CamR	This work	233496
pMTKM25	Level 0 plasmid MoClo with KM promoter TEF3	CamR	This work	233497
pMTKM26	Level 0 plasmid MoClo with KM promoter TDH3	CamR	This work	233498
pMTKM27	Level 0 plasmid MoClo with KM promoter HFF1	CamR	This work	233499
pMTKM28	Level 0 plasmid MoClo with KM promoter NC1	CamR	This work	233500
pMTKM29	Level 0 plasmid MoClo with KM promoter COX20	CamR	This work	233501
pMTKM30	<i>TDH3-tetO2p level 0 MoClo</i>	CamR	This work	233502
pMTKM31	<i>TEF3-tetO2p level 0 MoClo</i>	CamR	This work	233503

pMTKM32	<i>PGK1-tetO2p level 0 MoClo</i>	CamR	This work	233504
pMT029	pan(OPT)ARS level 0 YTK MoClo	CamR	This work	n/a
pAK214	One gRNA (sequence 1) targeting the URA3 locus	NatR, AmpR	This work	n/a
pJD005	Two gRNAs targeting the URA3 locus	NatR, AmpR	This work	n/a
pMT187	pJET plasmid containing the 450 bp 5' and 3' homology arms for URA3 integration flanking a PDC1p-mNeonGreen-PDCt cassette	AmpR	This work	n/a
pMT188	pJET plasmid containing the 900 bp 5' and 3' homology arms for URA3 integration flanking a PDC1p-mNeonGreen-PDCt cassette	Amp	This work	n/a
pUCC001	pCAS-GAPp-[gRNA:BsaI] backbone	HygR, AmpR	170	n/a
pAK001	pUCC001 with one gRNA targeting the DNL4 gene	HygR, AmpR	This work	n/a
pCD4	pJET plasmid containing a tRNAgly splicing unit together with a SpCas9 scaffold	AmpR	143	n/a
pAK219	pMT002 with one gRNA (sequence 2) targeting the URA3 locus	NatR, AmpR	This work	n/a
pAK014	<i>TetR-SV40 (K. lactis opt) level 0 MoClo</i>	CamR	This work	n/a
pMT189	Integration cassette into ARO1, gfp dropout level 2 MoClo	KanR, G418	This work	n/a
pAK20	<i>ConLS-Ssa3p-tetR-SV40(K. lactis opt)-Inu1t-ConR1 level 1 MoClo</i>	AmpR	This work	n/a
pAK21	<i>ConL1-TDH3-tetO2p-mNeonGreen-PDC1t-ConRE level 1 MoClo</i>	AmpR	This work	n/a
pAK22	<i>ConL1-TEF3-tetO2p-mNeonGreen-PDC1t-ConRE level 1 MoClo</i>	AmpR	This work	n/a
pAK23	<i>ConL1-PGK1-tetO2p-mNeonGreen-PDC1t-ConRE level 1 MoClo</i>	AmpR	This work	n/a
pMT193	Two gRNAs targeting the ARO1/SWF intergenic region (I3)	AmpR, NatR	This work	n/a
pKMK.T1	INU1t terminator from <i>K. marxianus</i> level 0 MoClo	CamR	101	125053
pKMK.T4	PDC1t terminator from <i>K. marxianus</i> level 0 MoClo	CamR	101	125056
pKMK.I3R	Right homology arm for SWF1/ARO1, <i>K. marxianus</i> level 0 MoClo	CamR	101	125065
pKMK.I3L	Left homology arm for SWF1/ARO1, <i>K. marxianus</i> level 0 MoClo	CamR	101	125032
pKmk.P2	pPDC1 level 0 Promtoer part YTK Moclo	CamR	101	125035

pKmK.P19	pLac4 level 0 Promoter part YTK Moclo	CamR	101	125052
PADH1-EGFP	K.m. CEN/ARS, PADH1-EGFP-TCYC1Sc, HIS3 marker	AmpR, -HIS3	100	n/a
PINU1-EGFP	K.m. CEN/ARS, PINU1-EGFP-TCYC1Sc, HIS3 marker	AmpR, -HIS3	100	n/a
PRPE1-EGFP	K.m. CEN/ARS, PRPE1-EGFP-TCYC1Sc, HIS3 marker	AmpR, -HIS3	100	n/a
PPIR1-EGFP	K.m. CEN/ARS, PPIR1-EGFP-TCYC1Sc, HIS3 marker	AmpR, -HIS3	100	n/a
PHSP26-EGFP	K.m. CEN/ARS, PHSP26-EGFP-TCYC1Sc, HIS3 marker	AmpR, -HIS3	100	n/a
PSSA3-EGFP	K.m. CEN/ARS, PSSA3-EGFP-TCYC1Sc, HIS3 marker	AmpR, -HIS3	100	n/a
PZWF -EGFP	K.m. CEN/ARS, PZWF-EGFP-TCYC1Sc, HIS3 marker	AmpR, -HIS3	100	n/a
PSCL1-EGFP	K.m. CEN/ARS, PSCL1-EGFP-TCYC1Sc, HIS3 marker	AmpR, -HIS3	100	n/a
PALD2-EGFP	K.m. CEN/ARS, PALD2-EGFP-TCYC1Sc, HIS3 marker	AmpR, -HIS3	100	n/a
PPST1-EGFP	K.m. CEN/ARS, PPST1-EGFP-TCYC1Sc, HIS3 marker	AmpR, -HIS3	100	n/a
PGLK1A-EGFP	K.m. CEN/ARS, PGLK1A-EGFP-TCYC1Sc, HIS3 marker	AmpR, -HIS3	100	n/a
PSOD1-EGFP	K.m. CEN/ARS, PSOD1-EGFP-TCYC1Sc, HIS3 marker	AmpR, -HIS3	100	n/a
PGPD1-EGFP	K.m. CEN/ARS, PGPD1-EGFP-TCYC1Sc, HIS3 marker	AmpR, -HIS3	100	n/a
PGLK1B-EGFP	K.m. CEN/ARS, PGLK1B-EGFP-TCYC1Sc, HIS3 marker	AmpR, -HIS3	100	n/a
PHSP60-EGFP	K.m. CEN/ARS, PHSP60-EGFP-TCYC1Sc, HIS3 marker	AmpR, -HIS3	100	n/a
PTDH3-EGFP	K.m. CEN/ARS, PTDH3-EGFP-TCYC1Sc, HIS3 marker	AmpR, -HIS3	100	n/a
PNC1-EGFP	K.m. CEN/ARS, PNC1-EGFP-TCYC1Sc, HIS3 marker	AmpR, -HIS3	100	n/a
PCOX20-EGFP	K.m. CEN/ARS, PCOX20-EGFP-TCYC1Sc, HIS3 marker	AmpR, -HIS3	100	n/a
PPGK-EGFP	K.m. CEN/ARS, PPGK-EGFP-TCYC1Sc, HIS3 marker	AmpR, -HIS3	100	n/a
PHHF1-EGFP	K.m. CEN/ARS, PHHF1-EGFP-TCYC1Sc, HIS3 marker	AmpR, -HIS3	100	n/a
PTEF3-EGFP	K.m. CEN/ARS, PTEF3-EGFP-TCYC1Sc, HIS3 marker	AmpR, -HIS3	100	n/a
pLN012	pUCC001 [gRNA: BsaI GFP dropout], parent plasmid for pLN026 and pLN028	HygR, AmpR	This work	n/a
pLN026	Two gRNAs targeting the KAT1 locus	HygR, AmpR	This work	n/a
pLN028	Two gRNAs targeting the $\alpha 3$ locus	HygR, AmpR	This work	n/a
pYTK001	Level 0 entry plasmid MoClo	CamR	128	65108
pYTK002	ConLS level 0 MoClo	CamR	128	65109
pYTK003	ConR1 level 0 MoClo	CamR	128	65110
pYTK067	ConR1 level 0 MoClo	CamR	128	65174
pYTK072	ConRE level 0 MoClo	CamR	128	65179

pYTK095	AmpR-ColE1 gfp-dropout plasmid level 1 MoClo	AmpR	128	65202
pYTK084	KanR-ColE1 as a Type 8 part YTK MoClo	KanR	128	65191
pYTK073	ConRE' as a Type 5 part YTK MoClo	CamR	128	65180
pYTK077	KanamycinR as a Type 6 part YTK MoClo	CamR, G418	128	65184
pYTK047	GFP dropout as a Type 234r part YTK MoClo	CamR	128	65154

3. Chapter 3: Tri-functional CRISPR system to enhance fumaric acid tolerance

Adapted from: Thornbury, M., Knoops A., Abushahin, R., Omran, R.P., Kumar, L., M., Whiteway, M., and Martin, V.J.J. (2025). Tri-Functional CRISPR Screen Reveals Overexpression of QDR2 and QDR3 Transporters Increase Fumaric Acid Production in *Kluyveromyces marxianus*. [Manuscript submitted for publication].

3.1. Abstract

Organic acids such as fumaric acid are widely used in the food and beverage industry as acidulants and preservatives, while also serving as versatile precursors for industrially relevant compounds. Fumaric acid is still predominantly produced through petroleum-derived processes. To enhance production efficiency and diversify supply, we are engineering *Kluyveromyces marxianus* as a biosynthetic platform from renewable feedstocks. In previous work, we have established *K. marxianus* Y-1190 as a host for lactose valorization based on its high growth rate on lactose and its tolerance for acid conditions. Here, we establish a trifunctional genome-wide library for *K. marxianus* using CRISPR activation, interference, and deletion to allow identification of gene expression perturbations that enhance tolerance to fumaric acid. We determined that deletion of *ATP7*, encoding a subunit of the mitochondrial F₁F₀ ATP synthase, and overexpression of *QDR2* and *QDR3*, two previously uncharacterized members of the 12-spanner H⁺ antiporter (DHA1) family in *K. marxianus*, can enhance fumaric acid tolerance. We also found that integrated overexpression of both *QDR2* and *QDR3* in a $\Delta FUM1$ background strain improved titers of fumaric acid production from 0.26 g L⁻¹ to 2.16 g L⁻¹. Together, these results highlight roles for membrane transport and mitochondrial function in enabling fumaric acid tolerance and production in *K. marxianus*.

3.2. Introduction

Organic acids have a long history of production through fermentation. Citric acid is primarily produced by the filamentous fungus *Aspergillus niger*, while lactic acid is typically produced by *Lactobacillus* species such as *Lb. amylovorus*, the fungus *Rhizopus oryzae*, and recombinant yeast strains^{1,3,4}. However, many other organic acids are not currently manufactured by fermentation at large scale due to high costs and/or low yields. Malic acid, fumaric acid, and adipic acid are examples of building blocks for plasticizers and polymers such as nylon 66^{12,15,171,172}, and research is ongoing to develop lower-cost fermentation methods for these acids. These approaches include using alternative feedstocks^{173–175} and non-conventional production organisms with inherent qualities that result in higher yields^{48,176–178}.

K. marxianus is a Crabtree-negative member of the *Saccharomycetaceae* family with many characteristics suited to biomanufacturing. It can metabolize diverse carbon sources⁶⁴, grows rapidly⁶¹, and is both heat- and acid-tolerant^{62,63}. Recent advances in genetic engineering methods for *K. marxianus* such as the development of synthetic biology toolkits^{101,108,179} and CRISPR genome editing^{72,106,115,180,181} have enhanced interest in using this host in biomanufacturing^{58,110,116}. The majority of this bioprocess development is done under glucose-fed growth conditions, with only a few studies leveraging *K. marxianus*'s substrate-utilizing diversity by using inulin or sucrose as a carbon source^{121,161}.

Lactose, another potential carbon source, is a prominent by-product of the dairy processing industry, and considerable research has focused on valorizing lactose-rich whey and milk permeate^{119,182}. Within the *K. marxianus* species, dairy strains utilize lactose more efficiently than non-dairy strains due to 13 amino acid variations in the lactose transporter Lac12p⁶⁵. We have previously characterized *K. marxianus* Y-1190 and found the variations in Lac12p associated with

lactose utilization, identifying it as a promising candidate for lactose valorization¹⁷⁹. This study investigates using lactose from dairy permeate as a feedstock for fumaric acid production by *K. marxianus*.

Fumaric acid is an attractive target for biomanufacturing due to its wide range of industrial applications and its potential for sustainable production. It is widely used in the food and beverage industry, both as an acidulant and in the production of unsaturated polyester resins¹⁸³. Fumaric acid is currently derived from petroleum-based processes therefore, bio-production offers a renewable alternative that reduces reliance on fossil resources¹³. Developing efficient microbial platforms for fumaric acid production will enable cost-effective and scalable bioprocesses, further driving its adoption as a sustainable chemical.

Previous fungal and yeast production of fumaric acid was achieved in *S. cerevisiae*, *R. oryzae*, *Scheffersomyces stipitis*, and *C. glabrata*^{30,44,184,185}. *R. oryzae* is the most prolific producer reaching a final titer of up to 103 g L⁻¹ in 20 L fermentation tanks¹⁸⁶. As alternatives to *R. oryzae*, recombinant strains were developed by expression of the reductive TCA genes of *R. oryzae* in *C. glabrata* and *S. cerevisiae* to achieve titers of 21.6 g L⁻¹ and 33.13 g L⁻¹, respectively^{30,44}. *Escherichia coli* studies have focused on intensive rewiring of central metabolism to achieve titers of 42.5 g L⁻¹³⁵. While these efforts have demonstrated the potential for high-titer production, challenges related to acid tolerance and utilization of substrates other than glucose highlight the need for alternative hosts such as *K. marxianus*.

In the present study, we deleted the fumarate hydratase *FUM1* gene in *K. marxianus* to increase fumaric acid titers. While *FUM1* deletion promoted fumaric acid accumulation, it also impaired acid tolerance. To address this trade-off, we used CRISPR-AID selection to uncover genetic perturbations that could restore growth under acid stress. This study adapts CRISPR-AID

for use in *K. marxianus* and identifies candidate genes that support both acid tolerance and fumaric acid production. The findings suggest that the proteins Qdr2p and Qdr3p likely function as fumaric acid exporters, mitigating intracellular stress and increasing fumaric acid production. This work provides valuable insights into engineering *K. marxianus* for organic acid production and industrial biotechnology applications.

3.3. Materials and methods

3.3.1. Plasmids, strains, and growth media

E. coli strain DH5 α , used to maintain and amplify plasmids, was cultured at 37 °C in Luria broth containing the appropriate antibiotics. *K. marxianus* strain Y-1190 and Y-1190 Δ FUM1 were used as the host for all genome engineering in this work. Unless otherwise stated, *K. marxianus* was cultivated and maintained in YPD (10 g L⁻¹ Bacto Yeast Extract, 20 g L⁻¹ Bacto peptone, 20 g L⁻¹ glucose). Plasmids, oligonucleotides, and engineered strains used in this study can be found in Table S3.1, Table S3.2, and Table S3.3, respectively.

3.3.2. K. marxianus transformations and CRISPR-Cas9 Editing

All sequences of *K. marxianus* Y-1190 were obtained from genome assembly ASM4656278v1. Unless otherwise stated, *K. marxianus* genetic modifications were facilitated using double-sgRNA CRISPR-Cas9-mediated deletions and integrations¹⁷⁹. The SpCas9 and sgRNAs were delivered by pMTKM01 (Addgene plasmid #233473) and edited strains were selected on YPD agar containing 100 μ g mL⁻¹ nourseothricin.

Transformations of *K. marxianus* were performed according to a previously published protocol with the following modifications⁹⁴. In brief, 1 mL of *K. marxianus* cells in exponential phase were centrifuged at 1860 g for 3 min. After washing once with 1 mL of sterile H₂O, the cells

were suspended in 10 μL of 10 mg mL^{-1} salmon sperm, and 500 ng of sgRNA plasmid was added, along with 1 μg of linearized donor DNA for genome integration. To the cells and DNA, 400 μl of transformation buffer (40 % polyethylene glycol 3350, 0.1 M lithium acetate, 10 mM Tris-HCl (pH 7.5), 1 mM EDTA, and 10 mM DTT) was added to a 1.5 mL microfuge tube and the solution was mixed by agitation. The transformation mix was then incubated at room temperature for 15 min followed by heat shock at 47 °C for 15 min. The cells were then centrifuged at 1860 g for 3 min and suspended in 400 μl of YPD to recover overnight at 30 °C without shaking. The next day the cells were plated on YPD agar with appropriate selection. If necessary, an outgrowth step was done where 50 μl of recovered transformation was grown in 500 μl of YPD with selection for 2-3 days. This outgrowth was diluted 1:1000-10000 before plating on YPD agar with selection.

3.3.3. Establishing CRISPR-AID in *K. marxianus*

CRISPR-AID (Activation, Interference, Deletion) is a tri-functional platform that enables simultaneous gene activation, repression, and deletion, providing a powerful tool for functional genomics and metabolic engineering, first built in *S. cerevisiae* to enhance furfural tolerance¹⁸⁷. To adapt and test the efficiency of CRISPR-AID in *K. marxianus*, the mNeonGreen (mNG) CDS was inserted at the *URA3* locus using CRISPR-Cas9 allowing *URA3* auxotrophy for sgRNA plasmid selections while providing a measurable phenotype. This strain (MTK019) also expressed the three CRISPR-AID Cas enzymes¹⁸⁸; however, we replaced the promoters driving Cas expression with *K. marxianus* promoters, and replaced the terminators with either *K. marxianus* or *S. cerevisiae* terminators known to be active in *K. marxianus*. The resulting plasmid (pMT195) was integrated into the I3 locus between *KmSWF1* and *KmARO1*¹⁰¹ using CRISPR-Cas9. Guides for activation, interference and deletion of the mNG cassette were designed on CCTop¹⁴⁰ and expressed from plasmids pMT149-pMT156. These plasmids were transformed into MTK019 and

three clones per sgRNA were picked into SC-URA media containing G418. Cultures were grown overnight and then back diluted to OD₆₀₀ of 0.05 and grown for an additional 24 h. The cultures were diluted 1:10 and fluorescence was measured at excitation 510 nm and emission 545 nm in a plate reader (Infinite200, Tecan).

3.3.4. *Genome-wide CRISPR-AID sgRNA library design and construction*

3.3.4.1. *Building the CRISPR-AID strain*

To integrate the CRISPR-AID cassette into *K. marxianus* Y-1190 Δ FUM1, we used a similar method as described above, however, we replaced the constitutive promoters with tetracycline inducible promoters resulting in plasmid pAK104. Other than the *Cas* coding sequences, all parts were sourced from the following tool kits: Lee *et al.*, 2015, Rajkumar *et al.*, 2019, and Thornbury *et al.*, 2025. The three *Cas* sequences and their scaffolds were amplified from CRISPR-AID plasmids from Lian *et al.* 2019 (Addgene¹⁸⁸ #136989, #136992-136994). A total of 5 ug of pAK104 was digested with NotI-HF (New England Biolabs) and the linear plasmid used as a donor. We used CCTop to generate sgRNAs targeting the I3 site between *SWF1* and *ARO1*, and selected sgRNAs displaying a CRISPRater score of 0.60 or higher¹⁴¹. We used Golden-Gate cloning as previously described to introduce two sgRNAs into the pMTKM01, a SpCas9 vector^{137,143,179}.

3.3.4.2. *CRISPR-AID sgRNA library design*

Genome-wide library sgRNA sequences were designed using CHOPCHOP via the command line^{189–191}. Promoter sequences were designated as 400 bp upstream of each transcriptional start site and were used to design activation and interference sgRNAs, while coding sequences were used for the deletion sgRNAs. *K. marxianus* preferentially uses non-homologous

end joining (NHEJ) for DNA repair¹⁰⁴, and therefore we did not design donor DNA for deletions. As proof of concept, we demonstrated that a single sgRNA was sufficient to disrupt the mNG reporter signal (Fig. 3.1 A), validating our approach.

Because CHOPCHOP generates all possible sgRNA sequences for a given DNA target sequence, we adapted a previously published ranking criterion to select sgRNAs likely to be the most efficient¹⁸⁸ (Table S3.4). We selected the top six sgRNAs for activation and interference, and the top four for deletion. In rare cases, no PAM sequence was available within the selected target region, or no unique sgRNAs could be designed resulting in less than four, or six, sgRNAs being ordered for that gene. Summary coverage of our library can be found in Supplementary Table 3.1. The final three libraries consisted of 26185 activation, 26988 interference and 18763 deletion sgRNAs; all sequences can be found in Table S3.5. When combined, the library contains a total of 71,936 guides covering 4,740 of the 4,853 annotated genes. The remaining 113 genes correspond to transposable elements with significant sequence overlap, which prevented unique guide design. While we aimed to include six guides per gene for activation and interference, limitations such as a lack of unique PAM sites or overlapping promoter regions prevented full coverage: For activation and interference, 83 % and 86 % of genes, respectively, had full six guide coverage, and 98.5 % and 99.6 % of genes were covered by at least one guide (Table 3.1). For deletion, 98 % of genes had complete four-guide coverage, and 99.5 % were targeted by at least one guide. Adapters were added to both ends to facilitate PCR amplification that corresponded to the flanking tRNA_{glycine} sequence within the plasmid, allowing all three sgRNA types to be amplified with one set of primers. All sgRNA libraries were synthesized by Twist Bioscience.

Table 3.1. Summary Coverage of the *K. marxianus* CRISPR-AID Library

	Number of Genes Targeted						Total Genes	Total sgRNAs
# of sgRNAs	1	2	3	4	5	6		
#of genes in ActLib	97	118	132	165	218	3951	4740	26185
#of genes in IntLib	29	74	105	167	236	4108	4740	26988
#of genes in DelLib	14	25	21	4659	0	0	4740	18763

3.3.4.3. *SgRNA library construction*

The pooled oligos from each of the three libraries were suspended in TE buffer to a final concentration of 10 ng μL^{-1} . An initial 50 μL Phusion high fidelity short PCR amplification of 1 ng of each library was done using blunt primers (MT736 & MT737). Because of the different sizes of oligos to be amplified, different amounts of amplification cycles were done. Activation sgRNAs for LbCas12a have shorter scaffolds and were amplified for 9 cycles, whereas SaCas9 and SpCas9 scaffolds are similar in length, as such, deletion and interference sgRNAs were amplified for 12 cycles. PCR amplicons were purified by kit with an isopropanol addition for small fragments (GeneJet, Thermofisher). The purified amplicons were used as template for a second PCR to add Golden Gate adapters for downstream cloning. Twelve 50 μL Phusion high fidelity polymerase PCR reactions were done per library with 10 ng of template and with PCR cycles kept to 9-12 as before. The 12 reactions were pooled, and gel purified (GeneJet, Thermofisher) using a 0.8 % agarose gel.

The purified sgRNA libraries were cloned into a GFP-dropout destination vector containing a *SNR52* promoter and a *SUP4* terminator (pMT351) using the Golden Gate cloning method¹³⁷. Briefly, 20 fmol of pMT351 and 40 fmol of each library were added to PCR tube with 10 U BsaI-HF, 200 U T4 ligase, 1.5 µl T4 ligase buffer, 1.5 µl BSA/PEG mix (1 mg mL⁻¹ BSA, 10 % PEG 3350) and ddH₂O added to a final volume of 15 µl. The libraries were ligated using a thermocycler and 45 cycles of 5 min 37 °C digest and 16 °C ligate with a final 1-h 37 °C digestion step. The Golden Gate ligations were purified by isopropanol DNA purification¹⁹² and 200 ng of each library was transformed into 40 µl of *E.coli* MegaX DH10B T1^R Electrocomp Cells (Thermofisher). Immediately after electroporation, 1 mL of pre-warmed SOC media was added before transferring the cells to a 15 mL conical tube containing 3 mL of pre-warmed SOC. The transformed library was recovered at 37 °C shaking for 1 h before plating on two 24.5 cm x 24.5 cm bioassay plates of LB-Carbenicillin. An aliquot of 50 µl was retained for serial dilutions to count colonies. We achieved over 10-fold transformation coverage for each library (Table S3.6). For each library, 25 mL of cold LB was used to scrape each bioassay plate resulting in ~45 mL of transformed *E. coli* culture per library, which was divided between two maxi prep columns (GeneJet, Thermofisher) for DNA extraction.

Construction of our CRISPR-AID sgRNA libraries in *K. marxianus* was done as previously described in Lian *et al.*, (2019) for *S. cerevisiae* with the following modifications. For each library, five transformations with 10 µg of DNA were done according to the protocol defined above. All five transformations were combined in 4 mL of YPD and recovered at 30 °C overnight. The next morning, the recovered cultures were topped up to 10 mL in YPD and 5 mL each were plated on SC-URA bioassay plates. Fifty µl were set aside to dilute 1:100 and 1:1000 for colony counting. After determining colony count for each library, the libraries were scraped into cold SC-URA and

pooled to ensure equal representation. Individual and pooled libraries were adjusted to 20 % total glycerol and stored at -80 °C.

3.3.4.4. *CRISPR-AID selection in fumaric acid*

A fumaric acid growth inhibition pilot experiment was done that identified 25 mM of fumaric acid as our starting concentration to select CRISPR-AID library clones (Fig. S3.2). For identifying clones with increased tolerance to fumaric acid, the pooled library was inoculated at an OD₆₀₀ of 0.1 in 50 mL of SC-URA with anhydrous tetracycline (aTC) at 500 ng µl⁻¹ with either 0 mM or 25 mM fumaric acid and grown in a shaking incubator at 30 °C and 220 rpm. Cultures were sampled when growth in the 25 mM condition had exceeded that of the control (EV, plasmid with no sgRNA). To increase the screening stringency, cultures selected on 25 mM fumaric acid were back-diluted to an OD₆₀₀ of 0.1 into the same medium supplemented with 30 mM fumaric acid, grown for an additional 24 hr and harvested. This back dilution step was repeated once more with a medium supplemented with 42 mM fumaric acid and a 30-hr growth period. The control cultures did not grow in 30 mM and 42 mM fumaric acid conditions.

3.3.5. *CRISPR-AID sgRNA enrichment analysis*

Cells from each library were harvested at an OD₆₀₀ unit of 0.5, corresponding to ~3.0x10⁷ cells, by centrifugation (1860 g for 3 min) and sgRNA plasmids were extracted from the cells using a Zymoprep yeast miniprep kit (Zymo Research). sgRNAs were amplified using Platinum SuperFi DNA polymerase (Invitrogen, 12351010) and primers containing heterogeneity spacers for 20 cycles from ~8 ng of plasmid template (Table S3.2). Following amplification, two 50 µl, 20 cycles, PCR reactions were performed per sample using Phusion high fidelity (New England Biolabs) and primers with Nextera index adapter (Illumina) sequences. The PCR products were purified (GeneJet, Thermofisher) and 130 ng of each sample were pooled. The pooled amplicons

were sequenced on an Illumina NextSeq platform using 300 paired-end reads (Génome Québec). Initial quality control was done using FastQC v0.11.5 and MultiQC v1.12^{193,194}. The paired reads were merged using PEAR v0.9.6 with a minimum read length (-t) of 20-bp and a quality score threshold (-q) of 25¹⁹⁵. We were able to merge 99.5 % of reads, for an average of 1.1×10^6 merged reads per condition (Table S3.7). We were able to maintain a minimum of 10-fold theoretical coverage for each condition, except for replicate B for the 42 mM fumaric acid condition. This replicate was included in downstream analysis despite lower coverage. These merged reads were binned into activation, interference, or deletion sgRNAs using the scaffold sequences and Cutadapt v5.0¹⁹⁶. Binned sgRNA sequence counts were analyzed using MAGeCK v0.5.8¹⁹⁷. sgRNA enrichment was measured over the 0 mM fumaric acid control condition. Targets with the highest enriched sgRNAs that had significantly low p-values and false discovery rates were selected for further testing.

3.3.6. Validation of CRISPR-AID hits

The top 10 over-represented sgRNAs from the library screen were all activation sgRNAs. To validate them as true hits, we overexpressed in *K. marxianus* Y-1190 Δ *FUM1* the target gene that corresponded to the enriched sgRNA and measured strain fitness in a medium supplemented with fumaric acid. The genes were amplified from *K. marxianus* Y-1190 genomic DNA and cloned into a vector with the strong promoter *KmNC1p* (Addgene plasmid #233500). As *K. marxianus* Y-1190 is a heterologous diploid, the PCR reaction had a chance of amplifying both the primary and the haplotig allele. In these cases, we included both in the validation screen.

3.3.7. Growth rate analysis

Three independent colonies of each transformed strain were inoculated into 96-well deep plate containing 700 μ l of Synthetic Complete Lactose (SC-Lactose; 6.7 g L⁻¹ yeast nitrogen base

with ammonium sulfate, 20 g L⁻¹ lactose, 1.92 g L⁻¹ yeast synthetic drop-out medium supplement without uracil, 0.08 g L⁻¹ uracil) media supplemented with 200 µg mL⁻¹ G418. To test strain fitness, cell growth was measured from overnight cultures back diluted to an OD₆₀₀ of 0.1 in the same medium but with concentrations of fumaric acid varying from 0 mM to the solubility limit of 42 mM. Cultures were incubated at 30 °C in clear, flat-bottomed 96-well microtiter plates and wrapped with Parafilm to minimize evaporation. Absorbance readings were taken at 595 nm with a Sunrise absorbance microplate reader (Tecan) every 20 min over the course of up to 3 days. Growth rates ($\mu_{\max}$) were calculated in R using the “growthrates” package¹³⁰.

3.3.8. Metabolite analysis

To assess the effect of increased strain fitness on fumaric acid production we measured substrate uptake and metabolites produced by our strains. Shake-flask cultures for fumaric acid production quantification were grown in SC-Lactose or 0.5x supplemented dairy permeate¹⁷⁹. Following overnight growth in SC-Lactose, cultures were diluted to an OD₆₀₀ of 0.2 in 20 mL SC-Lactose in a 125 mL baffled Erlenmeyer flask. Cultures were incubated at 30 °C, and shaking at 200 rpm for 72 h. At 24 h intervals, 250 µl of the culture supernatant was sampled and frozen at -20 °C. For HPLC analysis, supernatant samples were diluted 1:10 in distilled H₂O. Metabolites were measured using an Agilent 1290 Infinity II HPLC equipped with an Aminex HPX-87H column (Bio-Rad). Metabolites were resolved isocratically in 10 mM sulfuric acid at 65 °C at a flow rate of 0.6 ml min⁻¹. A UV-visible detector was used for the quantification of fumaric acid (260 nm), and a refractive index detector was used to quantify lactose. All measurements were compared to a linear standard curve and done in duplicate or triplicate.

3.3.9. *Fed-batch cultivation*

Controlled fed-batch fermentations were performed in 3 L BioBundle bioreactors (Applikon) at 30 °C. Batch fermentations were initially carried out without pH maintenance for the first 24 h. Thereafter, the pH was maintained at 4.0 by automated addition of 4 N NaOH throughout both batch and fed batch fermentations. Dissolved oxygen was maintained at 20 % of air saturation by automatically adjusting the agitation speed (400-600 rpm) and aeration rate (500-600 mL min⁻¹). Off-gas O₂ and CO₂ concentrations were monitored using a Tandem Multiplex Gas Analyzer (Magellan BioTech). Bioreactor inoculums were prepared in two 125 mL Erlenmeyer flasks containing 25 mL batch medium of 0.5x supplemented permeate, grown for 48 h at 30 °C and shaking at 200 rpm. Cultures were used to inoculate at OD₆₀₀ = ~0.1 in 1 L of batch medium (50 g Lactose, 8.361 g urea, 2.065 g KH₂PO₄, 1.15 g MgSO₄·7H₂O, 1 g NiSO₄·6H₂O, 5 mL vitamin stock, and 5 mL trace element solution per liter). The cultures were grown in batch mode until lactose depletion (49 h), followed by fed-batch operation initiated using a four-fold ultrafiltration-concentrated permeate (Agropur Cooperative) containing 200 g L⁻¹ lactose with 33.44 g urea, 8.26 g KH₂PO₄, 4.6 g MgSO₄·7H₂O, 4 g NiSO₄·6H₂O, 20 mL vitamin stock, and 20 mL trace element solution per liter. Samples were collected every 24 h for OD₆₀₀ and HPLC analysis for a total of 7 days. Biomass concentration (g L⁻¹) was calculated using a gravimetrically determined conversion factor of 0.22 g L⁻¹ per OD₆₀₀ unit. Process data were recorded and analyzed using BioXpert V2 (Applikon). Fermentation samples were diluted 1:5 to 1:20 with deionized water and analyzed by HPLC as described above. Fermentation profiles of dissolved oxygen, feed pump, and pH can be found in Figure S3.5.

3.3.10. Protein alignments

Alignment of proteins Qdr2 and Qdr3 from *K. marxianus* Y-1190 were done using Clustal Omega¹⁹⁸ and visualized using Jalview¹⁹⁹. Summary alignments were done using NCBI protein blast and transmembrane domains were predicted using DeepTMHMM²⁰⁰.

3.4. Results

3.4.1. Creation of a fumaric acid producing *K. marxianus*

To develop a fumaric acid production strain, we disrupted the TCA cycle by deleting *FUM1*, which encodes fumarate hydratase. This strategy was shown to enhance fumaric acid titers in *S. cerevisiae* and *E. coli*^{27,36}. In *K. marxianus*, *FUM1* deletion similarly increased fumaric titers from 0 to 0.22 g L⁻¹ but also compromised the strain's growth in fumaric acid (Fig. 3.1). The fumaric acid concentration-dependent growth defect observed in the *FUM1* deletion strain suggests an acid stress phenotype (Fig. 3.1 B). To restore acid tolerance, we designed and implemented a CRISPR-AID system to identify gene expression perturbations that enhance growth under fumaric acid stress.

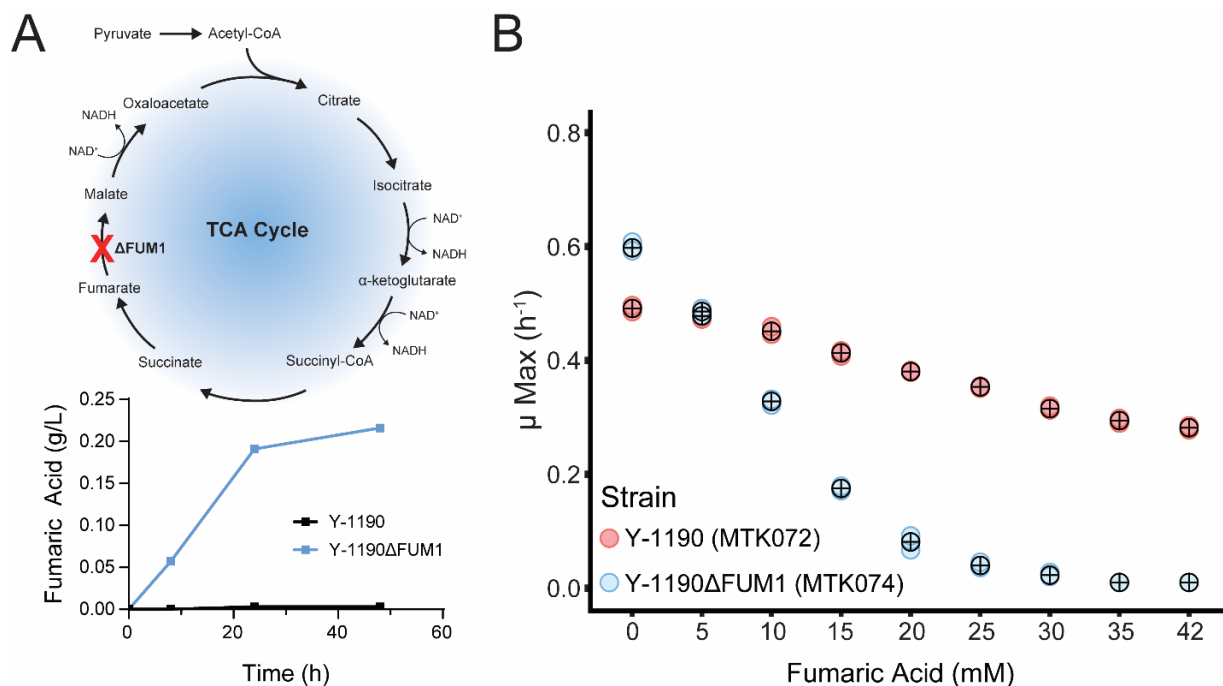


Figure 3.1. Effect of *FUM1* deletion on fumaric acid production and tolerance. **A.** Schematic of the TCA cycle and the *FUM1* deletion that enhanced fumaric acid production in *K. marxianus*. **B.** Growth rate of *K. marxianus* CRISPR-AID strain with (MTK074) and without (MTK072) *FUM1* deletion. Black circles are the mean of three biological replicates.

3.4.2. Validation of CRISPR-AID in *K. marxianus*

To assess CRISPR-based gene activation, interference, and deletion in *K. marxianus*, we measured GFP fluorescence of *K. marxianus* expressing the three CRISPR-AID Cas proteins, mNG and sgRNAs targeting mNG and its promoter region (MTK019). Two of the three sgRNAs (A1 and A3) designed for activation exhibited 4.6- and 2.6-fold higher fluorescence compared to the wild-type and control strains, indicating successful transcriptional activation (Fig. 3.2 A). All interference sgRNAs (I1, I2 and I3) resulted in a 2.9-to-5.7-fold decrease of fluorescence, and both

deletion sgRNAs (D1 and D2) resulted in complete elimination of fluorescence, demonstrating effective repression of gene expression from interference and deletion, respectively (Fig. 3.2 A).

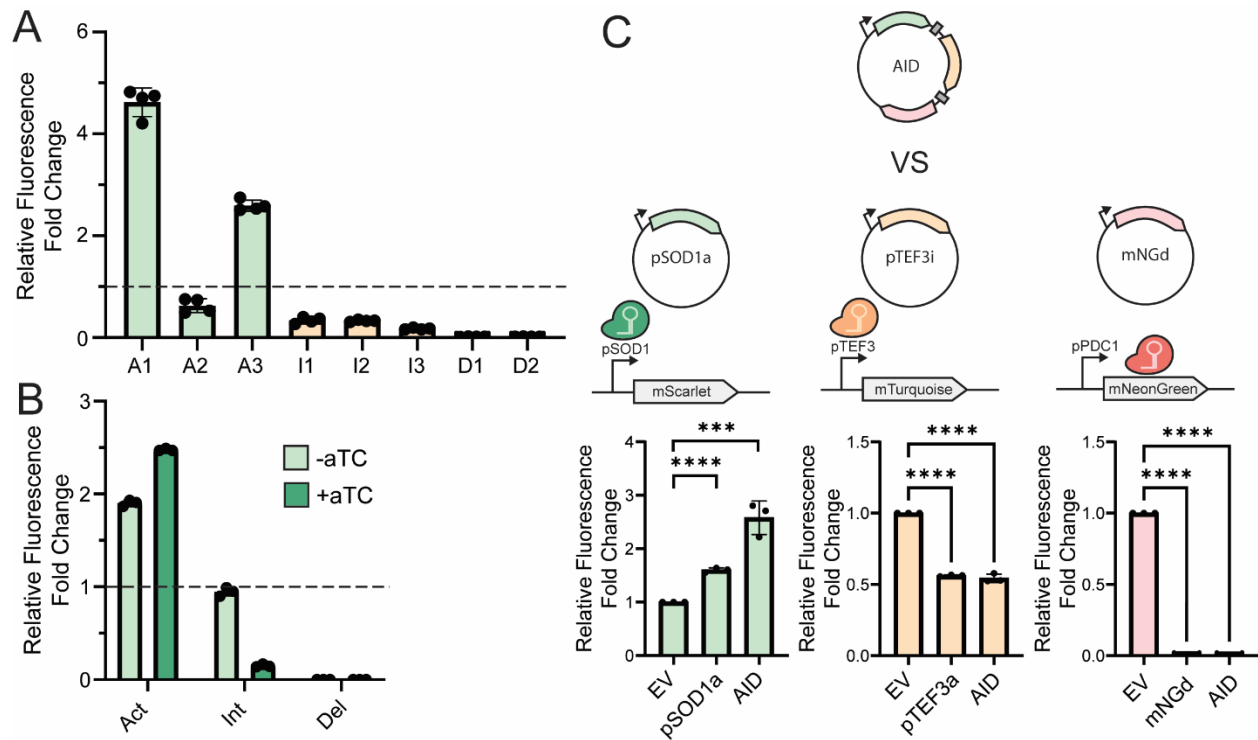


Figure 3.2. CRISPR-AID functionality in *K. marxianus*. **A.** Constitutive 3 Cas system integrated into Y-1190 transformed with sgRNA plasmids for each perturbation to mNG. The X axis denotes the guides tested, three sgRNA for activation (A), three sgRNA for interference (I) and two sgRNA for deletion (D) **B.** CRISPR-AID system with inducible promoters integrated into Y-1190 and transformed with sgRNAs for mNG with or without aTC inducer **C.** Single and combinatorial CRISPR-AID sgRNAs for activation of mScarlet, interference of mTurquoise, and deletion of mNG. Each black dot represents a biological replicate and error bars represent standard deviation.

High plasmid transformation efficiency is critical for constructing a genome-wide CRISPR library in yeast. We observed low transformation efficiency in MTK019, likely due to the constitutive expression of the three Cas proteins causing a burden on the cells (Fig. S3.1 A). To address this, we used a aTC inducible version of the three Cas genes (MTK072)¹⁷⁹. To ensure that the CRISPR-AID system maintained its function, we transformed the modified strain with the previously validated A1, I1, and D1 sgRNA plasmids and measured GFP fluorescence. We

observed an increase in fluorescence with the activation sgRNA, and complete repression of fluorescence with the deletion sgRNA irrespective of addition of the inducer suggesting some leakiness in the inducible promoters (Fig. 3.2 B). For the interference sgRNA, we only observed reduced fluorescence in the presence of aTC suggesting tighter regulation of dSpCas9-RD1152. Despite the leaky promoters for dLbCas12a-VP and SaCas9, the transformation efficiency of MTK072 increased by 5.4-fold ensuring adequate coverage of our sgRNA library (Fig. S3.1 A).

After validating the CRISPR-AID system in *K. marxianus*, we explored the potential for combinatorial perturbations by constructing arrayed sgRNA plasmids targeting multiple genes for simultaneous activation, interference, and deletion. While this combinatorial strategy was not applied in this study, our proof-of-concept demonstrates its feasibility and potential as a powerful tool for multiplexed strain engineering in non-conventional yeasts. To test this system, we modified the MTK019 strain to express two additional reporter fluorescent proteins, mTurquoise and mScarlet, generating MTK063. sgRNAs designed to activate mScarlet, interfere with mTurquoise expression, and delete mNG were expressed in MTK063 individually or in combination from a single plasmid and fluorescence of all three proteins was measured (Fig. 3.2 C). Both single-sgRNA and combination-sgRNA expression resulted in the expected genetic perturbation: mScarlet fluorescence increased 1.6-2.6-fold, mTurquoise decreased 1.8-1.9-fold, and mNG fluorescence was eliminated (Fig. 3.2 C). Notably, single-sgRNA activation constructs produced a moderate 1.6-fold increase in fluorescence over the control (EV) but showed enhanced 2.6-fold activation when expressed in the combinatorial plasmid. In contrast, interference and deletion strategies remained consistent between single- and combinatorial-sgRNA plasmids (Fig. 3.2 C). Together, these data confirm the effectiveness of CRISPR-based activation, interference, and deletion strategies in regulating gene expression in *K. marxianus*.

3.4.3. CRISPR-AID library screen

Enrichment of the $\Delta FUMI$ strain for increased tolerance to fumaric acid was initiated in 25 mM, a concentration determined through a pilot experiment, with 0 mM fumaric acid as the control (Fig. S3.2). After 144 hr, growth of the CRISPR-AID library surpassed the growth of the control strain. The culture was iteratively back-diluted into 30 mM and subsequently 42 mM fumaric acid, reaching the aqueous solubility limit of fumaric acid. At each concentration, samples were collected for Illumina sequencing to assess sgRNA enrichment (Fig. 3.3 A).

Across all three fumaric acid conditions, we identified 90, 52, and 38 significantly enriched sgRNAs at 25 mM, 30 mM and 42 mM fumaric acid, respectively, based on a p-value threshold of < 0.05 , an FDR value of < 0.05 , and a log2 fold-change of at least 1.5 (Fig. 3.3 B). In total, the enriched sgRNAs correspond to 86 unique genes across the three conditions (Fig. 3.3 C). The top 10 gene hits for each condition are listed in Table 3.2. A list of all significantly enriched sgRNA can be found in Table S3.8. A clear trend emerged in the distribution of enriched sgRNAs, with activation sgRNAs comprising the vast majority across all fumaric acid concentrations (Fig. 3.3 D). Activation sgRNAs constituted 98.98 %, 99.92 %, and 99.98 % of enriched sgRNAs in the 25 mM, 30 mM, and 42 mM fumaric acid conditions, respectively, compared to 62.3 % in the 0 mM control (Fig. 3.3 D). Among the enriched activation sgRNAs, guides targeting *QDR2* were highly overrepresented (Fig. 3.3 E). Activation sgRNAs for *QDR2* account for 90 %, 95 % and 98 % of enriched sgRNAs in the 25 mM, 30 mM and 42 mM conditions, respectively (Fig. 3 E). Of the six activation sgRNAs designed for *QDR2*, four were significantly upregulated at 25 mM (Fig. 3.3 B, green stars). While most of the highly enriched guides are activation sgRNAs, a small number corresponded to deletions (Table S3.8). Specifically, in the 25 mM condition, 81 of the

significantly enriched sgRNAs were activation sgRNAs, 9 targeted gene deletion, and none were observed for interference.

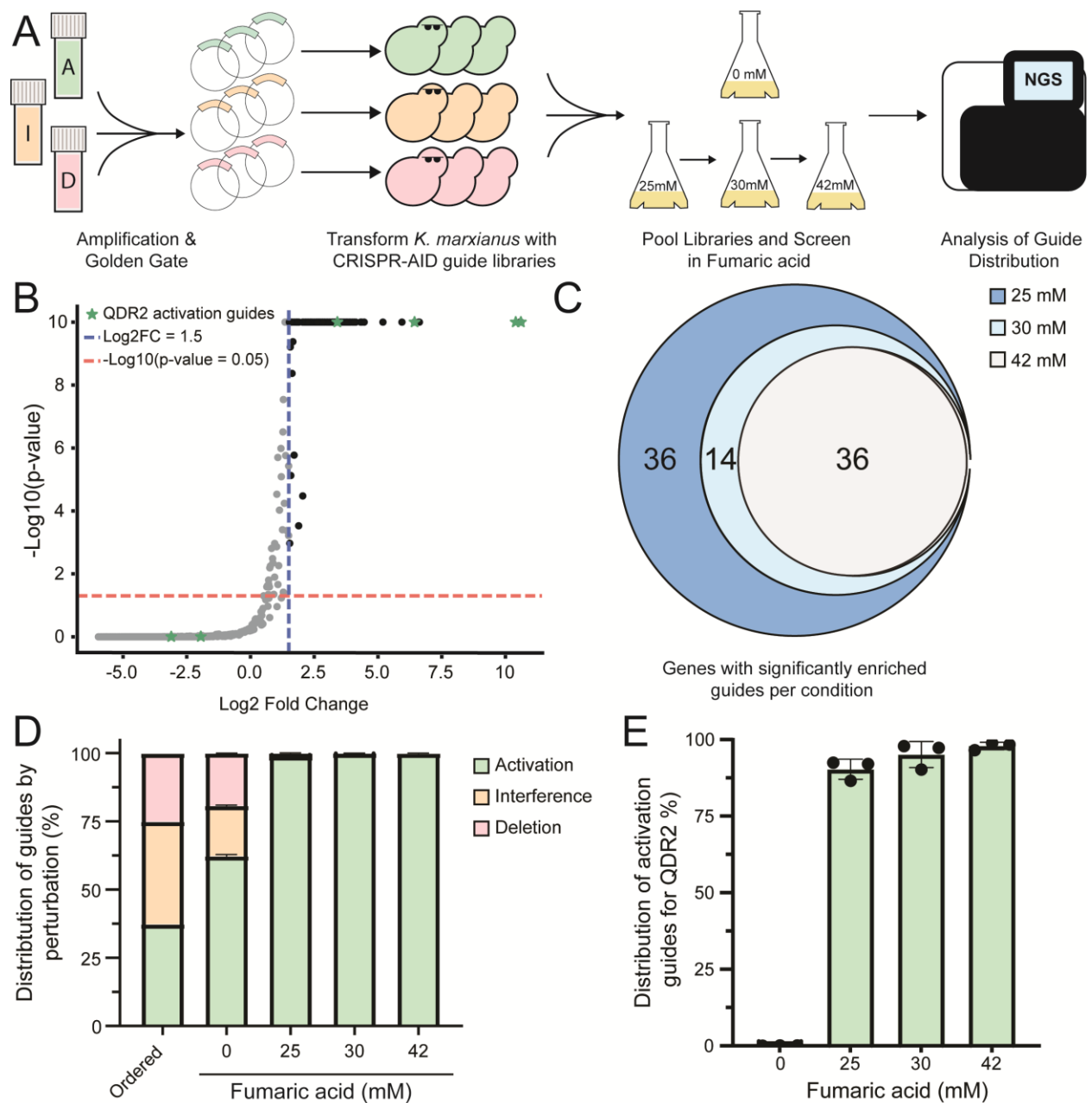


Figure 3.3. Library workflow and summary statistics. **A.** Schematic of library building and screening workflow. **B.** Volcano plots enriched sgRNAs in 25 mM fumaric acid condition. Black dots represent significantly enriched sgRNAs. **C.** Representation of genes with significantly enriched guides. **D.** Distribution of activation, interference, and deletion sgRNAs per condition. **E.** Percent distribution of *QDR2* activation sgRNAs in each fumaric acid condition. All conditions were done in triplicate.

Table 3.2. Significantly upregulated sgRNAs from CRISPR-AID Screen

Rank	Lib25	LFC	Lib30	LFC	Lib42	LFC
1	actQDR2	10.604	actQDR2	10.831	actQDR2	10.972
2	actQDR2	10.564	actQDR2	10.269	actQDR2	9.7333
3	actFAS1	6.6233	actUBP14	6.0777	actUBP14	5.5647
4	actUBP14	6.5932	actFAS1	5.3817	actIMD4	4.9176
5	actQDR2	6.4336	actIMD4	5.2305	actCSE1	4.5154
6	actIMD4	5.9393	actQDR2	4.785	actYPR1	3.7074
7	actYPR1	5.1933	actYPR1	4.0725	actLYS21	3.5609
8	delATP7	4.4575	actCSE1	4.0243	actFAS1	3.5103
9	actACO2a	4.3621	actLYS21	3.8844	actKLMA_20588	3.0591
10	actQDR3	4.3215	actRPN1	3.5368	actQDR2	3.0521

3.4.4. Validation of hits from CRISPR-AID enrichment

To validate the results of the CRISPR-AID library selection, we individually tested the top 10 candidate genes with enriched sgRNAs under fumaric acid stress from each library. This resulted in a list of 12 unique gene perturbations, of which 11 were activation, and one was a

deletion. To mimic activation in a CRISPR-Cas free system, the corresponding target genes were overexpressed from a plasmid using a strong promoter. For the single deletion target *ATP7* was deleted in the Y-1190 $\Delta FUMI$ strain using CRISPR-Cas9 and then cured of the CRISPR-Cas9 plasmid. Growth rate analysis revealed that, while most candidate genes generated no significant improvement in fumaric acid tolerance, overexpression of the DHA1-family multidrug efflux pumps *QDR2* and *QDR3*, as well as deletion of *ATP7*, the stator arm of the F_1F_0 ATPase, exhibited notably higher growth rates at 20 mM fumaric acid compared to their respective control strains (Fig. 3.4 AB). At 20 mM and 42 mM fumaric acid, both *QDR2* and *QDR3* overexpression resulted in increased growth over EV, however, with an overall lower growth rate in 42 mM. *QDR2* overexpression caused an 8.7-fold increase in the growth rate at 42 mM fumaric acid compared to EV, reaching 0.09 h⁻¹. A similar pattern was observed with *QDR3*, which had a growth rate of 0.14 h⁻¹ compared to 0.01 h⁻¹ for the EV (Fig. 3.4 A). Due to the greater fumaric acid tolerance conferred by *QDR2* and *QDR3* overexpression compared to *ATP7* deletion, we focused our efforts on these transporters. We tested a *QDR2* and *QDR3* overexpression strain to determine if co-expression of both QDR genes could promote growth closer to the solubility limit of fumaric acid. We integrated *QDR2* and *QDR3* at known safe harbour sites¹⁰¹ and tested the growth of the strains in SC-Lactose supplemented with fumaric acid. The highest growth rates were measured in the *QDR2* integrated strain and in the *QDR2* integrated strain with one copy of *QDR3* (Fig. 3.4 C). In both cases where *QDR3* was integrated into both alleles, maximum specific growth rate decreased in the absence of fumaric acid, but growth rate was still significantly higher than the $\Delta FUMI$ mutant in the presence of 42 mM fumaric acid (Fig. 3.4 C). To further explore the role of these two genes in fumaric acid tolerance, growth rates were measured from $\Delta QDR2$ and $\Delta QDR3$ mutants of Y-1190 $\Delta FUMI$ (Fig. 3.4 D). Deletion of *QDR2*, whether alone or in combination with *QDR3*, significantly decreased

the growth rate in 15 mM fumaric acid from 0.07 h^{-1} to 0.02 h^{-1} (Fig. 3.4 D), while deletion of *QDR3* alone did not significantly change the growth rate under these conditions (Fig. 3.4 D).

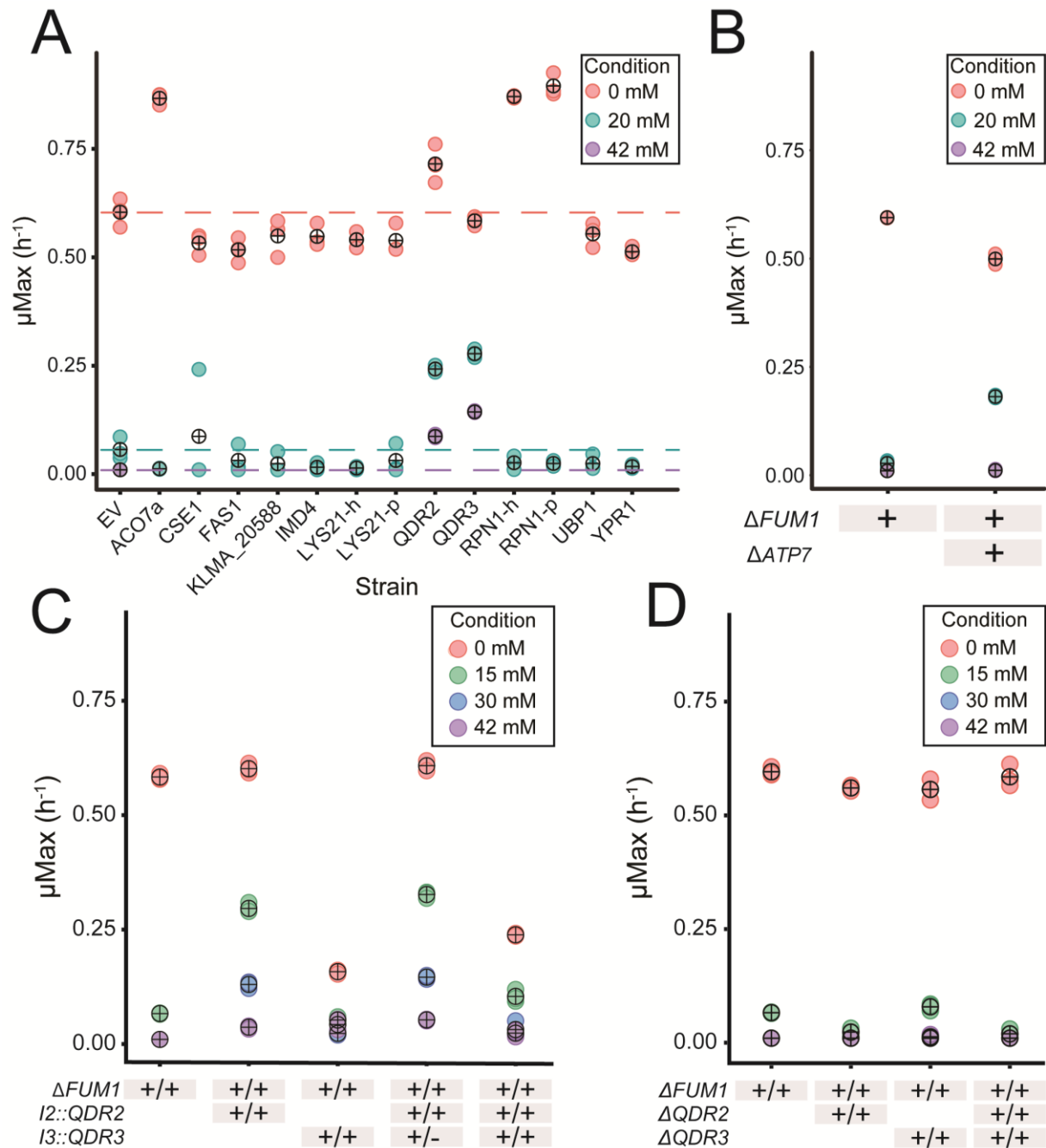


Figure 3.4. Validation of CRISPR-AID library hits in Y-1190 $\Delta FUM1$ strain. Growth rates of the top hits from the CRISPR-AID screen in fumaric acid for **A.** activation and **B.** deletion. **C.** Growth rates of the integrated and **D.** deletion strains for *QDR2* and *QDR3*. Each coloured circle represents a data point, the black circles with crosses represent the mean of those data points.

3.4.5. Fumaric acid production

One of the main goals of this work was to create a fumaric acid production strain that is resistant to high concentrations of fumaric acid. Due to their role as multidrug efflux pumps and H⁺ antiporters in *S. cerevisiae*²⁰¹, we hypothesized that in *K. marxianus* *QDR2* and *QDR3* function as organic acid efflux pumps and that their overexpression would lead to higher fumaric acid production. To test this, we grew our overexpression strains in both SC-Lactose and 0.5x supplemented permeate, a lactose-rich feedstock, and measured metabolite profiles. Our results demonstrated that overexpression of *QDR2* in strain Y-1190 Δ *FUM1* significantly increased fumaric acid production from 0.05 ± 0.00 g L⁻¹ OD₆₀₀⁻¹ to 0.32 ± 0.01 g L⁻¹ OD₆₀₀⁻¹, while *QDR3* overexpression resulted in slightly lower specific titers, reaching 0.13 ± 0.02 g L⁻¹ OD₆₀₀⁻¹ when grown in SC-Lactose media (Fig. 3.5 A). When grown in 0.5x permeate, we also saw significant increases in fumaric acid titers from 0.05 ± 0.01 g L⁻¹ OD₆₀₀⁻¹ to 0.26 ± 0.02 g L⁻¹ OD₆₀₀⁻¹ and 0.21 ± 0.00 g L⁻¹ OD₆₀₀⁻¹ for *QDR2* and *QDR3*, respectively (Fig. 3.5 A).

To confirm these findings in a plasmid-free system, we repeated the fumaric acid production experiment in our *QDR2* and *QDR3* integrated strains. The genomic integration of *QDR2* and *QDR3* both under the strong promoter pNC1 resulted in titers of 0.19 ± 0.00 g L⁻¹ OD₆₀₀⁻¹ and 0.26 ± 0.01 g L⁻¹ OD₆₀₀⁻¹, respectively when grown in 0.5x supplemented permeate (Fig. 3.5 B). When overexpressed together, we observed titers of 0.21 ± 0.00 g L⁻¹ OD₆₀₀⁻¹ for a single copy of *QDR3* and 0.29 ± 0.01 g L⁻¹ OD₆₀₀⁻¹ for the homozygous integration (Fig. 3.5 B). The trends were the same when these strains were grown in SC-Lactose, but the overall titers were lower with *QDR2* and *QDR3* co-overexpression producing 0.19 ± 0.01 g L⁻¹ OD₆₀₀⁻¹.

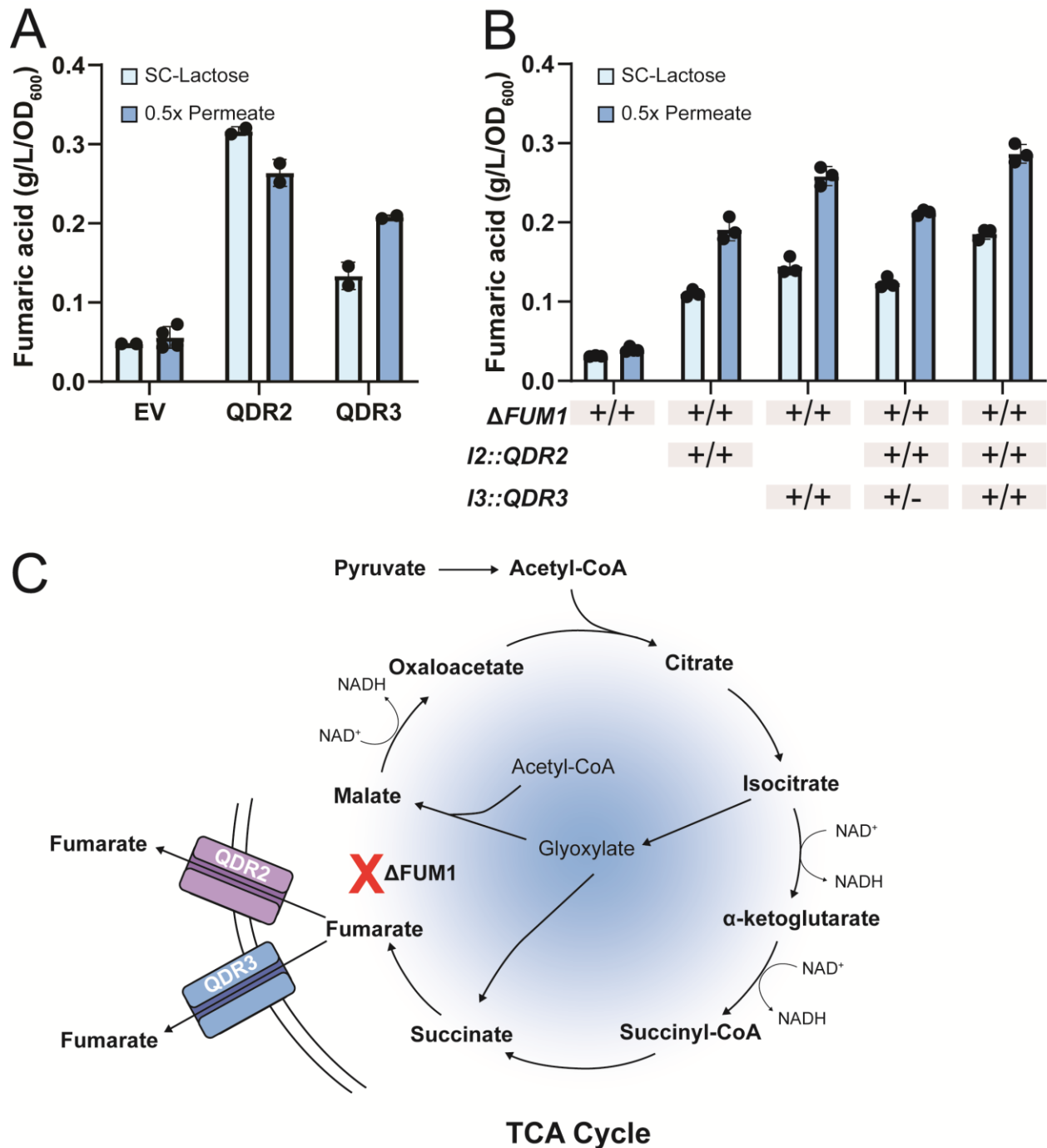


Figure 3.5. Effect of *QDR2* and *QDR3* on fumaric acid production in Y-1190 $\Delta FUM1$. **A.** Peak fumaric acid specific titers for *QDR2* and *QDR3* over expressed on plasmid **B.** Fumaric acid specific titers at 48 h with *QDR2* and/or *QDR3* integrated into I2 and I3 safe harbour sites. **C.** Schematic of increased fumaric acid production in *K. marxianus* with overexpression of *QDR2* and *QDR3*. Fumaric acid was analysis on an HPLC were in duplicate (A) or triplicate (B) and normalized to OD₆₀₀ measurements. This strain of *K. marxianus* is diploid. The + symbols represent how many copies were introduced into the genome. Error bars denote standard deviation.

While titers were the highest for the homozygous integration of *QDR2* and *QDR3* when normalized to biomass, the *QDR3* integrated strains produced less overall fumaric acid (Fig. S3.3). This is due to the overall lower growth of *QDR3* integrated strains (Fig. 3.4 C). The highest titer we achieved was in the *QDR2* and single-copy *QDR3* integrated strain at $0.89 \pm 0.003 \text{ g L}^{-1}$, which is $\sim 7.7 \text{ mM}$, much lower than 42 mM fumaric acid concentration we selected for (Fig. S3.3, Fig. 3.3 A). At 0 mM fumaric acid we observe a severe growth defect in *QDR3* integrated strains (Fig. 3.4 C), which suggests that while *QDR3* overexpression may enhance fumaric acid production and tolerance at high concentrations, it may also impose a burden that limits cell growth in the absence of product accumulation.

To evaluate the impact of the $\Delta FUMI$, *I2::QDR2*, *I3::QDR3* overexpression strain under controlled growth conditions, we scaled production up from 125 mL Erlenmeyer shake flasks to a 3 L fed-batch bioreactor. The *QDR2* and *QDR3* co-overexpression strain exhibited a striking improvement in fumaric acid production, reaching a maximum titer of 2.16 g L^{-1} at 89 hr , an 8.3 -fold increase over the control $Y-1190\Delta FUMI$ strain at 0.26 g L^{-1} under identical conditions (Fig. 3.6) Both strains produced substantial amounts of acetic acid, with maximum titers of 8.14 g L^{-1} in the $\Delta FUMI$ and 12.39 g L^{-1} in the $\Delta FUMI$, *I2::QDR2*, *I3::QDR3* overexpression strain (Fig. 3.6). Following the onset of exponential feeding at 49 h , neither strain exhibited long-term efficient lactose utilization, as evidenced by lactose accumulation commencing at $\sim 72 \text{ h}$ in the $\Delta FUMI$ strain (Fig. 3.6 A) and $\sim 90 \text{ h}$ in the *QDR* overexpression strain (Fig. 3.6 B). This result, together with the elevated acetic acid levels, suggests a potential metabolic bottleneck arising from the deletion of *FUMI*.

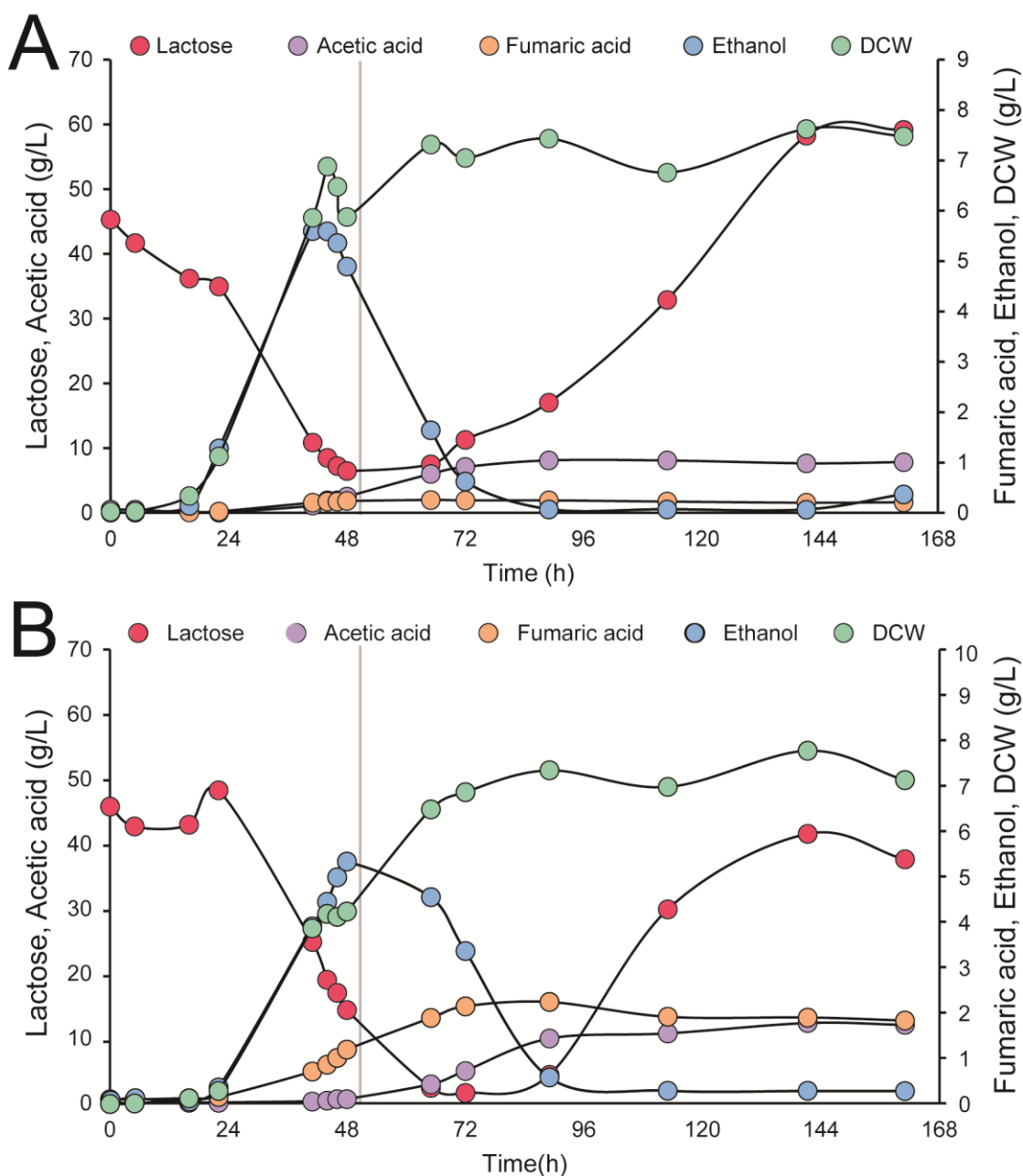


Figure 3.6. Growth and metabolite profiles of *K. marxianus* Y-1190. A. $\Delta FUM1$ and B. $\Delta FUM1$, I2::QDR2, I3::QDR3. Fed-batch fermentation in supplemented permeate with pH maintained at 4. The grey line represents the beginning of exponential feeding at 48 hr. Fermentation profiles of dissolved oxygen, feed pump, and pH can be found in Figure S3.5.

3.5. Discussion

In this study, we investigated *K. marxianus* for fumaric acid production using lactose-rich dairy permeate as a sustainable feedstock. Deletion of the fumarate hydratase *FUM1* increased fumaric acid titers but compromised acid tolerance (Fig. 3.1). To address this, we applied a tri-functional CRISPR-AID library to unbiasedly identify genetic perturbations that improve acid tolerance. Through this approach, we identified activation of *QDR2* and *QDR3* gene expression, and deletion of *ATP7*, as beneficial modifications. Notably, overexpression of *QDR2* and *QDR3* also enhanced fumaric acid production to 2.16 g L^{-1} , 8.3-fold over the *FUM1* deletion alone when grown in a fed-batch bioreactor (Fig. 3.6).

Much like in *S. cerevisiae* and *E. coli*, deletion of *FUM1* in *K. marxianus* increased fumaric acid accumulation^{27,36} (Fig. 3.1), however, this modification also led to a reduction in growth in the presence of fumaric acid stress, a phenotype not previously reported. Notably, Xu *et al.* predicted that a *FUM1* deletion might reduce growth rate in *S. cerevisiae* based on a flux balance analysis, but experimentally observed no significant difference between the wild type and *FUM1*-deleted strains³⁶. *FUM1* deletion represents only a preliminary step in the metabolic engineering of yeast for fumaric acid production. On its own, this perturbation resulted in modest titers of $0.610 \pm 0.031 \text{ g L}^{-1}$ in *S. cerevisiae*³⁶. Many *S. cerevisiae* studies have implemented more comprehensive strategies, including transporter overexpression, rerouting flux toward fumaric acid, downregulation of competing pathways such as ethanol and glycerol production, and DNA scaffolds to spatially organize enzymes in close proximity^{30,37,38}. These efforts have yielded fumaric acid titers ranging from 3.18 g L^{-1} to 33.13 g L^{-1} ^{30,37,38}. In *K. marxianus*, the reduced fumaric acid tolerance caused by the $\Delta FUM1$ deletion made enhancing tolerance a necessary prerequisite before advanced engineering strategies could be effectively implemented.

Unlike classical strain engineering approaches, CRISPR-based screens offer traceable interrogation of multiple genetic perturbations within one system. In addition, these functional approaches are especially powerful in non-conventional yeasts where much of the genome annotations are unknown or inferred from distant homology. While many CRISPR-based tools have been developed for *K. marxianus*^{72,101,108,109,116,179,180}, to date, only one genome-wide knockout library has been constructed in *K. marxianus*¹¹⁷. To determine CRISPR-AID functionality in *K. marxianus*, we targeted mNG for activation, interference and deletion and demonstrated all three expected gene expression perturbations (Fig. 3.2 A). Increased (2.6- to 4.6-fold) and decreased mNG fluorescence (2.9- to 5.7-fold) fell within the range reported in *S. cerevisiae*, with previous studies reporting 2.3- to 5-fold increases and 3- to 10-fold decreases in gene expression for activation and interference, respectively²⁰². Similarly in the non-conventional yeast *Pichia pastoris*, CRISPRa and CRISPRi have been shown to yield 3.5-fold increases and 3.3-fold decreases in eGFP fluorescence, respectively²⁰³. Prior CRISPRi studies in *K. marxianus* reported knockdown efficiencies between 1.5- and 5.6-fold¹¹⁶. Together, these results demonstrate that our CRISPR-AID system achieves gene modulation in *K. marxianus* at levels comparable to those reported in other yeast systems.

CRISPR-AID was originally applied in *S. cerevisiae* to enhance furfural tolerance in a genome-wide screen¹⁸⁸. In that study, most enriched guides corresponded to gene interference, with activation guides as the second most common and only a single deletion guide highlighted per condition. In contrast, our screen for fumaric acid tolerance revealed a strong enrichment for activation guides, comprising over ~99% of hits at all fumaric acid concentrations, compared to 62% in the 0 mM control (Fig. 3.3 D). Lian *et al.* identified hits that included interference of a negative regulator of the transporter *PDR5*, *PDR1*²⁰⁴. In our case, top hits included activation of

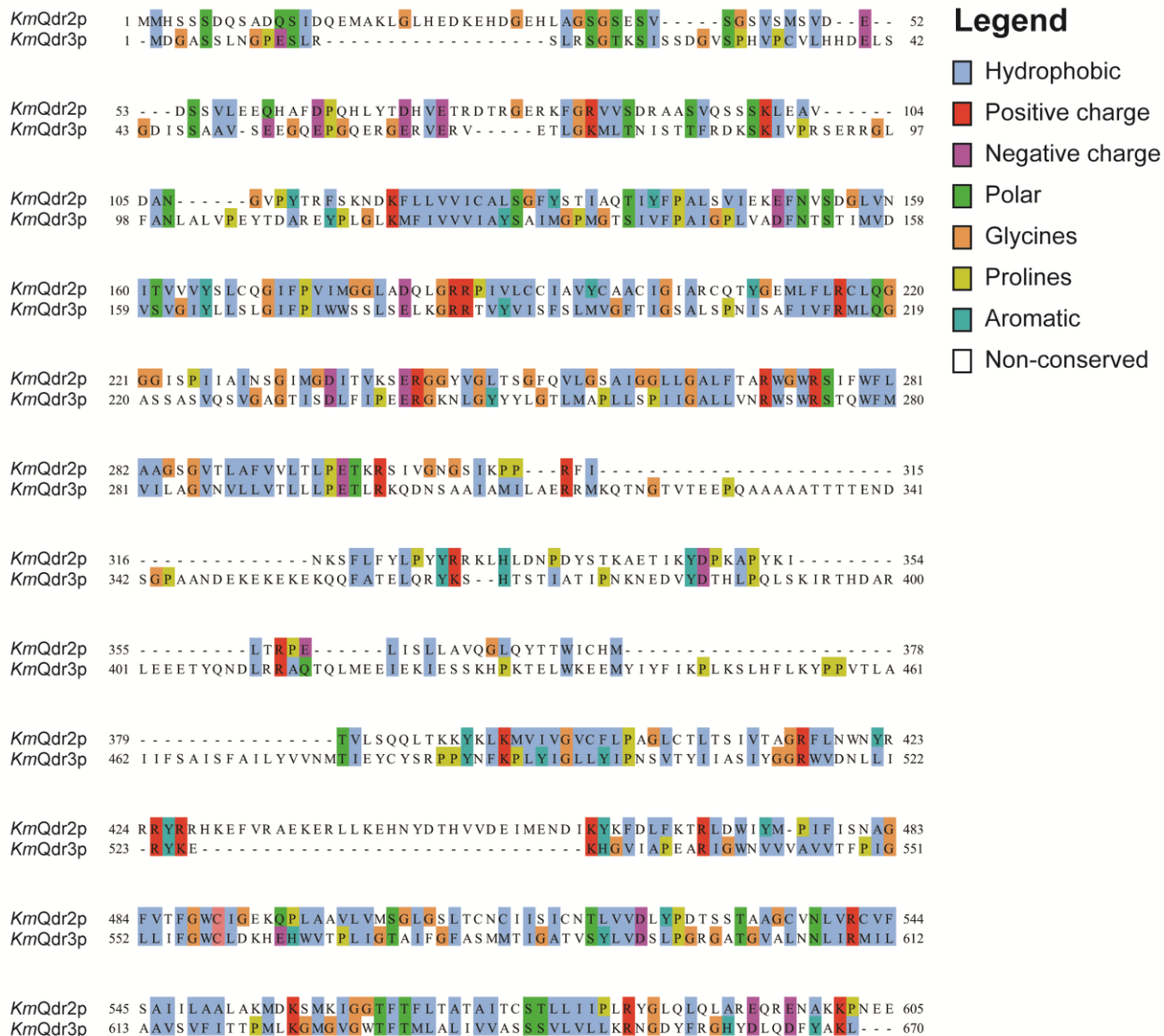
the multidrug transporters *QDR2* and *QDR3* (Fig. 3.4) whereas activation of *PDR5*, emerged in our screen at positions #30, #15, and #13 in the 25, 30, and 42 mM conditions, respectively (Table S3.8).

In our study, *QDR2* activation sgRNA accounted for 98% of all activation guide reads (Fig. 3.3) and was enriched ~350-fold over the next most abundant validated hit, a *QDR3* activation sgRNA (Fig. 3.4). A previous application of genome-wide CRISPR-AID screening also found that activation guides were consistently enriched across three iterative screening rounds²⁰⁵. For positive selection pooled library screens, it is common to have a narrow set of enriched hits. In an *E. coli* expression library screen for ethanol tolerance, three dominant hits were identified as being more abundant than the rest²⁰⁶, and in a *S. cerevisiae* CRISPR deletion screen for furfural tolerance, one primary guide was strongly enriched along with two additional, less abundant hits²⁰⁷. While this type of enrichment is typically desired, particularly under strong selection, it can also lead to population bottlenecks where a single dominant perturbation outcompetes others, potentially obscuring additional functionally relevant variants²⁰⁸. This dynamic likely contributed to our own findings, in which we validated only two additional hits outside of the dominant enrichment for *QDR2* activation (Fig. 3.4 A). It is possible that the fumaric acid concentrations applied for selection were overly stringent, leading to early bottlenecking of the library and overrepresentation of *QDR2* sgRNA. However, the CRISPR-AID study that relied solely on a fluorescent biosensor without applying direct selection pressure, still exhibited strong enrichment for a single dominant activation guide²⁰⁵. This suggests that factors beyond selection stringency may drive the emergence of dominant hits, even in the absence of overt selective constraints.

QDR2 and *QDR3* are genes encoding putative multidrug transporters previously characterized in *S. cerevisiae* as exporters involved in detoxification and efflux^{209,210}. Their

activation in *K. marxianus* improved acid tolerance up to the solubility limit of fumaric acid, but also enhanced fumaric acid production, suggesting that these transporters may reduce intracellular toxicity of fumaric acid through export (Fig. 3.4, Fig. 3.5 C). *QDR2* and *QDR3* are part of the Major Facilitator Superfamily (MFS), specifically, they are 12-spanner H⁺ antiporters in the *DHA1* family²¹¹. Despite their similar names, *QDR2* and *QDR3* are from different clusters of the DHA1 family and share only 23.4 % identity at the amino acid level, the majority of the matching sequence being across the transmembrane domains²¹² (Fig. 3.7). In *S. cerevisiae*, *QDR2* functions as a promiscuous cation/H⁺ antiporter with roles in potassium uptake²¹⁰, tolerance to cation stresses²¹³, and export of amino acids²¹⁴. Its broad substrate range, together with our finding that overexpression enhances growth in fumaric acid (Fig. 3.4 AC) while deletion impairs it (Fig. 3.4 D), suggests that *QDR2* may also export the anionic fumarate. *QDR3* similarly confers multi-drug resistance²⁰⁹, exports threonine²¹⁴, and is implicated in tolerance to several dicarboxylic acids²¹⁵, though paradoxically its deletion can improve acetic acid tolerance²¹⁶. The distinct yet overlapping substrate profiles and stress responses of *QDR2* and *QDR3* likely underly the enhanced fumaric acid tolerance we observed when both were overexpressed.

A



B

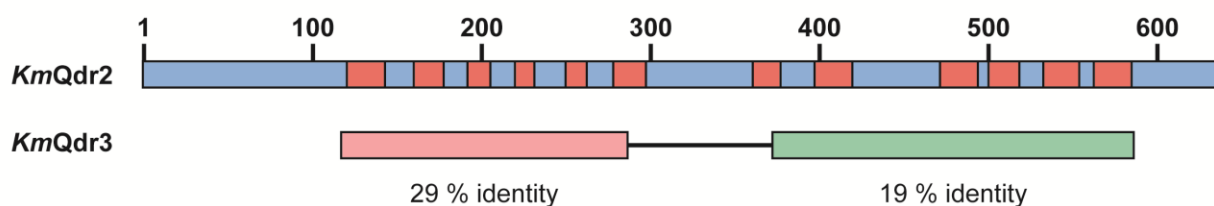


Figure 3.7 Alignment of *K. marxianus* Qdr2p and Qdr3p. A. Protein alignment at amino-acid level. **B.** A summary alignment with transmembrane domains highlighted in red, and areas of highest identity highlighted in pink and green.

In *S. cerevisiae*, acid efflux is a common response to acid stress, typically mediated by members of the ABC and MFS transporter superfamilies²¹⁷. Key ABC transporters include *SNQ2* for decanoic acid²¹⁸, and *PDR12* for sorbic, propionic, benzoic, octanoic, and levulinic acids^{218–220}. Among MFS transporters, several DHA1 and DHA2 members, including *AQR1*, *TPO1-4*, and *AZR1* export a range of short- to medium-chain acids^{218,221–224}. Within this framework, *QDR2* and *QDR3*, members of the DHA1 subfamily of MFS transporters, align well with other characterized acid efflux transporters, supporting their roles in mediating organic acid tolerance.

While deletion mutants leading to enhanced cellular function are less common than leading to reduced cellular function, examples such as deletion of *YBR238C* or *FOB1* in *S. cerevisiae* enhance yeast lifespan through enhanced mitochondrial function or replication of rDNA, respectively^{225–227}. In our case, the deletion of *ATP7*, which encodes a subunit of the stator arm of the mitochondrial F₁F₀ ATPase²²⁸, emerged as an unexpected but significant hit in our screen, conferring increased tolerance to 20 mM fumaric acid (Fig. 3.4 B). The F₁F₀ ATPase is the mitochondrial ATPase required for ATP generation through oxidative phosphorylation²²⁹. Disruption of ATPase complexes in general, including *ATP7*, are not classically associated with increasing acid tolerance. The identification of *ATP7* deletion in our study underscores the strength of an unbiased, genome-wide approach in revealing non-obvious genetic determinants of stress tolerance.

We observe high titers of fumaric acid in *QDR3* integrated strains when normalized to biomass (Fig. 3.5), but overall lower growth and final titers (Fig. S3.3). When grown in 0.5x permeate, but not SC media, our fumaric acid producing strains also produce high titers of acetic acid, reaching 5.6 g L⁻¹ at maximum (Fig. S3.4), possibly decreasing the fitness of the more susceptible *QDR3* overexpressed integrants. This effect may be linked to differences in nitrogen

sources: while both media contain $\sim 30 \text{ g L}^{-1}$ lactose, SC media uses ammonium sulphate for the nitrogen source, whereas 0.5x supplemented permeate uses urea¹⁷⁹. Urea catabolism releases CO_2 which can dissolve into the cytosol to increase the pH, rather than acidifying it, allowing for further accumulation of acetic acid before cell death^{158,230}. Supporting this idea, the combined *QDR2* and *QDR3* overexpression strain produced acetic acid at 12.39 g L^{-1} when grown in a bioreactor with the pH controlled to 4.0, likely mitigating some acid stress (Fig. 3.6).

The high concentrations of acetic acid observed in both shake flask and bioreactor experiments (Fig. S3.4, Fig. 3.6) indicate a metabolic imbalance, likely caused by the $\Delta FUMI$ mutation disrupting the TCA cycle (Fig. 3.1). Acetic acid accumulation may result from excess acetyl-CoA production or from overflow metabolism to generate additional $\text{NADH}^{231,232}$, and the reduced TCA flux due to *FUMI* deletion likely contributes to both mechanisms. While *QDR2* and *QDR3* overexpression enhances fumaric acid production, these transporters cannot fully mitigate the toxicity caused by accumulating acetic acid, highlighting a remaining metabolic limitation in the engineered strains.

With fumaric acid-tolerant *K. marxianus* strain in hand, we are now well-positioned to pursue targeted metabolic engineering strategies to further improve fumaric acid production. The identification and integration of fumaric acid tolerance-enhancing genes such as *QDR2* and *QDR3* provide a robust chassis for production. Future work will focus on reducing overflow metabolism and fine-tuning cofactor balance to maximize fumarate yield and reduce acetic acid. Together, these efforts will enable the construction of a high-performing production strain and expand the utility of *K. marxianus* as a platform for organic acid biosynthesis from industrial waste streams.

3.6. Supplementary figures and tables

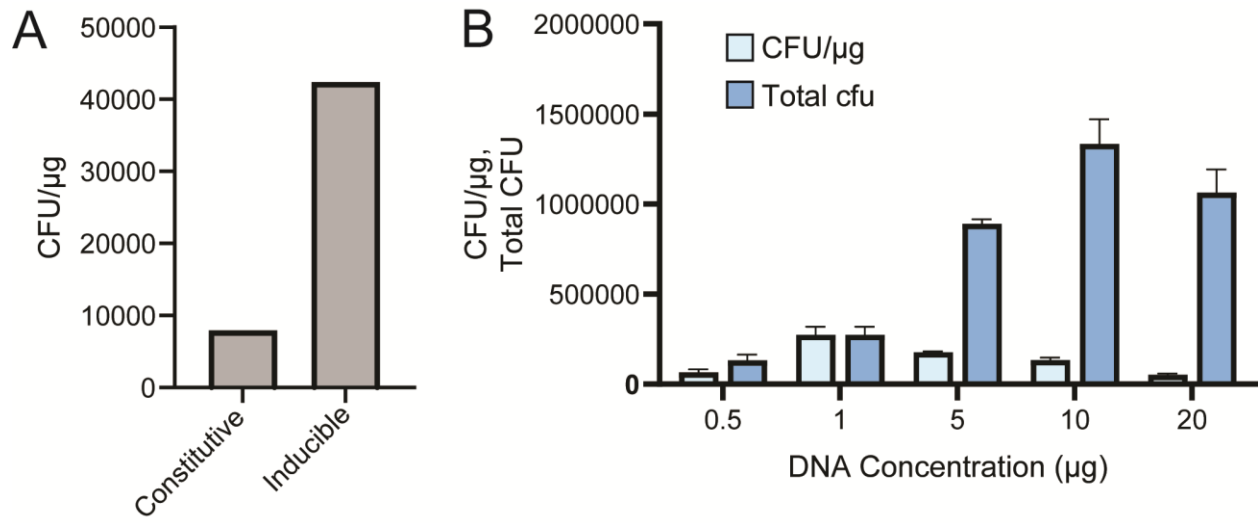


Figure S3.1. Transformation efficiency with CRISPR-AID system. A. Comparison of transformation efficiency of *K. marxianus* Y-1190 with the tri functional Cas cassette expressed constitutively (left), or under inducible promoters (right) B. Transformation efficiency in the inducible Cas system across multiple DNA concentrations. Error bars represent the standard deviation between two biological replicates.

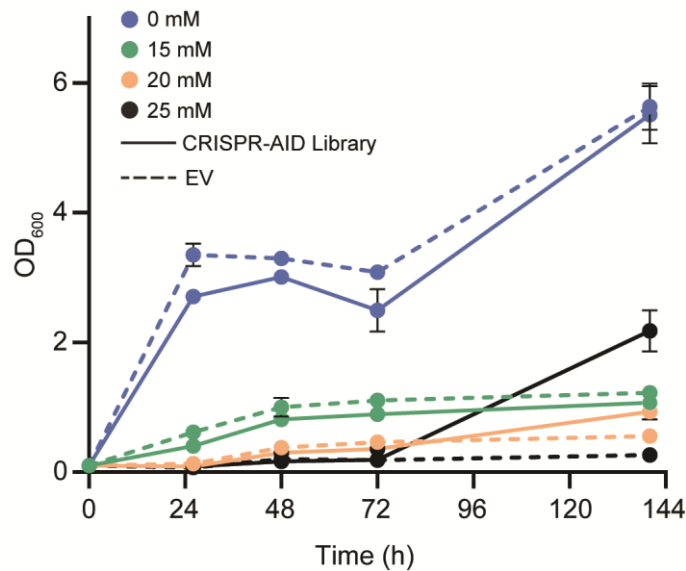


Figure S3.2. Testing fumaric acid concentrations to be used in CRISPR-AID enrichment experiments. In triplicate, MTK074 with an EV plasmid, or with the transformed genome-wide CRISPR-AID sgRNA library was grown in the presence of multiple concentrations of fumaric acid over six days. Each dot is the mean of three biological replicates and error bars represent standard deviation.

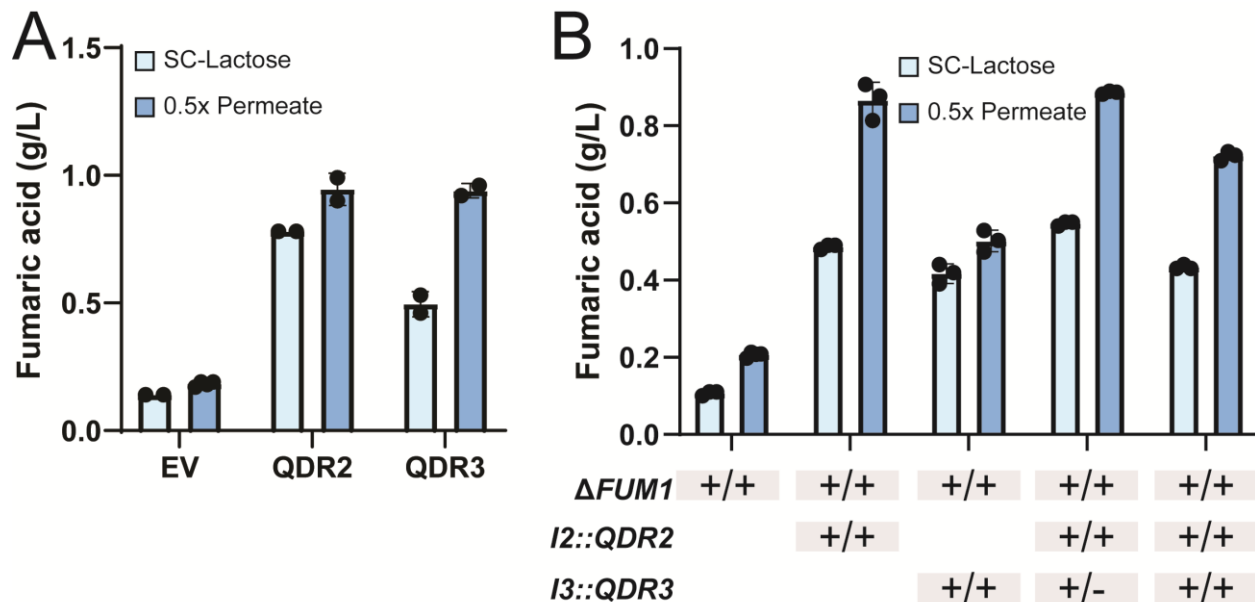


Figure S3.3. Fumaric acid titers in SC-Lactose and 0.5x Permeate with QDR2 and QDR3 overexpressed. **A.** on plasmid or **B.** by integration. Fumaric acid was measured on an HPLC using an Aminex 87-H column (Bio-Rad) in duplicate (A) or triplicate (B). This strain of *K. marxianus* is diploid. The + symbols represent how many copies were introduced into the genome. Error bars denote standard deviation.

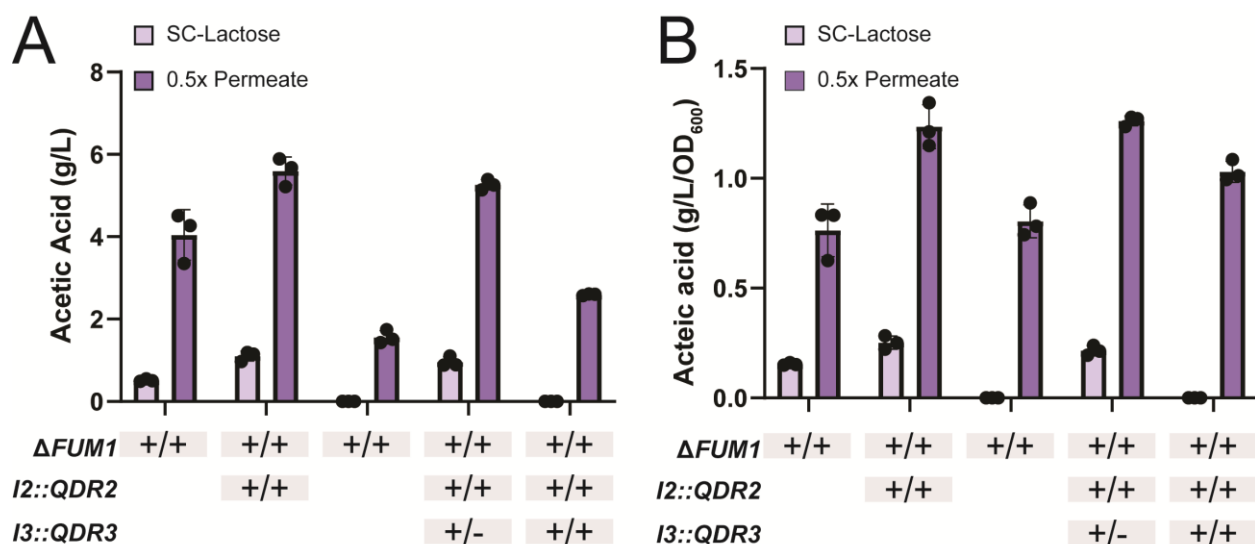


Figure S3.4. Production of acetic acid by QDR2 and QDR3 overexpression strains. **A.** Total titers and **B.** Specific production normalized to OD₆₀₀. Acetic acid was measured on an HPLC using an Aminex 87-H column (Bio-Rad) in triplicate. This strain of *K. marxianus* is diploid. The + symbols represent how many copies were introduced into the genome. Error bars denote standard deviation.

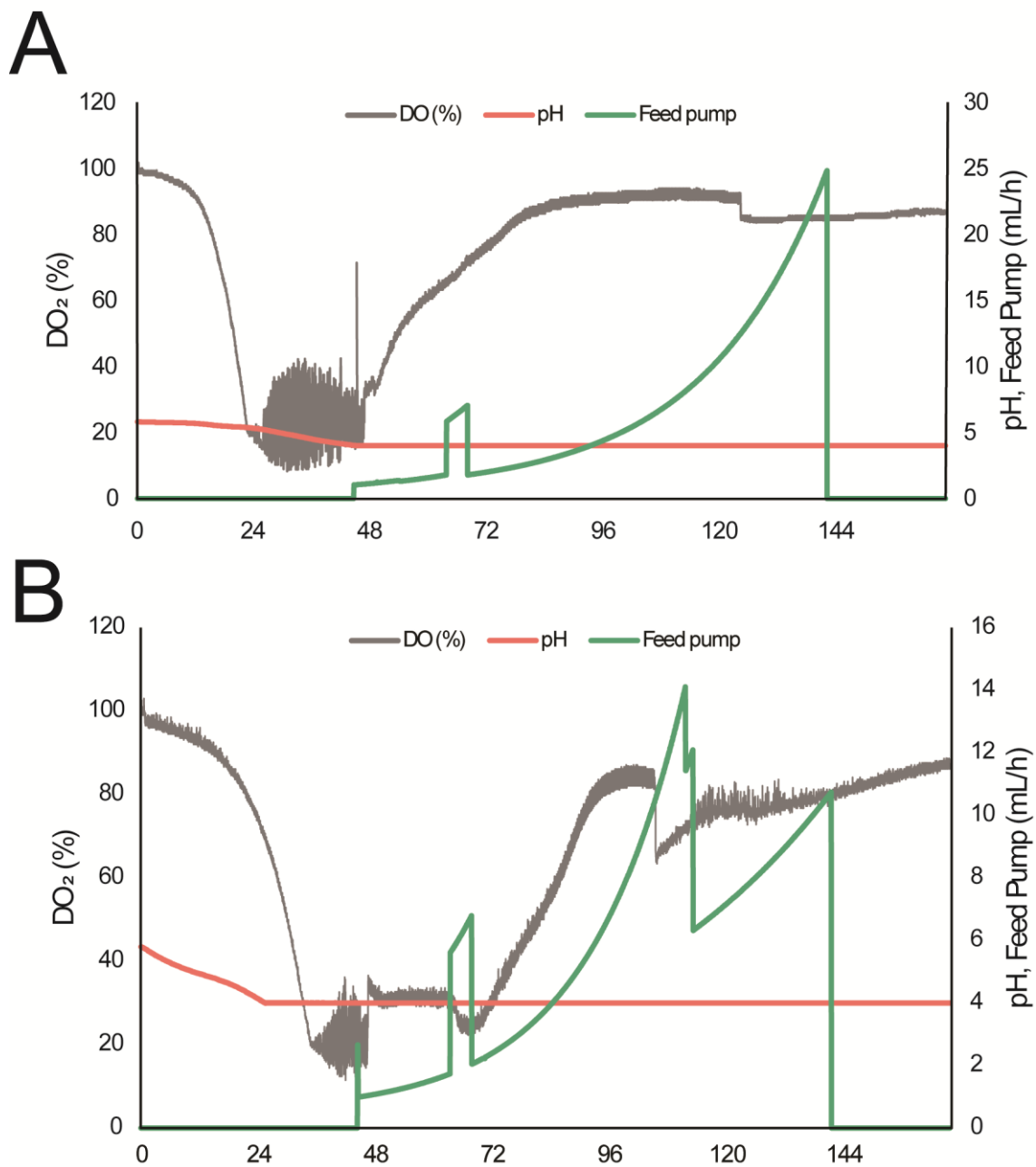


Figure S3.5 Fed-batch fermentation profiles of *K. marxianus* Y-1190. A. $\Delta FUM1$ and B. $\Delta FUM1$, I2::QDR2, I3::QDR3. Bioreactor was set up in batch mode until lactose was depleted at ~49 hr, at which time fed-batch was started using 200 g L⁻¹ lactose. pH was maintained at 4. Process data were recorded and analyzed using BioXpert V2 (Applikon).

Table S3.1. List of plasmids used in this study

Plasmid ID	Parental Plasmid	Characteristic	Selection	Reference	Addgene
Kmk.C1	-	Minimal ARS and centromeric sequence	Cam	101	125059
Kmk.I2L	-	Integration left arm downstream of KmBDH1 and KmBDH2	Cam	101	125031
Kmk.I2R	-	Integration right arm downstream of KmBDH1 and KmBDH2	Cam	101	125064
Kmk.I3L	-	Integration left arm between KmSWF1 and KmARO1	Cam	101	125032
Kmk.I3R	-	Integration right arm between of KmSWF1 and KmARO1	Cam	101	125065
Kmk.I4L	-	Integration left arm downstream of KmHSP104	Cam	101	125033
Kmk.I4R	-	Integration right arm downstream of KmHSP104	Cam	101	125066
Kmk.P2	-	CBS 6556 PDC1 promoter	Cam	101	125035
Kmk.T3	-	CBS 6556 KMXK_A03020 terminator	Cam	101	125055
Kmk.T4	-	CBS 6556 PDC1 terminator	Cam	101	125056
Kmk.T5	-	CBS 6556 PGK1 terminator	Cam	101	125057
p426*-Lb	-	Scaffold for LbCas12a for <i>S. cerevisiae</i>	Amp, - URA	188	136992
p426*-Sa	-	Scaffold for SaCas9 for <i>S. cerevisiae</i>	Amp, - URA	188	136994
p426*-Sp	-	Scaffold for SpCas9 for <i>S. cerevisiae</i>	Amp, - URA	188	136993
pAID6	-	CRISPR-AID Integration cassette for <i>S. cerevisiae</i>	Amp, G418	188	136989
pAK007	pYTK090	I2 GFP dropout landing pad	Kan	This work	-
pAK012	pMTKM01	SpCas9-sgRNA1-sgRNA2-double break in I2 landing pad	Amp, Nat	This work	-
pAK034	pYTK090	I4 Landing Pad	Kan	This work	-
pAK104	pMT189	Inducible CRISPR-AID Integration Cassette: KmTEF3tetO2prom-dLbCas12-VP-ScADH1term, KmTDH3tetO2prom-SaCas9-KmPDC1term, and KmPGK1tetO2prom-dSpCas9-RD1152- KmA3020term	Kan, G418	This work	-
pBB316	-	mScarlet CDS	Cam	Derivative of Bean <i>et al.</i> , 2022 ¹³⁹	-

pJD005	pMTKM01	URA3 double guide set	Amp, Nat	This work	-
pJet1.2	-	CloneJET PCR Cloning Kit	Amp	Thermofisher, K1231	-
pMT124	YTK083	Scaffold for SpCas9 in a level1 plasmid	Amp, - URA	This work	-
pMT149	pMT172	Sg4_mNeon_a guide in scaffold	Amp, - URA	This work	-
pMT150	pMT172	Sg5_mNeon_a guide in scaffold	Amp, - URA	This work	-
pMT151	pMT172	Sg6_mNeon_a guide in scaffold	Amp, - URA	This work	-
pMT152	pMT177	Sg7_mNEON guide in scaffold	Amp, - URA	This work	-
pMT153	pMT177	Sg8_mNeon guide in scaffold	Amp, - URA	This work	-
pMT154	pMT124	Sg1_mNeon_i guide in scaffold	Amp, - URA	This work	-
pMT155	pMT124	Sg2_mNeon_i guide in scaffold	Amp, - URA	This work	-
pMT156	pMT124	Sg3_mNeon_i guide in scaffold	Amp, - URA	This work	-
pMT172	YTK083	Scaffold for LbCas12 in a level1 plasmid, with GFP dropout	Amp, - URA	This work	-
pMT177	YTK083	Scaffold for SaCas9 in a level1 plasmid, with GFP dropout	Amp, - URA	This work	-
pMT188	pJet1.2	URA large::mNeon	Amp	This work	-
pMT189	-	I3 landing pad with selection	Kan, G418	This work	-
pMT193	pMTKM01	sgRNA plasmid targeting I3 locus	Amp, Nat	This work	-
pMT195	pMT189	Repair template to integrate the CRISPR-AID cassette: KmPGK1p-dLbCas12-VP-ScADH1t, KmTDH3p-SaCas9-KmPDC1t, and KmTEF3p-dSpCas9-RD1152-KmA3020t	Kan, G418	This work	-
pMT230	-	Destination plasmid for multiplexed guide and scaffold combinations. SNR52 promoter and SUP4 terminator	Amp, - URA	This work	-
pMT238	pYTK090	I3 Landing Pad without selection	Kan	This work	-
pMT249	pMT230	Combo CRISPR Test 1 (A)	Amp, - URA	This work	-

pMT250	pMT230	Combo CRISPR Test 2 (B)	Amp, - URA	This work	-
pMT251	pMT230	Combo CRISPR Test 3 (C)	Amp, - URA	This work	-
pMT252	pMT230	PDC1 activation with Sg4 act guide	Amp, - URA	This work	-
pMT253	pMT230	PDC1 inhibition with Sg2 Int guide	Amp, - URA	This work	-
pMT254	pMT230	mNeon Green deletion with Sg7 del guide	Amp, - URA	This work	-
pMT266	pAK034	I4 Donor: mTurquoise	Kan	This work	-
pMT267	pMT230	CRISPR-AID EV	Amp, - URA	This work	-
pMT273	pMT230	SOD1p guide act A	Amp, - URA	This work	-
pMT274	pMT230	SOD1p guide act B	Amp, - URA	This work	-
pMT275	pMT230	SOD1p guide act C	Amp, - URA	This work	-
pMT276	pMT230	TEF3p guide int A	Amp, - URA	This work	-
pMT277	pMT230	TEF3p guide int B	Amp, - URA	This work	-
pMT278	pMT230	TEF3p guide int C	Amp, - URA	This work	-
pMT321	YTK077	golden gate G418 destination plasmid	amp,G418	This work	-
pMT351	-	Destination plasmid for guide and scaffold combinations. SNR52 promoter and SUP4 terminator	Amp, - URA	This work	-
pMT418	pMTKM01	QDR2 CDS guide set 1	Amp, Nat	This work	-
pMT421	pMT321	pNC1_QDR2_ScADH1t	Amp, G418	This work	-
pMT422	pMT321	pNC1_LYS21_primary_ScADH1t	Amp, G418	This work	-
pMT423	pMT321	pNC1_LYS21_haplotig_ScADH1t	Amp, G418	This work	-
pMT424	pMT321	pNC1_IMD4_ScADH1t	Amp, G418	This work	-
pMT425	pMT321	pNC1_g3201_primary_ScADH1t	Amp, G418	This work	-

pMT426	pMT321	pNC1_g3201_haplotig_ScADH1t	Amp, G418	This work	-
pMT427	pMT321	pNC1_g2114_ScADH1t	Amp, G418	This work	-
pMT428	pMT321	pNC1_CSE1_ScADH1t	Amp, G418	This work	-
pMT429	pMT321	pNC1_FAS1_ScADH1t	Amp, G418	This work	-
pMT430	pMT321	pNC1_UBP14_ScADH1t	Amp, G418	This work	-
pMT432	pAK007	NC1-QDR2-ScADH1 I2 donor	Kan	This work	-
pMT433	pMT321	pNC1_QDR3_ScADH1t	Amp, G418	This work	-
pMT434	pMT321	spacer Empty vector (EV)	Amp, G418	This work	-
pMT435	pMT321	pNC1-YPR1-SCADH1t	Amp, G418	This work	-
pMT443	pMT238	I3 Donor: pNC1-QDR3-PGKt	Kan	This work	-
pMT444	pMTKM01	QDR3 CDS guide set 1	Amp, Nat	This work	-
pMT466	pMT321	pNC1-ACO7a-ScADH1t	Amp, G418	This work	-
pMT467	pMT321	pNC1-RPN1-ScADH1t	Amp, G418	This work	-
pMT468	pMTKM01	ATP7 guide 1	Amp, NAT	This work	-
pMTKM01	-	pCAS-Tyr-[gRNA: BsaI GFP dropout] with pan(OPT)ARS replication origin	Amp, Nat	Chapter 2 ¹⁷⁹	233473
pMTKM24	-	Level 0 plasmid MoClo with KM promoter PGK1	Cam	Chapter 2 ¹⁷⁹	233496
pMTKM25	-	Level 0 plasmid MoClo with KM promoter TEF3	Cam	Chapter 2 ¹⁷⁹	233497
pMTKM26	-	Level 0 plasmid MoClo with KM promoter TDH3	Cam	Chapter 2 ¹⁷⁹	233498
pMTKM28	YTK001	Y-1190 NC1 promoter	Cam	Chapter 2 ¹⁷⁹	233500
pMTKM30	-	TDH3-tetO2p level 0 Moclo part 2 (TetR repressed, with 2XtetO sequences)	Cam	Chapter 2 ¹⁷⁹	233502
pMTKM31	-	TEF3-tetO2p level 0 Moclo part 2 (TetR repressed, with 2XtetO sequences)	Cam	Chapter 2 ¹⁷⁹	233503
pMTKM32	-	PGK1-tetO2p level 0 Moclo part 2 (TetR repressed, with 2XtetO sequences)	Cam	Chapter 2 ¹⁷⁹	233504

pYTK001	-	Part Plasmid Entry Vector	Cam	128	65108
pYTK032	-	mTurquoise2 CDS	Cam	128	65139
pYTK053	-	ScADH1 terminator	Cam	128	65160
pYTK077	-	Encodes KanamycinR	Cam, G418	128	65184
pYTK083	-	Encodes AmpR-ColE1	Amp	128	65190
pYTK090	-	Encodes KanR-ColE1	Kan	128	65197

Table S3.2. List of oligonucleotides used in this study

Application	OligoID	Nucleotide sequence (5' – 3')
Amplify dSpCas9-RD1152 from pAID6	MT174	GCATCGTCTCATCGGTCTCATATGatggacgtccaaagaag
Amplify dSpCas9-RD1152 from pAID6	MT175	ATGCCGTCTCAGGTCTCAGGATttacgcaactggaacagatg
Amplify SaCas9 from pAID6	MT176	GCATCGTCTCATCGGTCTCATATGatggggaacggaactac
Amplify SaCas9 from pAID6	MT177	ATGCCGTCTCAGGTCTCAGGATTTAcgagactttctcttcttctg
Amplify dLbCas12a-VP from pAID6	MT178	GCATCGTCTCATCGGTCTCATATGccaccatggctcctcc
Amplify dLbCas12a-VP from pAID6	MT179	ATGCCGTCTCAGGTCTCAGGATtcacagcaaggctgagaaatc
SaCas9 diagnostic primer	MT191	tctctttgccggtgttc
SaCas9 diagnostic primer	MT192	aagatctgctgaataatccattc
dLbCas12a-VP diagnostic primer	MT193	gtggaatgccgagtatgac
dLbCas12a-VP diagnostic primer	MT194	gcaagctgtatatgttccag
dLbCas12a-VP diagnostic primer	MT195	gtatcagaagttcgagaagatg
dLbCas12a-VP diagnostic primer	MT196	cattggacgattttgatctgg
dSpCas9-RD1152 diagnostic primer	MT201	ttcagctcaatgattaagcg
dSpCas9-RD1152 diagnostic primer	MT202	gttcgacgataaagtgatgaag
dSpCas9-RD1152 diagnostic primer	MT203	atatctgaatgccgtggtg
dSpCas9-RD1152 diagnostic primer	MT204	caagtactcgacaccacc
Amplifies 900 bp upstream of URA3	MT293	GCAAGGGAATTTGCCACC
Amplifies 900 bp upstream of URA3, 30 bp homology to prPDC1	MT294	tccattattcacgctgtatattcgctggatCCTCCTTTGATTAGTTTATGGGC
Amplifies 900 bp downstream of URA3, 30 bp homology to PDC1term	MT297	caagttacaagttacaagttaccAGAGTTCTCCGAGACAAGC
Amplifies 900 bp downstream of URA3	MT298	TGGTATGCTTCTTGTGCATTC
Amplifies NeonGreen transcriptional unit with PDC1 prom&ter	MT295	atccagcgaatatacageg

Amplifies NeonGreen transcriptional unit with PDC1 promoter	MT296	Tggttaactgttaactgttaactgtg
mNeonGreen diagnostic primer	MT303	atcttcggttcattaacgg
mNeonGreen diagnostic primer	MT304	cgttaatgaaccgaagatg
short guide for inhibition of Δ URA::mNEON by CCTOP	MT324	TGATCGAGTGTCTGGATCAGTTTTG
short guide for inhibition of Δ URA::mNEON by CCTOP	MT325	AAACCAAACTGATCCAGACACTCG
short guide for inhibition of Δ URA::mNEON by CCTOP	MT326	TGATCATAATACAAATATACATTAA
short guide for inhibition of Δ URA::mNEON by CCTOP	MT327	AACTTAATGTATATTTGTATTATG
short guide for inhibition of Δ URA::mNEON by CCTOP	MT328	TGATCAATACAAATATACATTAAAG
short guide for inhibition of Δ URA::mNEON by CCTOP	MT329	AAACCTTTAATGTATATTTGTATTG
short guide for activation of Δ URA::mNEON by CCTOP	MT330	AGATGAATACTGAATATCGAATAT
short guide for activation of Δ URA::mNEON by CCTOP	MT331	AAAAATATTCGATATTCAGTATTC
short guide for activation of Δ URA::mNEON by CCTOP	MT332	AGATTATATAAGTGGAGTGTCTGG
short guide for activation of Δ URA::mNEON by CCTOP	MT333	AAAACCAGACACTCCACTTATATA
short guide for activation of Δ URA::mNEON by CCTOP	MT334	AGATATTTTATTCTGGTTCCTGCG
short guide for activation of Δ URA::mNEON by CCTOP	MT335	AAAACGCAGGAACCAGAATAAAAT
short guide for deletion of Δ URA::mNEON by CCTOP	MT336	TGATCTCAAGGGCGAAGCCCAAGTT
short guide for deletion of Δ URA::mNEON by CCTOP	MT337	AAACAACCTGGGCTTCGCCCTTGAG
short guide for deletion of Δ URA::mNEON by CCTOP	MT338	TGATCGAACCGAAGATGTGCAACTC
short guide for deletion of Δ URA::mNEON by CCTOP	MT339	AAACGAGTTGCACATCTTCGGTTCG
pSOD1 activation guide_1	MT557	taccactaaaccacttgcgcaTGGAGGAAGGCAATTGAAGAAgactacactta
pSOD1 activation guide_2	MT558	taccactaaaccacttgcgcaGAAAAGGAGCTGAAAATGACTTAatctacacttagtagaaatt
pSOD1 activation guide_3	MT559	taccactaaaccacttgcgcaAGGCGTAGCGCAACCCCTGTGGAatctacacttagtagaaatt
Universal forward primer for dLbCas12a guides	MT556	ttcgattccgggcttgcgcaaatcttactaagttagat
Universal reverse primer for dSpCas9 guides	MT560	taccactaaaccacttgcgcaaaagcaccgactcggtgccacttttcaagttgataacggactagccttatttaacttg
pTEF1 interference guide_1	MT561	ttcgattccgggcttgcgcaTCAATTTGAATAATTAGAGgttttagagctagaataagcaagttaaaataaggctagtc
pTEF1 interference guide_2	MT562	ttcgattccgggcttgcgcaTTTTCTTCTGTTTAGTACgttttagagctagaataagcaagttaaaataaggctagtc
pTEF1 interference guide_3	MT563	ttcgattccgggcttgcgcaAAAATTTACATTATCGAAAAgttttagagctagaataagcaagttaaaataaggctagtc

Universal reverse guide for SaCas9 guides	MT573	taccactaaaccacttgcgcatcaaaaaatctcgccaacaagttgac
mNeonGreen deletion guide_1	MT565	ttcgattccgggcttgcgcaTCAAGGGCGAAGCCCAAG
mNeonGreen deletion guide_2	MT566	ttcgattccgggcttgcgcaGAACCGAAGATGTGCAAC
Amplification of sgRNA library with heterogeneity spacers, universal reverse	MT788	TCGTCGGCAGCGTCagatgtgtataagagacagtatatgcatatatgcagtggtg
Amplification of sgRNA library with heterogeneity spacers, NGS1	MT789	GTCTCGTGGGCTCGGagatgtgtataagagacagCTAtctccgcagtgaagataa atgac
Amplification of sgRNA library with heterogeneity spacers, NGS2	MT790	GTCTCGTGGGCTCGGagatgtgtataagagacagTGACtctccgcagtgaagataa atgac
Amplification of sgRNA library with heterogeneity spacers, NGS3	MT791	GTCTCGTGGGCTCGGagatgtgtataagagacagGTCTAtctccgcagtgaagat aatgac
Amplification of sgRNA library with heterogeneity spacers, NGS4	MT792	GTCTCGTGGGCTCGGagatgtgtataagagacagACTGTGtctccgcagtgaag ataaatgac
Amplification of sgRNA library with heterogeneity spacers, NGS5	MT793	GTCTCGTGGGCTCGGagatgtgtataagagacagTAGTCAGtctccgcagtgaag ataaatgac
Amplification of sgRNA library with heterogeneity spacers, NGS6	MT794	GTCTCGTGGGCTCGGagatgtgtataagagacagTGCTctccgcagtgaagataa atgac
Amplification of sgRNA library with heterogeneity spacers, NGS7	MT795	GTCTCGTGGGCTCGGagatgtgtataagagacagCTGAtctccgcagtgaagataa atgac
Amplification of sgRNA library with heterogeneity spacers, NGS8	MT796	GTCTCGTGGGCTCGGagatgtgtataagagacagCTGAAtctccgcagtgaagat aatgac
Amplification of sgRNA library with heterogeneity spacers, NGS9	MT797	GTCTCGTGGGCTCGGagatgtgtataagagacagATCGACtctccgcagtgaag ataaatgac
Amplification of sgRNA library with heterogeneity spacers, NGS10	MT798	GTCTCGTGGGCTCGGagatgtgtataagagacagTAGTTGtctccgcagtgaag ataaatgac
Amplification of QDR2 (g4399) from K.marxianus Y-1190	MT812	GCATCGTCTCATCGGTCTCATatgatgcatagtagtagtgac
Amplification of QDR2 (g4399) from K.marxianus Y-1190	MT813	ATGCCGTCTCAGGTCTCAGGATctatagagctgagaaatttgattgc
Amplification of double guide set for QDR2	MT820	AAAAGGTCTCAGACTTTACTCTGACCCTGACCCCGCCgttttagagc tagaaatagcaagttaaaataag
Amplification of double guide set for QDR2	MT821	TACGGGTCTCTAAACttggcatccactgcttctagtgcgcaagcccgaatc
QDR2 diagnostic primer	MT855	atgatgcatagtagtagtgatcaaaag
QDR2 diagnostic primer	MT825	gccagacctccataataac
Amplification of IMD4 from K. marxianus Y-1190	MT835	GCATCGTCTCATCGGTCTCATATGccaggttaagctattgg
Amplification of IMD4 from K. marxianus Y-1190	MT836	ATGCCGTCTCAGGTCTCAGGATttagttgttagacgtttttcg
Amplification of QDR3 from K. marxianus Y-1190	MT837	GCATGAAGACAATCGGTCTCATatggacggggcaagttc
Amplification of QDR3 from K. marxianus Y-1190	MT858	ATGCGAAGACAAGGTCTCAGGATtattcaagcttagcatagaaatcttg
Amplification of LYS21 from K. marxianus Y-1190	MT839	GCATCGTCTCATCGGTCTCATATGtctgtgaactcaaatccg
Amplification of LYS21 from K. marxianus Y-1190	MT840	ATGCCGTCTCAGGTCTCAGGATtcataactccgaatggaaatttttg
Amplification of g2114 from K. marxianus Y-1190	MT841	GCATCGTCTCATCGGTCTCATatggatctagaggcagatatctc

Amplification of g2114 from K. marxianus Y-1190	MT842	ATGCCGTCTCAGGTCTCAGGATctactcgattgtttttgctttgg
Amplification of YPR1 from K. marxianus Y-1190	MT843	GCATCGTCTCATCGGGAAGACAAAtatgtcacattgaagccatc
Amplification of YPR1 from K. marxianus Y-1190	MT856	ATGCCGTCTCAGGTCTCAAGACAAGGATctattcaaatgtcttgaaagtacc
Amplification of CSE1 from K. marxianus Y-1190	MT845	ATCGCGTCTCATatggccgatttagagtcag
Amplification of CSE1 from K. marxianus Y-1190	MT846	ATCGCGTCTCAGGATctatcgagcacataattttgttaatg
Amplification of FAS1 from K. marxianus Y-1190	MT847	ATCGCGTCTCATatgagctcatctgcaagacc
Amplification of FAS1 from K. marxianus Y-1190	MT848	ATCGCGTCTCAGGATttattggttaacaatcttttcccag
Amplification of g3201 from K. marxianus Y-1190	MT849	GCATCGTCTCATCGGTCTCATatgaatccaacgagtgctc
Amplification of g3201 from K. marxianus Y-1190	MT850	ATGCCGTCTCAGGTCTCAGGATctagaccggactattatcaataaatttag
Removal of internal bsmbi sites in IMD4	MT851	ATTACGTCTCaggacggtaagcgtttgaag
Removal of internal bsmbi sites in IMD4	MT852	ATTACGTCTCagtcctgtgaaatattcacc
Amplification of UBP14 from K. marxianus Y-1190	MT853	ATCGCGTCTCATatgtctgactttgatattgtgac
Amplification of UBP14 from K. marxianus Y-1190	MT857	ATCGCGTCTCAGGATtcattccaatagggtgaagaag
QDR3 diagnostic primers	MT877	agaagaatattggaacaactactag
QDR3 diagnostic primers	MT878	catgaaccattgtgtcgatc
Amplification of double guide set for ATP7	MT892	AAAAGGTCTCAGACTTTCTTACCGGTCAACTTCAAAGgttttagagc tagaaatagcaagttaaaataag
Amplification of double guide set for ATP7	MT893	TACGGGTCTCTAAACctattgaggctttcgaaatcgcgcaagcccgaatc
ATP7 diagnostic primer	MT896	atgtctctagccaaatctgc
ATP7 diagnostic primer	MT897	acataacgggtcaaatcacgc
Amplification of RPN1 from K. marxianus Y-1190	MT898	GCATCGTCTCATCGGTCTCATatgggtgaagataaagtactaaag
Amplification of RPN1 from K. marxianus Y-1190	MT899	ATGCCGTCTCAGGTCTCAGGATttaattcacttctgtgtactttgg
Amplification of ACO7a from K. marxianus Y-1190	MT900	GCATCGTCTCATCGGTCTCATatgctagctgctagaagatc
Amplification of ACO7a from K. marxianus Y-1190	MT901	ATGCCGTCTCAGGTCTCAGGATtcatgtagatgagcggc
I2 diagnostic primers	AK831	AGGTATATAAGCGAGAAATTGGC
I2 diagnostic primers	AK832	TAATCACGGTGAAGTGAATGC
I4 diagnostic primers	AK950	tccttcctgttcaatgcag
I4 diagnostic primers	AK951	tgtgggtgtgtgtttttctc
I3 diagnostic primers	MT487	CTTTACGTGATCGAGACCG
I3 diagnostic primers	MT488	tctAGCCATTTGATATCAACAAAC

Blunt primers for first amplification of the AID library	MT736	attccgggcttgcgca
Blunt primers for first amplification of the AID library	MT737	taccactaaccacttgcg
Primers with overhangs for library insertion into entry vector	AK954	GCATGGTCTCAAATCCAACGTTGCCATCGTTGGGCCCGGTT CGATTCCGGGCTTGCGCA
Primers with overhangs for library insertion into entry vector	AK961	GCATGGTCTCACCCAACGATGGCAACGTTGGATTTTACCACT AAACCACTTGCGC
Amplification of double guide set for QDR3	MT863	AAAAGGTCTCAGACTTTGACGGGGCAAGTTCTTTAAAggttttagag ctagaatatgcaagttaaaataag
Amplification of double guide set for QDR3	MT864	TACGGGTCTCTAAACCGCATCTGTATACTCCGGGAtgcaagccc ggaatc

Table S3.3. List of yeast strains used in this study

Species	Strain Name	Parental Strain	Genetic modification	Details	Selection
K. marxianus	Y-1190	-	-	Diploid yeast	-
K. marxianus	ROK002	Y-1190	Δ FUM1		-
K. marxianus	MTK017	Y-1190	Δ URA3::mNeonGreen		-
K. marxianus	MTK019	Y-1190	Δ URA3::mNeonGreen, I3::CRISPR-AID	CRISPR-AID = KmPGK1p-dLbCas12-VP-ScADH1t, KmTDH3p-SaCas9-KmPDC1t, and KmTEF3p-dSpCas9-RD1152-KmA3020t	G418
K. marxianus	MTK056	MTK019	Δ URA3::mNeonGreen, I3::CRISPR-AID, I4::mturquoise (het)	CRISPR-AID = KmPGK1p-dLbCas12-VP-ScADH1t, KmTDH3p-SaCas9-KmPDC1t, and KmTEF3p-dSpCas9-RD1152-KmA3020t	G418
K. marxianus	MTK063	MTK056	Δ URA3::mNeonGreen, I3::CRISPR-AID, I4::mturquoise (het), I2::mScarlett ,	CRISPR-AID = KmPGK1p-dLbCas12-VP-ScADH1t, KmTDH3p-SaCas9-KmPDC1t, and KmTEF3p-dSpCas9-RD1152-KmA3020t	G418
K. marxianus	MTK072	MTK017	Δ URA3::mNeonGreen, I3::inducibleCRISPR-AID	inducibleCRISPR-AID=SSA3p-tetR-SV40-Inu1t-ConL1-TEF3tetO2p-LbCas12a-VP-ScADH1t-ConL2-TDH3tetO2p-SaCas9-PDC1t-ConL3-PGK1tetO2p-SpCas9-RD1152-KMX_A3020t-G418	G418

K. marxianus	MTK074	MTK072	Δ URA3::mNeonGreen, I3::inducibleCRISPR-AID, Δ FUM1	inducibleCRISPR- AID=SSA3p-tetR-SV40- Inu1t-ConL1- TEF3tetO2p-LbCas12a- VP-ScADH1t-ConL2- TDH3tetO2p-SaCas9- PDC1t-ConL3- PGK1tetO2p-SpCas9- RD1152-KMX_A3020t- G418	G418
K. marxianus	MTK115	ROK002	Δ FUM1, Δ QDR2		-
K. marxianus	MTK117	ROK002	Δ FUM1, I2:: pNC1-QDR2- ScADH1term		-
K. marxianus	MTK129	ROK002	Δ FUM1, I3::pNC1-QDR3- PGK1term		=
K. marxianus	MTK130het	MTK117	Δ FUM1, I2:: pNC1-QDR2- ScADH1term, I3::pNC1-QDR3- PGK1term (het)		-
K. marxianus	MTK130ho	MTK117	Δ FUM1, I2:: pNC1-QDR2- ScADH1term, I3::pNC1-QDR3- PGK1term (hom)		-
K. marxianus	MTK135	MTK115	Δ FUM1, Δ QDR2, Δ QDR3		-
K. marxianus	MTK139	ROK002	Δ FUM1, Δ QDR3		-
K. marxianus	MTK142	ROK002	Δ FUM1, Δ ATP7		-

Table S3.4. sgRNA scoring rubric

Scoring	Activation	Interference	Deletion
Efficiency (E)	Kim <i>et al.</i> , 2018 ²³³	Xu <i>et al.</i> , 2015 ²³⁴	Doench <i>et al.</i> , 2014 ²³⁵
Position (a)	$a = x-250 /250$	$a = X-125 /125$	If $x/CDS < 1/3$, $a=0$ If $1/3 \leq x/CDS \leq 2/3$, $a=0.2$ If $x/CDS > 2/3$, $a=0.5$

GC Score (b)	if 40-60% b=0 if 30-40% or 60-70% b=0.2 if 20-30% or 70-80% b=0.4 if 10-20% or 80-90% b=0.6 if 0-10% or 90-100% b=0.8
Off-target score (c)	$c = (SM+MM0+MM1+MM2+MM3)/20$
PolyT & PolyG Score (d & e)	If consecutive T>4, d =1, else d=0 If consecutive G>5, e =1, else e =0
Diversity Score (f)	If distance <10 bp, f=0 Else, f=1
Total score (S)	$S = (3+E-a-b-c)*d*e*f$

Table S3.5. List of sgRNA used in the CRISPR-AID genome-wide guide library. Please see attached spreadsheet.

Table S3.6. Genome-wide CRISPR-AID library coverage

	Library Size	<i>E. coli</i> transformants (total cfu)	<i>E. coli</i> library coverage	<i>K. marxianus</i> Transformants (total cfu)	<i>K. marxianus</i> Library Coverage
Activation	26185	4.84×10^5	18.5	5.4×10^5	20.6
Interference	26988	5.15×10^5	19.1	8.9×10^5	33.0
Deletion	18763	4.00×10^5	21.3	3.2×10^5	17.1
Total	71936	1.40×10^6	-	1.75×10^6	-

Table S3.7. Library sequencing read coverage per condition

Condition	Replicate	Merged reads	Percent Assembled	Theoretical read coverage per guide
0 mM	A	1080223	99.45	15.02
	B	1145441	99.44	15.92
	C	1027931	99.49	14.29
25 mM	A	1288488	99.49	17.91
	B	1248906	99.45	17.36
	C	1187174	99.48	16.50
30 mM	A	1115467	99.43	15.51
	B	1300580	99.47	18.08
	C	1229784	99.47	17.10
42 mM	A	1062307	99.49	14.77
	B	659048	99.37	9.16
	C	899214	99.44	12.50
Average		1.10E+06	99.46	15.34

Table S3.8. Significantly enriched sgRNA from the genome-wide CRISPR-AID screen in fumaric acid.

Condition	Gene name	sgRNA ID	Common Name	LFC	FDR
25 mM Fumaric acid	g4399	ACT21802	QDR2	10.604	0
25 mM Fumaric acid	g4399	ACT21806	QDR2	10.564	0
25 mM Fumaric acid	g3577	ACT16642	FAS1	6.6233	0
25 mM Fumaric acid	g3178	ACT14031	UBP14	6.5932	0
25 mM Fumaric acid	g4399	ACT21801	QDR2	6.4336	0
25 mM Fumaric acid	g592	ACT25286	IMD4	5.9393	0
25 mM Fumaric acid	g2142	ACT07423	YPR1	5.1933	0
25 mM Fumaric acid	g4077	DEL13335	ATP7	4.4575	4.63E-209
25 mM Fumaric acid	g902	ACT27254	ACO2a	4.3621	0
25 mM Fumaric acid	g547	ACT24985	QDR3	4.3215	0
25 mM Fumaric acid	g1501	ACT03264	LYS21	4.1276	0
25 mM Fumaric acid	g1123	ACT00802	NNK1	4.0638	4.71E-201

25 mM Fumaric acid	g292	ACT12438	KLMA_70317	4.0618	8.03E-282
25 mM Fumaric acid	g1640	ACT04144	RPN1	3.9984	0
25 mM Fumaric acid	g3201	ACT14193	KLMA_40132	3.9145	0
25 mM Fumaric acid	g3020	ACT13106	CSE1	3.8772	1.1226e-314
25 mM Fumaric acid	g3618	ACT16917	TOM1	3.8754	2.58E-268
25 mM Fumaric acid	g878	ACT27085	GIS2	3.8138	3.67E-154
25 mM Fumaric acid	g971	ACT27711	KLMA_10227	3.7693	5.21E-138
25 mM Fumaric acid	g2114	ACT07232	KLMA_20588	3.7518	0
25 mM Fumaric acid	g4278	ACT21032	GPI12	3.6949	9.50E-298
25 mM Fumaric acid	g2445	ACT09339	RRN3	3.6405	6.64E-167
25 mM Fumaric acid	g2043	ACT06768	PKP1	3.5818	2.60E-127
25 mM Fumaric acid	g1390	DEL01718	SNQ2	3.5325	5.70E-154
25 mM Fumaric acid	g1538	ACT03500	PEX32	3.5141	2.73E-196
25 mM Fumaric acid	g2642	ACT10606	CDA2	3.4765	4.36E-68
25 mM Fumaric acid	g267	ACT10790	KLMA_70292	3.4488	2.07E-71
25 mM Fumaric acid	g694	ACT25965	LAS21	3.4488	4.90E-153
25 mM Fumaric acid	g3308	ACT14880	PDR5	3.4192	3.08E-242
25 mM Fumaric acid	g4399	ACT21805	QDR2	3.4067	0
25 mM Fumaric acid	g659	ACT25730	GIP4	3.3736	1.02E-204
25 mM Fumaric acid	g4406	ACT21855	ZUO1	3.3221	1.14E-169
25 mM Fumaric acid	g3813	ACT18126	g3813	3.2926	1.67E-171
25 mM Fumaric acid	g2056	ACT06851	CAP2	3.285	5.19E-90
25 mM Fumaric acid	g2189	ACT07724	ARG81	3.279	3.12E-178
25 mM Fumaric acid	g1345	ACT02252	KLMA_10605	3.2759	4.94E-176
25 mM Fumaric acid	g3014	ACT13062	DUG1	3.2401	2.32E-148
25 mM Fumaric acid	g1291	ACT01896	HGT1	3.2201	1.21E-96
25 mM Fumaric acid	g2525	ACT09854	ATP16	3.1961	5.92E-187
25 mM Fumaric acid	g3700	ACT17407	VPS54	3.1134	5.24E-98
25 mM Fumaric acid	g407	ACT19751	PAC2	3.1113	1.06E-119
25 mM Fumaric acid	g547	ACT24990	QDR3	3.0754	1.87E-148
25 mM Fumaric acid	g2227	ACT07985	VMA2	3.019	3.84E-87
25 mM Fumaric acid	g1670	ACT04334	COX16	2.9979	2.77E-43
25 mM Fumaric acid	g4307	ACT21221	NSE1	2.9202	2.19E-83
25 mM Fumaric acid	g2521	ACT09831	RTG1	2.898	4.97E-177
25 mM Fumaric acid	g2989	ACT12888	SKI8	2.8409	8.11E-27
25 mM Fumaric acid	g3331	ACT15031	OCH1	2.8236	6.45E-68
25 mM Fumaric acid	g987	ACT27810	SDS23	2.7823	2.11E-52
25 mM Fumaric acid	g1362	ACT02359	YKT6	2.7424	9.25E-58
25 mM Fumaric acid	g2562	ACT10084	MEU1	2.7043	4.34E-35
25 mM Fumaric acid	g4513	DEL15196	MED1	2.6839	1.95E-32

25 mM Fumaric acid	g1522	ACT03398	MSY1	2.6518	4.27E-42
25 mM Fumaric acid	g2162	DEL05098	INP2	2.6218	3.23E-17
25 mM Fumaric acid	g1896	ACT05792	SIL1	2.6002	3.65E-84
25 mM Fumaric acid	g291	ACT12374	PRI1	2.564	1.33E-62
25 mM Fumaric acid	g1715	ACT04628	KLMA_20163	2.4962	5.00E-41
25 mM Fumaric acid	g624	ACT25500	JAC1	2.428	1.85E-62
25 mM Fumaric acid	g3434	ACT15698	MKS1	2.3908	2.56E-30
25 mM Fumaric acid	g916	ACT27343	TRM10	2.3586	4.65E-19
25 mM Fumaric acid	g1794	ACT05122	g1794	2.334	4.78E-33
25 mM Fumaric acid	g4379	ACT21672	KLMA_80156	2.2677	5.22E-26
25 mM Fumaric acid	g4788	ACT24165	GUD1	2.2552	4.16E-11
25 mM Fumaric acid	g1639	ACT04129	DAP2	2.246	1.76E-43
25 mM Fumaric acid	g3476	ACT15972	FSF1	2.1964	1.26E-33
25 mM Fumaric acid	g2753	ACT11340	SUR2	2.1823	3.96E-16
25 mM Fumaric acid	g680	ACT25875	DUG2	2.1479	1.63E-36
25 mM Fumaric acid	g3274	ACT14660	GCR2	2.0911	2.09E-15
25 mM Fumaric acid	g3585	ACT16696	VPS36	2.0559	8.25E-13
25 mM Fumaric acid	g138	DEL01669	TEL1	2.0487	0.0032932
25 mM Fumaric acid	g4078	DEL13339	HCS1	2.0468	3.13E-11
25 mM Fumaric acid	g45	ACT22460	PER33	1.8969	5.69E-18
25 mM Fumaric acid	g4568	DEL15433	RAD51	1.8957	5.64E-09
25 mM Fumaric acid	g1422	DEL01860	ATP22	1.8892	0.0075286
25 mM Fumaric acid	g583	ACT25227	CAX4	1.8493	6.35E-09
25 mM Fumaric acid	g3931	ACT18878	KLMA_60306	1.8193	2.19E-17
25 mM Fumaric acid	g3299	ACT14822	gcp	1.7935	2.74E-12
25 mM Fumaric acid	g423	ACT20770	IRE1	1.7135	2.69E-10
25 mM Fumaric acid	g3233	ACT14397	SAC6	1.7107	9.46E-04
25 mM Fumaric acid	g3110	ACT13601	STF2	1.6616	7.01E-07
25 mM Fumaric acid	g808	ACT26630	MGR1	1.6323	4.64E-09
25 mM Fumaric acid	g3149	ACT13847	fnx1	1.6301	1.70E-08
25 mM Fumaric acid	g1859	ACT05549	KLMA_20313	1.6207	1.29E-07
25 mM Fumaric acid	g4091	ACT19893	MNN2	1.617	6.59E-06
25 mM Fumaric acid	g263	ACT10523	MEP3	1.6084	4.94E-10
25 mM Fumaric acid	g461	ACT23182	CDC9	1.5807	0.0019247
25 mM Fumaric acid	g863	ACT26990	UFD1	1.5608	1.02E-06
25 mM Fumaric acid	g1308	ACT02009	RPC31	1.5414	7.66E-10
25 mM Fumaric acid	g1493	DEL02173	RAG2	1.5385	0.014018
25 mM Fumaric acid	g2957	ACT12684	NIT3	1.5231	3.16E-08
30 mM Fumaric acid	g4399	ACT21802	QDR2	10.831	0

30 mM Fumaric acid	g4399	ACT21806	QDR2	10.269	0
30 mM Fumaric acid	g3178	ACT14031	UBP14	6.0777	0
30 mM Fumaric acid	g3577	ACT16642	FAS1	5.3817	0
30 mM Fumaric acid	g592	ACT25286	IMD4	5.2305	0
30 mM Fumaric acid	g4399	ACT21801	QDR2	4.785	0
30 mM Fumaric acid	g2142	ACT07423	YPR1	4.0725	6.81E-203
30 mM Fumaric acid	g3020	ACT13106	CSE1	4.0243	0
30 mM Fumaric acid	g1501	ACT03264	LYS21	3.8844	0
30 mM Fumaric acid	g1640	ACT04144	RPN1	3.5368	6.15E-216
30 mM Fumaric acid	g1123	ACT00802	NNK1	3.4299	1.93E-76
30 mM Fumaric acid	g2114	ACT07232	KLMA_20588	3.3281	3.09E-289
30 mM Fumaric acid	g971	ACT27711	KLMA_10227	3.2804	1.79E-64
30 mM Fumaric acid	g4278	ACT21032	GPI12	3.2791	1.98E-156
30 mM Fumaric acid	g3308	ACT14880	PDR5	3.1379	2.77E-155
30 mM Fumaric acid	g694	ACT25965	LAS21	3.1375	2.57E-93
30 mM Fumaric acid	g878	ACT27085	GIS2	3.1304	3.95E-53
30 mM Fumaric acid	g2445	ACT09339	RRN3	3.078	6.37E-69
30 mM Fumaric acid	g902	ACT27254	ACO2a	2.9787	1.12E-39
30 mM Fumaric acid	g1538	ACT03500	PEX32	2.8542	1.09E-68
30 mM Fumaric acid	g1291	ACT01896	HGT1	2.854	3.95E-53
30 mM Fumaric acid	g292	ACT12438	KLMA_70317	2.8203	1.43E-40
30 mM Fumaric acid	g547	ACT24985	QDR3	2.8108	4.20E-226
30 mM Fumaric acid	g2227	ACT07985	VMA2	2.7874	4.00E-59
30 mM Fumaric acid	g4077	DEL13335	ATP7	2.7562	2.37E-14
30 mM Fumaric acid	g4406	ACT21855	ZUO1	2.6927	3.92E-61
30 mM Fumaric acid	g267	ACT10790	KLMA_70292	2.5663	2.43E-16
30 mM Fumaric acid	g3618	ACT16917	TOM1	2.5642	3.28E-33
30 mM Fumaric acid	g1345	ACT02252	KLMA_10605	2.4957	7.67E-49
30 mM Fumaric acid	g2525	ACT09854	ATP16	2.4669	4.43E-56
30 mM Fumaric acid	g987	ACT27810	SDS23	2.4595	8.75E-30
30 mM Fumaric acid	g3014	ACT13062	DUG1	2.4144	1.30E-37
30 mM Fumaric acid	g2043	ACT06768	PKP1	2.3343	6.77E-16
30 mM Fumaric acid	g2056	ACT06851	CAP2	2.2976	1.03E-16
30 mM Fumaric acid	g2189	ACT07724	ARG81	2.2743	3.25E-33
30 mM Fumaric acid	g1670	ACT04334	COX16	2.271	8.78E-12
30 mM Fumaric acid	g4307	ACT21221	NSE1	2.2326	3.23E-25
30 mM Fumaric acid	g659	ACT25730	GIP4	2.226	3.66E-30
30 mM Fumaric acid	g3813	ACT18126	g3813	2.1355	3.99E-24
30 mM Fumaric acid	g2521	ACT09831	RTG1	1.9937	2.21E-36

30 mM Fumaric acid	g1896	ACT05792	SIL1	1.9475	1.54E-25
30 mM Fumaric acid	g2989	ACT12888	SKI8	1.9466	1.28E-04
30 mM Fumaric acid	g624	ACT25500	JAC1	1.9314	1.22E-24
30 mM Fumaric acid	g680	ACT25875	DUG2	1.8156	9.72E-19
30 mM Fumaric acid	g3585	ACT16696	VPS36	1.8092	2.73E-07
30 mM Fumaric acid	g3700	ACT17407	VPS54	1.7763	2.21E-08
30 mM Fumaric acid	g407	ACT19751	PAC2	1.7106	4.10E-09
30 mM Fumaric acid	g3201	ACT14193	KLMA_40132	1.6967	1.80E-14
30 mM Fumaric acid	g291	ACT12374	PRI1	1.6857	2.05E-11
30 mM Fumaric acid	g3331	ACT15031	OCH1	1.6438	7.22E-07
30 mM Fumaric acid	g1390	DEL01718	SNQ2	1.5703	3.58E-04
30 mM Fumaric acid	g1522	ACT03398	MSY1	1.5204	3.40E-04
42 mM Fumaric acid	g4399	ACT21802	QDR2	10.972	0
42 mM Fumaric acid	g4399	ACT21806	QDR2	9.7333	0
42 mM Fumaric acid	g3178	ACT14031	UBP14	5.5647	0
42 mM Fumaric acid	g592	ACT25286	IMD4	4.9176	0
42 mM Fumaric acid	g3020	ACT13106	CSE1	4.5154	0
42 mM Fumaric acid	g2142	ACT07423	YPR1	3.7074	6.72E-107
42 mM Fumaric acid	g1501	ACT03264	LYS21	3.5609	2.30E-178
42 mM Fumaric acid	g3577	ACT16642	FAS1	3.5103	2.50E-187
42 mM Fumaric acid	g2114	ACT07232	KLMA_20588	3.0591	4.76E-166
42 mM Fumaric acid	g4399	ACT21801	QDR2	3.0521	2.82E-164
42 mM Fumaric acid	g694	ACT25965	LAS21	3.0118	1.39E-67
42 mM Fumaric acid	g2445	ACT09339	RRN3	2.9279	4.33E-48
42 mM Fumaric acid	g3308	ACT14880	PDR5	2.8817	2.36E-90
42 mM Fumaric acid	g1640	ACT04144	RPN1	2.8142	1.29E-60
42 mM Fumaric acid	g902	ACT27254	ACO2a	2.8104	7.51E-27
42 mM Fumaric acid	g4278	ACT21032	GPI12	2.6555	4.85E-50
42 mM Fumaric acid	g1123	ACT00802	NNK1	2.6533	4.36E-19
42 mM Fumaric acid	g878	ACT27085	GIS2	2.6403	8.58E-21
42 mM Fumaric acid	g2056	ACT06851	CAP2	2.1818	1.50E-11
42 mM Fumaric acid	g987	ACT27810	SDS23	2.0943	6.90E-13
42 mM Fumaric acid	g267	ACT10790	KLMA_70292	2.0179	1.35E-04
42 mM Fumaric acid	g971	ACT27711	KLMA_10227	2	2.33E-05
42 mM Fumaric acid	g3014	ACT13062	DUG1	1.938	5.92E-13
42 mM Fumaric acid	g1538	ACT03500	PEX32	1.8557	2.94E-09
42 mM Fumaric acid	g4406	ACT21855	ZUO1	1.8501	7.77E-11
42 mM Fumaric acid	g4077	DEL13335	ATP7	1.8171	0.0078687
42 mM Fumaric acid	g659	ACT25730	GIP4	1.7964	7.60E-11

42 mM Fumaric acid	g1291	ACT01896	HGT1	1.7945	7.03E-06
42 mM Fumaric acid	g3618	ACT16917	TOM1	1.7391	2.93E-05
42 mM Fumaric acid	g3813	ACT18126	g3813	1.7037	3.10E-08
42 mM Fumaric acid	g2525	ACT09854	ATP16	1.7002	9.27E-11
42 mM Fumaric acid	g624	ACT25500	JAC1	1.6952	1.06E-12
42 mM Fumaric acid	g292	ACT12438	KLMA_70317	1.6842	0.0012576
42 mM Fumaric acid	g2521	ACT09831	RTG1	1.6183	3.91E-14
42 mM Fumaric acid	g1670	ACT04334	COX16	1.5654	0.006548
42 mM Fumaric acid	g1345	ACT02252	KLMA_10605	1.5508	6.87E-06
42 mM Fumaric acid	g2227	ACT07985	VMA2	1.5303	6.18E-04
42 mM Fumaric acid	g4307	ACT21221	NSE1	1.5167	2.51E-04

4. Chapter 4: Metabolic engineering of primary metabolism for fumaric acid production

4.1. Abstract

Fumaric acid is a valuable C₄ dicarboxylic acid with broad industrial applications, yet achieving commercially viable microbial production remains challenging. In this study, we engineered the non-conventional yeast *K. marxianus* Y-1190 to enhance fumaric acid biosynthesis through a multi-pronged metabolic engineering approach. Building on previous work that identified acid tolerance improvements via *QDR2* and *QDR3* overexpression, we investigated transporter roles, heterologous pathway expression, and redox cofactor balancing. Deletion of the dicarboxylate importer *JEN2* confirmed its role in fumarate uptake, though it did not increase extracellular fumaric acid titers, indicating reuptake is not limiting under production conditions. Expression of the *R. oryzae* rTCA pathway, especially a modified fusion construct, doubled fumaric acid titers when combined with *QDR2* and *QDR3* transporter overexpression. To mitigate overflow metabolism byproducts, we overexpressed glutamate dehydrogenases from *S. cerevisiae*, which modestly improved fumarate yield and reduced glycerol and ethanol formation, while deletion of *PFK2* further decreased byproducts without boosting fumarate titers. Attempts to redirect pyruvate flux via phosphoenolpyruvate carboxykinase (PCK) overexpression were counterproductive, likely due to gluconeogenic flux. These results highlight the distinctive metabolic context of *K. marxianus*, including its respiratory metabolism, which may influence fumaric acid production differently than Crabtree-positive yeasts. Our findings emphasize the importance of integrating transporter function, pathway design, and cofactor balancing tailored to host physiology for successful organic acid biosynthesis.

4.2. Introduction

Fumaric acid is a high-value C₄ dicarboxylic acid with diverse industrial applications in food, pharmaceuticals and bioplastics, and its microbial production is a target for sustainable biomanufacturing²³⁶. Commercial scale production for bio-based chemicals such as fumaric acid generally requires titers of 50-100 g L⁻¹ to be economically viable^{237,238}. While this benchmark has not been reached, substantial progress has been made using recombinant model microorganisms, with titers reaching 41.5 g L⁻¹ and 33.13 g L⁻¹ in *E. coli*, and *S. cerevisiae*, respectively^{30,35}. However, model organisms often face limitations in acid tolerance, substrate range, or process conditions, prompting the exploration of non-conventional yeasts as alternative hosts. *K. marxianus* offers several advantages as a production host, including thermotolerance, rapid growth, and native lactose metabolism⁶⁷. This allows for the use of inexpensive, dairy waste-derived feedstocks such as milk permeate, which can significantly reduce production costs and improve process sustainability. Together, these traits position *K. marxianus* as a promising chassis for organic acid biosynthesis.

Initial work in this thesis focused on enabling fumaric acid accumulation in *K. marxianus* by deleting *FUM1*, thereby blocking the conversion of fumarate to malate in the TCA cycle. While this intervention successfully increased fumaric acid titers, it also sensitized cells to acid stress (Fig. 3.1). To overcome this limitation, a genome-wide CRISPR-AID screen identified three genetic perturbations that improved acid tolerance to 42 mM fumaric acid: activation of *QDR2* and *QDR3*, and deletion of *ATP7*. Among these, co-overexpression of *QDR2* and *QDR3* increased fumaric acid yield by 7.2-fold, reaching 0.89 g L⁻¹ in shake flasks (Fig. 3.5, Fig. S3.3) and 2.16 g L⁻¹ in fed-batch bioreactors (Fig. 3.6). With an acid-tolerant strain in hand, we next sought to rationally enhance fumaric acid titers through pathway engineering.

As discussed in Section 1.2, the rTCA has a higher theoretical yield than the oxidative pathway due to the net fixation of CO₂^[25] which previous studies have leveraged by using the *R. oryzae* rTCA pathway to achieve high fumaric acid titers in yeast, with up to 33.13 g L⁻¹ and 21.6 g L⁻¹ achieved in *S. cerevisiae* and *C. glabrata*, respectively^{30,44}. This approach typically involves heterologous expression of *RoPYC*, *RoMDH*, and *RoFUMI* coupled with a C₄ dicarboxylic acid transporter from *Schizosacchromyces pombe*, *SpMAEI* to facilitate fumaric acid transport (Fig. 4.1)^{25,30,44,239}.

Prior studies in both *S. cerevisiae* and *P. pastoris* have addressed this issue through cofactor pathway engineering. For example, in *S. cerevisiae*, overexpression of *GDH1* and *GDH2* enhanced NADH cycling, while deletion of *PFK2* and *PGII* restructured the pentose phosphate pathway (PPP) into a linear pathway, improving succinic acid production (Fig. 4.2)²⁴². Similarly, in *P. pastoris*, overexpression of four PPP enzymes alongside *PDC1* deletion increased NADPH availability and enhance production of human superoxide dismutase²⁴³. These approaches illustrate how metabolic rewiring of redox and cofactor systems can mitigate overflow metabolism and improve bioproduction efficiency.

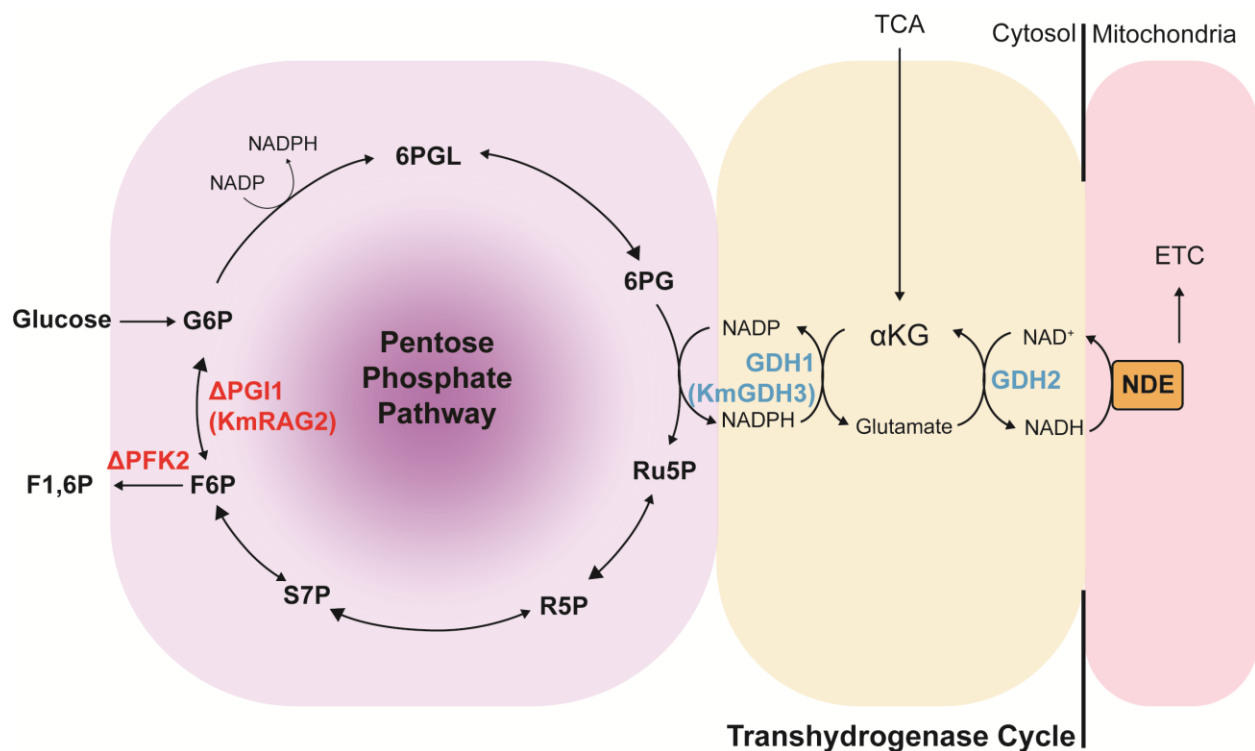


Figure 4.2 PPP engineering strategy for enhancing NADH pools. Genes represented in teal are overexpressed, while those represented in red are deleted. Adapted from Yu *et al.*, *Nat Metab*, 2022, <https://doi.org/10.1038/s42255-022-00654-1>, licensed under a [Creative Commons Attribution 4.0 International License](#).

In this chapter, I apply a set of rational metabolic engineering strategies to enhance fumaric acid production in the dairy isolate *K. marxianus* Y-1190. First, I investigate the role of acid transporters *JEN1* and *JEN2* in fumarate import. I then evaluate overexpression of the heterologous *R. oryzae* rTCA pathway for pathway amplification. To address high acetic acid accumulation, I implement NADH regeneration strategies using transhydrogenase cycling enzymes from both *K. marxianus* and *S. cerevisiae*. Finally, I attempt to rewire glycolytic flux directly toward oxaloacetate via overexpression of Phosphoenolpyruvate (PEP) carboxykinase (PCK), aiming to minimize overflow to acetate and ethanol (Fig. 4.1). Together, these efforts reflect a multifaceted approach to pathway optimization grounded in metabolic logic.

4.3. Materials and methods

4.3.1. Strains, media, and plasmids

E. coli strain DH5 α , used to maintain and amplify plasmids, was cultured in Luria broth at 37 °C containing the appropriate antibiotics. *K. marxianus* strain Y-1190 was used as the chassis for all genome engineering in this work, all strains can be found in Supplementary Table S4.1. *K. marxianus* was cultivated and maintained in YPD (10 g L⁻¹ Bacto Yeast Extract, 20 g L⁻¹ Bacto peptone, 20 g L⁻¹ glucose). YNB-FA was used to test the import of fumaric acid as the sole carbon source and contains 6.7 g L⁻¹ yeast nitrogen base (Sigma-Aldrich, Y0626), and 2 g L⁻¹ fumaric acid (Sigma-Aldrich, W248800). Ultrafiltered milk permeate was sourced directly from Agropur Cooperative (Québec, Canada) and supplemented with the following: 8.36 g L⁻¹ urea, 2.07 g L⁻¹ KH₂PO₄, 1.15 g L⁻¹ MgSO₄·7H₂O, 1 g L⁻¹ NiSO₄·6H₂O, 5 mL L⁻¹ vitamin stock, and 5 mL L⁻¹ microelements. The composition of the vitamin and microelement stock was as described previously¹³¹.

Plasmids in this study were made using Golden Gate cloning¹³⁷ using parts from the Yeast Toolkit¹²⁸, a biological parts kit for *K. marxianus*¹⁰¹, a promoter library for *K. marxianus* (Chapter 2)¹⁷⁹, or amplified in-house. All vectors were confirmed by Nanopore sequencing (Plasmidsaurus). Plasmids and oligonucleotides used in this chapter can be found in Supplementary Tables S4.2, and S4.3, respectively.

Over the course of this study, we use plasmids that have antibiotic resistance cassettes for hygromycin, G418, and nourseothricin; when growing transformed strains, we use those antibiotics at concentrations of 200 $\mu\text{g mL}^{-1}$, 200 $\mu\text{g mL}^{-1}$, and 100 $\mu\text{g mL}^{-1}$, respectively.

4.3.2. *Heterologous genes*

Genes for the rTCA pathway in *R. oryzae*, the *MAE1* transporter from *S. pombe*, and the *MDH2* from *C. glutamicum* were ordered from Twist Biosciences. Due to the lack of available species-specific tools, the genes were codon optimized to the highly related yeast *K. lactis* via the Integrated DNA Technologies codon optimization tool. Due to its size, *RoPYC* had to be split into two constructs, with BsmBI sites added for easy reassembly during the entry vector cloning into pYTK001. Additionally, the SKL peroxisome tripeptide coding sequence was removed from *RoMDH* to allow cytoplasmic localization²⁴⁴. *S. cerevisiae* genes *ScGDH1*, *ScGDH2*, and *ScPCK* were amplified in-house from CEN.PK 113-13D.

4.3.3. *K. marxianus* transformation

When transforming *K. marxianus* with plasmids, we follow a modified lithium acetate-based protocol designed for *S. cerevisiae*¹³². Briefly, 500 μL of log-phase culture (OD_{600} of 0.8-1.2) was harvested and transformed with 500 ng of plasmid DNA using 150 μL of a lithium acetate solution following the concentrations defined by Gietz *et al.* The transformations were incubated at room temperature for 15 min, and then 47 °C for 15 min. The cells were pelleted by

centrifugation at 1860 g for 3 min, the supernatant was replaced with 500 µl of YPD, and the cells were recovered overnight at 30 °C stationary. Twenty percent of the culture was plated on selection plates and incubated at 30 °C until the presence of colonies were observed (~2 days).

CRISPR-Cas9 deletions were performed using MTKM01 or MTKM03 which are SpCas9 plasmids, with different selection cassettes, modified from the Mylo Toolkit for use in *K. marxianus*^{139,179}. For each deletion, two separate sgRNA were designed per target and cloned into the SpCas9 plasmid using Golden Gate. This plasmid was then transformed into *K. marxianus* according to a high efficiency transformation protocol for *K. marxianus* with modifications⁹⁴. In brief, 1 mL of *K. marxianus* cells in exponential phase were centrifuged at 1860 g for 3 min. After washing once with 1 mL of sterile H₂O, the cells were suspended in 10 µL of 10 mg mL⁻¹ salmon sperm, and 1-1.5 µg of sgRNA plasmid was added. To the cells and DNA, 400 µl of transformation buffer (40 % polyethylene glycol 3350, 0.1 M lithium acetate, 10 mM Tris-HCl (pH 7.5), 1 mM EDTA, and 10 mM DTT) was added to a 1.5 mL microfuge tube and the solution was mixed by agitation. The transformation mix was then incubated at room temperature for 15 min followed by heat shock at 47 °C for 15 min. The cells were then centrifuged at 1860 g for 3 min and suspended in 400 µl of YPD to recover overnight at 30 °C without shaking. The next day the cells were plated on YPD agar with appropriate selection. If necessary, an outgrowth step was done where 50 µl of recovered transformation was grown in 500 µl of YPD with selection for 2-3 days. This outgrowth was diluted 1:1000-10000 before plating on YPD agar with selection.

4.3.4. Growth assays

Three independent colonies of each strain were inoculated into 96-well deep plate containing 700 µl of YPD. These strains were grown overnight to saturation and then washed twice with sterile H₂O before being reconstituted to an OD₆₀₀ of 2 in YNB-FA as a parental plate. A final

microtiter plate was seeded with 100 μ l of YNB-FA with the strains at a final OD₆₀₀ of 0.1. Cultures were incubated at 30 °C in clear, flat-bottom 96-well microtiter plates and wrapped with Parafilm to minimize evaporation. Absorbance readings were taken at 595 nm with a Sunrise absorbance microplate reader (Tecan) every 20 min over the course of up to 5 days. Maximum specific growth rates ($\mu_{\max}$) were calculated in R using the growthrates package¹³⁰.

4.3.5. Metabolite analysis

To assess metabolite production, strains of interest were grown in 5 mL of 0.5x supplemented permeate overnight to reach saturation. Following overnight growth, cultures were diluted to an OD₆₀₀ of 0.2 in 5 mL in a 50 mL glass culture tube. Cultures were incubated at 30 °C, and shaken at 200 rpm for 48-72 h. For sampling, 100 μ l of the culture supernatant was harvested by centrifugation at maximum speed (21000 g) and frozen at -20 °C. For HPLC analysis, supernatant samples were diluted 1:10 in distilled H₂O. Metabolites were measured using an Agilent 1290 Infinity II HPLC equipped with a Aminex HPX-87H column (Bio-Rad). Metabolites were resolved isocratically in 10 mM sulphuric acid at 65 °C at a flow rate of 0.6 ml min⁻¹. A UV-visible detector was used for the quantification of fumaric acid (260 nm) and acetic acid (210 nm), and a refractive index detector was used to quantify lactose, ethanol and glycerol. Analysis was done on Agilent OpenLab software. All measurements were compared to a linear standard curve and done in duplicate unless otherwise stated.

4.3.6. RT-qPCR

RNA from *K. marxianus* cells was extracted using the Direct-zol RNA miniprep kit (ZymoResearch) per manufacturer's instructions. Briefly, 3x10⁷ cells were harvested by centrifugation in 2 mL nuclease-free microfuge tubes. The supernatant was removed, and 1 mL of TRI Reagent Solution (AM9738, Invitrogen) was added along ~200 μ l glass beads (company).

The pellet was vortexed for 40 minutes on a vortex genie at maximum speed. The samples were then centrifuged for 2 min at 21000 x g and the supernatant was moved to a new nuclease free microcentrifuge tube. The samples were then processed as directed by the Direct-zol miniprep kit. After extraction, RNA was measured using a spectrophotometer (Infinite200, Tecan) so that the same amount of RNA was used for cDNA preparation (250 ng). The M-MuLV Reverse Transcriptase kit (NEB) was used to generate cDNA which was diluted 1:2.5 for the RT-qPCR reactions. All RT-qPCRs were done with Forget-Me-Not™ EvaGreen qPCR Master Mix (Biotium) and analyzed by a ViiA 7 Real-Time PCR System (Thermofisher). Primers were designed on Primer3^{245–247} with the following parameters: Optimal Tm of 60 °C, product size of 60-200 bp, and using the SantaLucia formula. These primers can be found in supplementary Table S4.3.

For analysis, we used the comparative Ct method ($-\Delta\Delta C_t$) to quantify relative gene expression levels²⁴⁸. Expression of the gene of interest was normalized to the reference gene *TAF10* to control for variation in cDNA input²⁴⁹. The experiment was performed using technical triplicates for each of three biological replicates, and fold changes were calculated relative to the control condition to assess differential expression.

4.4. Results

4.4.1. Deletions of putative carboxylic acid importers increase fumaric acid titers

To investigate the role of dicarboxylic acid transport in fumaric acid production, we focused on the *JEN* family of carboxylic acid transporters. *S. cerevisiae* encodes *JEN1*, a monocarboxylic acid transporter implicated in the uptake of organic acids such as pyruvate, lactate, and propionate²⁵⁰. Deletion of *JEN1* has been shown to improve lactic acid titers in *S. cerevisiae*²⁵¹, suggesting that preventing fumaric acid re-import might similarly enhance its production.

Kluyveromyces lactis, a close relative of *K. marxianus*, encodes both *JEN1* and *JEN2*. While *KIJEN1* is a functional homolog of *S. cerevisiae*, *KIJEN2* functions as a dicarboxylic acid importer for succinic and malic acids^{252,253}. In the genome of *K. marxianus* Y-1190, we identified homologs of both transporters g1806 (putative *JEN2*) and g4154 (putative *JEN1*). We aimed to characterize their roles in fumaric acid transport and evaluate their relevance to fumaric acid production.

To assess the involvement of JEN transporters in fumaric acid import, we deleted *JEN2* (MTK075) and *JEN1* (MTK076) individually in the wild-type Y-1190 background. These mutants were grown in minimal media containing 2 g L⁻¹ (~17 mM) fumaric acid as the sole carbon source (YNB-FA). As a negative control, we included a $\Delta FUM1$ mutant (ROK002), which cannot metabolize fumaric acid. While cultures of both MTK075 and MTK076 ultimately reached wild-type-level final OD₆₀₀ values, MTK075 exhibited a slower growth rate, with a maximum specific growth rate of 0.10 h⁻¹ compared to 0.20 h⁻¹ in the wild type and MTK076 (Fig. 4.3 AB). These results suggest that *JEN2* is not the sole mediator of fumaric acid uptake, but it contributes to the process.

Given this role for *JEN2* in fumaric acid import, we next investigated whether its deletion could improve fumaric acid production by limiting uptake. We deleted *JEN2* in the $\Delta FUM1$ production background and transformed that strain with the *QDR2* overexpression construct from Chapter 3 to assess whether enhanced export combined with reduced import could further boost titers. These strains were cultured in 0.5x supplemented milk permeate for 48 hours. Regardless of *QDR2* overexpression, deletion of *JEN2* did not impact fumaric acid titers or specific titers (Fig. 4.3 CD) defined as titer normalized to cell density (g L⁻¹ OD₆₀₀⁻¹). In the *QDR2*-expressing strains, titers ranged from 0.72–0.78 g L⁻¹ (Fig. 4.3 C), approximately 2.7-fold lower than the

concentration used in the fumaric acid import assay (Fig. 4.3 AB). It's possible that *JEN2*-mediated fumaric acid import may only be relevant at elevated concentrations of fumaric acid or under specific pH. While *JEN2* deletion had no effect under the current production conditions, its role may warrant re-evaluation if higher titers are achieved in future strains.

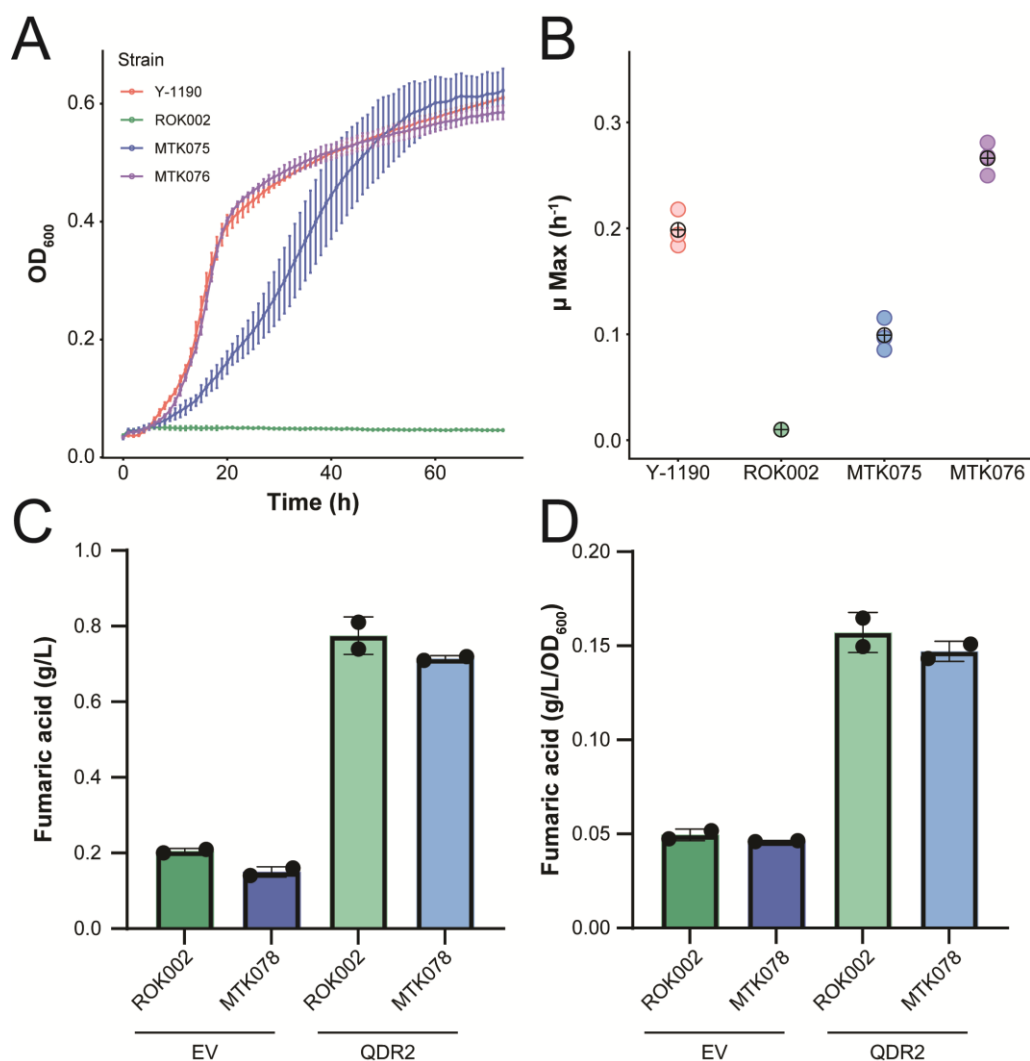


Figure 4.3 Elucidation of putative carboxylic transporters in fumaric acid import and production. **A.** Cell density and **B.** Growth rates of *K. marxianus* Y-1190 wildtype, ROK002 ($\Delta FUM1$), MTK075 ($\Delta JEN1$) and MTK076 ($\Delta JEN2$) mutants when grown on fumaric acid as the sole carbon source. Fumaric acid **C.** titers and **D.** specific titers when *JEN1* deletion is combined with $\Delta FUM1$ (MTK078) and *QDR2* is overexpressed on plasmid. Growth rate experiments were done in triplicate, with the trend lines following the mean with error bars representing standard deviation. For the fumaric acid production experiments, all experiments were done in duplicate, and the error bars represent standard deviation.

4.4.2. Engineering the TCA cycle for enhanced fumaric acid production

To enhance fumaric acid production, we explored the use of the *R. oryzae* reductive TCA pathway which diverts carbon from the oxidative TCA cycle via pyruvate (Fig. 4.1). This was achieved by recombinant expression of *RoPYC*, *RoMDH*, and *RoFUMI*, as described in previous studies^{30,38,44}. In parallel, we tested C₄ dicarboxylic acid transporter *SpMAEI* and a malate dehydrogenase from *Corynebacterium glutamicum* (*CgMDH*), based on the prior success in fumaric acid production and exhibiting low substrate inhibition, respectively^{30,254}. These constructs (Fig. 4.4 A) were transformed into Y-1190 Δ *FUMI* and tested in 0.5x supplemented milk permeate medium. At 48 hr, the rTCA2 construct containing the *R. oryzae* heterologous pathway and the *SpMAEI* transporter increased fumaric acid titers from 0.21 g L⁻¹ in the empty vector control to 0.32 g L⁻¹ (Fig. 4.4 B). The other two constructs showed modest improvements over the control at 24 hr, but by 48 hr their titers were comparable to the control. This trend was also observed in specific titers. At 24 hr, rTCA2 achieved the highest titer of 0.16 g L⁻¹ OD₆₀₀⁻¹, compared to 0.13 g L⁻¹ OD₆₀₀⁻¹ for the control, but by 48 hr, all strains converged to similar titers (Fig. 4.4C).

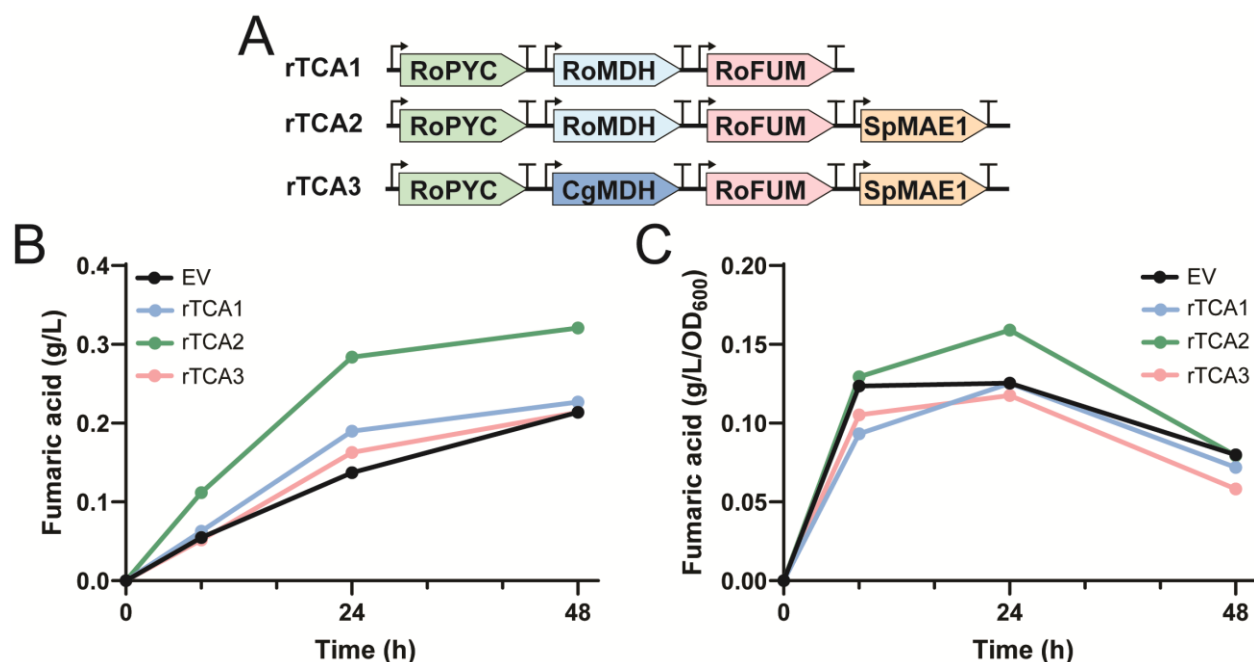


Figure 4.4 *R. oryzae* rTCA expression on fumaric acid production. **A.** Schematic of the three constructs used in this experiment. Fumaric acid **B.** titers and **C.** specific titers. All samples were taken in single replicate.

The expression of the *R. oryzae* heterologous rTCA pathway resulted in a modest improvement in fumaric acid titers, though levels were lower than anticipated based on previous studies. To further enhance flux through this pathway, we introduced a P160A point mutation in *RoFUM1*, previously shown to increase its affinity for malate over fumarate and thereby promoting its conversion to fumaric acid³⁰. In addition, we fused *RoFUM1*^{P160A} to *RoMDH*, a strategy previously reported to improve fumaric acid titers likely due to substrate channeling via enzyme colocalization²⁵⁵. The resulting modified rTCA^{mod} construct and an empty vector control was transformed into the *QDR2* overexpression background described in Chapter 3. Given the significant increase in fumaric acid production observed with *QDR2* overexpression (Fig. 3.5 AB), we chose to proceed with this transporter rather than *SpMAE1*, which had only a modest effect on

fumaric acid accumulation (Fig. 4.4). The resulting strains were cultured in 0.5× supplemented milk permeate medium for 72 hr.

Fumaric acid titers were not significantly different between the rTCA^{mod} strain and the EV control for most of the time course (Fig. 4.5 AB). However, a modest but statistically significant increase in yield was observed at 72 hr when the rTCA^{mod} strain reached a specific titer of 0.23 g L⁻¹ OD₆₀₀⁻¹ compared to 0.20 g L⁻¹ OD₆₀₀⁻¹ in the control (Fig. 4.5 B). This difference in titer is attributable to reduced biomass accumulation in the rTCA^{mod} strain at 72 hr, rather than an increase in fumaric acid titer, which remained comparable across both strains (Fig. 4.5 A).

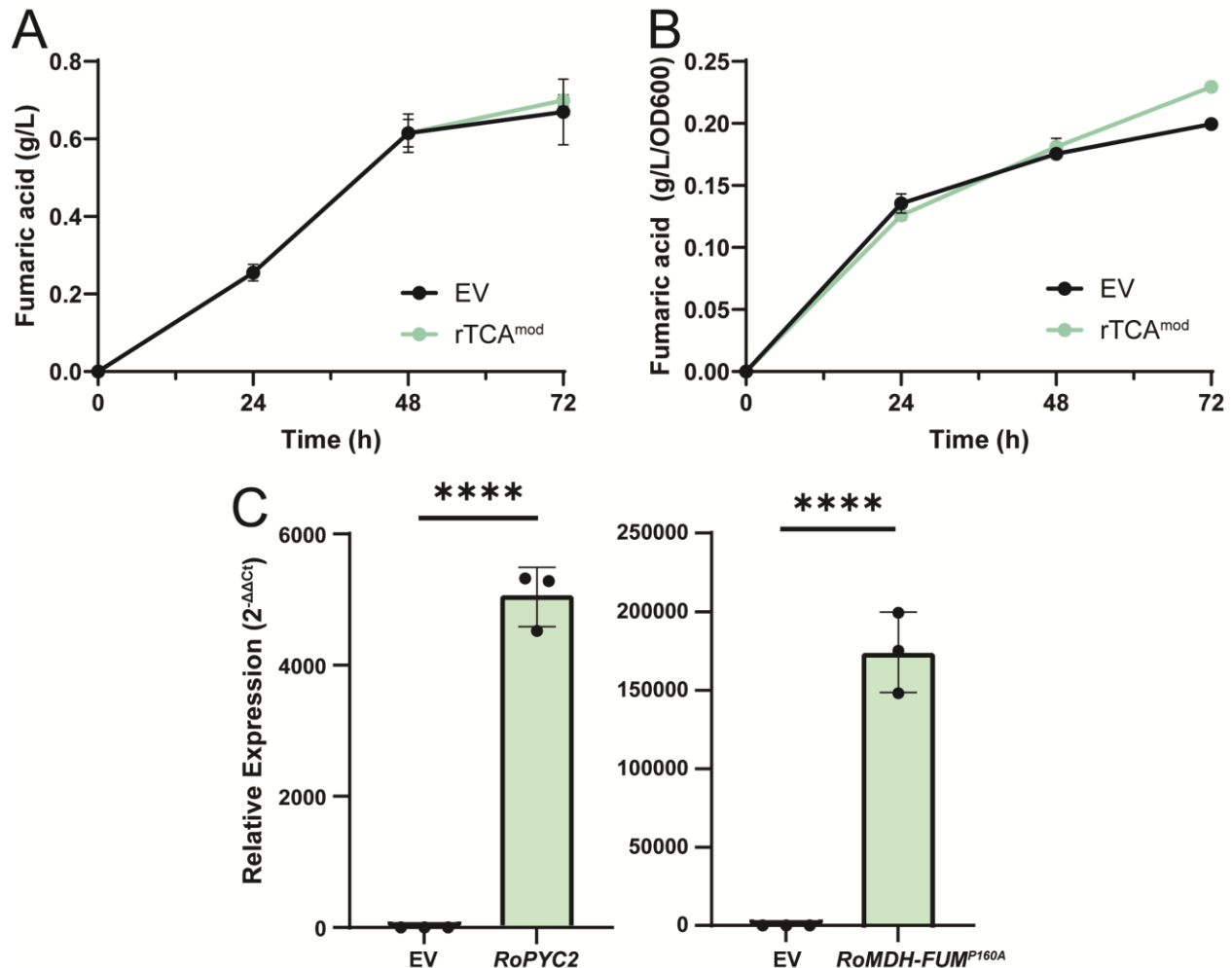


Figure 4.5 Effect of modified rTCA on fumaric acid production in *QDR2* overexpression strain. Fumaric acid titers **A.** and specific titers **B.** grown in 0.5x supplemented permeate and cultured for 72 hr. Each data point represents the mean of duplicate samples, and the error bars represent the standard deviation. **C.** Relative gene expression of *RoPYC2* and *RoMDH-FUM^{P160A}* cultured in 0.5x supplemented permeate for 24 hr as measured by RT-qPCR. The bars represent the mean of three independent samples and the error bars represent standard deviation.

Due to similar metabolite and growth profiles of the rTCA^{mod} construct to the EV control, we performed RT-qPCR analysis to ensure expression of our genes. The growth parameters for the metabolite experiment were replicated, and samples were harvested at 24 hr. RNA was extracted from these samples, followed by cDNA synthesis and qPCR. For the *R. oryzae* enzymes, we designed primers that targeted the *RoMDH-FUM^{P160A}* fusion, and another pair for the *RoPYC* gene, since both genes were being expressed from one construct. We detected strong expression of both genes, with *RoMDH-FUM^{P160A}* and *RoPYC* exhibiting 1.7×10^6 - and 5.0×10^4 -fold higher

expression, respectively, compared to the empty vector control (Fig. 4.5 C). While these values are arbitrary due to the absence of these exogenous genes in the control strain, they confirm robust transcription of the heterologous pathway.

With confirmation of the *R. oryzae* rTCA^{mod} pathway expression, we introduced this construct into our highest producing fumaric acid strain, MTK130hom, a *FUMI* deletion mutant overexpressing both *QDR2* and *QDR3*, characterized in Chapter 3 (Fig. 3.5). MTK170 (MTK130hom with an EV plasmid) and MTK172 (MTK30hom expressing the rTCA^{mod} plasmid) were cultured in 0.5x supplemented permeate and sampled every 24 hr for a total of 72 hr for metabolite analysis. At 72 hr, the strain expressing the *R. oryzae* rTCA^{mod} pathway produced 2-fold higher titers of fumaric acid while maintaining similar specific titers (Fig. 4.6). The specific titers remained the same because the growth of MTK172 was significantly higher than the empty vector control (Fig. 4.6). Analysis of by-products revealed comparable acetic acid levels between the strains, whereas ethanol and glycerol accumulation were significantly higher in MTK172 (Fig. 4.6).

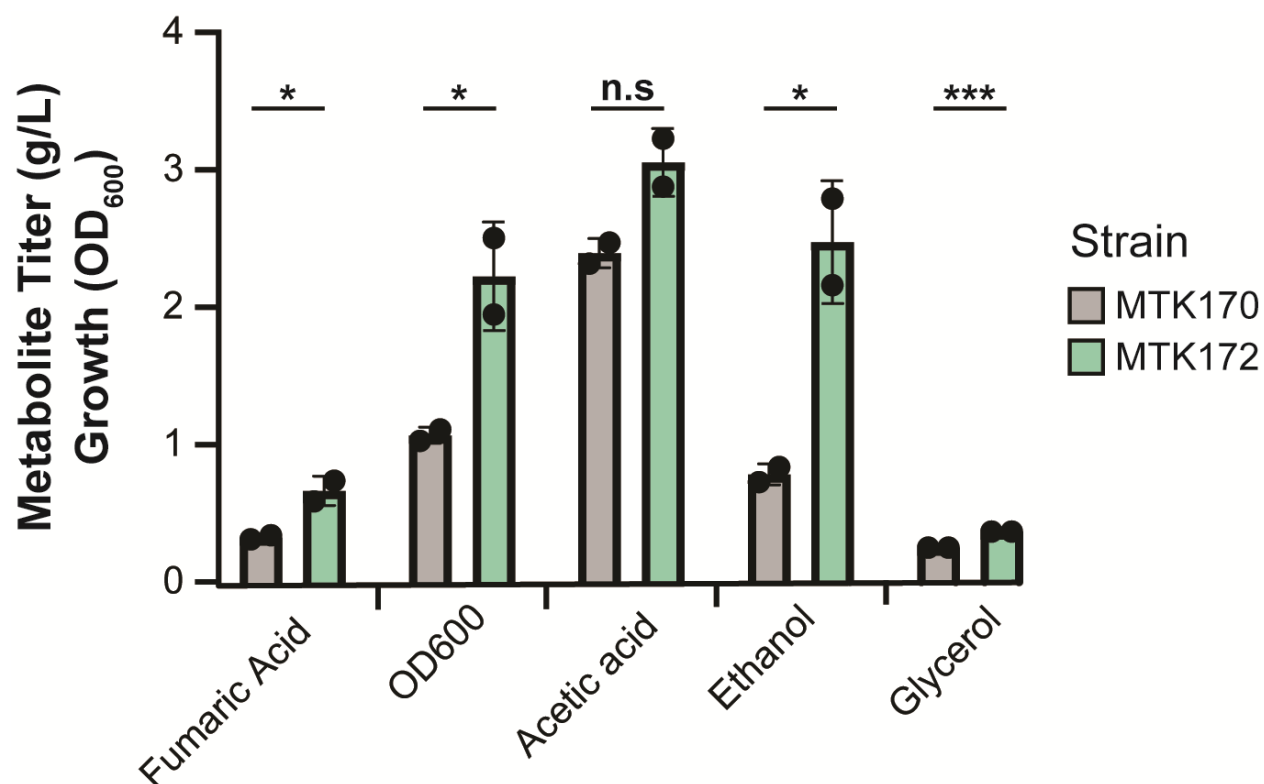


Figure 4.6 Metabolite and growth profiles for rTCA^{mod} in MTK130hom background. Strains were grown in 0.5x supplemented permeate in 50 mL culture tubes for 72 hr. The strains used are *K. marxianus* Y-1190 $\Delta FUM1$, I2::*QDR2*, I3::*QDR3* (hom) transformed with an EV plasmid (MTK170) or the rTCA^{mod} plasmid (MTK172). Each bar represents the mean of duplicate samples, and error bars represent the standard deviation. The asterisks denote statistical significance as defined by a p-value equal to or less than 0.05 as calculated by a two-tailed T-test.

4.4.3. Redox balancing fumaric acid production strains to optimize titers

High concentrations of acetic acid were previously observed in *QDR2* (MTK117) and *QDR2* and *QDR3* overexpression (MTK130hom) backgrounds (Fig. S3.4). Acetic acid may reduce fumaric acid by acting as a carbon sink and increasing acid stress to the cell. Its accumulation is associated with insufficient availability of NADH or NADPH in the cell^{241,256}. This redox imbalance is likely present in our production strains as *FUM1* deletion disrupts the TCA cycle impairing the reduction of NAD⁺ to NADH required for ATP generation and as a co-factor in many metabolic reactions, including the reductive branch of the TCA (Fig. 4.1).

To address this, we implemented the transhydrogenase cycle described in previous work, which increases NADH levels through the cycling of glutamate and α -ketoglutarate via *GDH1* and *GDH2*²⁴². In *S. cerevisiae*, *GDH1* and *GDH3* are paralogs that arose from a whole genome duplication event that occurred after the divergence of *K. marxianus*^{257,258}. As *K. marxianus* contains only a *GDH3* homolog, we amplified *KmGDH3* and *KmGDH2* from the genome and placed them under strong constitutive promoters. This construct was transformed into the *QDR2* overexpression strain and grown in 0.5x supplemented permeate for 72 hr, with sampling every 24 hr for metabolite analysis.

Overexpression of *KmGDH3* and *KmGDH2* did not improve fumaric acid titers or yields, nor did it reduce acetic acid titers or yields (Fig. 4.7 AB). At 72hr, fumaric acid titers for the EV control and *KmGDH* overexpression strain were 1.01 g L⁻¹ and 1.02 g L⁻¹ respectively, while specific titers were 0.23 g L⁻¹ OD₆₀₀⁻¹ and 0.21 g L⁻¹ OD₆₀₀⁻¹ (Fig. 4.7 A). Acetic acid levels had increased to 9.33 g L⁻¹ and 9.48 g L⁻¹ at 72 hr for the EV and *KmGDH*, respectively (Fig. 4.7 B).

To confirm construct expression, we repeated the growth experiment and harvested samples at 24 hr for RNA extraction and RT-qPCR. Gene expression analysis showed that *GDH3*

and *GDH2* were upregulated by 9.8-fold and 3.2-fold, respectively, compared to the empty vector control (Fig. 4.7 C). These results confirm successful overexpression but suggest that the *K. marxianus* system may behave differently from *S. cerevisiae* in the context of endogenous GDH enzyme function. *KmGDH3* and *KmGDH2* may be less effective at supporting NADH regeneration compared to their *S. cerevisiae* counterparts, or the broader metabolic context in *K. marxianus* may limit their impact.

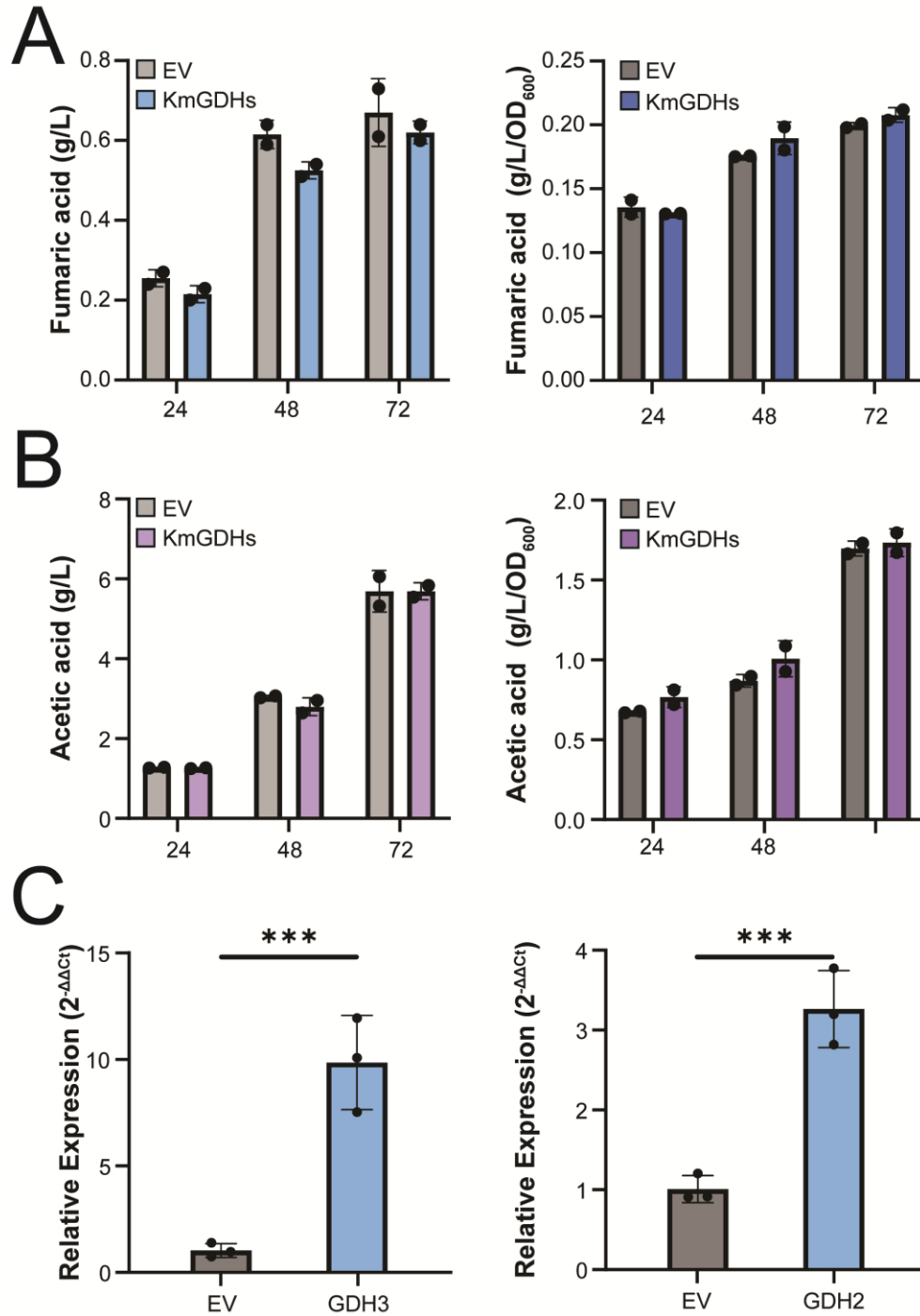


Figure 4.7 Effect of NADH recycling on fumaric acid production. A. Fumaric acid titers, **B.** Acetic acid titers. Bars represent the mean of duplicate experiments and error bars represent the standard deviation. **C.** Relative gene expression of *KmGDH3* and *KmGDH2* cultured in 0.5x supplemented permeate for 24 hr as measured by RT-qPCR. The bars represent the mean of cultures from three independent colonies and the error bars represent standard deviation.

To investigate whether *ScGDH1* and *ScGDH2* expression could enhance NADH availability in *K. marxianus*, we amplified the genes from *S. cerevisiae* strain CEN.PK113-7D and placed them under strong *K. marxianus* promoters on the same construct. We transformed MTK117 with an EV plasmid (MTK137) or the *ScGDH1* & 2 expressing plasmid (MTK160) and cultured them for 48 hr to assess fumaric acid and acetic acid production. At both 24 and 48 hr, fumaric acid titers were significantly higher in MTK160 compared to the empty vector control (Fig. 4.8 A). At 48 hr, fumaric acid reached 0.63 g L⁻¹ in MTK160 versus 0.53 g L⁻¹ in the control, while specific titers remained unchanged between strains (Fig. 4.8 A). Although acetic acid titers trended higher in MTK160 at 48 hours, the differences were not statistically significant (Fig. 4.8 B).

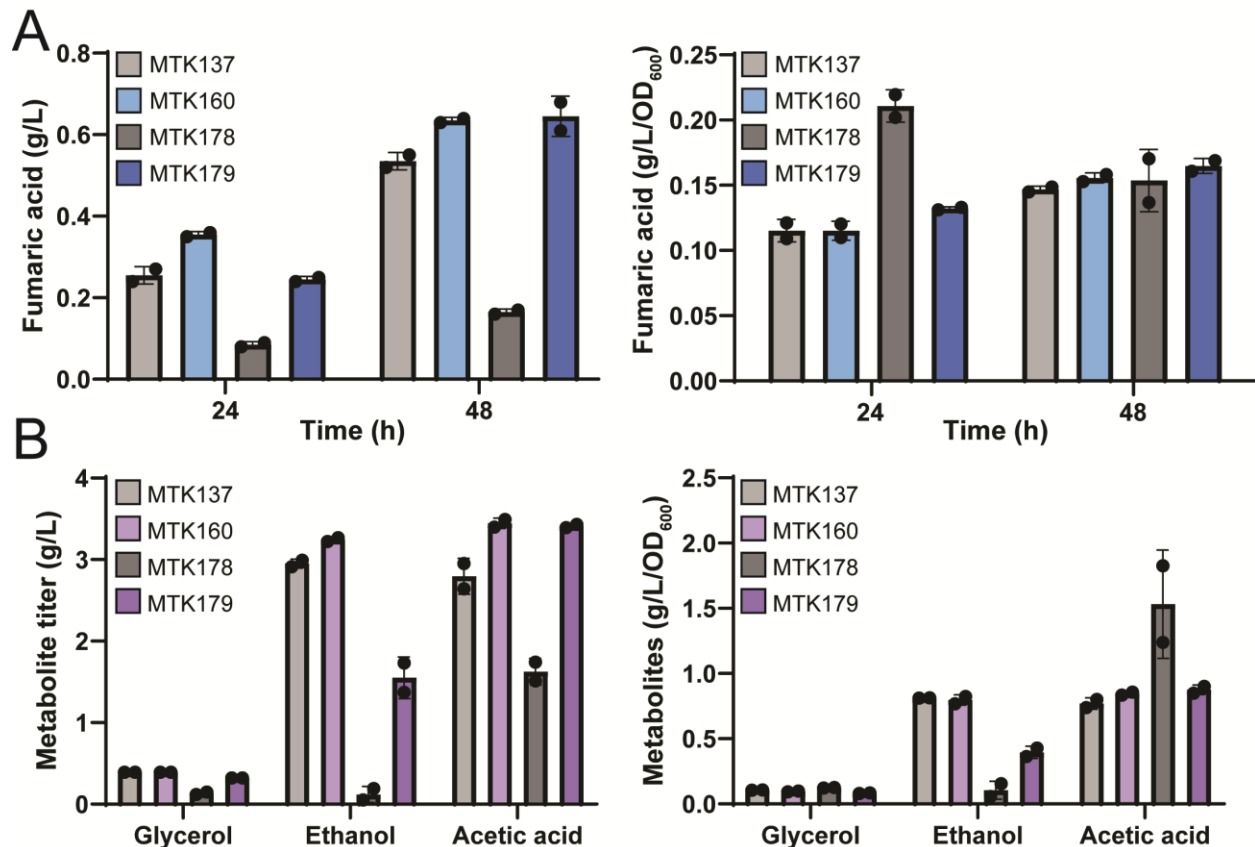


Figure 4.8 Exploring *ScGDH* enzyme expression in enhancing fumaric acid and reducing by-products. A. Titters of fumaric acid B. Titters of metabolite by-products, glycerol, ethanol and acetic acid. Strains used are *K. marxianus* Y-1190 $\Delta FUM1$, I2::*QDR2* transformed with an EV plasmid (MTK137) or with an *ScGDH1* and *ScGDH2* co-expression plasmid (MTK160) and *K. marxianus* Y-1190 $\Delta FUM1\Delta PFK2$, I2::*QDR2* transformed with an EV plasmid (MTK178) or with an *ScGDH1* and *ScGDH2* co-expression plasmid (MTK179). Each bar represents the mean of cultures from two independent colonies, and the error bars represent standard deviation.

The *S. cerevisiae* transhydrogenase cycle was only the first step in enhancing NADH availability in the previous study²⁴². To further increase cofactor abundance, we deleted *PFK2* from MTK117 (creating MTK159) to disrupt the reversible nature of the PPP and force flux towards the transhydrogenase cycle (Fig. 4.2). MTK159 was transformed with an EV (MTK178) or *ScGDH1* and *ScGDH2* (MTK179) expressing construct and strains were grown in 0.5 x permeate media for 48 hr before metabolites analysis. Deletion of *PFK2* significantly decreased fumaric acid titers to 0.16 g L⁻¹ in MTK178, compared to 0.53 g L⁻¹ produced by MTK137 without the deletion (Fig. 4.8 A). However, co-expression of *ScGDH1* and *ScGDH2* in the *PFK2* deletion

mutant recovered the fumaric acid titers to 0.64 g L⁻¹, similar to titers in the MTK160 strain that expresses the same genes (0.63 g L⁻¹). While the *PFK2* deletion seemed to have no effect on fumaric acid production in the presence of *ScGDH1* and *ScGDH2*, it significantly reduced ethanol and glycerol accumulation, but acetic acid titers remained unchanged from MTK160 (Fig. 4.8 B).

Attempts were also made to generate a double $\Delta FUM1\Delta PDC1$ mutant to block carbon flux toward acetate and ethanol, as previously demonstrated in the NADH regenerating strategies for *P. pastoris* and *S. cerevisiae*^{242,243}. However, only a single colony emerged three weeks after transformation. Although genotyping confirmed the correct mutation, its severely impaired growth prevented the establishment of a viable glycerol stock or further experimental use.

4.4.4. Reducing carbon loss by redirecting flux around the pyruvate node

With limited success in regenerating NADH to reduce acetic acid production, we explored alternative strategies to minimize carbon loss through overflow metabolism. Pyruvate sits at the intersection of the TCA cycle, rTCA pathway, and ethanol and acetate production, we therefore aimed to bypass this metabolic node by directing carbon from glycolysis directly to oxaloacetate (Fig. 4.1). To achieve this, we investigated overexpressing *PCK*, which encodes PEPCK, an enzyme that catalyzes the conversion of oxaloacetate (OAA) to PEP during gluconeogenesis^{259,260}. This enzyme is also capable of catalyzing the reverse reaction (PEP to OAA), which has been leveraged in *E. coli* engineering for fumaric acid production by overexpressing *ppc*, a gene for phosphoenolpyruvate carboxylase which similarly catalyzes PEP to OAA²⁸.

We amplified both the *K. marxianus* and *S. cerevisiae* *PCK* genes and placed them under strong promoters. We transformed MTK117 with an EV plasmid (MTK147), *KmPCK* (MTK148), and *ScPCK* (MTK149) tested metabolite production in 0.5 x supplemented permeate medium. Despite overexpression mediated by a strong constitutive promoter, fumaric acid production

decreased, while acetic acid, ethanol, and glycerol levels increased (Fig. 4.9). At 72 hours, the EV control produced 0.84 g L⁻¹ fumaric acid, compared to 0.68 g L⁻¹ in both MTK148 and MTK149 (Fig. 4.9 A). In contrast, glycerol, ethanol, and acetic acid levels increased by approximately 1.3–1.4-fold, 2.4-fold, and 1.2–1.3-fold, respectively in the *PCK* expressing strains compared to the control (Fig. 4.9 B).

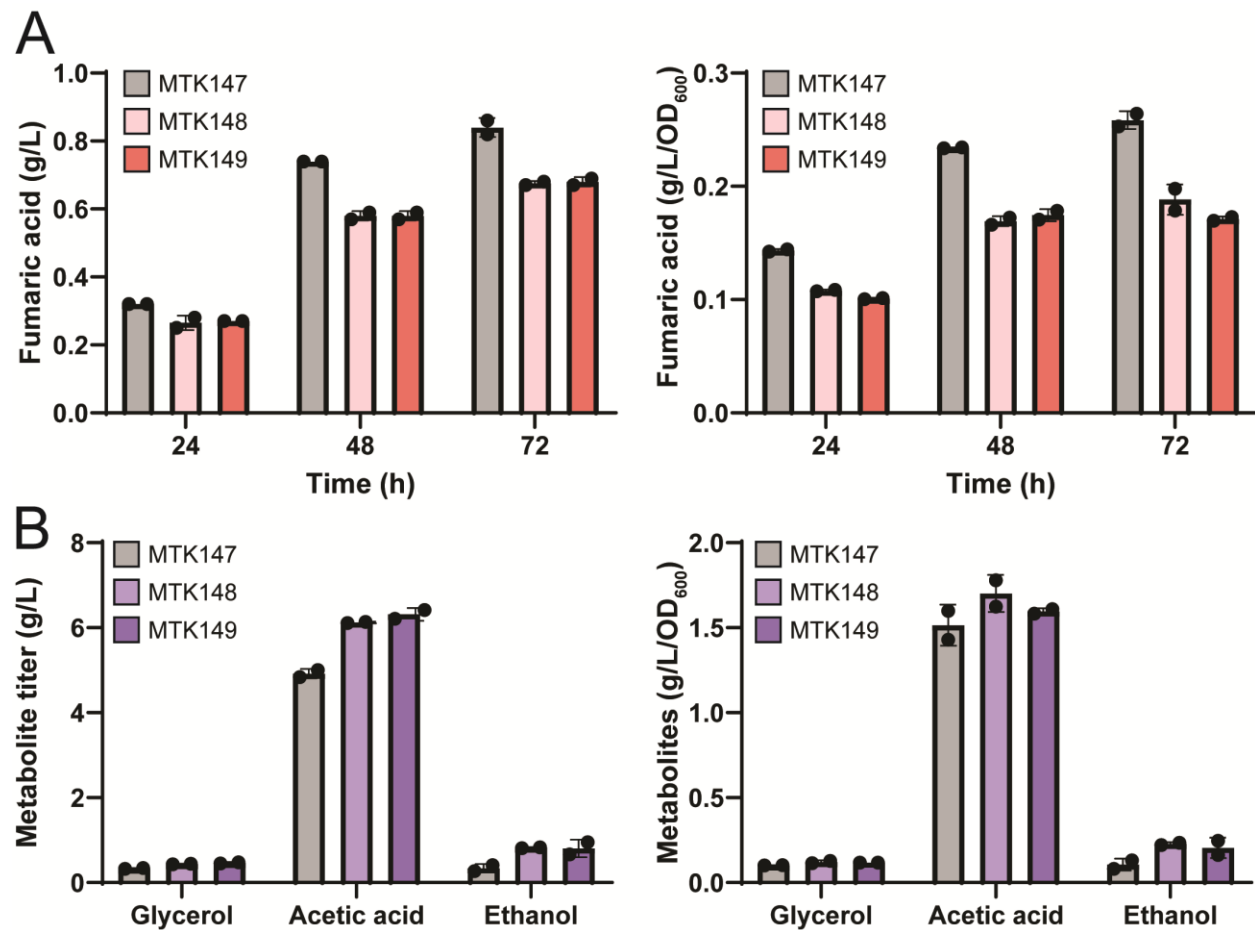


Figure 4.9 Effect of Overexpression of *KmPCK* and *ScPCK* on fumaric acid titer. A. Titrers and yield of fumaric acid **B.** Titrers and yield of metabolite by-products, glycerol, ethanol and acetic acid. Strains used are *K. marxianus* Y-1190 $\Delta FUM1$, I2::QDR2 transformed with an EV plasmid (MTK147), a *KmPCK* expression plasmid (MTK148) and a *ScPCK* expression plasmid (MTK149). Each bar represents the mean of two independent colonies and the error bars represent standard deviation.

4.5. Discussion

In this study, we used a multi-pronged metabolic engineering strategy to improve fumaric acid production in *K. marxianus*. We confirmed that *JEN2*, rather than *JEN1*, mediates fumarate uptake, though its deletion did not enhance titers, indicating reuptake is not limiting under current conditions (Fig. 4.3). Introducing a modified *R. oryzae* rTCA pathway in a *QDR2* and *QDR3* co-overexpression background increased fumaric acid titers 2-fold, underscoring the value of pathway optimization (Fig. 4.6). Redox balancing through heterologous GDH expression and *PFK2* deletion modestly boosted titers and reduced ethanol/glycerol, though acetate accumulation persisted (Fig. 4.8). Finally, diverting flux via *PCK* overexpression decreased fumarate yields, likely due to futile cycling (Fig. 4.9). Together, these results highlight the interplay of transport, cofactor balance, and carbon partitioning as key levers for optimizing fumaric acid biosynthesis in *K. marxianus*.

Overexpression of the heterologous pathway for the *R. oryzae* rTCA was the most successful of the rational metabolic engineering strategies explored in this chapter, raising fumaric acid titers 2-fold (Fig. 4.6). However, both the fold-change and the absolute titers were lower than expected based on previous studies. In *S. cerevisiae*, for example, overexpressing the rTCA pathway in a $\Delta FUM1\Delta THI2$ background led to titers of 3.92 g L⁻¹, compared to their $\Delta FUM1\Delta THI2$ control which produced 0.25 g L⁻¹ of fumaric acid, a ~16-fold increase³⁸. One key difference in our strain is the absence of the *THI2* deletion, which was originally done to produce more pyruvate as a pool for use in synthesis of other organic acids²⁶¹. However, we attempted this strategy in *K. marxianus* but could never produce a combined *THI2* and *FUM1* deletion mutant. Notably, even without the *THI2* mutant, *C. glabrata* overexpressing just *RoFUM*, *RoMDH*, and,

RoPYC yielded fumaric acid titers of 3.06 g L⁻¹ compared to the control strain which produced ~0.18 g L⁻¹, a 17-fold increase⁴⁴.

An important difference among these organisms is their metabolic classification: both *S. cerevisiae* and *C. glabrata* are Crabtree positive, while *K. marxianus* is Crabtree negative^{76,77,262}. The Crabtree effect refers to the phenomenon where fermentative metabolism is favoured over respiration under high glucose and high oxygen conditions, despite respiration producing 9 times more ATP per glucose molecule²⁶³. This apparent paradox has been explained as a strategy for optimizing growth rate since oxidative phosphorylation produces more ATP per molecule of glucose, but the fermentative pathway produces more ATP per unit of protein made^{264,265}. Under low glucose, Crabtree-positive yeast rely on the oxidative TCA cycle; when glucose is abundant, they shift to fermentation to maximize ATP output per proteome investment²⁶⁴. Crabtree negative yeast like *K. marxianus* do not undergo this metabolic shift, likely due to higher protein content, preventing the need to prioritize ribosomal resources. It is estimated that 72% of *K. marxianus* biomass is protein, compared to 46 % in *S. cerevisiae*^{81,266}.

These divergent proteomic and metabolic profiles likely contribute to the reduced effectiveness of the rTCA pathway in *K. marxianus*. In Crabtree positive organisms, glycolytic flux results in a net accumulation of cytosolic NADH, which can drive the reverse reaction of malate dehydrogenase required for fumaric acid production by the rTCA (Fig. 4.1)^{267,268}. In contrast, the $\Delta FUM1$ background of *K. marxianus* disrupts the oxidative TCA in a strain optimized for respiratory metabolism, likely creating a net loss of NADH as flux through the TCA has slowed (Fig. 4.1). Previous work has shown that *K. marxianus* allocates a greater proportion of its proteome to ATP synthase, ETC, and the TCA while dedicating less to glycolysis proteins than *S. cerevisiae*²⁶⁹. As a result, *K. marxianus* is less able to regenerate ATP and NADH through

glycolysis alone, and instead accumulates acetate to replenish the pools of NADH required for ATP synthesis²³¹.

To address the potential limitation of intracellular NADH, we overexpressed GDHs to increase NADH availability (Fig. 4.2). Both *K. marxianus* GDH genes (*GDH2* and *GDH3*) and the *S. cerevisiae* *GDH1* and *GDH2* were tested. Only the *S. cerevisiae* GDHs showed a modest but significant improvement in fumaric acid titers. We also deleted *PFK2* to divert carbon flux toward the PPP; however, when combined with *ScGDH1* and *ScGDH2* overexpression, this intervention decreased glycerol and ethanol accumulation without further improving fumaric acid titers beyond the effect of *ScGDHs* alone or decreasing acetic acid titers (Fig. 4.8).

While studies of GDH enzymes in *K. marxianus* are limited, insights from the closely related *K. lactis* may provide context. *K. lactis* also contains only one NADP-dependent GDH and its deletion could only be rescued by *ScGDH3*, not *ScGDH1*²⁷⁰. Additionally, when specific activity of *ScGDH1* and *ScGDH3* were measured in *K. lactis*, *ScGDH1* was ~2-fold lower than *ScGDH3*²⁷⁰. This suggests that the modest improvement in fumaric acid production observed upon *ScGDH1* and *ScGDH2* overexpression in *K. marxianus* may be attributable more to increased α -ketoglutarate (aKG) formation feeding into the TCA via *ScGDH2* than to enhanced NADH generation (Figs. 4.2 and 4.8). Since *ScGDH2* produces both aKG and NADH, its expression could be contributing to fumarate production via multiple mechanisms.

K. marxianus exhibits high flux towards the TCA cycle under respiratory conditions, especially when compared to *S. cerevisiae*^{269,271}. The increased acetic acid levels observed could result from mitochondrial acetyl-CoA accumulation due to the disruption of the TCA by the *FUM1* deletion. This accumulation of acetate due to the disruption of the TCA is not seen in *S. cerevisiae*³⁶ but has been reported in the Crabtree negative yeast *Y. lipolytica*²³². In one study, *Y. lipolytica*

engineered for succinate production accumulated 7.5 g L⁻¹ of acetic acid, which was attributed to the activity of acetyl-CoA hydrolase (*ACH1*) activity²³². Deletion of *YLACH1* led to a 37.5-fold reduction in acetic acid and a 27 % increase in succinic acid²³². These findings suggest that acetate accumulation in *K. marxianus* may similarly result from mitochondrial acetyl-CoA overflow and point to *ACH1* as a potential deletion target to mitigate acetate production and improve carbon flux toward dicarboxylic acid synthesis.

To divert carbon away from fermentative metabolism and reduce by-product formation, we overexpressed *PCK* to enhance carbon flux toward the TCA cycle. However, this resulted in the opposite phenotype than anticipated: lower fumaric acid titers and increased production of by-products such as acetate, ethanol, and glycerol (Fig. 4.9). This outcome is likely due to PEPCK functioning in the gluconeogenic direction, converting OAA to PEP. This diversion could explain the elevated acetate and ethanol levels, as carbon is redirected to pyruvate from PEP, which may also limit OAA availability for fumarate synthesis (Fig. 4.9). Additionally, increased PEP levels may drive flux toward DHAP and G3P, precursors for glycerol, resulting in its accumulation²⁷². Notably, experimental evidence suggests that the reverse carboxylation reaction of PEPCK is thermodynamically unfavorable under most conditions, only proceeding when PEPCK is the sole anaplerotic enzyme and CO₂ is abundant²⁷³.

A potential strategy to reroute PEP to OAA could be to express the *E. coli* PEP carboxylase (*ppc*), which catalyzes this reaction independently of ATP, therefore making it more thermodynamically preferable²⁷⁴. Overexpression of *ppc* has been widely used in *E. coli* fumaric acid production strains, contributing to titers of up to 41.5 g L⁻¹ ^{27,28,35}. When expressed in *S. cerevisiae*, *ppc* increased respiratory activity relative to control strains²⁷⁵ and was also successfully used in metabolic engineering strategies to enhance propionic acid production²⁷⁶; this suggests that

the enzyme retains functional activity when heterologously expressed. Given these observations, *ppc* expression represents a promising route to reinforce anaplerotic flux in *K. marxianus* and alleviate overflow metabolism during fumaric acid production.

Overall, this work demonstrates the importance of host-specific metabolic context when engineering central carbon metabolism. While many strategies derived from Crabtree-positive yeasts proved at least partially effective, the distinct proteomic allocation, redox architecture, and carbon partitioning strategies of *K. marxianus* present unique challenges. Future work should focus on decreasing acetate titers by preventing production or employing reassimilation strategies to reduce carbon loss and acid stress in fumaric acid production strains.

4.6. Supplemental tables and figures

Table S4.1. List of strains used in this study

Species	Strain Name	Parental Strain	Genetic modification	Details	Selection
<i>S. cerevisiae</i>	CEN.PK 113-13D	-	Δ URA3		URA3
<i>K. marxianus</i>	Y-1190	-	-	Diploid yeast	-
<i>K. marxianus</i>	ROK002	Y-1190	Δ FUM1		-
<i>K. marxianus</i>	MTK075	Y-1190	Δ JEN2		-
<i>K. marxianus</i>	MTK076	Y-1190	Δ JEN1		-
<i>K. marxianus</i>	MTK078	MTK075	Δ FUM1, Δ JEN2		-
<i>K. marxianus</i>	MTK117	ROK002	Δ FUM1, I2::pNC1-QDR2-ScADH1		-
<i>K. marxianus</i>	MTK126	MTK078	Δ FUM1, Δ JEN2	Transformed with EV plasmid pMT434	
<i>K. marxianus</i>	MTK127	MTK078	Δ FUM1, Δ JEN2	Transformed with QDR2 overexpression plasmid pMT421	
<i>K. marxianus</i>	MTK130het	MTK117	Δ FUM1, I2::pNC1-QDR2-ScADH1, I3::pNC1-QDR3-PGKt (het)		-
<i>K. marxianus</i>	MTK130hom	MTK117	Δ FUM1, I2::pNC1-QDR2-ScADH1, I3::pNC1-QDR3-PGKt (hom)		-
<i>K. marxianus</i>	MTK136	MTK117	Δ FUM1, I2::pNC1-QDR2-ScADH1	transformed with KmGDH	G418

				plasmid pMT452	
<i>K. marxianus</i>	MTK137	MTK117	Δ FUM1, I2::pNC1-QDR2-ScADH1	transformed with EV plasmid pMT453	G418
<i>K. marxianus</i>	MTK147	MTK117	Δ FUM1, I2:: pNC1-QDR2-ScADH1	transformed with EV plasmid pMT442	Hyg
<i>K. marxianus</i>	MTK148	MTK117	Δ FUM1, I2:: pNC1-QDR2-ScADH1	transformed with EV plasmid pMT289	Hyg
<i>K. marxianus</i>	MTK149	MTK117	Δ FUM1, I2:: pNC1-QDR2-ScADH1	transformed with EV plasmid pMT291	Hyg
<i>K. marxianus</i>	MTK159	MTK117	Δ FUM, I2::pNC1-QDR2-ScADH1, Δ PFK2		-
<i>K. marxianus</i>	MTK160	MTK117	Δ FUM1, I2::pNC1-QDR2-ScADH1	transformed with ScGDH plasmid pMT461	G418
<i>K. marxianus</i>	MTK170	MTK130hom	Δ FUM1, I2:: pNC1-QDR2-ScADH1, I3::pNC1-QDR3-PGKt (hom)	transformed with EV plasmid pMT453	G418
<i>K. marxianus</i>	MTK172	MTK130hom	Δ FUM1, I2:: pNC1-QDR2-ScADH1, I3::pNC1-QDR3-PGKt (hom)	transformed with rTCAmoD plasmid pRO022	G418
<i>K. marxianus</i>	MTK178	MTK117	Δ FUM, I2::pNC1-QDR2-ScADH1, Δ PFK2	transformed with EV plasmid pMT453	G418
<i>K. marxianus</i>	MTK179	MTK117	Δ FUM, I2::pNC1-QDR2-ScADH1, Δ PFK2	transformed with ScGDH plasmid pMT461	G418

Table S4.2. List of plasmids used in this study

Plasmid ID	Parental Plasmid	Details	Selection	Reference	Addgene
KmK.C1	-	Level 0 plasmid MoClo with KM minimal ARS and centromeric sequence CenD/C1	Cam	¹⁰¹	#125059
Kmk.P2	-	Level 0 plasmid MoClo with KM promoter PDC1	Cam	¹⁰¹	#125035
Kmk.T3	-	Level 0 plasmid MoClo with KM terminator KMXXK_A03020	Cam	¹⁰¹	#125055
Kmk.T4	-	Level 0 plasmid MoClo with KM terminator PDC1	Cam	¹⁰¹	#125056
KmK.T5	-	Level 0 plasmid MoClo with KM terminator PGK1	Cam	¹⁰¹	#125057
pMT064	-	Part 1 of RoPYC from Twist	Amp	This work	
pMT065	-	Part 2 of RoPYC from Twist	Amp	This work	
pMT066	-	RoMDH from twist	Cam	This work	

pMT067	-	RoFH from twist	Cam	This work	
pMT068		CgMDH from Twist	Cam	This work	-
pMT069	-	SpMAE1 from Twist	Cam	This work	
pMT070	YTK001	Assembled RoPYC	Cam	This work	
pMT121	pYTK084	Km destination plasmid with G418 selection	Kan, G418	This work	
pMT261	pMT121	pTEF3-RoPYC-ScADH1t,pPGK1-RoMDH-Kmk.T3,pTDH3-RoFUM-PDC1t	Kan, G418	This work	
pMT262	pMT121	pTEF3-RoPYC-ScADH1t,pPGK1-RoMDH-Kmk.T3,pTDH3-RoFUM-PDC1t, pPDC1-SpMAE1-ScPGK1t	Kan, G418	This work	
pMT263	pMT121	pTEF3-RoPYC-ScADH1t,pPGK1-CgMDH-Kmk.T3,pTDH3-RoFUM-PDC1t, pPDC1-SpMAE1-ScPGK1t	Kan, G418	This work	
pMT289	pYTK083	pTEF3-ScPCK-PGK1t	Amp, Hyg	This work	
pMT291	pYTK083	pTEF3-KmPCK1-PGK1t	Amp, Hyg	This work	
pMT313	pMTKM01	SpCas9 plasmid with dual guides against JEN2	Amp, Nat	This work	
pMT314	pMTKM01	SpCas9 plasmid with dual guides against JEN1	Amp, Nat	This work	
pMT321	-	golden gate destination plasmid (ConLS, ConRE)	Amp, G418	Chapter 3	
pMT323	-	golden gate destination plasmid (ConLS, ConRE)	Amp, Hyg	This work	
pMT421	pMT321	pNC1_QDR2_ScADH1t	Amp, G418	Chapter 3	
pMT434	pMT321	Empty Vector Spacer using pYTK048	Amp, G418	Chapter 3	
pMT442	pMT323	Hygromycin Empty Vector with pYTK048 spacer	Amp, Hyg	This work	
pMT452	pMT121	pPGK1-KmGDH3-Kmk.T3, pTEF3-KmGDH2-PGKt	Kan, G418	This work	
pMT453	pMT121	G418 empty vector with pYTK048 spacer	Kan, G418	This work	
pMT461	pMT121	pPGK1-ScGDH3-ScADH1t, pTEF3-ScGDH2-PGKt	Amp, G418	This work	
pMT479	pMTKM03	SpCas9 plasmid with dual guides against PFK2	Amp, G418	This work	
pMTKM01	-	SpCAS9-Tyr-[gRNA: BsaI GFP dropout] with pan(OPT)ARS replication origin and Nat selection	Amp, Nat	Chapter 2 ¹⁷⁹	#233473
pMTKM03	-	SpCAS9-Tyr-[gRNA: BsaI GFP dropout] with pan(OPT)ARS replication origin and G418 selection.	Amp, G418	Chapter 2 ¹⁷⁹	#233475
pMTKM24	pYTK001	Level 0 plasmid MoClo with KM promoter PGK	Cam	Chapter 2 ¹⁷⁹	#233496
pMTKM25	pYTK001	Level 0 plasmid MoClo with KM promoter TEF3	Cam	Chapter 2 ¹⁷⁹	#233497
pMTKM26	pYTK001	Level 0 plasmid MoClo with KM promoter TDH3	Cam	Chapter 2 ¹⁷⁹	#233498
pMTKM28	pYTK001	Level 0 plasmid MoClo with KM promoter NC1	Cam	Chapter 2 ¹⁷⁹	#233500
pRO022	pMT121	pTEF3-RoPYC-ScADH1t, pNC1-RoMDHRoP160A-PDC1t	Kan, G418	This work	
pYTK001	-	Part Plasmid Entry Vector to be used in the Dueber YTK system.	Cam	¹²⁸	#65108
pYTK002	-	Encodes ConLS as a Type 1 part to be used in the Dueber YTK system	Cam	¹²⁸	#65109
pYTK008	-	Encodes ConLS' as a Type 1 part to be used in the Dueber YTK system	Cam	¹²⁸	#65115
pYTK047	-	Encodes GFP dropout as a Type 234r part to be used in the Dueber YTK system	Cam	¹²⁸	#65154

pYTK048	-	Encodes Spacer as a Type 234 part to be used in the Dueber YTK system	Cam	128	#65155
pYTK053	-	Level 0 plasmid MoClo with Sc terminator ADH1	Cam	128	#65160
pYTK054	-	Level 0 plasmid MoClo with Sc terminator PGK1	Cam	128	#65161
pYTK072	-	Encodes ConRE as a Type 5 part to be used in the Dueber YTK system	Cam	128	#65179
pYTK073	-	Encodes ConRE' as a Type 5 part to be used in the Dueber YTK system	Cam	128	#65180
pYTK077	-	Encodes KanamycinR as a Type 6 part to be used in the Dueber YTK system	Cam, G418	128	#65184
pYTK079	-	Encodes HygromycinR as a Type 6 part to be used in the Dueber YTK system	Cam, Hyg	128	#65186
pYTK083	-	Encodes AmpR-ColE1 as a Type 8 part to be used in the Dueber YTK system	Amp	128	#65190
pYTK084	-	Encodes KanR-ColE1 as a Type 8 part to be used in the Dueber YTK system	Kan	128	#65191

Table S4.3. List of oligonucleotides and sgRNA used in this study

Application	OligoID	Nucleotide sequence (5' – 3')
Forward primer to amplify ScPCK1 from CEN.PK 113-13D into golden gate destination vector	MT 380	GCATCGTCTCATCGGTCTCATATGTCCCCT TCTAAAATGAATG
Reverse primer to amplify ScPCK1 from CEN.PK 113-13D into golden gate destination vector	MT 381	ATGCCGTCTCAGGTCTCAGTTACTCGAAT TGAGGACCAG
Forward primer to amplify KmPCK1 from K.marxianus Y-1190 into golden gate destination vector	MT 384	GCATCGTCTCATCGGTCTCATatgtcccaagcaa aatc
Reverse primer to amplify KmPCK1 from K.marxianus Y-1190 into golden gate destination vector	MT 385	ATGCCGTCTCAGGTCTCAGtcagttgaggaccag c
Forward primer for creating the double guide targeting JEN2, amplify of pCD4	MT 657	AAAAGGTCTCAGACTTTTGTCTTGTAGGA AACAACGTgttttagagctagaaatagcaagttaaataag
Reverse primer for creatine the double guide targeting JEN2, amplify of pCD4	MT 658	TACGGGTCTCTAAACTTTGGAACACGACA CCCAAGtgcgcaagcccggaatc
Forward primer for creating the double guide targeting JEN1, amplify of pCD4	MT 659	AAAAGGTCTCAGACTTTGTATGCGCTGAA GAACAAACgttttagagctagaaatagcaagttaaataag
Reverse primer for creatine the double guide targeting JEN1, amplify of pCD4	MT 660	TACGGGTCTCTAAACCGGTGGTATCTGCG GTGGTTtgcgcaagcccggaatc
Diagnostic forward primer for JEN2	MT 667	atggcagaatcaatcgc
Diagnostic reverse primer for JEN2	MT 668	ttatatccgttctttgtattggac
Diagnostic forward primer for JEN1	MT 669	atgaacagctcggactc
Diagnostic reverse primer for JEN1	MT 670	actaagcgttctcgacttg
Forward primer to amplify KmGDH3 from K.marxianus Y-1190 into golden gate destination vector	MT 867	GCATCGTCTCATCGGTCTCATatgtcatcogaacc atagc
Reverse primer to amplify KmGDH3 from K.marxianus Y-1190 into golden gate destination vector	MT 868	ATGCCGTCTCAGGTCTCAGGATttagaacacgtc accttgatc

Forward primer to amplify KmGDH2 from K.marxianus Y-1190 into golden gate destination vector	MT 869	GCATCGTCTCATCGGTCTCATatgttgtaactcca gtacc
Reverse primer to amplify KmGDH2 from K.marxianus Y-1190 into golden gate destination vector	MT 870	ATGCCGTCTCAGGTCTCAGGATttaagaatctct aggttgtagc
Forward primer to amplify ScGDH1 from CEN.PK 113-13D into golden gate destination vector	MT 871	GCATCGTCTCATCGGTCTCATATGTCAGA GCCAGAATTTCAAC
Reverse primer to amplify ScGDH1 from CEN.PK 113-13D into golden gate destination vector	MT 872	ATGCCGTCTCAGGTCTCAGGATTTAAAAT ACATCACCTTGGTCAAAC
Forward primer to amplify ScGDH2 from CEN.PK 113-13D into golden gate destination vector	MT 873	GCATCGTCTCATCGGTCTCATatgcttttgataaca aaaatcgc
Reverse primer to amplify ScGDH2 from CEN.PK 113-13D into golden gate destination vector	MT 874	ATGCCGTCTCAGGTCTCAGGATtcaagcacttgc ctcc
Diagnostic forward primer for PFK2	MT 888	cgcacacacacaatcataaag
Diagnostic reverse primer for PFK2	MT 889	agcttcttttaccaccaag
qPCR forward primer for KmGDH3	MT 906	ccttggttacacgttggtca
qPCR reverse primer for KmGDH3	MT 907	gtttgcgacaagtgctcag
qPCR forward primer for KmGDH2	MT 908	tcccaggaaaagcagcacia
qPCR reverse primer for KmGDH2	MT 909	gacattggcgatcgtttccg
qPCR forward primer for RoPYC	MT 910	GTTTCGATGTGGCTATGCGTT
qPCR reverse primer for RoPYC	MT 911	GGCCCCCTCTCAATAGCATCT
qPCR forward primer for RoMDHFUM	MT 912	TGGAGGGGATGAAGTTGTCC
qPCR reverse primer for RoMDHFUM	MT 913	ACTCCCTTCTTTCCACCGAC
SpCas9 SgRNA Sequence for Guide 1 in the JEN1 dual guide vector	n/a	GTATGCGCTGAAGAACAAAC
SpCas9 SgRNA Sequence for Guide 2 in the JEN1 dual guide vector	n/a	AACCACCGCAGATACCACCG
SpCas9 SgRNA Sequence for Guide 1 in the JEN2 dual guide vector	n/a	TGTCTTGTAGGAAACAACGT
SpCas9 SgRNA Sequence for Guide 2 in the JEN2 dual guide vector	n/a	CTTGGGTGTCGTGTTCCAAA
SpCas9 SgRNA Sequence for Guide 1 in the PFK2 dual guide vector	n/a	gtagctttcaacagcaaga
SpCas9 SgRNA Sequence for Guide 2 in the PFK2 dual guide vector	n/a	gtgtacttgaaggacctag

5. Chapter 5: Conclusion and Perspectives

5.1. Preamble

The goal of this thesis was to expand the potential of *K. marxianus* as a versatile microbial chassis for sustainable organic acid production, with a particular focus on fumaric acid. At the core of this work lies the intersection of synthetic biology and strain domestication: how we can reprogram the metabolism of a non-conventional, industrially promising yeast to valorize low-cost feedstocks and overcome limitations associated with acid production, tolerance, and export.

Chapter 2 established the foundation of this work by sequencing and characterizing a dairy isolate of *K. marxianus*, providing the basis for its genetic tractability and synthetic biology potential. To this end we characterized 23 different promoters across two strains of *K. marxianus*, the lactose metabolizing strain Y-1190 and the popular CBS 6556, in both lactose and glucose media (Fig. 2.3). We constructed a CRISPR protocol using a two-guide system that increased deletion and integration efficiencies in both a NHEJ competent and NHEJ mutant background enabling metabolic engineering efforts (Fig. 2.5). This chapter also demonstrated media optimization of the dairy waste product milk permeate for both shake flask and bioreactor use (Fig. 2.6). Overall, Chapter 2 positions *K. marxianus* as a chassis organism with potential to metabolize dairy byproducts into useful biochemicals through reliable genetic engineering techniques.

Building on this platform, Chapter 3 developed and applied a genome-wide CRISPR-AID enrichment strategy to investigate fumaric acid tolerance genotypes. We demonstrated the power of titratable gene perturbation and phenotypic enrichment to enhance fumaric acid tolerance, identifying both known and novel genes involved in dicarboxylic acid stress response (Table 3.3, Fig. 3.4). This approach yielded key hits such as activation of *QDR2* and *QDR3*, members of the

DHA1 transporter family, whose overexpression enhanced growth under acid stress and increased fumaric acid titers suggesting acid efflux as the mechanism of action (Fig. 3.5, Fig S3.3).

Chapter 4 advanced fumaric acid production through rational metabolic engineering strategies. A heterologous rTCA pathway was introduced, putative fumaric acid importers were investigated, and carbon flux bottlenecks were addressed through NADH balancing and targeted pathway modulation. We identified *KmJEN2* as a putative fumaric acid importer, however its deletion did not improve titers at the concentrations of current production (Fig. 4.3). Notably, fumaric acid titers increased upon expression of *ScGDH1* and *ScGDH2* and with implementation of the rTCA^{mod} into the *QDR2* and *QDR3* co-overexpression background strain (Fig. 4.6, Fig. 4.8). This work revealed both the opportunities and constraints of engineering organic acid production in *K. marxianus*, a non-conventional yeast with distinct respiratory metabolism, redox cofactor usage, and regulatory networks compared to traditional Crabtree-positive hosts.

Together, these chapters illustrate an integrated strategy of strain domestication, high-throughput screening, and metabolic engineering to unlock the biosynthetic potential of *K. marxianus*. The following perspectives summarize key findings, remaining challenges, and future directions for deploying non-conventional yeasts in industrial biotechnology.

5.2. *K. marxianus* as a biomanufacturing chassis

5.2.1. Industrial relevance

Due to innate high growth rate, thermotolerance, GRAS status, and acid tolerance explored in Chapter 1.4, *K. marxianus* has been leveraged in industry and academia to produce bioethanol, xylitol and aroma compounds from lignocellulosic biomass⁸⁶ as well as bioethanol, aroma compounds, and enzymes from whey and dairy products^{277–280}. These examples inspired our

selection of *K. marxianus* as a chassis to valorize ultrafiltered milk permeate into the valuable chemical fumaric acid. Starting with 19 publicly available strains of *K. marxianus*, we screened for growth on lactose, milk permeate and on fumaric acid, ultimately selecting strain Y-1190 as our production host (Fig. 2.1). Efficient bioreactor growth on milk permeate was then achieved through a targeted supplementation strategy and fermentation optimization (Fig. 2.6).

Despite successfully increasing fumaric acid titers from undetectable levels to 2.16 g L⁻¹ (Fig. 3.6), we remain well below the ~50 g L⁻¹ benchmark generally considered necessary for economically viable industrial production of bulk chemicals^{237,238}. Nonetheless, this represents a meaningful first step in establishing *K. marxianus* as a potential production host. Our results underscore the importance of considering host-specific physiology when adapting rational engineering strategies developed in *S. cerevisiae*. Several well-documented interventions in *S. cerevisiae* and *C. glabrata*, such as pathway overexpression for TCA rewiring, did not produce the expected outcomes in *K. marxianus*. This divergence likely stems from fundamental differences in overflow metabolism, respiratory capacity, and cofactor dynamics, all of which are shaped by Crabtree status⁷⁷. In contrast, our unbiased CRISPR-AID enrichment screen, which did not assume prior knowledge of the underlying metabolic architecture, had a much greater impact on fumaric acid titers (Fig. 3.3, Fig. 3.5).

There is growing momentum in the exploration of non-conventional organisms for biomanufacturing, however it contends with decades of *S. cerevisiae*-focused research, and centuries of its use in industrial production. While *S. cerevisiae* remains a powerful tool for genetic manipulation and biomanufacturing, it represents only a narrow slice of the available yeast biodiversity. Many non-conventional yeasts possess inherent traits that could expand the range of feasible bioprocesses and improve production performance^{281,282}. For instance, most *S. cerevisiae*

strains cannot grow above 34 °C, or below pH 3, which limits the range of biochemicals that can be efficiently produced²⁸³. Engineering these complex traits into *S. cerevisiae* is often challenging due to the wide range of pathways involved^{284,285}. While historically it has been beneficial to choose the most relevant *S. cerevisiae* strain for a given application²⁸⁶, we are now at a stage where selecting the most appropriate species may be even more impactful^{287,288}. Our findings emphasize the need to select rational engineering approaches that align to the biological and metabolic context or explore unbiased, systems-level screening, or ideally, a combination of both.

5.2.2. Dual nature of diploidy in *K. marxianus* Y-1190

The diploid nature of *K. marxianus* Y-1190 introduces both opportunities and challenges for metabolic engineering. On one hand, diploidy can enhance robustness during industrial fermentation by providing genetic redundancy. This buffering effect may support stress tolerance, rapid growth, and viability under harsh conditions, aligning with observations in polyploid *S. cerevisiae* strains used in ethanol production²⁸⁹. In our context, we observed reduced growth when we lost the heterogeneity of the diploid (Fig. S2.7). Conversely, increased ploidy complicates genome editing, as full gene deletions or modifications require editing all alleles which reduces efficiency as ploidy increases^{138,290}. However, partial editing can produce intermediate genotypes that offer built-in gradient of gene function that may be useful for some phenotypes. For example, our strain with *QDR2* integrated into both I2 alleles and *QDR3* inserted into only one I3 allele exhibited an intermediate phenotype between the *QDR2* overexpression strain alone and the strain where *QDR2* and *QDR3* were integrated into both I3 alleles, maintaining high growth rate and high fumaric acid production (Fig. 3.4, Fig. S3.5).

Uncertainty in allele variant also arises during gene cloning. When amplifying endogenous genes for overexpression, it is typically unclear which allele serves as the PCR template. If the

alleles differ in functionally relevant ways due to SNPs, indels, or regulatory elements, this can influence the performance of cloned constructs without our knowledge. In this study, the genes we have chosen to overexpress from *K. marxianus*, *KmPCK*, *KmGDH3*, and *KmGDH2*, do not have a haplotig variant, so we have not had to contend with this issue. However, both *KmJEN1* and *KmJEN2* have haplotig variants that have 98.7 % and 99 % identity at the nucleotide level, therefore care was taken to design guides that targeted shared sequence identity between them (Table S4.4).

Together, these findings underscore the importance of accounting for ploidy and allelic variation in *K. marxianus* strain development. Advances in long-read sequencing are reshaping our understanding of heterozygosity in this species, revealing that allelic diversity in diploid strains may be more common than first realized^{164,291}. As such, engineering strategies must consider the potential for divergent alleles, gene dosage effects, and unintended variation in construct design. In our work, this awareness guided the design of our CRISPR-AID library to ensure that the guides designed targeted both alleles, if possible (Table S3.5). As genome engineering in non-conventional yeasts continues to advance, recognizing and addressing the consequences of ploidy will be essential for unlocking their full biotechnological potential.

5.3. CRISPR-AID as a functional genomics tool

5.3.1. Establishing CRISPR-AID in K. marxianus

A key innovation of this dissertation is the development and first application of CRISPR-AID in *K. marxianus*, addressing a gap in functional genomics tools. Prior to this work, there has only been one genome-wide CRISPR-based screen in *K. marxianus*, which used CRISPR deletion¹¹⁷. While CRISPR-Cas tools have been adapted for gene disruption, integration, and interference in *K. marxianus*²⁹², no previous studies had developed CRISPR activation. Our

platform uniquely leverages inducible promoters developed in Chapter 2 to mitigate growth burdens associated with strong constitutive Cas expression^{138,293,294}. This regulatory control was critical for achieving high transformation efficiency and sufficient coverage of our genome-wide library in the CRISPR-AID strain (Table S3.6, Fig. S3.1).

The power of our CRISPR-AID screen lies in its unbiased, semi-titratable control of gene expression. This unbiased approach is essential in a non-model organism like *K. marxianus*, where mechanistic pathways are less well-characterized. Transporters are often implicated in acid stress tolerance, and as anticipated, our screen identified activation of *QDR2* and *QDR3*. Critically, it also uncovered *ATP7* as a novel tolerance determinant, an insight that rational engineering alone may have overlooked (Fig. 3.4).

Interestingly, most significant hits in our enrichment screen were activation guides, despite using a pooled library that normalized representation across the activation, interference, and deletion constructs. A similar bias toward activation was observed in a CRISPR-AID screen for SAM production in *S. cerevisiae*²⁰⁵. We hypothesize that early enrichment of *QDR2* activation strains outcompeted the rest of the library, leading to an overwhelming proportion of sequencing reads mapping to just four *QDR2* activation guides (Fig. 3.3). In future applications, separating the A, I, and D libraries could yield more diverse hits while also improving sequencing depth per guide. Moreover, iterative screening by reintroducing the library into strains co-overexpressing *QDR2* and *QDR3* could uncover additional layers of tolerance, as demonstrated in prior CRISPR-AID studies^{188,205}.

To expand the system's versatility, we also developed a multiplexing framework using tRNA_{Gly} spacers to separate guide RNAs (Fig. 3.2 C). This enabled successful combinatorial application of CRISPR-A, -I, and -D within a single construct. Although our project did not require

combinatorial perturbation since individual guides provided strong improvements in fumaric acid tolerance, this system lays the groundwork for future exploration of additive or synergistic effects in *K. marxianus*.

In conclusion, our CRISPR-AID system introduces CRISPR activation to *K. marxianus* for the first time and combines it with inducible Cas expression and flexible guide multiplexing. It represents a powerful, scalable platform for functional genomics. This system offers graded gene expression control, surpassing binary knockouts, and will be foundational for future metabolic engineering and stress-tolerance efforts in *K. marxianus*.

5.3.2. Expanding CRISPR-AID utility beyond growth-coupled screens in *K. marxianus*

While the *K. marxianus* CRISPR-AID library is a powerful tool for identifying genetic drivers of desired phenotypes, its current utility is limited to traits that enhance growth. In this work, we applied it to fumaric acid tolerance, but the system could be leveraged to explore a wide range of stress-related traits relevant to bioproduction and beyond such as osmotic stress, ethanol and thermal tolerance, acid resistance, and inhibitor tolerance²⁹⁵. The latter is especially relevant for lignocellulosic biomass processing, where breakdown byproducts like furfural, acetic acid, and HMF continue to limit industrial adoption⁸⁶.

To broaden the scope of screenable phenotypes beyond growth, metabolite biosensors could be incorporated into the CRISPR-AID strain. These tools enable high-throughput estimation of metabolite production, reducing the reliance on costly and time-consuming analytical methods²⁹⁶. When combined with genome-wide perturbation libraries, biosensors can identify top-producing strains for further validation. In *S. cerevisiae*, biosensors have been used to enhance production of naringin^{297,298}, *cis,cis*-muconic acid²⁹⁹, vanillic acid, and protocatechuic acid³⁰⁰. The majority of these biosensors are derived from bacterial allosteric transcription factors (aTF),

DNA binding proteins that regulate gene expression in response to a wide variety of chemical signals³⁰¹. While a dicarboxylic acid-responsive aTF has recently been adapted for use in *S. cerevisiae*³⁰², its broad cross-reactivity with TCA intermediates like succinate and malate limits its specificity for fumaric acid.

In *K. marxianus*, biosensor implementation remains rare. One recent study engineered a plant hormone receptor (*PYRI*) as a biosensor to increase geranyl production in a CRISPR deletion screen, achieving a 2-fold improvement¹¹⁷. Interestingly, the biosensor used in their study is plant-derived rather than a bacterial aTF. This highlights a technical gap and an opportunity for future innovation for evaluating the performance of bacterial aTFs as biosensors in *K. marxianus*. Such advances could unlock more versatile, high-throughput screening platforms for diverse metabolites, particularly when integrated with genome-wide perturbation tools like CRISPR-AID, thereby accelerating both strain engineering and functional genomics in this emerging industrial yeast.

5.4. Metabolic engineering insights and future directions

This work applied a rational metabolic engineering strategy to enhance fumaric acid production in *K. marxianus*, combining *FUM1* deletion, *QDR2* and *QDR3* co-overexpression, heterologous rTCA pathway expression, transporter characterization, and redox balancing through *ScGDH1* and *ScGDH2* overexpression. These interventions yielded modest titer improvements but as acetic acid concentrations remained high, further optimization is required to enhance cell health and fumaric acid production.

5.4.1. Decreasing acetic acid through metabolic engineering

The biggest hurdle in our metabolic engineering efforts has been the accumulation of acetate following the deletion of *FUM1* in our fumaric acid production strains (Fig. S3.4, Fig. 4.8

B). Acetic acid not only diverts carbon away from our product of interest, but it also imposes additional stress on the cells³⁰³. Initially, we hypothesized that acetate accumulation resulted from a lack of NADH production due to disrupted flux through the TCA²⁴¹. However, attempts to regenerate NADH through the PPP and transhydrogenase cycle²⁴² increased fumaric acid titers without reducing acetic acid (Fig. 4.8). This suggests NADH may not be the bottleneck for this pathway.

Another reason acetate could be accumulating is due to a backup of acetyl-CoA resulting from the strong flux towards the TCA in Crabtree negative yeast like *K. marxianus*^{271,304}. In the *FUM1* mutant, flux through the TCA is slowed, but acetyl-CoA cannot freely exit through the mitochondrial membrane. This may lead to increased reliance on the *ACH1* shuttle that uses succinate to mediate CoA transfer to create acetate and succinyl-CoA²³². Future work should focus on deleting *ACH1* in the *K. marxianus* $\Delta FUM1$ strain to evaluate its impact on acetate production.

Should the accumulation of acetyl-CoA be the driving factor behind acetic acid formation, a potential solution would be to increase flux through the oxidative TCA to use up the mitochondrial backlog. Most of the metabolic engineering work in this thesis has been focused around implementing the heterologous rTCA pathway from *R. oryzae* (Chapter 4.4.2). This cytosolic pathway draws carbon through pyruvate to the “reverse” direction to OAA, malate, and finally fumarate (Fig. 4.1). However, because of its localization in the cytosol, it does not draw away from potential acetyl-CoA pools in the mitochondria. To do this, we could overexpress the TCA cycle genes α -ketoglutarate dehydrogenase (*KGD2*), succinyl-CoA ligase (*LSC*), and succinate dehydrogenase (*SDHI*) drawing carbon towards fumaric acid. This strategy has been used to enhance fumaric acid or succinic acid production in *S. cerevisiae*^{239,305} and *Y. lipolytica*^{232,306}, giving us confidence in the strategy for both Crabtree positive and negative yeasts.

If acetyl-CoA is not the main precursor of acetate under our conditions, an alternative approach would be to divert carbon away from the pyruvate node, reducing acetate production through overflow metabolism (Fig. 4.1). This strategy was explored in Chapter 4 by overexpressing *KmPCK* and *ScPCK* to convert PEP into OAA. However, this approach decreased fumaric acid titers, likely due to the strong thermodynamic preference of the enzyme to operate in the OAA to PEP direction²⁷³. Another method to increase PEP flux towards OAA would be to express a codon optimized *ppc* gene from *E. coli*. This enzyme catalyzes the ATP-independent conversion of PEP to OAA and has been successfully used in *S. cerevisiae* to increase propionate production, demonstrating effective heterologous expression and the expected metabolic outcome²⁷⁶. Employing the two complementary strategies of targeting acetyl-CoA conversion and pyruvate node flux, we may accelerate the resolution of acetate overproduction in our system.

5.5. Conclusion

In this thesis, I established *K. marxianus* as a promising chassis for fumaric acid production from dairy byproducts by addressing genetic, physiological, and metabolic barriers. For my first objective, I expanded the genetic engineering toolkit by developing and standardizing a modular promoter library compatible with the Yeast Toolkit (MoClo) and improved CRISPR-based methods providing a flexible platform for future strain engineering in Chapter 2. Additionally, I developed CRISPR-AID for *K. marxianus* and designed a genome-wide guide library in Chapter 3. For my second objective, I applied these tools in a genome-wide CRISPR enrichment screen to uncover determinants of fumaric acid tolerance, identifying *QDR2* and *QDR3* overexpression that enhanced growth of *K. marxianus* Y-1190 Δ *FUM1* to the solubility limit of fumaric acid, at 42 mM. For my third and final objective, I demonstrated proof-of-concept metabolic rewiring strategies that improved fumaric acid yields, including fumarase deletion, transporter engineering,

rTCA pathway expression, and redox balancing, while also revealing persistent challenges such as elevated acetic acid accumulation. The most successful engineering effort was the co-overexpression of *QDR2* and *QDR3* that enhanced fumaric acid production to 2.16 g L⁻¹ in a fed-batch bioreactor. Collectively, this work provides both foundational genetic resources and applied engineering strategies to accelerate the biotechnological deployment of *K. marxianus*.

Bibliography

1. Sauer, M., Porro, D., Mattanovich, D. & Branduardi, P. Microbial production of organic acids: expanding the markets. *Trends Biotechnol.* **26**, 100–108 (2008).
2. Max, B. *et al.* Biotechnological production of citric acid. *Braz. J. Microbiol.* **41**, 862–875 (2010).
3. Abedi, E., Mohammad, S. & Hashemi, B. Lactic acid production - producing microorganisms and substrates sources-state of art. *Heliyon* **6**, e04974 (2020).
4. FoodIngredientsFirst. Cargill Awarded for Innovation in Lactic Acid Production. *FoodIngredientsFirst* <https://www.foodingredientsfirst.com/news/cargill-awarded-for-innovation-in-lactic-acid-production.html> (2010).
5. Kim, J., Kim, Y.-M., Reddy Lebaka, V. & Wee, Y.-J. Lactic Acid for Green Chemical Industry: Recent Advances in and Future Prospects for Production Technology, Recovery, and Applications. *Fermentation* **8**, 609 (2022).
6. Yu, J. *et al.* PLA bioplastic production: From monomer to the polymer. *Eur. Polym. J.* **193**, 112076 (2023).
7. Balla, E. *et al.* Poly(lactic Acid): A Versatile Biobased Polymer for the Future with Multifunctional Properties—From Monomer Synthesis, Polymerization Techniques and Molecular Weight Increase to PLA Applications. *Polymers (Basel)*. **13**, 1822 (2021).
8. Cok, B., Tsiropoulos, I., Roes, A. L. & Patel, M. K. Succinic acid production derived from carbohydrates: An energy and greenhouse gas assessment of a platform chemical toward a bio-based economy. *Biofuels, Bioprod. Biorefining* **8**, 16–29 (2014).
9. Denny, J. Biobased Succinic Acid: Is the Sustainable Route Becoming Economically Viable Once More? *ResourceWise* <https://www.resourcewise.com/blog/chemicals-blog/biobased-succinic-acid-is-the-sustainable-route-becoming-ecomically-viable-once-more> (2022).
10. Dickson, R. *et al.* Sustainable bio-succinic acid production: superstructure optimization, techno-economic, and lifecycle assessment. *Energy Environ. Sci.* **14**, 3542–3558 (2021).
11. Kumar, V. *et al.* Recent advances in bio-based production of top platform chemical, succinic acid: an alternative to conventional chemistry. *Biotechnol. Biofuels Bioprod.* **2024 171** **17**, 1–39 (2024).
12. Roa Engel, C. A., Straathof, A. J. J., Zijlmans, T. W., Van Gulik, W. M. & Van Der Wielen, L. A. M. Fumaric acid production by fermentation. *Appl. Microbiol. Biotechnol.* **78**, 379–389 (2008).
13. Werpy, T. & Petersen, G. *Top Value Added Chemicals from Biomass Volume I. Us Nrel* <http://www.osti.gov/scitech//servlets/purl/15008859-s6ri0N/native/> (2004) doi:10.2172/15008859.
14. Yang, S. T., Zhang, K. & Zhang, B. *3.16 Fumaric Acid*. (2011).

15. Xu, Q., Li, S., Huang, H. & Wen, J. Key technologies for the industrial production of fumaric acid by fermentation. *Biotechnol. Adv.* **30**, 1685–1696 (2012).
16. Atwan, A. *et al.* Oral fumaric acid esters for psoriasis: abridged Cochrane systematic review including GRADE assessments. *Br. J. Dermatol.* **175**, 873–881 (2016).
17. Smith, D. Fumaric acid esters for psoriasis: a systematic review. *Ir. J. Med. Sci.* **186**, 161–177 (2017).
18. Gold, R., Linker, R. A. & Stangel, M. Fumaric acid and its esters: An emerging treatment for multiple sclerosis with antioxidative mechanism of action. *Clinical Immunology* vol. 142 44–48 (2012).
19. Barnett, J. A. A history of research on yeasts 6: the main respiratory pathway. *Yeast* **20**, 1015–1044 (2003).
20. Gest, H. Evolution of the citric acid cycle and respiratory energy conversion in prokaryotes. *FEMS Microbiol. Lett.* **12**, 209–215 (1981).
21. Kornberg, H. L. The role and control of the glyoxylate cycle in *Escherichia coli*. *Biochem. J.* **99**, 1 (1966).
22. Kornberg, H. L. & Krebs, H. A. Synthesis of Cell Constituents from C2-Units by a Modified Tricarboxylic Acid Cycle. *Nat. 1957 1794568* **179**, 988–991 (1957).
23. Trevisson, E. *et al.* Functional Complementation in Yeast Allows Molecular Characterization of Missense Argininosuccinate Lyase Mutations. *J. Biol. Chem.* **284**, 28926 (2009).
24. Lee, P. & Colman, R. F. Expression, purification, and characterization of stable, recombinant human adenylosuccinate lyase. *Protein Expr. Purif.* **51**, 227–234 (2007).
25. Zelle, R. M. *et al.* Malic acid production by *Saccharomyces cerevisiae*: Engineering of pyruvate carboxylation, oxaloacetate reduction, and malate export. *Appl. Environ. Microbiol.* **74**, 2766–2777 (2008).
26. Zou, X., Mo, Z., Wang, L., Chen, S. & Lee, S. Y. Overcoming Bacteriophage Contamination in Bioprocessing: Strategies and Applications. *Small Methods* **9**, (2025).
27. Song, C. W., Kim, D. I., Choi, S., Jang, J. W. & Lee, S. Y. Metabolic engineering of *Escherichia coli* for the production of fumaric acid. *Biotechnol. Bioeng.* **110**, 2025–2034 (2013).
28. Song, C. W. & Lee, S. Y. Combining rational metabolic engineering and flux optimization strategies for efficient production of fumaric acid. *Appl. Microbiol. Biotechnol.* **99**, 8455–8464 (2015).
29. Chen, X., Ma, D., Liu, J., Luo, Q. & Liu, L. Engineering the transmission efficiency of the noncyclic glyoxylate pathway for fumarate production in *Escherichia coli*. *Biotechnol. Biofuels* **13**, 1–10 (2020).
30. Chen, X., Zhu, P. & Liu, L. Modular optimization of multi-gene pathways for fumarate production. *Metab. Eng.* **33**, 76–85 (2016).

31. Zhang, T. *et al.* Pull-in urea cycle for the production of fumaric acid in *Escherichia coli*. *Appl. Microbiol. Biotechnol.* **99**, 5033–5044 (2015).
32. Skorokhodova, A. Y., Gulevich, A. Y. & Debabov, V. G. Engineering *Escherichia coli* for efficient aerobic conversion of glucose to fumaric acid. *Biotechnol. Reports* **33**, e00703 (2022).
33. Liu, H. *et al.* Genetic manipulation of *Escherichia coli* central carbon metabolism for efficient production of fumaric acid. *Bioresour. Technol.* **270**, 96–102 (2018).
34. Zhang, T. *et al.* Regulating C4-dicarboxylate transporters for improving fumaric acid production. *RSC Adv.* **7**, 2897–2904 (2017).
35. Li, N. *et al.* Engineering *Escherichia coli* for fumaric acid production from glycerol. *Bioresour. Technol.* **174**, 81–87 (2014).
36. Xu, G. *et al.* Fumaric Acid Production in *Saccharomyces cerevisiae* by In Silico Aided Metabolic Engineering. *PLoS One* **7**, 1–10 (2012).
37. Xu, G., Liu, L. & Chen, J. Reconstruction of cytosolic fumaric acid biosynthetic pathways in *Saccharomyces cerevisiae*. *Microb. Cell Fact.* **11**, 1–10 (2012).
38. Xu, G., Chen, X., Liu, L. & Jiang, L. Fumaric acid production in *Saccharomyces cerevisiae* by simultaneous use of oxidative and reductive routes. *Bioresour. Technol.* **148**, 91–96 (2013).
39. Karim, A., Gerliani, N. & Aider, M. *Kluyveromyces marxianus*: An emerging yeast cell factory for applications in food and biotechnology. *Int. J. Food Microbiol.* **333**, (2020).
40. Heisteringer, L., Gasser, B. & Mattanovich, D. Microbe Profile: *Komagataella phaffii*: a methanol devouring biotech yeast formerly known as *Pichia pastoris*. *Microbiology* **166**, 614–616 (2020).
41. Ledesma-Amaro, R. & Nicaud, J. M. *Yarrowia lipolytica* as a biotechnological chassis to produce usual and unusual fatty acids. *Prog. Lipid Res.* **61**, 40–50 (2016).
42. Suominen, P. *et al.* Genetically modified yeast of the species *Issatchenkia orientalis* and closely relates species, and fermentation processes using same. (2012).
43. Chen, X. *et al.* Fumaric acid production by *Torulopsis glabrata*: Engineering the urea cycle and the purine nucleotide cycle. *Biotechnol. Bioeng.* **112**, 156–167 (2015).
44. Chen, X., Li, Y., Tong, T. & Liu, L. Spatial modulation and cofactor engineering of key pathway enzymes for fumarate production in *Candida glabrata*. *Biotechnol. Bioeng.* **116**, 622–630 (2019).
45. Chen, X. *et al.* Fumarate Production by *Torulopsis glabrata*: Engineering Heterologous Fumarase Expression and Improving Acid Tolerance. (2016)
doi:10.1371/journal.pone.0164141.
46. Zhou, J., Liu, L. & Chen, J. Improved ATP supply enhances acid tolerance of *Candida glabrata* during pyruvic acid production. *J. Appl. Microbiol.* **110**, 44–53 (2011).

47. Zhang, T., Ge, C., Deng, L., Tan, T. & Wang, F. C4-dicarboxylic acid production by overexpressing the reductive TCA pathway. *FEMS Microbiol. Lett.* **362**, (2015).
48. Pyne, M. E. *et al.* Screening non-conventional yeasts for acid tolerance and engineering *Pichia occidentalis* for production of muconic acid. *Nat. Commun.* **14**, (2023).
49. Sun, W. *et al.* Metabolic engineering of an acid-tolerant yeast strain *Pichia kudriavzevii* for itaconic acid production. *Metab. Eng. Commun.* **10**, (2020).
50. Park, H. J. *et al.* Low-pH production of d-lactic acid using newly isolated acid tolerant yeast *Pichia kudriavzevii* NG7. *Biotechnol. Bioeng.* **115**, 2232–2242 (2018).
51. Carruthers, D. N. & Lee, T. S. Translating advances in microbial bioproduction to sustainable biotechnology. *Front. Bioeng. Biotechnol.* **10**, 968437 (2022).
52. Leong, H. Y. *et al.* Waste biorefinery towards a sustainable circular bioeconomy: a solution to global issues. *Biotechnol. Biofuels* **14**, 1–15 (2021).
53. Segers, B. *et al.* Lignocellulosic biomass valorisation: a review of feedstocks, processes and potential value chains and their implications for the decision-making process. *RSC Sustain.* **2**, 3730–3749 (2024).
54. Chandel, A. K., Garlapati, V. K., Singh, A. K., Antunes, F. A. F. & da Silva, S. S. The path forward for lignocellulose biorefineries: Bottlenecks, solutions, and perspective on commercialization. *Bioresour. Technol.* **264**, 370–381 (2018).
55. Varela, J. A. *et al.* Origin of Lactose Fermentation in *Kluyveromyces lactis* by Interspecies Transfer of a Neo-functionalized Gene Cluster during Domestication. *Curr. Biol.* **29**, 4284-4290.e2 (2019).
56. Market Research Future. *Whey Permeate Market Size Worth USD 970.6 Million by 2028*. <https://www.globenewswire.com/en/news-release/2022/05/04/2435465/0/en/Whey-Permeate-Market-Size-Worth-USD-970-6-Million-by-2028-Witnessing-a-CAGR-of-4-05-Report-by-Market-Research-Future-MRFR.html> (2022).
57. Kurtzman, C. P., Fell, J. W. & Boekhout, T. *The Yeasts, a Taxonomy Study*. (Elsevier Science, 2011).
58. Goshima, T. *et al.* Bioethanol Production from Lignocellulosic Biomass by a Novel *Kluyveromyces marxianus* Strain. *Biosci. Biotechnol. Biochem.* **77**, 1505–1510 (2013).
59. Rajkumar, A. S. & Morrissey, J. P. Rational engineering of *Kluyveromyces marxianus* to create a chassis for the production of aromatic products. *Microb. Cell Fact.* **19**, (2020).
60. Bilal, M. *et al.* Bioprospecting *Kluyveromyces marxianus* as a Robust Host for Industrial Biotechnology. *Front. Bioeng. Biotechnol.* **0**, 562 (2022).
61. Groeneveld, P., Stouthamer, A. H., Westerhoff, H. V & Westerhoff, H. V. Super life-how and why ‘cell selection’ leads to the fastest-growing eukaryote. *Authors J. Compil.* **a 276**, 254–270 (2009).
62. Chang, J. J. *et al.* A thermo- and toxin-tolerant kefir yeast for biorefinery and biofuel production. *Appl. Energy* **132**, 465–474 (2014).

63. Anderson, P. J., McNeil, K. & Watson, K. High-efficiency carbohydrate fermentation to ethanol at temperatures above 40°C by *Kuyveromyces marxianus* var. *marxianus* isolated from sugar mills. *Appl. Environ. Microbiol.* **51**, 1314–1320 (1986).
64. Nonklang, S. *et al.* High-Temperature Ethanol Fermentation and Transformation with Linear DNA in the Thermotolerant Yeast *Kluyveromyces marxianus* DMKU3-1042. *Appl. Environ. Microbiol.* **74**, 7514 (2008).
65. Varela, J. A. *et al.* Polymorphisms in the LAC12 gene explain lactose utilisation variability in *Kluyveromyces marxianus* strains. *FEMS Yeast Res.* **17**, (2017).
66. Ortiz-Merino, R. A. *et al.* Ploidy variation in *Kluyveromyces marxianus* separates dairy and non-dairy isolates. *Front. Genet.* **9**, (2018).
67. Lane, M. M. & Morrissey, J. P. *Kluyveromyces marxianus*: A yeast emerging from its sister's shadow. *Fungal Biol. Rev.* **24**, 17–26 (2010).
68. Llorente, B. *et al.* Genomic Exploration of the Hemiascomycetous Yeasts: 12. *Kluyveromyces marxianus* var. *marxianus*. *FEBS Lett.* **487**, 71–75 (2000).
69. Lane, M. M. *et al.* Physiological and metabolic diversity in the yeast *Kluyveromyces marxianus*. *Antonie van Leeuwenhoek, Int. J. Gen. Mol. Microbiol.* **100**, 507–519 (2011).
70. Barsoum, E., Martinez, P. & Åström, S. U. $\alpha 3$, a transposable element that promotes host sexual reproduction. *Genes Dev.* **24**, 33 (2010).
71. Rajaei, N., Chiruvella, K. K., Lin, F. & Åström, S. U. Domesticated transposase Kat1 and its fossil imprints induce sexual differentiation in yeast. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 15491–15496 (2014).
72. Cernak, P. *et al.* Engineering *kluyveromyces marxianus* as a robust synthetic biology platform host. *MBio* **9**, (2018).
73. Tittarelli, F. *et al.* Development and implementation of multilocus sequence typing to study the diversity of the yeast *Kluyveromyces marxianus* in Italian cheeses. doi:10.1099/mgen.0.000153.
74. Sellick, C. A., Campbell, R. N. & Reece, R. J. Galactose metabolism in yeast-structure and regulation of the leloir pathway enzymes and the genes encoding them. *Int. Rev. Cell Mol. Biol.* **269**, 111–150 (2008).
75. Du, C. *et al.* Co-utilization of multiple lignocellulose-derived sugars and inhibitors by the robust mutant *Kluyveromyces marxianus* through adaptive laboratory evolution. *Fuel* 122816 (2021) doi:10.1016/J.FUEL.2021.122816.
76. De Deken, R. H. The Crabtree effect: a regulatory system in yeast. *J. Gen. Microbiol.* **44**, 149–156 (1966).
77. Van Urk, H., Voll, W. S. L. L., Scheffers, W. A. & Van Dijken, J. P. Transient-State Analysis of Metabolic Fluxes in Crabtree-Positive and Crabtree-Negative Yeasts. **56**, (1990).
78. Guo, Y. *et al.* The advances in creating Crabtree-negative *Saccharomyces cerevisiae* and

- the application for chemicals biosynthesis. *FEMS Yeast Res.* **25**, 14 (2025).
79. Administration, U. S. F. and D. *Microorganisms & Microbial-Derived Ingredients Used in Food (Partial List)*. <https://www.fda.gov/food/generally-recognized-safe-gras/microorganisms-microbial-derived-ingredients-used-food-partial-list> (2018).
 80. EFSA Panel on Biological Hazards & Allende, A., Alvarez-Ordóñez, A., Bortolaia, V., Bover-Cid, S., De Cesare, A., Dohmen, W., Guillier, L., Jacxsens, L., Nauta, M., Mughini-Gras, L., Ottoson, J., Peixe, L., Perez-Rodriguez, F., Skandamis, P., Suffredini, E., Cocconcelli, P. S., Fernández E, L. *Update of the list of qualified presumption of safety (QPS) recommended microbiological agents intentionally added to food or feed as notified to EFSA 22: Suitability of taxonomic units notified to EFSA until March 2025*. <https://doi.org/10.2903/j.efsa.2025.9510> (2025) doi:10.2903/j.efsa.2025.9510.
 81. Van Hoek, P., Van Dijken, J. P. & Pronk, J. T. Effect of specific growth rate on fermentative capacity of baker's yeast. *Appl. Environ. Microbiol.* **64**, 4226–4233 (1998).
 82. Shemesh, P. & Fishman, A. Optimal fermentation conditions for growth and recombinant protein production in *Pichia pastoris*: Strain selection, ploidy level and carbon source. *Curr. Res. Food Sci.* **9**, 100840 (2024).
 83. Koutinas, A. A., Chatzifragkou, A., Kopsahelis, N., Papanikolaou, S. & Kookos, I. K. Design and techno-economic evaluation of microbial oil production as a renewable resource for biodiesel and oleochemical production. *Fuel* **116**, 566–577 (2014).
 84. Prado, C. D. *et al.* Physiological characterization of a new thermotolerant yeast strain isolated during Brazilian ethanol production, and its application in high-temperature fermentation. *Biotechnol. Biofuels* **13**, 1–15 (2020).
 85. Paulino de Souza, J. *et al.* Improvement of Brazilian bioethanol production – Challenges and perspectives on the identification and genetic modification of new strains of *Saccharomyces cerevisiae* yeasts isolated during ethanol process. *Fungal Biol.* **122**, 583–591 (2018).
 86. Baptista, M. & Domingues, L. *Kluyveromyces marxianus* as a microbial cell factory for lignocellulosic biomass valorisation. *Biotechnol. Adv.* **60**, 108027 (2022).
 87. de Souza, C. J. A. *et al.* The influence of presaccharification, fermentation temperature and yeast strain on ethanol production from sugarcane bagasse. *Bioresour. Technol.* **109**, 63–69 (2012).
 88. Chapman, J. W., Musters, W., Rouwenhorst, Robert J., Toschka, H. Y. & Verbakel, J. M. A. The use of the *kluyveromyces marxianus* inulinase gene promoter for protein production. (1994).
 89. Gazaloğlu, M., Camarasa, C. & Nevoigt, E. Exploring pectinolytic yeast diversity: toward effective polygalacturonase producers for applications in wine-making. *FEMS Yeast Res.* **25**, 33 (2025).
 90. Rollero, S. *et al.* *Kluyveromyces marxianus* Secretes a Pectinase in Shiraz Grape Must That Impacts Technological Properties and Aroma Profile of Wine. *J. Agric. Food Chem.* **66**, 11739–11747 (2018).

91. Schwan, R. F., Cooper, R. M. & Wheals, A. E. Endopolygalacturonase secretion by *Kluyveromyces marxianus* and other cocoa pulp-degrading yeasts. *Enzyme Microb. Technol.* **21**, 234–244 (1997).
92. Pasotti, L. *et al.* Fermentation of lactose to ethanol in cheese whey permeate and concentrated permeate by engineered *Escherichia coli*. *BMC Biotechnol.* **17**, (2017).
93. Nilsen, D. The Whey Forward — The Beverage Makers Innovating to Solve a Major Food Waste Issue — Good Beer Hunting. *Good Beer Hunting*
<https://www.goodbeerhunting.com/blog/2022/7/6/the-whey-forward-the-beverage-makers-innovating-to-solve-a-major-food-waste-issue> (2022).
94. Antunes, D. F., De Souza, J. & De Moraes, J. A simple and rapid method for lithium acetate-mediated transformation of *Kluyveromyces marxianus* cells. *World J. Microbiol. Biotechnol.* **16**, 653–654 (2000).
95. Iborra, F. High efficiency transformation of *Kluyveromyces marxianus* by a replicative plasmid. *Curr. Genet.* **24**, 181–183 (1993).
96. Abdel-Banat, B. M. A., Hoshida, H. & Akada, R. Various short autonomously replicating sequences from the yeast *Kluyveromyces marxianus* seemingly without canonical consensus. *Curr. Res. Microb. Sci.* **2**, 100053 (2021).
97. Liachko, I. & Dunham, M. J. An autonomously replicating sequence for use in a wide range of budding yeasts. *FEMS Yeast Res.* **14**, 364–367 (2014).
98. Lee, K. S. *et al.* Characterization of *Saccharomyces cerevisiae* promoters for heterologous gene expression in *Kluyveromyces marxianus*. *Appl. Microbiol. Biotechnol.* **97**, 2029–2041 (2013).
99. Kumar, P., Sahoo, D. K. & Sharma, D. The identification of novel promoters and terminators for protein expression and metabolic engineering applications in *Kluyveromyces marxianus*. *Metab. Eng. Commun.* **12**, e00160 (2021).
100. Lang, X., Besada-Lombana, P. B., Li, M., Da Silva, N. A. & Wheeldon, I. Developing a broad-range promoter set for metabolic engineering in the thermotolerant yeast *Kluyveromyces marxianus*. *Metab. Eng. Commun.* **11**, (2020).
101. Rajkumar, A. S., Varela, J. A., Juergens, H., Daran, J.-M. G. & Morrissey, J. P. Biological Parts for *Kluyveromyces marxianus* Synthetic Biology. *Front. Bioeng. Biotechnol.* **7**, 97 (2019).
102. Ren, H. *et al.* Establishment of a Cre-loxP System Based on a Leaky LAC4 Promoter and an Unstable panARS Element in *Kluyveromyces marxianus*. *Microorganisms* **10**, 1240 (2022).
103. Wu, L. *et al.* A protocol of rapid laboratory evolution by genome shuffling in *Kluyveromyces marxianus*. *MethodsX* **7**, (2020).
104. Abdel-Banat, B. M. A., Nonklang, S., Hoshida, H. & Akada, R. Random and targeted gene integrations through the control of non-homologous end joining in the yeast *Kluyveromyces marxianus*. *Yeast* **27**, 29–39 (2010).

105. Chang, J. J. *et al.* PGASO: A synthetic biology tool for engineering a cellulolytic yeast. *Biotechnol. Biofuels* **5**, 1–12 (2012).
106. Juergens, H. *et al.* Genome editing in *Kluyveromyces* and *Ogataea* yeasts using a broad-host-range Cas9/gRNA co-expression plasmid. *FEMS Yeast Res.* **18**, (2018).
107. Rajkumar, A. S. & Morrissey, J. P. Protocols for marker-free gene knock-out and knock-down in *Kluyveromyces marxianus* using CRISPR/Cas9. *FEMS Yeast Res.* **22**, 67 (2022).
108. Nambu-Nishida, Y., Nishida, K., Hasunuma, T. & Kondo, A. Development of a comprehensive set of tools for genome engineering in a cold- and thermo-tolerant *Kluyveromyces marxianus* yeast strain. *Sci. Rep.* **7**, (2017).
109. Bever, D., Wheeldon, I. & Da Silva, N. RNA polymerase II-driven CRISPR-Cas9 system for efficient non-growth-biased metabolic engineering of *Kluyveromyces marxianus*. *Metab. Eng. Commun.* 141267 (2022) doi:10.1016/j.mec.2022.e00208.
110. Li, M. *et al.* CRISPR-mediated multigene integration enables Shikimate pathway refactoring for enhanced 2-phenylethanol biosynthesis in *Kluyveromyces marxianus*. *Biotechnol. Biofuels* **14**, 3 (2021).
111. Löbs, A. K., Engel, R., Schwartz, C., Flores, A. & Wheeldon, I. CRISPR-Cas9-enabled genetic disruptions for understanding ethanol and ethyl acetate biosynthesis in *Kluyveromyces marxianus*. *Biotechnol. Biofuels* **10**, 164 (2017).
112. Ren, L. *et al.* Improving glycerol utilization during high-temperature xylitol production with *Kluyveromyces marxianus* using a transient clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated protein 9 system. *Bioresour. Technol.* **365**, (2022).
113. Nambu-Nishida, Y., Nishida, K., Hasunuma, T. & Kondo, A. Genetic and physiological basis for antibody production by *Kluyveromyces marxianus*. *AMB Express* **8**, (2018).
114. Lee, M. H. *et al.* Constructing a human complex type N-linked glycosylation pathway in *Kluyveromyces marxianus*. *PLoS One* **15**, e0233492 (2020).
115. Wang, W. *et al.* CRISPR-Cas9 mediated genome editing of *Kluyveromyces marxianus* for iterative, multiplexed gene disruption and pathway integration. (2024) doi:10.1002/bit.28736.
116. Löbs, A. K., Schwartz, C., Thorwall, S. & Wheeldon, I. Highly Multiplexed CRISPRi Repression of Respiratory Functions Enhances Mitochondrial Localized Ethyl Acetate Biosynthesis in *Kluyveromyces marxianus*. *ACS Synth. Biol.* **7**, 2647–2655 (2018).
117. Robertson, N. R. *et al.* PYR1 Biosensor-Driven Genome-Wide CRISPR Screens for Improved Monoterpene Production in *Kluyveromyces marxianus*. *ACS Synth. Biol.* (2025) doi:10.1021/ACSSYNBIO.4C00797.
118. Ha-Tran, D. M. *et al.* Construction of engineered RuBisCO *Kluyveromyces marxianus* for a dual microbial bioethanol production system. *PLoS One* **16**, e0247135 (2021).
119. Christensen, A. D., Kádár, Z., Oleskowicz-Popiel, P. & Thomsen, M. H. Production of

- bioethanol from organic whey using *Kluyveromyces marxianus*. *J. Ind. Microbiol. Biotechnol.* **38**, 283–289 (2011).
120. Hensing, M., Vrouwenvelder, H., Hellinga, C., Baartmans, R. & van Dijken, H. Production of extracellular inulinase in high-cell-density fed-batch cultures of *Kluyveromyces marxianus*. *Appl. Microbiol. Biotechnol.* **42**, 516–521 (1994).
 121. Zhang, Y. *et al.* Inulinase hyperproduction by *kluyveromyces marxianus* through codon optimization, selection of the promoter, and high-cell-density fermentation for efficient inulin hydrolysis. *Ann. Microbiol.* **69**, 647–657 (2019).
 122. Zhou, H. X., Xin, F. H., Chi, Z., Liu, G. L. & Chi, Z. M. Inulinase production by the yeast *Kluyveromyces marxianus* with the disrupted MIG1 gene and the over-expressed inulinase gene. *Process Biochem.* **49**, 1867–1874 (2014).
 123. Aggelopoulos, T. *et al.* Solid state fermentation of food waste mixtures for single cell protein, aroma volatiles and fat production. *Food Chem.* **145**, 710–716 (2014).
 124. Alasmar, R. *et al.* Isolation of a Novel *Kluyveromyces marxianus* Strain QKM-4 and Evidence of Its Volatilome Production and Binding Potentialities in the Biocontrol of Toxigenic Fungi and Their Mycotoxins. *ACS Omega* **5**, 17637–17645 (2020).
 125. Inokuma, K. *et al.* Complete Genome Sequence of *Kluyveromyces marxianus* NBRC1777, a Nonconventional Thermotolerant Yeast. *Genome Announc.* **3**, (2015).
 126. Limtong, S., Sringiew, C. & Yongmanitchai, W. Production of fuel ethanol at high temperature from sugar cane juice by a newly isolated *Kluyveromyces marxianus*. *Bioresour. Technol.* **98**, 3367–3374 (2007).
 127. Teofilo, H., Miguel, U. & Irma, F. Descripcion De Una Especie Nueva de Hansenula Y una Variedad Nueva de Candida Parapsilosis Aisladas del Pozol. *Bol. Soc. Mex. Mic* **7**, (1973).
 128. Lee, M. E., DeLoache, W. C., Cervantes, B. & Dueber, J. E. A Highly Characterized Yeast Toolkit for Modular, Multipart Assembly. *ACS Synth. Biol.* **4**, 975–986 (2015).
 129. Tran, V. G. *et al.* An end-to-end pipeline for succinic acid production at an industrially relevant scale using *Issatchenkia orientalis*. *Nat. Commun.* **2023 141 14**, 1–14 (2023).
 130. Petzoldt, T. growthrates: Estimate Growth Rates from Experimental Data. *CRAN* (2022).
 131. Pyne, M. E. *et al.* An engineered Aro1 protein degradation approach for increased cis,cis-muconic acid biosynthesis in *Saccharomyces cerevisiae*. *Appl. Environ. Microbiol.* **84**, (2018).
 132. Gietz, R. D. & Schiestl, R. H. High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nat. Protoc.* **2007 21 2**, 31–34 (2007).
 133. Pacific Biosciences. *SMRT Link: Explore and Analyze Your Data With Confidence*. <https://www.pacb.com/wp-content/uploads/Product-Brochure-SMRT-Link-Explore-and-Analyze-Your-Data-with-Confidence.pdf> (2020).
 134. Smit, A., Hubley, R. & Green, P. RepeatMasker Open-4.0. <http://www.repeatmasker.org>

- (2015).
135. Hoff, K. J. & Stanke, M. Predicting Genes in Single Genomes with AUGUSTUS. *Curr. Protoc. Bioinforma.* **65**, (2019).
 136. Buchfink, B., Reuter, K. & Drost, H. G. Sensitive protein alignments at tree-of-life scale using DIAMOND. *Nat. Methods* **18**, 366–368 (2021).
 137. Engler, C., Kandzia, R. & Marillonnet, S. A one pot, one step, precision cloning method with high throughput capability. *PLoS One* **3**, (2008).
 138. Ryan, O. W. *et al.* Selection of chromosomal DNA libraries using a multiplex CRISPR system. *Elife* **3**, 1–15 (2014).
 139. Bean, B. D. M., Whiteway, M. & Martin, V. J. J. The MyLO CRISPR-Cas9 toolkit: a markerless yeast localization and overexpression CRISPR-Cas9 toolkit. *G3 Genes|Genomes|Genetics* **12**, jkac154 (2022).
 140. Stemmer, M., Thumberger, T., Del Sol Keyer, M., Wittbrodt, J. & Mateo, J. L. CCTop: An Intuitive, Flexible and Reliable CRISPR/Cas9 Target Prediction Tool. *PLoS One* **10**, e0124633 (2015).
 141. Labuhn, M. *et al.* Refined sgRNA efficacy prediction improves large- and small-scale CRISPR–Cas9 applications. *Nucleic Acids Res.* **46**, 1375 (2017).
 142. Zhang, Y. *et al.* A gRNA-tRNA array for CRISPR-Cas9 based rapid multiplexed genome editing in *Saccharomyces cerevisiae*. *Nat. Commun.* **10**, 1–10 (2019).
 143. Dykstra, C. B., Pyne, M. E. & Martin, V. J. J. CRAPS: Chromosomal-Repair-Assisted Pathway Shuffling in Yeast. *ACS Synth. Biol.* **12**, 2578–2587 (2023).
 144. Todd, R. T., Braverman, A. L. & Selmecki, A. Flow Cytometry Analysis of Fungal Ploidy. *Curr. Protoc. Microbiol.* **50**, e58 (2018).
 145. Dubencovs, K. *et al.* Optimization of synthetic media composition for *Kluyveromyces marxianus* fed-batch cultivation. *Fermentation* **7**, 1–21 (2021).
 146. Heijnen, J. J. & Kleerebezem, R. Bioenergetics of Microbial Growth. *Encycl. Ind. Biotechnol.* 1–66 (2010) doi:10.1002/9780470054581.EIB084.
 147. Krzystek, L. & Ledakowicz, S. Stoichiometric analysis of *Kluyveromyces fragilis* growth on lactose. *J. Chem. Technol. Biotechnol. J Chem Technol Biotechnol* **75**, 1110–1118 (2000).
 148. Kuznetsov, D. *et al.* OrthoDB v11: annotation of orthologs in the widest sampling of organismal diversity. *Nucleic Acids Res.* **51**, D445–D451 (2023).
 149. Shaner, N. C. *et al.* A bright monomeric green fluorescent protein derived from *Branchiostoma lanceolatum*. *Nat. Methods* **10**, 407–409 (2013).
 150. Wilson, T. E., Grawunder, U. & Lieber, M. R. Yeast DNA ligase IV mediates non-homologous DNA end joining. *Nat.* 1997 3886641 **388**, 495–498 (1997).
 151. Willmore, E. *et al.* A novel DNA-dependent protein kinase inhibitor, NU7026, potentiates

- the cytotoxicity of topoisomerase II poisons used in the treatment of leukemia. *Blood* **103**, 4659–4665 (2004).
152. Yang, H. *et al.* Methods Favoring Homology-Directed Repair Choice in Response to CRISPR/Cas9 Induced-Double Strand Breaks. *Int. J. Mol. Sci.* **21**, 6461 (2020).
 153. Joberty, G. *et al.* A Tandem Guide RNA-Based Strategy for Efficient CRISPR Gene Editing of Cell Populations with Low Heterogeneity of Edited Alleles. *Cris. J.* **3**, 123 (2020).
 154. Acosta, S., Fiore, L., Carota, I. A. & Oliver, G. Use of two gRNAs for CRISPR/Cas9 improves bi-allelic homologous recombination efficiency in mouse embryonic stem cells. *Genesis* **56**, e23212 (2018).
 155. Hickman, M. A. *et al.* The ‘obligate diploid’ *Candida albicans* forms mating-competent haploids. *Nature* **494**, 55 (2013).
 156. Walker, G. M. Metals in Yeast Fermentation Processes. *Adv. Appl. Microbiol.* **54**, 197–229 (2004).
 157. Löser, C., Urit, T., Gruner, E. & Bley, T. Efficient growth of *Kluyveromyces marxianus* biomass used as a biocatalyst in the sustainable production of ethyl acetate. *Energy. Sustain. Soc.* **5**, 2 (2015).
 158. Hensing, M. C. M. *et al.* Effects of cultivation conditions on the production of heterologous α -galactosidase by *Kluyveromyces lactis*. *Appl. Microbiol. Biotechnol.* **43**, 58–64 (1995).
 159. Rech, R., Cassini, C. F., Secchi, A. & Ayub, M. A. Z. Utilization of protein-hydrolyzed cheese whey for production of β -galactosidase by *Kluyveromyces marxianus*. *J. Ind. Microbiol. Biotechnol.* **23**, 91–96 (1999).
 160. Bellaver, L. H., Barbosa De Carvalho, N. M., Abrahão-Neto, J. & Gombert, A. K. Ethanol formation and enzyme activities around glucose-6-phosphate in *Kluyveromyces marxianus* CBS 6556 exposed to glucose or lactose excess. *FEMS Yeast Res.* **4**, 691–698 (2004).
 161. Rouwenhorst, R. J., Visser, L. E., Van Der Baan, A. A., Scheffers, W. A. & Van Dijken, J. P. Production, Distribution, and Kinetic Properties of Inulinase in Continuous Cultures of *Kluyveromyces marxianus* CBS 6556. *Appl. Environ. Microbiol.* **54**, 1131–1137 (1988).
 162. Lertwattanasakul, N. *et al.* Genetic basis of the highly efficient yeast *Kluyveromyces marxianus*: Complete genome sequence and transcriptome analyses. *Biotechnol. Biofuels* **8**, (2015).
 163. Mukherjee, P. K. *et al.* Socio-economic sustainability with circular economy — An alternative approach. *Sci. Total Environ.* **904**, 166630 (2023).
 164. Gómez-Márquez, C. *et al.* Diploid genome assembly of *Kluyveromyces marxianus* NRRL Y-50883 (SLP1). *G3 Genes|Genomes|Genetics* **12**, (2022).
 165. Mercier, G. *et al.* A haploid-specific transcriptional response to irradiation in

- Saccharomyces cerevisiae*. *Nucleic Acids Res.* **33**, 6635–6643 (2005).
166. Galitski, T., Saldanha, A. J., Styles, C. A., Lander, E. S. & Fink, G. R. Ploidy regulation of gene expression. *Science* (80-.). **285**, 251–254 (1999).
 167. Keren, L. *et al.* Promoters maintain their relative activity levels under different growth conditions. *Mol. Syst. Biol.* **9**, 701 (2013).
 168. Banerjee, S., Smith, S. & Myung, K. Suppression of gross chromosomal rearrangements by yKu70-yKu80 heterodimer through DNA damage checkpoints. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 1816–1821 (2006).
 169. Boulton, S. J. & Jackson, S. P. Identification of a *Saccharomyces cerevisiae* Ku80 homologue: Roles in DNA double strand break rejoining and in telomeric maintenance. *Nucleic Acids Res.* **24**, 4639–4648 (1996).
 170. Varela, J. A., Puricelli, M., Montini, N. & Morrissey, J. P. Expansion and Diversification of MFS Transporters in *Kluyveromyces marxianus*. *Front. Microbiol.* **10**, 3330 (2019).
 171. Ilica, R. A., Kloetzer, L., Galaction, A. I. & Cașcaval, D. Fumaric acid: production and separation. *Biotechnol. Lett.* **41**, 47–57 (2019).
 172. Van De Vyver, S. & Román-Leshkov, Y. Emerging catalytic processes for the production of adipic acid. *Catal. Sci. Technol.* **3**, 1465–1479 (2013).
 173. Thakker, C., San, K. Y. & Bennett, G. N. Production of succinic acid by engineered *E. coli* strains using soybean carbohydrates as feedstock under aerobic fermentation conditions. *Bioresour. Technol.* **130**, 398–405 (2013).
 174. Zheng, Z., Chen, T., Zhao, M., Wang, Z. & Zhao, X. Engineering *Escherichia coli* for succinate production from hemicellulose via consolidated bioprocessing. *Microb. Cell Fact.* **11**, 1–11 (2012).
 175. Liu, R. *et al.* Succinate production by metabolically engineered *Escherichia coli* using sugarcane bagasse hydrolysate as the carbon source. *Bioresour. Technol.* **135**, 574–577 (2013).
 176. Ju, J. H. *et al.* Production of adipic acid by short- and long-chain fatty acid acyl-CoA oxidase engineered in yeast *Candida tropicalis*. *Bioprocess Biosyst. Eng.* **43**, 33–43 (2020).
 177. Jiang, M. *et al.* Succinic acid production from cellobiose by *Actinobacillus succinogenes*. *Bioresour. Technol.* **135**, 469–474 (2013).
 178. Zhang, B., Skory, C. D. & Yang, S. T. Metabolic engineering of *Rhizopus oryzae*: Effects of overexpressing *pyc* and *pepc* genes on fumaric acid biosynthesis from glucose. *Metab. Eng.* **14**, 512–520 (2012).
 179. Thornbury, M. *et al.* Sequencing of a Dairy Isolate Unlocks *Kluyveromyces marxianus* as a Host for Lactose Valorization. *ACS Synth. Biol.* **14**, 2667–2680 (2025).
 180. Lee, M. H. *et al.* Genome-wide prediction of CRISPR/Cas9 targets in *Kluyveromyces marxianus* and its application to obtain a stable haploid strain. *Sci. Rep.* **8**, (2018).

181. Zhou, H. *et al.* Efficient and markerless gene integration with SlugCas9-HF in *Kluyveromyces marxianus*. *Commun. Biol.* 2024 71 **7**, 1–10 (2024).
182. Ryan, M. P. & Walsh, G. The biotechnological potential of whey. *Rev. Environ. Sci. Bio/Technology* 2016 153 **15**, 479–498 (2016).
183. Martin-Dominguez, V., Estevez, J., De Borja Ojembarrena, F., Santos, V. E. & Ladero, M. Fumaric acid production: A biorefinery perspective. *Fermentation* **4**, (2018).
184. Huang, L., Wei, P., Zang, R., Xu, Z. & Cen, P. High-throughput screening of high-yield colonies of *Rhizopus oryzae* for enhanced production of fumaric acid. *Ann. Microbiol.* 2010 602 **60**, 287–292 (2010).
185. Wei, L., Liu, J., Qi, H. & Wen, J. Engineering *Scheffersomyces stipitis* for fumaric acid production from xylose. *Bioresour. Technol.* **187**, 246–254 (2015).
186. Rhodes, R. A. *et al.* Production of Fumaric Acid in 20-Liter Fermentors. *Appl. Microbiol.* **10**, 9–15 (1962).
187. Lian, J., Hamedirad, M., Hu, S. & Zhao, H. Combinatorial metabolic engineering using an orthogonal tri-functional CRISPR system. *Nat. Commun.* **8**, (2017).
188. Lian, J., Schultz, C., Cao, M., Hamedirad, M. & Zhao, H. Multi-functional genome-wide CRISPR system for high throughput genotype–phenotype mapping. *Nat. Commun.* **10**, 1–10 (2019).
189. Montague, T. G., Cruz, J. M., Gagnon, J. A., Church, G. M. & Valen, E. CHOPCHOP: a CRISPR/Cas9 and TALEN web tool for genome editing. *Nucleic Acids Res.* **42**, W401–W407 (2014).
190. Labun, K., Montague, T. G., Gagnon, J. A., Thyme, S. B. & Valen, E. CHOPCHOP v2: a web tool for the next generation of CRISPR genome engineering. *Nucleic Acids Res.* **44**, W272–W276 (2016).
191. Labun, K. *et al.* CHOPCHOP v3: expanding the CRISPR web toolbox beyond genome editing. *Nucleic Acids Res.* **47**, W171–W174 (2019).
192. Green, M. R. & Sambrook, J. Precipitation of DNA with Isopropanol. (2017) doi:10.1101/pdb.prot093385.
193. Cock, P. J. A., Fields, C. J., Goto, N., Heuer, M. L. & Rice, P. M. The Sanger FASTQ file format for sequences with quality scores, and the Solexa/Illumina FASTQ variants. *Nucleic Acids Res.* **38**, 1767 (2010).
194. Ewels, P., Magnusson, M., Lundin, S. & Käller, M. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* **32**, 3047–3048 (2016).
195. Zhang, J., Kobert, K., Flouri, T. & Stamatakis, A. PEAR: a fast and accurate Illumina Paired-End reAd mergeR. *Bioinformatics* **30**, 614–620 (2014).
196. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**, 10–12 (2011).

197. Li, W. *et al.* MAGECK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens. *Genome Biol.* **15**, 554 (2014).
198. Madeira, F. *et al.* The EMBL-EBI Job Dispatcher sequence analysis tools framework in 2024. *Nucleic Acids Res.* **52**, W521–W525 (2024).
199. Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M. & Barton, G. J. Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**, 1189–1191 (2009).
200. Hallgren, J. *et al.* DeepTMHMM predicts alpha and beta transmembrane proteins using deep neural networks. *bioRxiv* 2022.04.08.487609 (2022)
doi:10.1101/2022.04.08.487609.
201. Gbelska, Y., Krijger, J. J. & Breunig, K. D. Evolution of gene families: the multidrug resistance transporter genes in five related yeast species. *FEMS Yeast Res.* **6**, 345–355 (2006).
202. Jensen, E. D. *et al.* Transcriptional reprogramming in yeast using dCas9 and combinatorial gRNA strategies. *Microb. Cell Fact.* **16**, 1–16 (2017).
203. Liao, X., Li, L., Jameel, A., Xing, X. H. & Zhang, C. A versatile toolbox for CRISPR-based genome engineering in *Pichia pastoris*. *Appl. Microbiol. Biotechnol.* **105**, 9211–9218 (2021).
204. Nishida-Aoki, N., Mori, H., Kuroda, K. & Ueda, M. Activation of the mitochondrial signaling pathway in response to organic solvent stress in yeast. *Curr. Genet.* **61**, 153–164 (2015).
205. Dong, C. *et al.* Identification of novel metabolic engineering targets for S-adenosyl-L-methionine production in *Saccharomyces cerevisiae* via genome-scale engineering. *Metab. Eng.* **66**, 319–327 (2021).
206. Woodruff, L. B. A., Boyle, N. R. & Gill, R. T. Engineering improved ethanol production in *Escherichia coli* with a genome-wide approach. *Metab. Eng.* **17**, 1–11 (2013).
207. Bao, Z. *et al.* Genome-scale engineering of *Saccharomyces cerevisiae* with single-nucleotide precision. *Nat. Biotechnol.* 2018 366 **36**, 505–508 (2018).
208. Bock, C. *et al.* High-content CRISPR screening. *Nat. Rev. Methods Prim.* **2**, 9 (2022).
209. Tenreiro, S., Vargas, R. C., Teixeira, M. C., Magnani, C. & Sá-Correia, I. The yeast multidrug transporter Qdr3 (Ybr043c): localization and role as a determinant of resistance to quinidine, barban, cisplatin, and bleomycin. *Biochem. Biophys. Res. Commun.* **327**, 952–959 (2005).
210. Vargas, R. C. *et al.* *Saccharomyces cerevisiae* Multidrug Resistance Transporter Qdr2 Is Implicated in Potassium Uptake, Providing a Physiological Advantage to Quinidine-Stressed Cells. *Eukaryot. Cell* **6**, 134 (2006).
211. Goffeau, A. U. *et al.* Multidrug-Resistant Transport Proteins in Yeast: Complete Inventory and Phylogenetic Characterization of Yeast Open Reading Frames within the Major

- Facilitator Superfamily. **13**, 43–54 (1997).
212. Dias, P. J., Seret, M. L., Goffeau, A., Correia, I. S. & Baret, P. V. Evolution of the 12-spanner drug:H⁺ antiporter DHA1 family in hemiascomycetous yeasts. *OMICS* **14**, 701–710 (2010).
 213. Ríos, G. *et al.* Role of the yeast multidrug transporter Qdr2 in cation homeostasis and the oxidative stress response. *FEMS Yeast Res.* **13**, 97–106 (2013).
 214. Kapetanakis, G. C., Gournas, C., Prévost, M., Georis, I. & André, B. Overlapping Roles of Yeast Transporters Aqr1, Qdr2, and Qdr3 in Amino Acid Excretion and Cross-Feeding of Lactic Acid Bacteria. *Front. Microbiol.* **12**, (2021).
 215. Pereira, R. *et al.* Adaptive laboratory evolution of tolerance to dicarboxylic acids in *Saccharomyces cerevisiae*. *Metab. Eng.* **56**, 130–141 (2019).
 216. Ma, C. *et al.* Improvement of acetic acid tolerance of *Saccharomyces cerevisiae* using a zinc-finger-based artificial transcription factor and identification of novel genes involved in acetic acid tolerance. *Appl. Microbiol. Biotechnol.* **99**, 2441–2449 (2015).
 217. Sá-Correia, I. & Godinho, C. P. Exploring the biological function of efflux pumps for the development of superior industrial yeasts. *Curr. Opin. Biotechnol.* **74**, 32–41 (2022).
 218. Legras, J. L. *et al.* Activation of Two Different Resistance Mechanisms in *Saccharomyces cerevisiae* upon Exposure to Octanoic and Decanoic Acids. *Appl. Environ. Microbiol.* **76**, 7526 (2010).
 219. Holyoak, C. D., Bracey, D., Piper, P. W., Kuchler, K. & Coote, P. J. The *Saccharomyces cerevisiae* weak-acid-inducible ABC transporter Pdr12 transports fluorescein and preservative anions from the cytosol by an energy-dependent mechanism. *J. Bacteriol.* **181**, 4644–4652 (1999).
 220. Nygård, Y. *et al.* The diverse role of Pdr12 in resistance to weak organic acids. *Yeast* **31**, 219–232 (2014).
 221. Tenreiro, S. *et al.* AQR1 Gene (ORF YNL065w) Encodes a Plasma Membrane Transporter of the Major Facilitator Superfamily That Confers Resistance to Short-Chain Monocarboxylic Acids and Quinidine in *Saccharomyces cerevisiae*. *Biochem. Biophys. Res. Commun.* **292**, 741–748 (2002).
 222. Godinho, C. P. *et al.* Yeast response and tolerance to benzoic acid involves the Gcn4- and Stp1-regulated multidrug/multixenobiotic resistance transporter Tpo1. *Appl. Microbiol. Biotechnol.* **101**, 5005–5018 (2017).
 223. Zhang, X., Nijland, J. G. & Driessen, A. J. M. Combined roles of exporters in acetic acid tolerance in *Saccharomyces cerevisiae*. *Biotechnol. Biofuels Bioprod.* **15**, 1–13 (2022).
 224. Tenreiro, S., Rosa, P. C., Viegas, C. A. & Sa Â-Correia, I. Expression of the AZR1 gene (ORF YGR224w), encoding a plasma membrane transporter of the major facilitator superfamily, is required for adaptation to acetic acid and resistance to azoles in *Saccharomyces cerevisiae*. doi:10.1002/1097-0061(200012)16:16.

225. Alfatah, M. *et al.* Uncharacterized yeast gene YBR238C, an effector of TORC1 signaling in a mitochondrial feedback loop, accelerates cellular aging via HAP4- and RMD9-dependent mechanisms. *Elife* **12**, RP92178 (2024).
226. Kaeberlein, M., Kirkland, K. T., Fields, S. & Kennedy, B. K. Genes determining yeast replicative life span in a long-lived genetic background. *Mech. Ageing Dev.* **126**, 491–504 (2005).
227. Defossez, P.-A. *et al.* Elimination of Replication Block Protein Fob1 Extends the Life Span of Yeast Mother Cells. *Mol. Cell* **3**, 447–455 (1999).
228. Norais, N., Prome, D. & Velours, J. ATP Synthase of Yeast Mitochondria: CHARACTERIZATION OF SUBUNIT d AND SEQUENCE ANALYSIS OF THE STRUCTURAL GENE ATP7. *J. Biol. Chem.* **266**, 16541–16549 (1991).
229. Weber, J. & Senior, A. E. Catalytic mechanism of F1-ATPase. *Biochim. Biophys. Acta* **1319**, 19–58 (1997).
230. Yang, X., Yang, Y., Huang, J., Man, D. & Guo, M. Comparisons of urea or ammonium on growth and fermentative metabolism of *Saccharomyces cerevisiae* in ethanol fermentation. *World J. Microbiol. Biotechnol.* **37**, (2021).
231. Pentjuss, A. *et al.* Model-based biotechnological potential analysis of *Kluyveromyces marxianus* central metabolism. *J. Ind. Microbiol. Biotechnol.* **44**, 1177–1190 (2017).
232. Cui, Z. *et al.* Engineering of unconventional yeast *Yarrowia lipolytica* for efficient succinic acid production from glycerol at low pH. *Metab. Eng.* (2017) doi:10.1016/j.ymben.2017.06.007.
233. Kim, H. K. *et al.* Deep learning improves prediction of CRISPR-Cpf1 guide RNA activity. *Nat. Biotechnol.* **36**, 239–241 (2018).
234. Xu, H. *et al.* Sequence determinants of improved CRISPR sgRNA design. *Genome Res.* **25**, 1147–1157 (2015).
235. Doench, J. G. *et al.* Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. *Nat. Biotechnol.* **32**, 1262–1267 (2014).
236. Guo, F. *et al.* Current advances on biological production of fumaric acid. *Biochem. Eng. J.* **153**, 107397 (2020).
237. Wang, J., Lin, M., Xu, M. & Yang, S. T. Anaerobic Fermentation for Production of Carboxylic Acids as Bulk Chemicals from Renewable Biomass. *Adv. Biochem. Eng. Biotechnol.* **156**, 323–363 (2016).
238. Warnecke, T. & Gill, R. T. Organic acid toxicity, tolerance, and production in *Escherichia coli* biorefining applications. *Microb. Cell Fact.* **4**, (2005).
239. Babaei, M. *et al.* Engineering Oleaginous Yeast as the Host for Fermentative Succinic Acid Production From Glucose. *Front. Bioeng. Biotechnol.* **7**, 1–14 (2019).
240. Pronk, J. T. Pyruvate metabolism in *Saccharomyces cerevisiae*. *Yeast* **12**, 1607–33 (1996).

241. Mohamed, S., Khattab, R. & Watanabe, T. Replacing Glycerol-3-Phosphate Dehydrogenase with NADH Oxidase: Effects on Glucose Fermentation and Product Formation in *Saccharomyces cerevisiae*. *Arch. Microbiol.* **207**, 3 (2025).
242. Yu, T. *et al.* Metabolic reconfiguration enables synthetic reductive metabolism in yeast. *Nat. Metab.* **4**, (2022).
243. Nocon, J. *et al.* Model based engineering of *Pichia pastoris* central metabolism enhances recombinant protein production. *Metab. Eng.* **24**, 129–138 (2014).
244. Kabran, P., Rossignol, T., Gaillardin, C., Nicaud, J. M. & Neuvglise, C. Alternative Splicing Regulates Targeting of Malate Dehydrogenase in *Yarrowia lipolytica*. *DNA Res.* **19**, 231–244 (2012).
245. Kõressaar, T. *et al.* Primer3_masker: integrating masking of template sequence with primer design software. *Bioinformatics* **34**, 1937–1938 (2018).
246. Koressaar, T. & Remm, M. Enhancements and modifications of primer design program Primer3. *Bioinformatics* **23**, 1289–1291 (2007).
247. Untergasser, A. *et al.* Primer3—new capabilities and interfaces. *Nucleic Acids Res.* **40**, e115 (2012).
248. Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* **25**, 402–408 (2001).
249. Teste, M. A., Duquenne, M., François, J. M. & Parrou, J. L. Validation of reference genes for quantitative expression analysis by real-time RT-PCR in *Saccharomyces cerevisiae*. *BMC Mol. Biol.* **10**, 99 (2009).
250. Peetermans, A., Foulquié-Moreno, M. R. & Thevelein, J. M. Mechanisms underlying lactic acid tolerance and its influence on lactic acid production in *Saccharomyces cerevisiae*. *Microb. Cell* **8**, 111 (2021).
251. Pacheco, A. *et al.* Lactic acid production in *Saccharomyces cerevisiae* is modulated by expression of the monocarboxylate transporters Jen1 and Ady2. *FEMS Yeast Res.* **12**, 375–381 (2012).
252. Lodi, T., Fontanesi, F., Ferrero, I. & Donnini, C. Carboxylic acids permeases in yeast: two genes in *Kluyveromyces lactis*. *Gene* **339**, 111–119 (2004).
253. Queirós, O., Pereira, L., Paiva, S., Moradas-Ferreira, P. & Casal, M. Functional analysis of *Kluyveromyces lactis* carboxylic acids permeases: Heterologous expression of KIJEN1 and KIJEN2 genes. *Curr. Genet.* **51**, 161–169 (2007).
254. Ahn, J. H. *et al.* Enhanced succinic acid production by *Mannheimia* employing optimal malate dehydrogenase. *Nat. Commun.* **11**, 1–12 (2020).
255. Zhou, Y. J. *et al.* Modular pathway engineering of diterpenoid synthases and the mevalonic acid pathway for miltiradiene production. *J. Am. Chem. Soc.* **134**, 3234–3241 (2012).
256. Grabowska, D. & Chelstowska, A. The ALD6 Gene Product Is Indispensable for

- Providing NADPH in Yeast Cells Lacking Glucose-6-phosphate Dehydrogenase Activity. *J. Biol. Chem.* **278**, 13984–13988 (2003).
257. Wolfe, K. H. *et al.* Clade- and species-specific features of genome evolution in the Saccharomycetaceae. *FEMS Yeast Res.* **15**, fov035 (2015).
 258. Avendano, A., Deluna, A., Olivera, H., Valenzuela, L. & Gonzalez, A. GDH3 encodes a glutamate dehydrogenase isozyme, a previously unrecognized route for glutamate biosynthesis in *Saccharomyces cerevisiae*. *J. Bacteriol.* **179**, 5594–5597 (1997).
 259. Valdés-Hevia, M. D., de la Guerra, R. & Gancedo, C. Isolation and characterization of the gene encoding phosphoenolpyruvate carboxykinase from *Saccharomyces cerevisiae*. *FEBS Lett.* **258**, 313–316 (1989).
 260. HAARASILTA, S. & OURA, E. On the activity and regulation of anaplerotic and gluconeogenic enzymes during the growth process of baker's yeast. *Eur. J. Biochem.* **52**, 1–7 (1975).
 261. Xu, G., Hua, Q., Duan, N., Liu, L. & Chen, J. Regulation of thiamine synthesis in *Saccharomyces cerevisiae* for improved pyruvate production. *Yeast* **29**, 209–217 (2012).
 262. Kumar, K., Askari, F., Sahu, M. S. & Kaur, R. *Candida glabrata*: A Lot More Than Meets the Eye. *Microorganisms* **7**, 39 (2019).
 263. van Gulik, W. M. & Heijnen, J. J. A metabolic network stoichiometry analysis of microbial growth and product formation. *Biotechnol. Bioeng.* **48**, 681–698 (1995).
 264. Nilsson, A. & Nielsen, J. Metabolic Trade-offs in Yeast are Caused by F1F0-ATP synthase. *Sci. Reports 2016 61* **6**, 1–11 (2016).
 265. Molenaar, D., Van Berlo, R., De Ridder, D. & Teusink, B. Shifts in growth strategies reflect tradeoffs in cellular economics. *Mol. Syst. Biol.* **5**, 323 (2009).
 266. Fonseca, G. G., Gombert, A. K., Heinzle, E. & Wittmann, C. Physiology of the yeast *Kluyveromyces marxianus* during batch and chemostat cultures with glucose as the sole carbon source. *FEMS Yeast Res.* **7**, 422–435 (2007).
 267. van Dijken, J. P. & Scheffers, W. A. Redox balances in the metabolism of sugars by yeasts. *FEMS Microbiol. Rev.* **1**, 199–224 (1986).
 268. Imura, M. *et al.* Comparison of metabolic profiles of yeasts based on the difference of the Crabtree positive and negative. (2019) doi:10.1016/j.jbiosc.2019.07.007.
 269. Malina, C., Yu, R., Björkeroth, J., Kerkhoven, E. J. & Nielsen, J. Adaptations in metabolism and protein translation give rise to the Crabtree effect in yeast. *Proc. Natl. Acad. Sci.* **118**, e2112836118 (2021).
 270. Campero-Basaldúa, C. *et al.* Diversification of the kinetic properties of yeast NADP-glutamate-dehydrogenase isozymes proceeds independently of their evolutionary origin. *Microbiologyopen* **6**, e00419 (2016).
 271. Sakihama, Y., Hidese, R., Hasunuma, T. & Kondo, A. Increased flux in acetyl-CoA synthetic pathway and TCA cycle of *Kluyveromyces marxianus* under respiratory

- conditions. *Sci. Reports* 2019 91 **9**, 1–10 (2019).
272. Yang, J., Kalhan, S. C. & Hanson, R. W. What Is the Metabolic Role of Phosphoenolpyruvate Carboxykinase? *J. Biol. Chem.* **284**, 27025 (2009).
 273. Zelle, R. M., Trueheart, J., Harrison, J. C., Pronk, J. T. & Van Maris, A. J. A. Phosphoenolpyruvate carboxykinase as the sole anaplerotic enzyme in *saccharomyces cerevisiae*. *Appl. Environ. Microbiol.* **76**, 5383–5389 (2010).
 274. Sabe, H. *et al.* Molecular cloning of the phosphoenolpyruvate carboxylase gene, *ppc*, of *Escherichia coli*. *Gene* **31**, 279–283 (1984).
 275. Flores, C. L. & Gancedo, C. Expression of PEP carboxylase from *Escherichia coli* complements the phenotypic effects of pyruvate carboxylase mutations in *Saccharomyces cerevisiae*. *FEBS Lett.* **412**, 531–534 (1997).
 276. Ding, W. *et al.* Metabolic engineering of threonine catabolism enables *Saccharomyces cerevisiae* to produce propionate under aerobic conditions. *Biotechnol. J.* **17**, (2022).
 277. Alves, É. de P., Morioka, L. R. I. & Suguimoto, H. H. Comparison of bioethanol and beta-galactosidase production by *Kluyveromyces* and *Saccharomyces* strains grown in cheese whey. *Int. J. Dairy Technol.* **72**, 409–415 (2019).
 278. Tesfaw, A., Oner, E. T. & Assefa, F. Evaluating crude whey for bioethanol production using non-*Saccharomyces* yeast, *Kluyveromyces marxianus*. *SN Appl. Sci.* **3**, 1–8 (2021).
 279. Alonso-Vargas, M. *et al.* Optimization of 2-Phenylethanol Production from Sweet Whey Fermentation Using *Kluyveromyces marxianus*. *Ferment.* 2022, Vol. 8, Page 39 **8**, 39 (2022).
 280. Braga, A. R. C., Gomes, P. A. & Kalil, S. J. Formulation of Culture Medium with Agroindustrial Waste for β -Galactosidase Production from *Kluyveromyces marxianus* ATCC 16045. *Food Bioprocess Technol.* **5**, 1653–1663 (2012).
 281. Geijer, C., Ledesma-Amaro, R. & Tomas-Pejo, E. Unraveling the potential of non-conventional yeasts in biotechnology. *FEMS Yeast Res.* **22**, foab071 (2022).
 282. Navarrete, C., Martínez, J. L., Navarrete, C. & Martínez, J. L. Non-conventional yeasts as superior production platforms for sustainable fermentation based bio-manufacturing processes. *AIMS Bioeng.* 2020 4289 **7**, 289–305 (2020).
 283. Lacerda, M. P., Oh, E. J. & Eckert, C. The Model System *Saccharomyces cerevisiae* Versus Emerging Non-Model Yeasts for the Production of Biofuels. *Life* **10**, 299 (2020).
 284. Deparis, Q., Claes, A., Foulquié-Moreno, M. R. & Thevelein, J. M. Engineering tolerance to industrially relevant stress factors in yeast cell factories. *FEMS Yeast Res.* **17**, fox036 (2017).
 285. Mo, W. *et al.* *Kluyveromyces marxianus* developing ethanol tolerance during adaptive evolution with significant improvements of multiple pathways. *Biotechnol. Biofuels* **12**, 1–15 (2019).
 286. Costa, C. E., Romaní, A., Cunha, J. T., Johansson, B. & Domingues, L. Integrated

- approach for selecting efficient *Saccharomyces cerevisiae* for industrial lignocellulosic fermentations: Importance of yeast chassis linked to process conditions. *Bioresour. Technol.* **227**, 24–34 (2017).
287. Korhola, M. *et al.* Production of folate in oat bran fermentation by yeasts isolated from barley and diverse foods. *J. Appl. Microbiol.* **117**, 679–689 (2014).
 288. Löser, C., Urit, T. & Bley, T. Perspectives for the biotechnological production of ethyl acetate by yeasts. *Appl. Microbiol. Biotechnol.* **98**, 5397–5415 (2014).
 289. Piškur, J., Rozpedowska, E., Polakova, S., Merico, A. & Compagno, C. How did *Saccharomyces* evolve to become a good brewer? *Trends Genet.* **22**, 183–186 (2006).
 290. Ikushima, S., Fujii, T. & Kobayashi, O. Efficient gene disruption in the high-ploidy yeast *Candida utilis* using the Cre-loxP system. *Biosci. Biotechnol. Biochem.* **73**, 879–884 (2009).
 291. Fasoli, G., Barrio, E., Tofalo, R., Suzzi, G. & Belloch, C. Multilocus analysis reveals large genetic diversity in *Kluyveromyces marxianus* strains isolated from Parmigiano Reggiano and Pecorino di Farindola cheeses. *Int. J. Food Microbiol.* **233**, 1–10 (2016).
 292. Reina-Posso, D. & Gonzales-Zubiate, F. A. Expanding Horizons: The Untapped Potential of *Kluyveromyces marxianus* in Biotechnological Applications. *Ferment.* **2025**, Vol. 11, Page 98 **11**, 98 (2025).
 293. Shaw, W. M. *et al.* Inducible expression of large gRNA arrays for multiplexed CRISPRai applications. *Nat. Commun.* **2022** *131* **13**, 1–10 (2022).
 294. Laughery, M. F. *et al.* New vectors for simple and streamlined CRISPR–Cas9 genome editing in *Saccharomyces cerevisiae*. *Yeast* **32**, 711–720 (2015).
 295. Radecka, D. *et al.* Looking beyond *Saccharomyces*: the potential of non-conventional yeast species for desirable traits in bioethanol fermentation. *FEMS Yeast Res.* **15**, (2015).
 296. Rogers, J. K. & Church, G. M. Genetically encoded sensors enable real-time observation of metabolite production. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 2388–2393 (2016).
 297. Skjoedt, M. L. *et al.* Engineering prokaryotic transcriptional activators as metabolite biosensors in yeast. *Nat. Chem. Biol.* **12**, 951–958 (2016).
 298. Lv, Y., Gu, Y., Xu, J., Zhou, J. & Xu, P. Coupling metabolic addiction with negative autoregulation to improve strain stability and pathway yield. *Metab. Eng.* **61**, 79–88 (2020).
 299. Wang, G. *et al.* Improvement of cis,cis-Muconic Acid Production in *Saccharomyces cerevisiae* through Biosensor-Aided Genome Engineering. (2020) doi:10.1021/acssynbio.9b00477.
 300. D’Ambrosio, V. *et al.* Regulatory control circuits for stabilizing long-term anabolic product formation in yeast. *Metab. Eng.* **61**, 369–380 (2020).
 301. Fernandez-López, R., Ruiz, R., de la Cruz, F. & Moncalián, G. Transcription factor-based biosensors enlightened by the analyte. *Front. Microbiol.* **6**, 648 (2015).

- 302. Pham, C. *et al.* Functional and structural characterization of an IclR family transcription factor for the development of dicarboxylic acid biosensors. *FEBS J.* **291**, 3481–3498 (2024).
- 303. Acosta-Cuevas, C. K. *et al.* Tolerance of *Kluyveromyces marxianus* Under Acetic Acid-, Isoamyl Alcohol-, Hydrogen Peroxide-, and Ethanol-Induced Stress. *Fermentation* **11**, 171 (2025).
- 304. Zeng, H. & Yang, A. Modelling overflow metabolism in *Escherichia coli* with flux balance analysis incorporating differential proteomic efficiencies of energy pathways. *BMC Syst. Biol.* **13**, 1–18 (2019).
- 305. Chen, X., Dong, X., Wang, Y., Zhao, Z. & Liu, L. Mitochondrial engineering of the TCA cycle for fumarate production. *Metab. Eng.* **31**, 62–73 (2015).
- 306. Cui, Z. *et al.* Reconfiguration of the reductive TCA cycle enables high-level succinic acid production by *Yarrowia lipolytica*. *Nat. Commun.* **2023 141** **14**, 1–11 (2023).