UNIVERSITY OF CALIFORNIA
RIVERSIDE

Understanding and Engineering Ester Biosynthesis Pathways in the Yeast *Kluyveromyces marxianus*

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Chemical and Environmental Engineering

by

Ann-Kathrin Loebs

June 2018

Dissertation Committee:
Dr. Ian Wheeldon, Chairperson
Dr. Xin Ge
Dr. Charles Wyman

The Dissertation of Ann-Kathrin Loebs is approved:

_____

_____

_____

Committee Chairperson

University of California, Riverside

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my advisor Dr. Ian Wheeldon for his guidance, support and confidence in me and my work throughout the past five years. Our weekly research discussions, and your willingness to support my ideas about research directions, have made me the independent researcher I am today.

Thank you, also, to my lab mates who have made the last 5 years a time to cherish and look back on. Thank you for all your help, the countless discussions about research and a friendly atmosphere.

The text of this dissertation, in part, is a reprint of the material as it appears in the journals of Synthetic and Systems Biotechnology (2017, chapter 2), Biotechnology Journal (2016, chapter 3) and Biotechnology for Biofuels (2017, chapter 4). The co-author Ian Wheeldon, listed in those publications, directed and supervised the research, which forms the basis for this dissertation. Thank you to Cory Schwartz, who co-wrote the review listed in chapter 2 and his contribution to the work in chapter 4. Thank you also to Megan Cook, Andrew Flores and Ronja Engel, who performed experiments that contributed to the success of the work in chapter 3 and 4. I would also like to acknowledge UC Riverside and the National Science Foundation for supporting the work in this thesis.

Thanks to my LaLaFam, who kept me sane and helped me enjoy what LA and the surrounding area has to offer. I also couldn't have made it with the continuing support of my friends back in Germany.

I want to thank my family for their support and love. Without having such great parents I would not have made it so far. You always stand behind me and support my plans. Thanks Philipp for being the brother that you are.

Last but not least, I want to express my gratitude to my husband, Tyler Lamb, for your patience and continuing support. Your positive attitude and your ability to cheer me up and motivate me whenever necessary was essential in getting me to this point. Thank you for taking me on many adventures during our time together. I am looking forward to see what the future has in store for us!

ABSTRACT OF THE DISSERTATION

Understanding and Engineering Ester Biosynthesis Pathways in the Yeast *Kluyveromyces marxianus*

by

Ann-Kathrin Loebs

Doctor of Philosophy, Graduate Program in Chemical and Environmental Engineering
University of California, Riverside, June 2018
Dr. Ian Wheeldon, Chairperson

The threat of climate change and a recent sway in popular opinion towards sustainable energy and chemicals have fueled the field of biotechnology for renewable chemicals production. To achieve process feasibility and compete with fossil-based processes, microbial production units are required to synthesize products at high titers and productivities while utilizing cheap and sustainable substrates. To fulfill these requirements careful selection of the host organism is essential. The emergence of efficient genome editing tools has enabled engineering of, thus far, intractable organisms, and allows for the selection of a host organism based on a desired phenotype that is beneficial for the process. The yeast *Kluyveromyces marxianus* was chosen because of its natural capacity to produce high amounts of ethyl acetate. Other characteristics such as fast growth kinetics, thermotolerance and the ability to metabolize various carbon sources make this host especially interesting for industrial applications. The development of an efficient CRISPR-Cas9 system in *K. marxianus* allowed us to interrogate the role of alcohol acetyltransferases and alcohol dehydrogenases in volatile metabolite production. We identified Eat1 as the critical enzyme for acetate ester production, and found that mitochondrial localization of Eat1 is essential for high ester production yields. Overexpression of Eat1 significantly

increased ester production indicating that this step is the bottleneck of the reaction. To furthermore increase ester production TCA cycle flux was slowed down through CRISPR interference-mediated knockdown of the TCA cycle and electron transport chain.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1: Introduction

1.1 Background

Amid climate change and the finite availability of fossil fuels researchers have been striving towards chemical production from renewable resources. The production of sustainable chemicals can be catalyzed using microorganisms such as yeasts and bacteria.[1] While conventional synthetic chemistry processes often require high-energy input and harsh conditions, microbial fermentations can be performed at mild or ambient conditions.

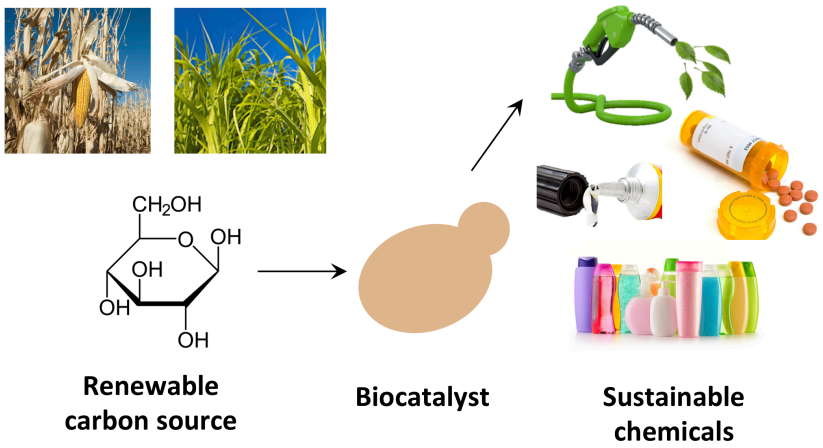


Figure 1.1: Microorganisms serve as biocatalysts to convert renewable carbon sources to sustainable chemicals

Historically, processes for biotechnological chemical production were developed based on the capacity of different microorganisms to produce a certain product. Some prominent examples include the productions of penicillin using the fungus *Penicillium chrysogenum,* glutamic acid by the bacterium *Corynebacterium glutamicum* or ethanol using the yeast *S. cerevisiae*.[2-4]

Because of its importance in ethanol production *S. cerevisiae* has been extensively studied and developed as a model organisms. In depth knowledge about the metabolism and physiology, and the availability of synthetic biology tools has promoted the development of this yeast for advanced biofuels and specialty chemicals production. Because of its high fermentative capacity and its crabtree-positive character, *S. cerevisiae* is the perfect candidate for ethanol production; however, alternative processes are oftentimes hampered by low productivity due to detrimental physiology and metabolism. One caveat of biotechnological chemicals production is the need for high titers to achieve economic feasibility and compete with conventional processes.

In recent years, research groups and companies have fanned out to interrogate and engineer myriads of organisms for production of advanced biofuels and specialty chemicals from cheap and renewable substrates such as biomass or whey. Here the safety of a model organism such as *S. cerevisiae* is left behind to choose organisms that excel at a certain task or posses a certain beneficial phenotype critical for high productivity.[5] For example, the oleaginous yeast *Yarrowia lipolytica* is chosen for biotechnological production of lipids and fatty acid-derived chemicals.[6, 7] The native ability of *Y. lipolytica* to shuttle a large amount of carbon in the form of acetyl-CoA away from cell growth pathways and towards chemicals production has been harnessed to produce several commercial products.[8, 9]

Yeasts from the *Saccharomycetaceae* family have also gained interest. On particular species *Kluyveromyces marxianus* has a high native ability to synthesize ethyl acetate and is thermotolerant to temperatures upward of 50°C.[10] One wild type strain of *K. marxianus* is able to produce upward of 2 g L^{-1} h^{-1} of ethyl acetate from waste whey on a pilot scale.[11] In addition to its high ester production and thermotolerance, *K. marxianus* has fast growth

kinetics and is able to metabolize a variety of substrates.[12, 13] However, the metabolism and ester production pathways have not been fully elucidated.[12, 14, 15]

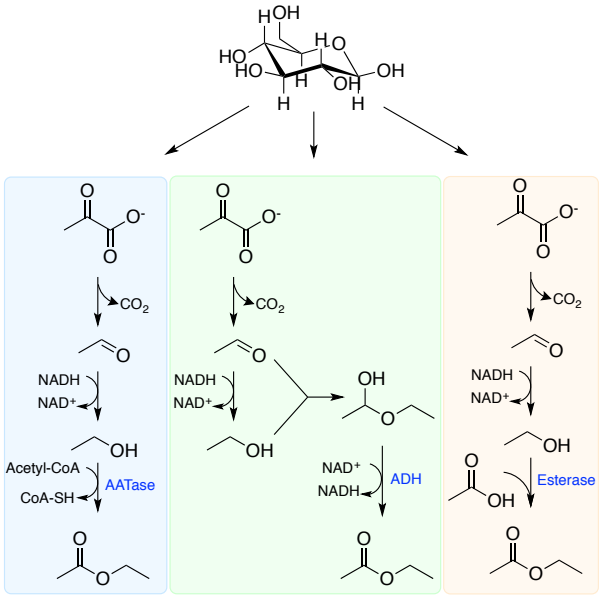


Figure 1.2: Pathways associated with ester biosynthesis in yeast. Ethyl acetate production may be synthesized through three enzymatic pathways. The first pathway relies on the activity of an alcohol acetyltransferase for condensation of ethanol with acetyl-coA. Ethyl acetate may also be synthesized through oxidation of a spontaneously formed hemiacetal to ethyl acetate. Further, reverse esterase activity may condense ethanol with acetate to ethyl acetate.

Ester production in yeast can occur through 3 different pathways as shown in Figure 1.2.[16] Ester production in *S. cerevisiae* has mainly been attributed to activities of the alcohol acetyltransferases (AATase) Atf1 and 2, where Atf1 is the main contributor to ethyl acetate formation. Atf1 and 2, are localized to the endoplasmic reticulum and lipid droplets and localization has been shown to be essential for high activity.[17-19] Double knockouts of Atf1 and 2 lead to a reduction of ethyl acetate by 50% hinting to the presence of alternative production routes.[21] It has been found that different alcohol dehydrogenases (Adh) possess the ability to oxidize a spontaneously formed hemiacetal to an ester. These activities have been observed for a variety of yeast Adh's including *S. cerevisiae* and *Candida utilis*.[21, 22] Additionally, esters can be produced through reverse esterase or lipase activity through condensation of ethanol with acetate. This reaction however is thermodynamically unfavorable.[16]

Recent work led to the discovery of a new alcohol acetyltransferase named ethanol acetyltransferase Eat1 in the yeast *Wickerhamomyces anomalus*.[23] Expression of this Eat1 along with homologs from other yeast species led to significant ethyl acetate production and disruption of Eat1 in *S. cerevisiae* and *K. lactis* led to a decrease in ethyl acetate by 50% and 80%, respectively.[23]

Studies on ester biosynthesis in *K. marxianus* are limited and while prior research suggests the importance of Atf in ester biosynthesis, we previously showed that Atf only marginally contributes to ester production in *K. marxianus*.[12, 14] Alternatively, Eat1 from *K. marxianus* showed the highest ethyl acetate production among the newly discovered Eat1's when expressed in *S. cerevisiae*, thus presenting itself as a possible candidate to contribute to ethyl acetate production in *K. marxianus*.[23]

The type II bacterial CRISPR system (clustered regularly interspaced palindromic repeats) is an efficient genome editing system in a variety of organisms, and has enabled genetic engineering in less genetically tractable organisms.[12, 24, 25] The use of a CRISPR-Cas9 system for genome editing is especially desirable in organisms that favor DNA repair by nonhomologous end-joining over homologous directed repair.[2] *K. marxianus* genome editing has thus far relied on random integration, the screening of a large number of transformants, or the use of a KU70/80 disruption strain to favor directed genome editing.[26-28] In these cases, genome editing is contingent on genetic marker integration and thus necessitates marker recovery for subsequent engineering efforts.[2, 29]

Previous studies suggested that trace metal limitation induces high ethyl acetate production.[14] This effect was thought to be caused by a decrease in carbon flux through the TCA cycle caused by a lack of iron or copper metal centers that are essential for TCA cycle and electron transport chain enzymes.[14] The decreased flux was hypothesized to

increase acetyl-coA availability and thus ethyl acetate production. This hypothesis was tested by addition of electron transport chain inhibitors and while moderate inhibition was beneficial, too much inhibition deemed detrimental to ester production. Similarly, we found that disruption of the TCA cycle decreased growth and ethyl acetate production significantly indicating the necessity for a decreased but not fully disrupted flux through TCA cycle and electron transport chain.

Recently, researchers have utilized the DNA targeting and binding capacity of Cas9 for transcriptional regulation. Two point mutations in the Cas9 nuclease eliminate its ability to cleave DNA, while leaving its DNA binding ability intact.[30] Targeting of this catalytically dead version of the Cas9 (dCas9) to different regions of the promoter or open reading frame (ORF) of the gene of interest can be used to regulate gene expression and thus fine tune metabolic fluxes.[30-32] Development of a CRISPR interference system enabled the study of TCA cycle and ETC knockdowns on ethyl acetate.

1.2 Thesis Organization

The scope of this thesis includes the development of synthetic biology tools, pathway elucidation and engineering of the non-conventional yeasts *Kluyveromyces marxianus* for enhanced ethyl acetate production.

In chapter 2 we discuss the advantages of utilizing non-model yeasts for biotechnological processes and highlight the challenges of engineering non-conventional yeasts.

To facilitate ester screening and improve throughput we developed a colorimetric high-throughput assay that is described in chapter 3. The assay was used to evaluate ester production of different *K. marxianus* strains on a variety of C5, C6 and C12 carbon sources.

The lack of genetic engineering tools and knowledge about metabolic pathways necessitated the development of a CRISPR-Cas9 system for ethyl acetate biosynthesis pathways interrogation in *K. marxianus* and is described in chapter 4. We applied a synthetic RNA polymerase III sgRNA expression system to efficiently disrupt different alcohol dehydrogenases (Adh) and alcohol acetyltransferase (Atf) in *K. marxianus*. This work also evaluated the activity of these enzymes towards ethyl acetate production and found that Adh7 is capable of producing ethyl acetate from hemiacetal. However, Adh7 disruption did not affect ethyl acetate production in *K. marxianus*, suggesting that Adh7 activity has a minimal effect on *in vivo* ethyl acetate production. Further we found that Adh1, 2 and 3 are important for ethanol and ethyl acetate production and that disruption of Adh2 leads to accumulation of acetaldehyde. The disruption of Atf only marginally decreased ethyl acetate biosynthesis suggesting the presence of a yet unknown enzyme for ethyl acetate production.

Chapter 5 describes work on elucidating ester production in *K. marxianus* along with a metabolic engineering approach for increased ester production through overexpressions and transcriptional control. Disruptions showed ethanol alcoholtransferase (Eat1) as critical enzyme for acetate ester formation. Furthermore, mitochondrial localization of Eat1 was found to be critical for high ethyl acetate production. Overexpression of Eat1 increased ethyl acetate production yields and further improvement was achieved by transcriptional repression of TCA cycle and electron transport chain enzymes using a CRISPRi knockdown strategy

1.3 References

[1] Nielsen J, Larsson C, van Maris A, Pronk J: **Metabolic engineering of yeast for production of fuels and chemicals**. *Current opinion in biotechnology* 2013, **24**(3):398-404.

[2] Lobs AK, Schwartz C, Wheeldon I: **Genome and metabolic engineering in non-conventional yeasts: Current advances and applications**. *Synth Syst Biotechnol* 2017, **2**(3):198-207.

[3] Kinoshita S, Udaka S, Shimono M: **Studies on the amino acid fermentation - Part I. Production of L-glutamic acid by various microorganisms**. *J Gen Appl Microbiol* 2004, **50**(6):331-343.

[4] van den Berg MA: **Functional characterisation of penicillin production strains**. *Fungal Biology Reviews* 2010, **24**(1):73-78.

[5] Wagner JM, Alper HS: **Synthetic biology and molecular genetics in non-conventional yeasts: Current tools and future advances**. *Fungal Genet Biol* 2016, **89**:126-136.

[6] Hussain MS, Rodriguez GM, Gao DF, Spagnuolo M, Gambill L, Blenner M: **Recent advances in bioengineering of the oleaginous yeast Yarrowia lipolytica**. *Aims Bioeng* 2016, **3**(4):493-514.

[7] Blazeck J, Hill A, Liu LQ, Knight R, Miller J, Pan A, Otoupal P, Alper HS: **Harnessing Yarrowia lipolytica lipogenesis to create a platform for lipid and biofuel production**. *Nature communications* 2014, **5**.

[8] Qiao KJ, Wasylenko TM, Zhou K, Xu P, Stephanopoulos G: **Lipid production in Yarrowia lipolytica is maximized by engineering cytosolic redox metabolism**. *Nature biotechnology* 2017, **35**(2):173-177.

[9] Xue Z, Sharpe PL, Hong SP, Yadav NS, Xie D, Short DR, Damude HG, Rupert RA, Seip JE, Wang J *et al*: **Production of omega-3 eicosapentaenoic acid by metabolic engineering of Yarrowia lipolytica**. *Nature biotechnology* 2013, **31**(8):734-740.

[10] Banat IM, Nigam P, Marchant R: **Isolation of Thermotolerant, Fermentative Yeasts Growing at 52-Degrees-C and Producing Ethanol at 45-Degrees-C and 50-Degrees-C**. *World J Microb Biot* 1992, **8**(3):259-263.

[11] Loser C, Urit T, Stukert A, Bley T: **Formation of ethyl acetate from whey by Kluyveromyces marxianus on a pilot scale**. *Journal of biotechnology* 2013, **163**(1):17-23.

[12] Lobs AK, Engel R, Schwartz C, Flores A, Wheeldon I: **CRISPR-Cas9-enabled genetic disruptions for understanding ethanol and ethyl acetate biosynthesis in Kluyveromyces marxianus**. *Biotechnology for biofuels* 2017, **10**:164.

[13] Lobs AK, Lin JL, Cook M, Wheeldon I: **High throughput, colorimetric screening of microbial ester biosynthesis reveals high ethyl acetate production from Kluyveromyces marxianus on C5, C6, and C12 carbon sources**. *Biotechnology journal* 2016, **11**(10):1274-1281.

[14] Loser C, Urit T, Keil P, Bley T: **Studies on the mechanism of synthesis of ethyl acetate in Kluyveromyces marxianus DSM 5422**. *Applied microbiology and biotechnology* 2014.

[15] Kallelmhiri H, Miclo A: **Mechanism of Ethyl-Acetate Synthesis by Kluyveromyces-Fragilis**. *Fems Microbiol Lett* 1993, **111**(2-3):207-212.

[16] Loser C, Urit T, Bley T: **Perspectives for the biotechnological production of ethyl acetate by yeasts**. *Applied microbiology and biotechnology* 2014, **98**(12):5397-5415.

[17] Lin JL, Wheeldon I: **Dual N- and C-Terminal Helices Are Required for Endoplasmic Reticulum and Lipid Droplet Association of Alcohol Acetyltransferases in Saccharomyces cerevisiae**. *PloS one* 2014, **9**(8).

[18] Nagasawa N, Bogaki T, Iwamatsu A, Hamachi M, Kumagai C: **Cloning and nucleotide sequence of the alcohol acetyltransferase II gene (ATF2) from Saccharomyces cerevisiae Kyokai No. 7**. *Biosci Biotech Bioch* 1998, **62**(10):1852-1857.

[19] Fujii T, Nagasawa N, Iwamatsu A, Bogaki T, Tamai Y, Hamachi M: **Molecular cloning, sequence analysis, and expression of the yeast alcohol acetyltransferase gene**. *Applied and environmental microbiology* 1994, **60**(8):2786-2792.

[20] Verstrepen KJ, Van Laere SDM, Vanderhaegen BMP, Derdelinckx G, Dufour JP, Pretorius IS, Winderickx J, Thevelein JM, Delvaux FR: **Expression levels of the yeast alcohol acetyltransferase genes ATF1, Lg-ATF1, and ATF2 control the formation of a broad range of volatile esters**. *Applied and environmental microbiology* 2003, **69**(9):5228-5237.

[21] Kusano M, Sakai Y, Kato N, Yoshimoto H, Tamai Y: **A novel hemiacetal dehydrogenase activity involved in ethyl acetate synthesis in Candida utilis**. *J Biosci Bioeng* 1999, **87**(5):690-692.

[22] Kusano M, Sakai Y, Kato N, Yoshimoto H, Sone H, Tamai Y: **Hemiacetal dehydrogenation activity of alcohol dehydrogenases in Saccharomyces cerevisiae**. *Biosci Biotechnol Biochem* 1998, **62**(10):1956-1961.

[23] Kruis AJ, Levisson M, Mars AE, van der Ploeg M, Garces Daza F, Ellena V, Kengen SWM, van der Oost J, Weusthuis RA: **Ethyl acetate production by the elusive alcohol acetyltransferase from yeast**. *Metab Eng* 2017, **41**:92-101.

[24] Schwartz CM, Hussain MS, Blenner M, Wheeldon I: **Synthetic RNA Polymerase III Promoters Facilitate High-Efficiency CRISPR-Cas9-Mediated Genome Editing in Yarrowia lipolytica**. *ACS synthetic biology* 2016, **5**(4):356-359.

[25] DiCarlo JE, Norville JE, Mali P, Rios X, Aach J, Church GM: **Genome engineering in Saccharomyces cerevisiae using CRISPR-Cas systems**. *Nucleic acids research* 2013, **41**(7):4336-4343.

[26] Abdel-Banat BM, 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* 2010, **27**(1):29-39.

[27] Choo JH, Han C, Kim JY, Kang HA: **Deletion of a KU80 homolog enhances homologous recombination in the thermotolerant yeast Kluyveromyces marxianus**. *Biotechnology letters* 2014, **36**(10):2059-2067.

[28] Nonklang S, Abdel-Banat BMA, Cha-Aim K, Moonjai N, Hoshida H, Limtong S, Yamada M, Akada R: **High-Temperature Ethanol Fermentation and Transformation with Linear DNA in the Thermotolerant Yeast Kluyveromyces marxianus DMKU3-1042**. *Applied and environmental microbiology* 2008, **74**(24):7514-7521.

[29] Pecota DC, Rajgarhia V, Da Silva NA: **Sequential gene integration for the engineering of Kluyveromyces marxianus**. *Journal of biotechnology* 2007, **127**(3):408-416.

[30] Smith JD, Suresh S, Schlecht U, Wu M, Wagih O, Peltz G, Davis RW, Steinmetz LM, Parts L, St Onge RP: **Quantitative CRISPR interference screens in yeast identify chemical-genetic interactions and new rules for guide RNA design**. *Genome Biol* 2016, **17**:45.

[31] Schwartz C, Frogue K, Ramesh A, Misa J, Wheeldon I: **CRISPRi repression of nonhomologous end-joining for enhanced genome engineering via homologous recombination in Yarrowia lipolytica**. *Biotechnol Bioeng* 2017, **114**(12):2896-2906.

[32] Deaner M, Mejia J, Alper HS: **Enabling Graded and Large-Scale Multiplex of Desired Genes Using a Dual-Mode dCas9 Activator in Saccharomyces cerevisiae**. *ACS synthetic biology* 2017, **6**(10):1931-1943.

Chapter 2: Genome and metabolic engineering in non-conventional yeasts: current advances and applications

2.1 Abstract

Microbial production of chemicals and proteins from biomass-derived and waste sugar streams is a rapidly growing area of research and development. While the model yeast *Saccharomyces cerevisia*e is an excellent host for the conversion of glucose to ethanol, production of other chemicals from alternative substrates often require extensive strain engineering. To avoid complex and intensive engineering of *S. cerevisiae,* other yeasts are often selected as hosts for bioprocessing based on their natural capacity to produce a desired product: for example, the efficient production and secretion of proteins, lipids, and primary metabolites that have value as commodity chemicals. Even when using yeasts with beneficial native phenotypes, metabolic engineering to increase yields, titer, and production rate is essential. The non-conventional yeasts *Kluyveromyces lactis, K. marxianus, Scheffersomyces stipitis, Yarrowia lipolytica, Hansenula polymorpha and Pichia pastoris* have been developed as eukaryotic hosts because of their desirable phenotypes, including thermotolerance, assimilation of diverse carbon sources, and high protein secretion. However, advanced metabolic engineering in these yeasts has been limited. This review outlines the challenges of using non-conventional yeasts for strain and pathway engineering, and discusses the developed solutions to these problems and the resulting applications in industrial biotechnology.

2.2 Introduction

The microbial production of fuels and chemicals from biomass and other renewable carbon sources is an attractive alternative to petroleum-derived products. One of the largest scale example of this is ethanol production by the yeast *Saccharomyces cerevisiae*— in 2015, over 25 billion gallons were produced worldwide from starch, waste sugar streams, and biomass-derived sugars. (www.afdc.energy.gov/data/10331) *S. cerevisiae* is the organism of choice because of its high rate of production and tolerance to ethanol titers upwards of 120 g L^{-1} [1, 2]. These phenotypes, among others, have led to the widespread study of *S. cerevisiae* and its development as a model eukaryotic host for chemical biosynthesis. A valuable approach to metabolic engineering is identifying organisms with desirable phenotypes and developing new synthetic biology tools to enhance these phenotypes. Bioethanol production in *S. cerevisiae* is a good example of this, and illustrates the potential of identifying other hosts and phenotypes to synthesize bioproducts other than ethanol. A number of examples of this strategy already exist in industry, where non-conventional yeasts with unique and advantageous phenotypes are used to produce proteins, lipids, and commodity chemicals. Metabolic engineering in these yeasts is, however, more challenging in comparison with *S. cerevisiae*, because less is known about their metabolism and genomics, and advanced genetic engineering tools are limited.

In this review, we focus on six specific non-conventional yeasts (Table 2.1): *Kluyveromyces lactis, K. marxianus, Scheffersomyces (Pichia) stipitis, Yarrowia lipolytica, Hansenula polymorpha, and Pichia pastoris.* In contrast to *S. cerevisiae,* these yeasts are Crabtree negative and favor respiration over fermentation; phenotypes that are particularly useful for protein production as well as the biosynthesis of chemicals other than ethanol [3]. *K. lactis* is discussed here because of its capacity to metabolize inexpensive substrates such as

waste whey and because of its use as a host for heterologous protein production in the food, feed, and pharmaceutical industries [4]. The *Kluyveromyces* species *K. marxianus* is also industrially relevant because of its wide substrate spectrum, fast growth characteristics, and thermotolerance to ~50°C [5, 6]. Native strains of *K. marxianus* are also known to synthesize ethyl acetate at rates above 2 g L^{-1} h^{-1} in aerated bioreactors [7, 8]. *S. stipitis* is capable of fermenting xylose at high rates compared to other yeasts and has been widely studied for ethanol production from biomass-derived sugars [9, 10]. *Y. lipolytica* is a well-studied oleaginous yeast, and has attracted interest due to its ability to synthesize and accumulate high levels of intracellular lipids [11-13]. The methylotrophic yeast *H. polymorpha* has been studied as a model system for peroxisome function as well as for its methanol and nitrate assimilation pathways [14, 15]. Significant efforts have gone into heterologous protein production in *H. polymorpha* due to its efficient secretion pathways, effective glycosylation machinery, and tightly controlled expression systems [16]. *H. polymorpha* is also thermotolerant to temperatures comparable to *K. marxianus* and can assimilates various substrates, thus making it a potential alternative host for ethanol production [17]. The methylotrophic yeast *P. pastoris* has similar protein secretion and glycosylation capabilities to *H. polymorpha* and has been widely used for heterologous protein production [18]. Its capacity to grow to extremely high cell densities and high capacity for membrane protein expression also provide inherent advantages over other yeast hosts [19, 20].

Table 2.1: Overview of non-conventional yeast species, their industrially-relevant phenotypes, common uses in biotechnology, and comparison with S. cerevisiae.

Yeast	Beneficial Phenotype	Products	Ref.
K. lactis	High protein secretion Growth on lactose	Proteins for food and feed industry Pharmaceutical enzymes	[4]
K. marxianus	Thermotolerance Fast growth characteristics High ethyl acetate production Growth on a range of sugars	Ethanol and volatile acetate esters	[5]
S. stipitis	High ethanol production from xylose	Ethanol fermentation from biomass derived carbohydrates	[21]
Y. lipolytica	Efficient production of lipids Growth on glycerol and alkanes	Lipids and oleochemicals	[12]
H. polymorpha	Thermotolerance Tightly regulated expression system Beneficial glycosylation for therapeutics	Heterologous protein High temperature ethanol fermentation	[17, 18]
P. pastoris	Tightly regulated expression system High cell density on minimal media Beneficial glycosylation for therapeutics Efficient production of membrane proteins	Pharmaceuticals and industrial enzymes	[18]
S. cerevisiae	High ethanol production High HR capacity Well known genomics and physiology Advanced synthetic biology tools	Ethanol in fermented beverages and as biofuel Commodity and specialty chemicals Pharmaceuticals	[2, 22]

Despite these many advantages, metabolic engineering of non-conventional yeasts is limited by a lack of sophisticated genome editing tools and an incomplete understanding of their genetics, metabolism, and cellular physiology. In this review, we discuss the challenges and solutions that have arisen in engineering non-conventional yeasts for metabolic engineering and synthetic biology applications. We begin our review with a discussion of the challenges to genetic engineering, followed by a discussion of strategies for improving genome and pathway engineering. Finally, we discuss representative examples of metabolic engineering in each of the selected yeasts. While the presented examples are not exhaustive, they are exemplative of current and past research efforts that exploit the yeasts'

advantageous phenotypes. Reviews that provide comprehensive discussions on engineering each of the non-conventional yeasts described here are available elsewhere [4, 12, 17, 21, 23, 24].

2.3 Genetic Engineering Challenges in Non-Conventional Yeasts

A basic requirement for metabolic engineering is the ability to express a gene (native or heterologous) from an expression cassette. Most strategies for heterologous gene expression in yeasts utilize auxotrophic markers to provide selective pressure for the maintenance of heterologous DNA containing the expression cassette (Figure 2.1). Gene expression of a native or heterologous genes is most often accomplished through episomal vectors or by integration of the gene(s)-of-interest into the host genome. In *S. cerevisiae*, transformation and expression from replicating plasmids is widely used due to the availability of stable and high copy number vectors [25, 26]. In non-conventional yeasts, options for stable plasmids are more limited. Plasmids are initially generated by combining centromeric regions of the genome organisms and autonomous replicating sequences with a selectable auxotrophic marker [27]. While functional plasmids are available for most non-conventional yeasts, they tend to be low copy number and also tend to show variable expression across cells in a single population, an effect that is due to imperfect partitioning of plasmids upon cell division [28, 29].

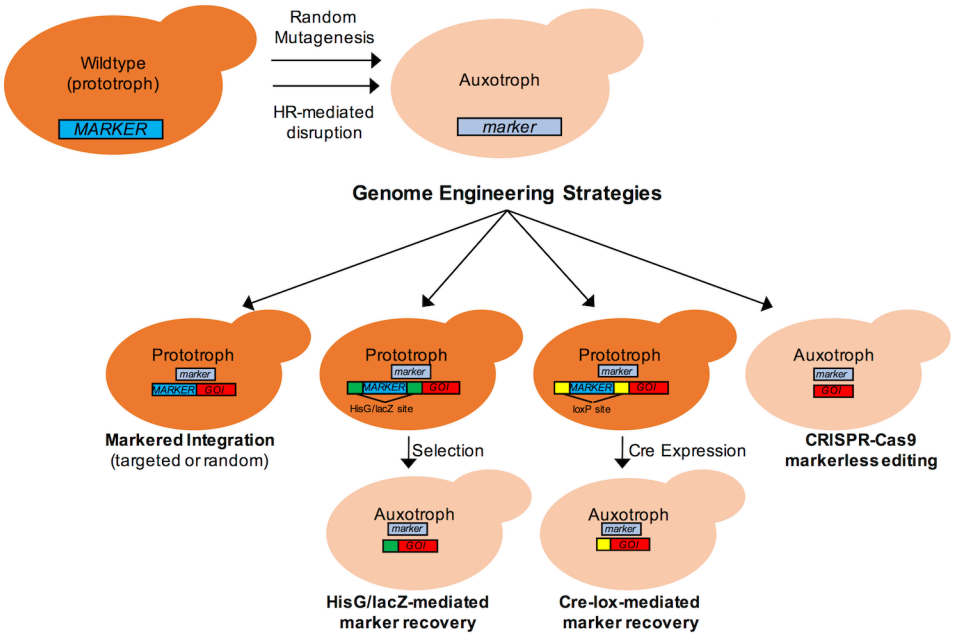


Figure 2.1: Schematic diagram of the generation and utilization of auxotrophic markers for engineering yeast. Random mutagenesis of host DNA or homologous recombination of a cassette that inactivates an essential gene for nutrient synthesis can be used to produce stable auxotrophic strains. The presence of an auxotrophy allows more advanced genome editing and pathway engineering tools to be applied in the yeast species of interest. Shown here are 1) targeted and random integration using a selectable marker (bottom, left), 2) HisG/LaZ mediated marker recovery (bottom, middle), 3) Cre-Lox marker recovery (bottom, middle), and 4) Markerless editing by CRISPR-Cas9 (bottom, right).

The preferred strategy for heterologous expression in industrial strains is integration into the host's genome. Genomic integration leads to more homogenous expression levels across the population, increases the stability of the expression cassette over extended culture times, and eliminates the need for constant selection of a genetic marker [30]. Transformation with a linear DNA fragment containing an expression cassette and a selectable marker results in genomic integration in one of two ways: heterologous DNA is either incorporated into the genome at a random locus (often called illegitimate recombination [31]), or the cassette is targeted to a specific site in the genome by homology to the site of interest (Figure 2.1). Both types of integrations are performed by native DNA

repair pathways. Random integration proceeds through nonhomologous end-joining (NHEJ), while targeted integration occurs by homologous recombination (HR) [32]. Integration via HR is often preferred, because it enables control over the integration loci, avoids disrupting essential genes, and allows for integration into a site with a consistent expression profile [33]. Integration via HR can also be used to knockout native genes.

In *S. cerevisiae*, HR is the dominant DNA repair pathway. The high capacity for HR makes genome engineering relatively efficient and has facilitated the development of a wide range of *in vivo* DNA assembly tools [34, 35]. This is not the case in most other yeasts, where NHEJ is the favored DNA repair pathway and genome engineering by HR is inefficient. As a result, engineering of non-conventional yeasts is frequently accomplished by random integration. The random integration of the transformed expression cassette can lead to unwanted disruptions of open reading frames or other genomic elements. In addition, expression levels of heterologous cassettes have been shown to be highly dependent on the integration site, and so random integration can result in variable expression across transformants [33, 36].

An additional challenge to engineering multi-gene pathways is the limited number of viable selectable markers. To overcome this experimental challenge, researchers have developed several techniques for marker recovery (Figure 2.1). The most commonly used systems are Cre-*loxP*, *hisG*, and *LacZ* [37, 38]. In these cases, the selectable marker (*e.g.*, an antibiotic resistance gene or auxotrophic marker) is surrounded by *hisG*, *lacZ*, or *loxP* sequences. After genome integration of an expression or knockout cassette, the marker is excised by spontaneous HR or Cre recombinase activity. While the *hisG* and *lacZ* systems are effective, the Cre-*loxP* method is more common because *hisG* and *lacZ* systems require a counter-selection such as growth on media supplemented with 5-fluoroorotic acid (5-FOA)

for *URA3* excision [39-43]. While Cre-*loxP* systems are available for use in the non-conventional yeasts discussed in this review, this marker recovery technique does not solve the challenge of random, unknown integration sites that is problematic with illegitimate recombination.

2.4 Enhancing HR in Non-Conventional Yeasts

A widely used strategy to enhance HR in non-conventional yeasts is disruption of genes essential for the NHEJ pathway, such as *KU70* or *KU80*. *K. lactis* provides an early example of this strategy, where disruption of *KU80* produced a strain capable of integrating heterologous DNA via HR at a rate of 97% [44]. Similarly, disruption of *KU80* in *S. stipitis* resulted in an increase in the rate of HR-mediated integration of transformed linear donors [45]. In *Y. lipolytica*, disruption of *KU70* or *KU80* produced significant increases in HR rates, and allowed HR to occur with homology regions down to 50 bp [46, 47]. In *H. polymorpha*, *KU80* knockout gave an increase in alcohol oxidase gene knockout rates (AOX2-8) from an average of 19% in the wildtype background to 76% in the *KU80* deficient strain [48]. In *P. pastoris*, knockout of *KU70* enabled HR rates as high as 90% [49]. A similar result was found in *K. marxianus,* where *KU70* knockout increased HR rates to as high as 95% [50]. *KU80* disruption in *K. marxianus* was similarly effective, with HR rates increasing to upwards of 70% [51].

A second strategy that has had success in increasing HR is cell cycle synchronization. Natively, the activity of the HR DNA repair pathway is dependent on cell cycle [52]. When a single copy of chromosomal DNA is present, as in G1 phase, NHEJ is favored. Genes required for HR tend to only be expressed during phases of the cell cycle when multiple copies of chromosomes are available, *i.e.*, S phase and G2 phase. Cell cycle synchronization has been

widely used for fundamental biochemistry studies, and a recent work took advantage of this strategy to stall cells in S phase with the intent of increasing HR [53]. By adding hydroxyurea to cultures undergoing exponential growth, the authors demonstrated that S phase stalling resulted in enhanced HR in *Y. lipolytica*, *K. lactis*, and *P. pastoris*.

An alternative strategy to achieve efficient HR is the introduction of a genomic double strand break (DSB) using a programmable endonuclease in the presence of a homologous repair template [54]. Due to the deleterious effects of DSBs on cell viability, native repair pathways attempt to repair the cut. If a repair template with adequate homology to the region flanking the break is present, the host may use the template as a donor for HR. This strategy has the added benefit of not requiring a selectable marker on the integrated DNA fragment. Several programmable tools exist for targeted DSB, including dimeric meganucleases, zinc finger nucleases, transcription activator-like effector nucleases (TALENs), and clustered regularly interspaced short palindromic repeats (CRISPR) and CRISPR-associated 9 (CRISPR-Cas9) [54]. The first three of these have primarily been developed and applied in *S. cerevisiae*, although TALENs were recently used in *Y. lipolytica* [55, 56]. CRISPR-Cas9, however, has been widely applied in non-conventional yeasts as described in the following section.

2.5 CRISPR-Cas9 Genome Editing and Transcriptional Control

In recent years, the application of CRISPR-Cas9 technology has revolutionized genome editing. Specifically, the type II CRISPR-Cas9 system from *Streptococcus pyogenes* has been widely adopted to enable targeted DSB generation in a wide number of organisms [57, 58]. Functional expression of CRISPR-Cas9 in yeast has two main requirements. First, a codon-optimized Cas9 expression cassette is generated, with a nuclear localization tag fused

to its C-terminus. A nuclear localization tag is needed because *S. pyogenes* is a bacterium, and so unmodified Cas9 would localize to the cytosol in yeast. The second component of CRISPR-Cas9 systems (as commonly applied for genome editing) is a short (or single) guide RNA (sgRNA) [59]. The sgRNA has two main roles. The first 20 bp at the 5' end are known as the spacer and are responsible for genome targeting through complementation to the desired locus. The sgRNA also contains a structural region encoded downstream of the spacer that facilitates the interaction of the sgRNA and Cas9. Upon formation of the CRISPR-Cas9 ribonucleoprotein, the complex unwinds double stranded DNA and begins scanning for a sequence complementary to the spacer region of the sgRNA. When a complementary sequence is found, and if there is an appropriate protospacer adjacent motif (PAM; for *S. pyogenes* a genomic "NGG" found immediately 3' of the targeted sequence), the nuclease domains of Cas9 cleave both strands of the DNA [57, 59]. The introduction of this DSB must then be repaired to avoid host cell death (Figure 2.2). Repair of the DSB by NHEJ commonly results in indel mutations and gene inactivation. Providing a homology repair template induces repair of the break by HR and allows for a desired sequence to be inserted at the cut site.

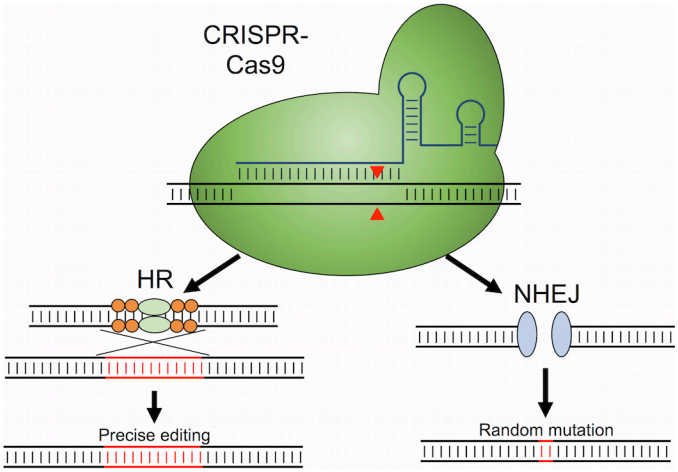


Figure 2.2: CRISPR-Cas9-mediated genome editing. The Cas9-sgRNA complex scans DNA until finding a complementary sequence. Upon binding, endonuclease domains cleave both DNA strands 3 bases upstream of the PAM sequence. The double strand break is then repaired either by homologous recombination (HR) if an appropriate homology donor is present, or by nonhomologous end-joing (NHEJ). Repair via HR allows for precise genome editing at the target site, while NHEJ introduces short insertions or deletions.

To date, CRISPR-Cas9 systems have been developed to allow gene disruptions and/or markerless integrations in all 6 of the non-conventional yeasts discussed in this review (Table 2.2). While most systems use a similar strategy for Cas9 expression and nuclear localization (commonly an SV40 C-terminal tag), a variety of strategies for sgRNA expression have been developed. In *S. cerevisiae*, the native SNR52 RNA polymerase III promoter is often used, as it allows for proper 5' and 3' maturation of the expressed sgRNA [57]. A similar sgRNA strategy was used in *K. lactis* and the resulting CRISPR-Cas9 system was demonstrated by simultaneously introducing three DSBs for multiplexed HR-mediated gene integration, successfully engineering a six gene pathway in a single transformation [35]. A native SNR52 promoter was also used to enable functional CRISPR-Cas9 genome editing in *S. stipitis*, where gene disruption rates upwards of 80% were achieved [27]. Two different CRISPR-Cas9 systems have been used in *Y. lipolytica*: the first relies on synthetic RNA polymerase III promoters for sgRNA expression, while the second used an RNA polymerase II promoter with ribozymes flanking the sgRNA to ensure proper 5' and 3' maturation [60, 61]. Both systems have been shown to achieve efficient gene disruption and gene integration rates. An analogous synthetic RNA polymerase III strategy was recently used for sgRNA expression to adapt the CRISPR-Cas9 system for use in *K. marxianus*, where gene disruption rates of 66% have been reported [62]. Successful adaptation of CRISPR-Cas9 to *P. pastoris* required the use of an RNA polymerase II promoter and ribozymes flanking the sgRNA, achieving efficiencies up to 100% for disruptions [63]. In *H. polymorpha*, a CRISPR-Cas9 system was developed by using $tRNA^{Leu}$ as a promoter to drive sgRNA expression[64]. This system uses the endogenous tRNA processing system for proper sgRNA maturation and disruption efficiencies of up to 71% were achieved.

Table 2.2: CRISPR-Cas9 systems for genome editing in non-conventional yeasts.

Yeast	Cas9 expression	sgRNA expression	Gene disruption rate	HR rate	Ref.
K. lactis	ScFBA1 promoter Genome integrated	SNR52 promoter	N/A	2% (3 integrations simultaneously) NHEJ deficient strain	[35]
K. marxianus	ScTEF1 promoter codon-optimized	RPR1'-tRNA^{Gly}	66%	N/A	[62]
S. stipitis	eno1 promoter codon-optimized	SNR52 promoter	80%	N/A	[27]
Y. lipolytica	TEFintron promoter codon-optimized	TEFintron promoter, flanked by hammerhead and hepatitis delta virus ribozymes	85%	11% in wildtype up to 100% in NHEJ deficient strain	[60]
Y. lipolytica	UAS1B8-TEF promoter codon-optimized	SCR1'-tRNA^{Gly} promoter	92%	64% in wildtype up to 100% in NHEJ deficient strain	[61]
H. polymorpha	DH3 promoter human codon-optimized	tRNA^{Leu}	71%	47% (marker integration with selection)	[64]
P. pastoris	HTX1 promoter human codon-optimized	HTX1 promoter, flanked by hammerhead and hepatitis delta virus	100%	20%	[63]

To further extend the yeast CRISPR-Cas9 toolbox, Cas9 can be mutated to deactivate its endonuclease activity while retaining its DNA targeting and binding capacity (dCas9). Targeting dCas9 to the promoter region of a gene can sterically block the RNA polymerase machinery from assembling, thus suppressing transcription; a technology referred to as CRISPR interference (CRISPRi) [65]. Fusion of a transcriptional repressor to dCas9 can result in a more effective CRISPRi system. In yeasts, the Mxi1 protein domain has proven most effective to date [66, 67]. In the context of non-conventional yeasts, CRISPRi has so far only been demonstrated in *Y. lipolytica* [68]. In this case, the synthetic RNA polymerase III system of sgRNA expression and Mxi1 fusion to dCas9 reduced target gene expression to as low as 10% of native expression levels. Finally, the fusion of a transcriptional activator to

Cas9 has enabled CRISPR activation (CRISPRa), where native genes can be overexpressed by targeting CRISPRa to a gene's promoter [69]. To date, however, CRISPRa has not been demonstrated in the yeasts discussed in this review. Targeted transcriptional control represents a novel experimental ability in non-conventional yeasts, where well-characterized promoters and tightly tunable inducible promoters are less common relative to model organisms.

2.6 Bioprocessing and Metabolic Engineering with Non-Conventional Yeasts

Despite the challenges of engineering non-conventional yeasts in comparison to the model host *S. cerevisiae*, a variety of successful bioprocesses have been developed. Here we present selected metabolic engineering examples that exploit desirable phenotypes expressed by *K. lactis*, *K. marxianus*, *S. stipitis*, *Y. lipolytica*, *H. polymorpha*, and *P. pastoris* and discuss the genetic engineering tools used to create new strains of these yeasts. Exemplative products produced from these hosts are presented in Table 2.3.

Table 2.3: Exemplative list of non-conventional yeast products

Yeast	Products	Reference
K. lactis	*Proteins*	
	Native β-galactosidase	Reviewed in [4]
	Chymosin	
	Brazzein	
	Human serum albumin (HSA)	
	Human interleukin 1-β	
	Interferon-α f	
	Chemicals	
	Glycolic acid	Reviewed in [3]
	Lactic acid	
K. marxianus	*Proteins*	
	Native inulinase	Reviewed in [24]
	Native β-galactosidases	
	Native pectinases	
	Chemicals	
	Ethanol from dairy waste or lignocellulosic feeds	Reviewed in [5]
	2-Phenylethanol/ 2-phenyl ethyl acetate	[70]
	Ethyl acetate	[7]
Y. lipolytica	*Proteins*	
	Lipases	Reviewed in [71]
	Proteases	
	α-amylases	
	β-mannases	
	Chemicals	
	Lipids	Reviewed in[12]and [3]
	α-Ketoglutaric acid (KGA)	Reviewed in [71]
	Lycopene	[36]
	Omega-3 eicosapentaenoic acid (EPA)	[72]
	Citric acid	Reviewed in [3]
S. stipitis	*Chemicals*	
	Ethanol from lignocellulosic feeds	Reviewed in [21]
	Fumaric acid	[73]
	Lactic acid	[74]
	Xylitol	[75]
H. polymorpha	*Proteins*	
	Hepatitis B surface antigen (HBsAg)	Reviewed in [76]
	Insulin	
	IFNα-2a	
	Hexose oxidase	
	Phytase	
	Chemicals	
	Ethanol from various carbon sources	Reviewed in [17]
P. pastoris	*Proteins*	
	Ecallantide	Reviewed in [23] and [77]
	Ocriplasmin	
	Phytase	
	Trypsin	
	Phospholipase C	
	Chemicals	
	(+)-nootkatone	[78]
	violacein	[79]
	b-carotene	[79]

2.6.1 Kluyveromyces lactis

Over the past three decades, considerable efforts have been put towards developing *K. lactis* as a yeast host for heterologous protein expression. To date, over 100 proteins have been produced, with more than 20% of these demonstrations occurring in the past five years [4]. Examples, such as b-galactosidase and the endopeptidase chymosin, sold by DSM under the trade names of Maxilact and Maxiren (DSM), respectively, have been produced at industrial-scale [4, 80]. The economic success of these processes is, in part, due to the ability of *K. lactis* to secrete high titers of protein and the ability metabolize inexpensive carbon sources such as waste whey streams produced in the dairy industry.

Strain development for *K. lactis* bioprocesses is most often achieved through an established and commercially available gene integration technology, pKLac2. The plasmid contains an acetamidase selection marker that allows for growth on acetamide as sole nitrogen source and facilitates multiple integrations into the genome to *K. lactis*. A mutant variant of the strongly inducible *Lac4* promoter eliminates recognition of the promoter by *E. coli* and thus enables cloning of constructs toxic to *E. coli*. Efficient protein secretion is achieved using a *K. lactis* α-mating factor secretion domain [81]. One example of the successful use of the pKLac2 system was the production of cardosin B chymosin, a coagulant essential for cheese production, from galactose media [82]. Another example is the sweetener brazzein, which was produced from galactose with protein titers reaching 104 mg L^{-1} [83].

In comparison to protein synthesis, the use of *K. lactis* as a host for chemical biosynthesis has been limited. New CRISPR-Cas9 genome editing systems are, however, enabling multiplexed engineering and driving the field forward. For example, a recent work engineered a synthetic muconic acid pathway by simultaneous integrating six heterologous

genes into three different *K. lactis* loci by HR. While triple integration efficiency was low at 2.1%, the desired strain was constructed in a time-efficient manner and produced ~0.9 g L^{-1} muconic acid [35].

2.6.2 Kluyveromyces marxianus

One of the reasons that *K. marxianus* has attracted interest is its high capacity to produce the volatile short chain ester ethyl acetate [6]. Wild type strains of *K. marxianus* have been shown to produce ethyl acetate at yields of 0.265 g g^{-1} glucose (51.4% of maximum) and pilot-scale plants with productivity upwards of 2 g L^{-1} h^{-1} using waste whey as a feed stock have been demonstrated [7, 8]. In addition to ethyl acetate, *K. marxianus* is able to produce fusel alcohols and their corresponding acetate esters. This capacity has been harnessed for 2-phenylethanol production from phenylalanine at industrial scale [5, 70]. Biosynthesis of 2-phenylethanol from glucose and the synthesis of phenylethyl acetate from phenylalanine feeds have also been demonstrated [70, 84].

K. marxianus has also been considered as a host for bioethanol production from lignocellulosic biomass hydrolysates and crude waste whey streams [85, 86]. Commercial production plants have been built or are under consideration in the United States, Ireland, and New Zealand, and rely on production from dairy waste streams [5]. Ethanol fermentation has also been engineered using metabolic engineering and cofactor balancing approaches [87, 88]. In one study, ethanol production from xylose was enhanced by 1) overexpressing heterologous xylose reductase (XYL1) and xylitol dehydrogenase (XYL2), and 2) increasing the capacity of the pentose phosphate pathway and flux towards ethanol through the overexpression of native xylulokinase (XYL3), L-ribulose-5-phosphate 4-epimerase (RPE1), ribose-5-phosphate isomerase (RKI1), transketolase (TKL1),

transaldolase (TAL1) genes as well as pyruvate decarboxylase (PDC1) and alcohol dehydrogenase (ADH2) [87]. The heterologous XYL1 and XYL2 genes were selected due to their preference of NADP(H) over NAD(H), thus helping to rectify an imbalance in co-factors when grown on xylose. Further improvement of fermentation efficiencies was achieved by eliminating glycerol production through the disruption of glycerol-3-phosphate dehydrogenase (GPD1). The resulting strain contained disruptions to three native genes and overexpression of two heterologous and seven native genes, and was able to produce ethanol from xylose at rates of 2.49 g L^{-1} h^{-1}. In this case, strain engineering was achieved by markered gene disruption via HR and sequential random integration facilitated by a *URA3* marker. Prior to each gene integration the marker was inactivated by HR with a truncated *URA3* cassette and selection on 5-FOA containing media [87, 89].

K. marxianus' high capacity for NHEJ can, for many applications, limit genome editing. However, some researchers have exploited this capacity for multiplexed gene integration. For example, a five-gene pathway for the production of hexanoic acid was integrated in a single transformation by selection on URA3 dropout media [90]. In this case, each integrated gene was accompanied by a URA3 selectable marker, resulting in a 50% success rate for full pathway integration. Random integration still proved to be problematic as hexanoic acid production varied widely between successful transformants, likely due to gene integration at critical genomic loci.

2.6.3 Scheffersomyces stipitis

A primary advantage of *S. stipitis* over other yeasts is its ability to ferment xylose at high rates [9, 10]. This phenotype has been exploited for ethanol production from biomass-derived and pure xylose streams. *S. stipitis* has also been engineered for higher ethanol

tolerance as well as growth inhibiters present in biomass hydrolysates [91]. Due to a lack of efficient genome editing tools, engineering of *S. stipitis* has been limited to random mutagenesis through UV radiation, adaptive evolution, protoplast fusion, and genome shuffling [92-95]. Single gene deletions have also been achieved through genomic integration using a selectable genetic marker. For example, a HR method was used to create a HXK1 deficient strain lacking glucose repression and a XYL2 deficient strain that produces xylitol from xylose [75]. In another example, *S. stipitis* was engineered to efficiently produce lactic acid through random integration of a heterologous LDH gene, with engineered strains producing up to 58 g L^{-1} lactate from 100 g L^{-1} xylose [74].

More complex pathway engineering has also been achieved. In one case, the deletion of two genes coupled with the overexpression of four heterologous genes led to the production of 4.67 g L^{-1} fumaric acid from 20 g L^{-1} xylose [73]. The plasmid-based pathway was comprised of a fumaric acid biosynthesis steps from *Rhizopus oryzae* and a fumaric acid transporter from *Schizosaccharomyces pombe.* Disruption of reaction steps competing with fumaric acid production, such as fumarase genes FUM1 and FUM2, was achieved via HR with a *URA3* selectable marker and marker recovery by Cre-*loxP*. A critical lesson from this work was the need for codon optimization of the heterologous genes, as *S. stipitis* has an unusual usage of CTG, which it uses to code for serine instead of leucine, as in other yeasts [42].

2.6.4 Yarrowia lipolytica

The oleaginous nature of *Y. lipolytica* has made it the focus of considerable efforts to convert a range of carbon sources into neutral lipids and lipid-derived compounds [11, 96, 97]. These efforts have been extensively reviewed elsewhere (see [12, 13] and references

therein). *Y. lipolytica* has also been used for heterologous protein production, but we focus here on its capacity for lipid biosynthesis [71]. In one recent work, high levels of triacylglycerides were engineered [98]. By engineering the conversion of glycolytic NADH to lipid precursors, specifically NADPH and acetyl-CoA, lipid production was increased by ~25% to 0.27 g g^{-1} glucose while reducing oxygen requirements of the strain. These improvements, along with a resulting high rate of lipid production (1.2 g $L^{-1}h^{-1}$), help to move this process closer to industrial feasibility.

In another example, researchers from DuPont used *Y. lipolytica* as a host for the biosynthesis of the nutritional supplement omega-3 eicosapentaenoic acid (EPA) [72]. The resulting strain gave rise to two commercial products, Newharvest™ EPA oil, a supplement for human consumption, and Verlasso®, a salmon feed with the high EPA biomass. Random integration of 30 copies of nine homologous and heterogeneous genes along with the disruption of b-oxidation resulted in an industrial production strain capable of producing EPA at 15% of dry cell weight and 57% of the total fatty acid content by weight. The aforementioned project relied on genome editing by random integration, thus necessitating marker recovery at each integration step. The recent adaptation of CRISPR-Cas9 for use in *Y. lipolytica* has alleviated this challenge by enabling site specific, markerless integration [36].

2.6.5 *Hansenula polymorpha*

The methylotrophic yeast *H. polymorpha (*previously *Pichia angusta or Ogataea polymorpha)* was first studied as model organism for peroxisome function as well as nitrate assimilation [14, 15, 76, 99]. The availability of a strong inducible expression system coupled with effective protein secretion and glycosylation has also made *H. polymorpha* a

successful host for protein production. While *S. cerevisiae* is able to N-glycosylate proteins, it tends to hyperglycosylate with alpha-1,3-linked mannose residues, which triggers immunogenicity in humans [100]. *H. polymorpha's* glycosylation machinery does not produce alpha-1,3-linked residues and is less prone to hyperglycosylation [76]. Moreover, significant efforts have been put towards optimization of human-like glycosylation [101]. Industrially produced biopharmaceutical examples include, but are not limited to, insulin under the trade name AgB, IFNα-2a sold as Wosulin, and proteins for hepatitis B vaccine HepaVax Gene [16, 18, 102]. With respect to the methylotrophic nature of *H. polymorpha*, the compartmentalized methanol-assimilation pathway has been exploited for the overexpression of peroxisome-dependent pathways. For example, penicillin production is localized, in part, to the peroxisomes. Growth on methanol promotes peroxisome proliferation, efficient heterologous protein expression, and consequently penicillin production [103].

H. polymorpha is also a good candidate for chemical biosynthesis. Thermotolerance, broad substrate utilization, and resistance to a variety of growth inhibitors match well with lignocellulosic as well as crude substrate streams [17]. Ethanol biosynthesis has been engineered with glycerol, cellulose hydrolysate, and starch-derived sugars as process inputs [104-108]. Most commonly, pathway engineering has been achieved by random integration or integration into the telomeric regions of the *H. polymorpha* genome by HR. While *H. polymorpha* easily accepts integration of heterologous genes, high NHEJ capacity makes it hard to disrupt genes. In a recent study, disruption of a *CAT8,* a transcriptional activator that is involved in in gluconeogenesis, respiration, the glyoxylic cycle and ethanol catabolism, increased ethanol yields to 12.5 g/L at 45°C. This work produced disruption efficiencies of 2.5% and below using a homology donor with a resistance marker [109]. The

recent development of a CRISPR-Cas9 in *H. polymorpha* has enabled gene disruption rates of up to 71%, significantly facilitating future metabolic engineering approaches [64].

2.6.6 *Pichia pastoris*

Similar to *H. polymorpha*, early interest in *P. pastoris* was driven by its ability to grow on methanol as sole carbon source, with research and development focusing on the production of single-cell protein [18]. Its use in bioprocessing is also similar to *H. polymorpha*, as *P. pastoris* is a common yeast host for protein production for the pharmaceutical and feed and food industries [18, 77]. Processes benefit from effective protein secretion as well as strong constitutive and inducible promoters engineered from the methanol assimilation pathway [18, 110]. Glycosylation pathways have been extensively engineered, thus facilitating mammalian protein production [111-113]. Bioprocesses also benefit from *P. pastoris*' ability to grow to high cell density and efficiently produce membrane proteins [18-20]. These characteristics have been exploited for the industrial production of several proteins including ecallantide (trade name Kalbitor® produced by Dyax), a recombinant protein inhibitor of the plasma protease kallikrein, and ocriplasmin (trade name Jetrea® produced by ThromboGenics), a truncated recombinant form of human plasmin [22, 23].

Due to low expression from plasmids, heterologous genes are usually integrated into the genome [18]. These traditional techniques have also been used for metabolic pathway engineering. For example, a biosynthetic pathway for terpenoid (+)-nootkatone was engineered in *P. pastoris* in a KU70 deficient strain [78]. The pathway required the integration of four heterologous and one homologous overexpression cassettes, which was achieved by targeted integration. Central to the success of this pathway was expression of

two membrane-associated cytochrome P450 enzymes. The resulting strain produced upwards of 200 mg L^{-1} of (+)-nootkatone in a high cell density fermentation. In a different example, *P. pastoris* was used to construct a nine gene polycistronic pathway using a 2A sequence that causes a ribosomal skip that terminates translation at the final proline codon of its C-terminally located conserved sequence ''NPGP'. This allows for production of multiple proteins form a single mRNA [79]. The system was used to produce the pigments violacein and, β-carotene. This study served as a proof-of-concept for stable and balanced multi enzyme pathway expression using a single promoter, and will facilitate future metabolic engineering approaches in *P. pastoris*.

2.7 Perspectives

Non-conventional yeasts have been extensively used for a range of biotechnological applications. So far, wild type strains and straightforward pathway engineering that leverages advantageous phenotypes native to the host have been the focus. With the increasing availability of next generation sequencing, genome editing tools, and the development of system wide –omics studies, more advanced understanding of the unique metabolisms and physiologies of non-conventional yeast has become attainable. Future engineering efforts will need to leverage these emerging systems and synthetic biology tools to address a critical lack of fundamental biochemical information, maximize the desired phenotypes, and increase productivity to reach industrially relevant production yields of new products.

Non-conventional yeast engineering will also be advanced by the application of genome-wide engineering tools. Tools such as yeast oligo-mediated genome engineering (YOGE, a recombineering strategy) and the yeast deletion collection in *S. cerevisiae*

demonstrate the power that functional genomics studies can have in yeast [116, 117]. While neither YOGE nor a full deletion collection are feasible in each non-conventional yeast of interest, an alternative strategy seems poised to fill this niche. Genome-wide CRISPR-Cas9 loss of function screens will allow researchers to perform analogous functional genomics studies by transforming pooled plasmids to introduce an indel into each gene in the genome separately[118, 119]. Already widely used and validated in mammalian studies, the application of genome-wide CRISPR-Cas9 screens will greatly advance engineering in non-conventional yeasts, and will allow for further enhancement of desirable phenotypes.

While this review mainly focuses on genome and pathway engineering, other methods and techniques, such as genome-scale modeling and metabolic flux balance analysis, have been used to guide strain engineering. For example, such models and analyses were used to optimize lipid production in *Y. lipolytica* and assess the biotechnological potential of *P. pastoris* and *S. stipitis* [114, 115]. Culture condition optimization has also been prominently featured in process development with non-conventional yeast. For example, low iron content media and *in situ* product removal strategies have led to the high rate production of ethyl acetate and 2-phenylethanol in wild type strains of *K. marxianus* [7, 70].

Most often, the limits of metabolic engineering and synthetic biology have been pushed using common lab strains of *S. cerevisiae* and *E. coli*. At the same time, many industrial biotechnology efforts have relied on wild type strains and traditional mutagenesis methods to create viable bioprocesses from non-conventional yeasts. As new systems and synthetic biology methods and tools are adapted for use in non-conventional yeasts, we expect that new bioprocesses that exploit desired phenotypes in non-

conventional yeasts will be developed and that these yeasts will become new model strain on their own merits.

2.8 Abbreviations

HR, homologous recombination; NEHJ, nonhomologous end-joining, DSB, double stand break; CRISPR, Clustered regularly interspaced short palindromic repeats; TALEN, transcription activator-like effector nucleases; sgRNA, short (or single) guide RNA; PAM, protospacer adjacent motif;

2.9 Acknowledgements

This work was supported by NSF CBET-1510697 and -1403264.

2.10 References

[1] Qiu Z, Jiang R: **Improving Saccharomyces cerevisiae ethanol production and tolerance via RNA polymerase II subunit Rpb7**. *Biotechnol Biofuels* 2017, **10**:125.

[2] Nielsen J, Larsson C, van Maris A, Pronk J: **Metabolic engineering of yeast for production of fuels and chemicals**. *Curr Opin Biotechnol* 2013, **24**(3):398-404.

[3] Wagner JM, Alper HS: **Synthetic biology and molecular genetics in non-conventional yeasts: Current tools and future advances**. *Fungal Genet Biol* 2016, **89**:126-136.

[4] Spohner SC, Schaum V, Quitmann H, Czermak P: **Kluyveromyces lactis: An emerging tool in biotechnology**. *J Biotechnol* 2016, **222**:104-116.

[5] Varela JA, Gethins L, Stanton C, Ross P, Morrissey JP: **Applications of Kluyveromyces marxianus in Biotechnology**. In: *Yeast Diversity in Human Welfare.* Edited by Satyanarayana T, Kunze G. Singapore: Springer Singapore; 2017: 439-453.

[6] Lobs AK, Lin JL, Cook M, Wheeldon I: **High throughput, colorimetric screening of microbial ester biosynthesis reveals high ethyl acetate production from Kluyveromyces marxianus on C5, C6, and C12 carbon sources**. *Biotechnol J* 2016, **11**(10):1274-1281.

[7] Loser C, Urit T, Stukert A, Bley T: **Formation of ethyl acetate from whey by Kluyveromyces marxianus on a pilot scale**. *J Biotechnol* 2013, **163**(1):17-23.

[8] Loser C, Urit T, Keil P, Bley T: **Studies on the mechanism of synthesis of ethyl acetate in Kluyveromyces marxianus DSM 5422**. *Appl Microbiol Biotechnol* 2014.

[9] Agbogbo FK, Coward-Kelly G: **Cellulosic ethanol production using the naturally occurring xylose-fermenting yeast, Pichia stipitis**. *Biotechnol Lett* 2008, **30**(9):1515-1524.

[10] Dupreez JC, Bosch M, Prior BA: **The Fermentation of Hexose and Pentose Sugars by Candida-Shehatae and Pichia-Stipitis**. *Appl Microbiol Biotechnol* 1986, **23**(3-4):228-233.

[11] Blazeck J, Hill A, Liu LQ, Knight R, Miller J, Pan A, Otoupal P, Alper HS: **Harnessing Yarrowia lipolytica lipogenesis to create a platform for lipid and biofuel production**. *Nat Commun* 2014, **5**.

[12] Hussain MS, Rodriguez GM, Gao DF, Spagnuolo M, Gambill L, Blenner M: **Recent advances in bioengineering of the oleaginous yeast Yarrowia lipolytica**. *Aims Bioeng* 2016, **3**(4):493-514.

[13] Beopoulos A, Cescut J, Haddouche R, Uribelarrea JL, Molina-Jouve C, Nicaud JM: **Yarrowia lipolytica as a model for bio-oil production**. *Prog Lipid Res* 2009, **48**(6):375-387.

[14] van der Klei IJ, Yurimoto H, Sakai Y, Veenhuis M: **The significance of peroxisomes in methanol metabolism in methylotrophic yeast**. *Bba-Mol Cell Res* 2006, **1763**(12):1453-1462.

[15] Siverio JM: **Assimilation of nitrate by yeasts**. *Fems Microbiol Rev* 2002, **26**(3):277-284.

[16] Gellissen G: **Heterologous protein production in methylotrophic yeasts**. *Appl Microbiol Biotechnol* 2000, **54**(6):741-750.

[17] Dmytruk K, Kurylenko O, Ruchala J, Ishchuk O, Sibirny A: **Development of the Thermotolerant Methylotrophic Yeast Hansenula polymorpha as Efficient Ethanol Producer**. In: *Yeast Diversity in Human Welfare.* Springer; 2017: 257-282.

[18] Gellissen G, Kunze G, Gaillardin C, Cregg JM, Berardi E, Veenhuis M, van der Klei I: **New yeast expression platforms based on methylotrophic Hansenula polymorpha and Pichia pastoris and on dimorphic Arxula adeninivorans and Yarrowia lipolytica - A comparison**. *Fems Yeast Res* 2005, **5**(11):1079-1096.

[19] Jahic M, Veide A, Charoenrat T, Teeri T, Enfors SO: **Process technology for production and recovery of heterologous proteins with Pichia pastoris**. *Biotechnol Prog* 2006, **22**(6):1465-1473.

[20] Byme B: **Pichia pastoris as an expression host for membrane protein structural biology**. *Curr Opin Struc Biol* 2015, **32**:9-17.

[21] Dmytruk KV, Kurylenko OO, Ruchala J, Abbas CA, Sibirny AA: **Genetic Improvement of Conventional and Nonconventional Yeasts for the Production of First-and Second-Generation Ethanol**. In: *Biotechnology of Yeasts and Filamentous Fungi.* Springer; 2017: 1-38.

[22] Walsh G: **Biopharmaceutical benchmarks 2014**. *Nat Biotechnol* 2014, **32**(10):992-1000.

[23] Ahmad M, Hirz M, Pichler H, Schwab H: **Protein expression in Pichia pastoris: recent achievements and perspectives for heterologous protein production**. *Appl Microbiol Biotechnol* 2014, **98**(12):5301-5317.

[24] Lane MM, Morrissey JP: **Kluyveromyces marxianus: a yeast emerging from its sister's shadow**. *Fungal Biol Rev* 2010, **24**(1):17-26.

[25] Mead DJ, Gardner DCJ, Oliver SG: **The Yeast 2-Mu Plasmid - Strategies for the Survival of a Selfish DNA**. ***Mol*** *Mol Gen Genet* 1986, **205**(3):417-421.

[26] Lee ME, DeLoache WC, Cervantes B, Dueber JE: **A Highly Characterized Yeast Toolkit for Modular, Multipart Assembly**. *ACS Synth Biol* 2015, **4**(9):975-986.

[27] Cao M, Gao M, Lopez-Garcia CL, Wu Y, Seetharam AS, Severin AJ, Shao Z: **Centromeric DNA Facilitates Nonconventional Yeast Genetic Engineering**. *ACS Synth Biol* 2017.

[28] Jensen NB, Strucko T, Kildegaard KR, David F, Maury J, Mortensen UH, Forster J, Nielsen J, Borodina I: **EasyClone: method for iterative chromosomal integration of multiple genes in Saccharomyces cerevisiae**. *Fems Yeast Res* 2014, **14**(2):238-248.

[29] Vernis L, Poljak L, Chasles M, Uchida K, Casaregola S, Kas E, Matsuoka M, Gaillardin C, Fournier P: **Only centromeres can supply the partition system required for ARS function in the yeast Yarrowia lipolytica**. *J Mol Biol* 2001, **305**(2):203-217.

[30] Da Silva NA, Srikrishnan S: **Introduction and expression of genes for metabolic engineering applications in Saccharomyces cerevisiae**. *Fems Yeast Res* 2012, **12**(2):197-214.

[31] Kegel A, Martinez P, Carter SD, Astrom SU: **Genome wide distribution of illegitimate recombination events in Kluyveromyces lactis**. *Nucleic Acids Res* 2006, **34**(5):1633-1645.

[32] Lieber MR: **The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway**. *Annu Rev Biochem* 2010, **79**:181-211.

[33] Flagfeldt DB, Siewers V, Huang L, Nielsen J: **Characterization of chromosomal integration sites for heterologous gene expression in Saccharomyces cerevisiae**. *Yeast* 2009, **26**(10):545-551.

[34] Shao Z, Zhao H, Zhao H: **DNA assembler, an in vivo genetic method for rapid construction of biochemical pathways**. *Nucleic Acids Res* 2009, **37**(2):e16.

[35] Horwitz AA, Walter JM, Schubert MG, Kung SH, Hawkins K, Platt DM, Hernday AD, Mahatdejkul-Meadows T, Szeto W, Chandran SS *et al*: **Efficient Multiplexed Integration of Synergistic Alleles and Metabolic Pathways in Yeasts via CRISPR-Cas**. *Cell Syst* 2015, **1**(1):88-96.

[36] Schwartz C, Shabbir-Hussain M, Frogue K, Blenner M, Wheeldon I: **Standardized Markerless Gene Integration for Pathway Engineering in Yarrowia lipolytica**. *ACS Synth Biol* 2017, **6**(3):402-409.

[37] Sauer B: **Functional expression of the cre-lox site-specific recombination system in the yeast Saccharomyces cerevisiae**. *Mol Cell Biol* 1987, **7**(6):2087-2096.

[38] Alani E, Cao L, Kleckner N: **A Method for Gene Disruption That Allows Repeated Use of Ura3 Selection in the Construction of Multiply Disrupted Yeast Strains**. *Genetics* 1987, **116**(4):541-545.

[39] Cheon SA, Choo J, Ubiyvovk VM, Park JN, Kim MW, Oh DB, Kwon O, Sibirny AA, Kim JY, Kang HA: **New selectable host-marker systems for multiple genetic manipulations**

based on TRP1, MET2 and ADE2 in the methylotrophic yeast Hansenula polymorpha. *Yeast* 2009, **26**(9):507-521.

[40] Pecota DC, Rajgarhia V, Da Silva NA: **Sequential gene integration for the engineering of Kluyveromyces marxianus**. *J Biotechnol* 2007, **127**(3):408-416.

[41] Pan RQ, Zhang J, Shen WL, Tao ZQ, Li SP, Yan X: **Sequential deletion of Pichia pastoris genes by a self-excisable cassette**. *Fems Yeast Res* 2011, **11**(3):292-298.

[42] Laplaza JM, Torres BR, Jin YS, Jeffries TW: **Sh ble and Cre adapted for functional genomics and metabolic engineering of Pichia stipitis**. *Enzyme Microb Tech* 2006, **38**(6):741-747.

[43] Steensma HY, Ter Linde JJM: **Plasmids with the Cre-recombinase and the dominant nat marker, suitable for use in prototrophic strains of Saccharomyces cerevisiae and Kluyveromyces lactis**. *Yeast* 2001, **18**(5):469-472.

[44] Kooistra R, Hooykaas PJ, Steensma HY: **Efficient gene targeting in Kluyveromyces lactis**. *Yeast* 2004, **21**(9):781-792.

[45] Maassen N, Freese S, Schruff B, Passoth V, Klinner U: **Nonhomologous end joining and homologous recombination DNA repair pathways in integration mutagenesis in the xylose-fermenting yeast Pichia stipitis**. *Fems Yeast Res* 2008, **8**(5):735-743.

[46] Verbeke J, Beopoulos A, Nicaud JM: **Efficient homologous recombination with short length flanking fragments in Ku70 deficient Yarrowia lipolytica strains**. *Biotechnol Lett* 2013, **35**(4):571-576.

[47] Kretzschmar A, Otto C, Holz M, Werner S, Hubner L, Barth G: **Increased homologous integration frequency in Yarrowia lipolytica strains defective in non-homologous end-joining**. *Curr Genet* 2013, **59**(1-2):63-72.

[48] Saraya R, Krikken AM, Kiel JA, Baerends RJ, Veenhuis M, van der Klei IJ: **Novel genetic tools for Hansenula polymorpha**. *Fems Yeast Res* 2012, **12**(3):271-278.

[49] Naatsaari L, Mistlberger B, Ruth C, Hajek T, Hartner FS, Glieder A: **Deletion of the Pichia pastoris KU70 homologue facilitates platform strain generation for gene expression and synthetic biology**. *PloS one* 2012, **7**(6):e39720.

[50] Abdel-Banat BM, 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* 2010, **27**(1):29-39.

[51] Choo JH, Han C, Kim JY, Kang HA: **Deletion of a KU80 homolog enhances homologous recombination in the thermotolerant yeast Kluyveromyces marxianus**. *Biotechnol Lett* 2014, **36**(10):2059-2067.

[52] Chapman JR, Taylor MRG, Boulton SJ: **Playing the End Game: DNA Double-Strand Break Repair Pathway Choice**. *Mol Cell* 2012, **47**(4):497-510.

[53] Tsakraklides V, Brevnova E, Stephanopoulos G, Shaw AJ: **Improved Gene Targeting through Cell Cycle Synchronization**. *PloS one* 2015, **10**(7):e0133434.

[54] Liu Z, Liang Y, Ang EL, Zhao H: **A New Era of Genome Integration-Simply Cut and Paste!** *ACS Synth Biol* 2017, **6**(4):601-609.

[55] David F, Siewers V: **Advances in yeast genome engineering**. *Fems Yeast Res* 2015, **15**(1):1-14.

[56] Rigouin C, Gueroult M, Croux C, Dubois G, Borsenberger V, Barbe S, Marty A, Daboussi F, Andre I, Bordes F: **Production of Medium Chain Fatty Acids by Yarrowia lipolytica: Combining Molecular Design and TALEN to Engineer the Fatty Acid Synthase**. *ACS Synth Biol y* 2017.

[57] DiCarlo JE, Norville JE, Mali P, Rios X, Aach J, Church GM: **Genome engineering in Saccharomyces cerevisiae using CRISPR-Cas systems**. *Nucleic Acids Res* 2013, **41**(7):4336-4343.

[58] Stovicek V, Holkenbrink C, Borodina I: **CRISPR/Cas system for yeast genome engineering: advances and applications**. *Fems Yeast Res* 2017.

[59] Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E: **A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity**. *Science* 2012, **337**(6096):816-821.

[60] Gao S, Tong Y, Wen Z, Zhu L, Ge M, Chen D, Jiang Y, Yang S: **Multiplex gene editing of the Yarrowia lipolytica genome using the CRISPR-Cas9 system**. *J Ind Microbiol* 2016, **43**(8):1085-1093.

[61] Schwartz CM, Hussain MS, Blenner M, Wheeldon I: **Synthetic RNA Polymerase III Promoters Facilitate High-Efficiency CRISPR-Cas9-Mediated Genome Editing in Yarrowia lipolytica**. *ACS Synth Biol* 2016, **5**(4):356-359.

[62] Lobs AK, 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* 2017, **10**:164.

[63] Weninger A, Hatzl AM, Schmid C, Vogl T, Glieder A: **Combinatorial optimization of CRISPR/Cas9 expression enables precision genome engineering in the methylotrophic yeast Pichia pastoris**. *J Biotechnol* 2016, **235**:139-149.

[64] Numamoto M, Maekawa H, Kaneko Y: **Efficient genome editing by CRISPR/Cas9 with a tRNA-sgRNA fusion in the methylotrophic yeast Ogataea polymorpha**. *J Biosci Bioeng* 2017.

[65] Qi LS, Larson MH, Gilbert LA, Doudna JA, Weissman JS, Arkin AP, Lim WA: **Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression**. *Cell* 2013, **152**(5):1173-1183.

[66] Smith JD, Suresh S, Schlecht U, Wu M, Wagih O, Peltz G, Davis RW, Steinmetz LM, Parts L, St Onge RP: **Quantitative CRISPR interference screens in yeast identify chemical-genetic interactions and new rules for guide RNA design**. *Genome Biol* 2016, **17**:45.

[67] Gilbert LA, Larson MH, Morsut L, Liu Z, Brar GA, Torres SE, Stern-Ginossar N, Brandman O, Whitehead EH, Doudna JA *et al*: **CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes**. *Cell* 2013, **154**(2):442-451.

[68] Schwartz C, Frogue K, Misa J, Ramesh A, Wheeldon I: **CRISPRi repression of nonhomologous end-joining for enhanced genome engineering via homologous recombination in Yarrowia lipolytica**. *Biotechnol Bioeng* 2017, **submitted**.

[69] Gilbert LA, Larson MH, Morsut L, Liu ZR, Brar GA, Torres SE, Stern-Ginossar N, Brandman O, Whitehead EH, Doudna JA *et al*: **CRISPR-Mediated Modular RNA-Guided Regulation of Transcription in Eukaryotes**. *Cell* 2013, **154**(2):442-451.

[70] Etschmann MMW, Schrader J: **An aqueous-organic two-phase bioprocess for efficient production of the natural aroma chemicals 2-phenylethanol and 2-phenylethylacetate with yeast**. *Appl Microbiol Biotechnol* 2006, **71**(4):440-443.

[71] Madzak C: **Yarrowia lipolytica: recent achievements in heterologous protein expression and pathway engineering**. *Appl Microbiol Biotechnol* 2015, **99**(11):4559-4577.

[72] Xue Z, Sharpe PL, Hong SP, Yadav NS, Xie D, Short DR, Damude HG, Rupert RA, Seip JE, Wang J *et al*: **Production of omega-3 eicosapentaenoic acid by metabolic engineering of Yarrowia lipolytica**. *Nat Biotechnol* 2013, **31**(8):734-740.

[73] Wei L, Liu J, Qi HS, Wen JP: **Engineering Scheffersomyces stipitis for fumaric acid production from xylose**. *Bioresour Technol* 2015, **187**:246-254.

[74] Ilmen M, Koivuranta K, Ruohonen L, Suominen P, Penttila M: **Efficient production of L-lactic acid from xylose by Pichia stipitis**. *Appl Environ Microbiol* 2007, **73**(1):117-123.

[75] Dashtban M, Wen X, Bajwa PK, Ho CY, Lee H: **Deletion of hxk1 gene results in derepression of xylose utilization in Scheffersomyces stipitis**. *J Ind Microbiol* 2015, **42**(6):889-896.

[76] Kunze G, Kang HA, Gellissen G: **Hansenula polymorpha (Pichia angusta): biology and applications**. In: *Yeast Biotechnology: Diversity and Applications.* Springer; 2009: 47-64.

[77]Spohner SC, Muller H, Quitmann H, Czermak P: **Expression of enzymes for the usage in food and feed industry with Pichia pastoris**. *J Biotechnol* 2015, **202**:118-134.

[78] Wriessnegger T, Augustin P, Engleder M, Leitner E, Muller M, Kaluzna I, Schurmann M, Mink D, Zellnig G, Schwab H *et al*: **Production of the sesquiterpenoid (+)-nootkatone by metabolic engineering of Pichia pastoris**. *Metab Eng* 2014, **24**:18-29.

[79] Geier M, Fauland P, Vogl T, Glieder A: **Compact multi-enzyme pathways in P. pastoris**. *Chem Commun* 2015, **51**(9):1643-1646.

[80] Vandenberg JA, Vanderlaken KJ, Vanooyen AJJ, Renniers TCHM, Rietveld K, Schaap A, Brake AJ, Bishop RJ, Schultz K, Moyer D *et al*: **Kluyveromyces as a Host for Heterologous Gene-Expression - Expression and Secretion of Prochymosin**. *Bio-Technol* 1990, **8**(2):135-139.

[81] Colussi PA, Taron CH: **Kluyveromyces lactis LAC4 promoter variants that lack function in bacteria but retain full function in K. lactis**. *Appl Environ Microbiol* 2005, **71**(11):7092-7098.

[82] Almeida CM, Gomes D, Faro C, Simoes I: **Engineering a cardosin B-derived rennet for sheep and goat cheese manufacture**. *Appl Microbiol Biotechnol* 2015, **99**(1):269-281.

[83] Jo HJ, Noh JS, Kong KH: **Efficient secretory expression of the sweet-tasting protein brazzein in the yeast Kluyveromyces lactis**. *Protein Expres Purif* 2013, **90**(2):84-89.

[84] Kim TY, Lee SW, Oh MK: **Biosynthesis of 2-phenylethanol from glucose with genetically engineered Kluyveromyces marxianus**. *Enzyme Microb Technol* 2014, **61-62**:44-47.

[85] Ballesteros M, Oliva JM, Negro MJ, Manzanares P, Ballesteros I: **Ethanol from lignocellulosic materials by a simultaneous saccharification and fermentation process (SFS) with Kluyveromyces marxianus CECT 10875**. *Process Biochem* 2004, **39**(12):1843-1848.

[86] Zafar S, Owais M: **Ethanol production from crude whey by Kluyveromyces marxianus**. *Biochem Eng J* 2006, **27**(3):295-298.

[87] Zhang J, Zhang B, Wang DM, Gao XL, Sun LH, Hong J: **Rapid ethanol production at elevated temperatures by engineered thermotolerant Kluyveromyces marxianus via the NADP(H)-preferring xylose reductase-xylitol dehydrogenase pathway**. *Metab Eng* 2015, **31**:140-152.

[88] Yanase S, Hasunuma T, Yamada R, Tanaka T, Ogino C, Fukuda H, Kondo A: **Direct ethanol production from cellulosic materials at high temperature using the thermotolerant yeast Kluyveromyces marxianus displaying cellulolytic enzymes**. *Appl Microbiol Biotechnol* 2010, **88**(1):381-388.

[89] Wang RL, Li LL, Zhang B, Gao XL, Wang DM, Hong J: **Improved xylose fermentation of Kluyveromyces marxianus at elevated temperature through construction of a xylose isomerase pathway**. *J Ind Microbiol* 2013, **40**(8):841-854.

[90] Cheon Y, Kim JS, Park JB, Heo P, Lim JH, Jung GY, Seo JH, Park JH, Koo HM, Cho KM *et al*: **A biosynthetic pathway for hexanoic acid production in Kluyveromyces marxianus**. *J Biotechnol* 2014, **182**:30-36.

[91] Slininger PJ, Shea-Andersh MA, Thompson SR, Dien BS, Kurtzman CP, Balan V, da Costa Sousa L, Uppugundla N, Dale BE, Cotta MA: **Evolved strains of Scheffersomyces stipitis achieving high ethanol productivity on acid- and base-pretreated biomass hydrolyzate at high solids loading**. *Biotechnol Biofuels* 2015, **8**:60.

[92] Shi J, Zhang M, Zhang L, Wang P, Jiang L, Deng H: **Xylose-fermenting Pichia stipitis by genome shuffling for improved ethanol production**. *Microb Biotechnol* 2014, **7**(2):90-99.

[93] Pereira SR, Sanchez INV, Frazao CJ, Serafim LS, Gorwa-Grauslund MF, Xavier AM: **Adaptation of Scheffersomyces stipitis to hardwood spent sulfite liquor by evolutionary engineering**. *Biotechnol Biofuels* 2015, **8**:50.

[94] Zhang W, Geng AL: **Improved ethanol production by a xylose-fermenting recombinant yeast strain constructed through a modified genome shuffling method**. *Biotechnol Biofuels* 2012, **5**.

[95] Hughes SR, Gibbons WR, Bang SS, Pinkelman R, Bischoff KM, Slininger PJ, Qureshi N, Kurtzman CP, Liu S, Saha BC *et al*: **Random UV-C mutagenesis of Scheffersomyces (formerly Pichia) stipitis NRRL Y-7124 to improve anaerobic growth on lignocellulosic sugars**. *J Ind Microbiol* 2012, **39**(1):163-173.

[96] Rodriguez GM, Hussain MS, Gambill L, Gao DF, Yaguchi A, Blenner M: **Engineering xylose utilization in Yarrowia lipolytica by understanding its cryptic xylose pathway**. *Biotechnol Biofuels* 2016, **9**.

[97] Xu J, Liu N, Qiao K, Vogg S, Stephanopoulos G: **Application of metabolic controls for the maximization of lipid production in semicontinuous fermentation**. *Proc Natl Acad Sci U S A* 2017, **114**(27):E5308-E5316.

[98] Qiao KJ, Wasylenko TM, Zhou K, Xu P, Stephanopoulos G: **Lipid production in Yarrowia lipolytica is maximized by engineering cytosolic redox metabolism**. *Nat Biotechnol* 2017, **35**(2):173-177.

[99] Gellissen G, Veenhuis M: **The methylotrophic yeast Hansenula polymorpha: its use in fundamental research and as a cell factory**. *Yeast* 2001, **18**(3):i-iii.

[100] Ballou CE: **Isolation, characterization, and properties of Saccharomyces cerevisiae mnn mutants with nonconditional protein glycosylation defects**. *Methods Enzymol* 1990, **185**:440-470.

[101] Cheon SA, Kim H, Oh DB, Kwon O, Kang HA: **Remodeling of the glycosylation pathway in the methylotrophic yeast Hansenula polymorpha to produce human hybrid-type N-glycans**. *J Microbiol* 2012, **50**(2):341-348.

[102] Muller II F, Tieke A, Waschk D, Muhle C, Muller I F, Seigelchifer M, Pesce A, Jenzelewski V, Gellissen G: **Production of IFN alpha-2a in Hansenula polymorpha**. *Process Biochem* 2002, **38**(1):15-25.

[103] Gidijala L, Kiel JA, Douma RD, Seifar RM, van Gulik WM, Bovenberg RA, Veenhuis M, van der Klei IJ: **An engineered yeast efficiently secreting penicillin**. *PloS one* 2009, **4**(12):e8317.

[104] Kurylenko OO, Ruchala J, Hryniv OB, Abbas CA, Dmytruk KV, Sibirny AA: **Metabolic engineering and classical selection of the methylotrophic thermotolerant yeast Hansenula polymorpha for improvement of high-temperature xylose alcoholic fermentation**. *Microb Cell Fact* 2014, **13**.

[105] Kata I, Semkiv MV, Ruchala J, Dmytruk KV, Sibirny AA: **Overexpression of the genes PDC1 and ADH1 activates glycerol conversion to ethanol in the thermotolerant yeast Ogataea (Hansenula) polymorpha**. *Yeast* 2016, **33**(8):471-478.

[106] Voronovsky AY, Rohulya OV, Abbas CA, Sibirny AA: **Development of strains of the thermotolerant yeast Hansenula polymorpha capable of alcoholic fermentation of starch and xylan**. *Metab Eng* 2009, **11**(4-5):234-242.

[107] Ryabova OB, Chmil OM, Sibirny AA: **Xylose and cellobiose fermentation to ethanol by the thermotolerant methylotrophic yeast Hansenula polymorpha**. *Fems Yeast Res* 2003, **4**(2):157-164.

[108] Hong WK, Kim CH, Heo SY, Luo L, Oh BR, Seo JW: **Enhanced production of ethanol from glycerol by engineered Hansenula polymorpha expressing pyruvate decarboxylase and aldehyde dehydrogenase genes from Zymomonas mobilis**. *Biotechnol Lett* 2010, **32**(8):1077-1082.

[109] Ruchala J, Kurylenko OO, Soontorngun N, Dmytruk KV, Sibirny AA: **Transcriptional activator Cat8 is involved in regulation of xylose alcoholic fermentation in the thermotolerant yeast Ogataea (Hansenula) polymorpha**. *Microb Cell Facts* 2017, **16**.

[110] Macauley-Patrick S, Fazenda ML, McNeil B, Harvey LM: **Heterologous protein production using the Pichia pastoris expression system**. *Yeast* 2005, **22**(4):249-270.

[111] Jacobs PP, Geysens S, Vervecken W, Contreras R, Callewaert N: **Engineering complex-type N-glycosylation in Pichia pastoris using GlycoSwitch technology**. *Nat Protoc* 2009, **4**(1):58-70.

[112] Hamilton SR, Gerngross TU: **Glycosylation engineering in yeast: the advent of fully humanized yeast**. *Curr Opin Biotechnol* 2007, **18**(5):387-392.

[113] Hamilton SR, Davidson RC, Sethuraman N, Nett JH, Jiang YW, Rios S, Bobrowicz P, Stadheim TA, Li HJ, Choi BK *et al*: **Humanization of yeast to produce complex terminally sialylated glycoproteins**. *Science* 2006, **313**(5792):1441-1443.

[114] Kavscek M, Bhutada G, Madl T, Natter K: **Optimization of lipid production with a genome-scale model of Yarrowia lipolytica**. *BMC Syst Biol* 2015, **9**.

[115] Caspeta L, Shoaie S, Agren R, Nookaew I, Nielsen J: **Genome-scale metabolic reconstructions of Pichia stipitis and Pichia pastoris and in silico evaluation of their potentials**. *BMC Syst Biol* 2012, **6**.

[116] DiCarlo JE, Conley AJ, Penttila M, Jantti J, Wang HH, Church GM: **Yeast oligo-mediated genome engineering (YOGE)**. *ACS Synth Biol* 2013, **2**(12):741-749.

[117] Giaever G, Nislow C: **The yeast deletion collection: a decade of functional genomics**. *Genetics* 2014, **197**(2):451-465.

[118] Shalem O, Sanjana NE, Hartenian E, Shi X, Scott DA, Mikkelsen TS, Heckl D, Ebert BL, Root DE, Doench JG *et al*: **Genome-scale CRISPR-Cas9 knockout screening in human cells**. *Science* 2014, **343**(6166):84-87.

[119] Wang T, Wei JJ, Sabatini DM, Lander ES: **Genetic screens in human cells using the CRISPR-Cas9 system**. *Science* 2014, **343**(6166):80-84.

Chapter 3: High throughput, colorimetric screening of microbial ester biosynthesis reveals high ethyl acetate production from *Kluyveromyces marxianus* on C5, C6, and C12 carbon sources.

3.1 Abstract

Advances in genome and metabolic pathway engineering have enabled large combinatorial libraries of mutant microbial hosts for chemical biosynthesis. Despite these advances, strain development is often limited by the lack of high throughput functional assays for effective library screening. Recent synthetic biology efforts have engineered microbes that synthesize acetyl and acyl esters and many yeasts naturally produce esters to significant titers. Short and medium chain volatile esters have value as fragrance and flavor compounds, while long chain acyl esters are potential replacements for diesel fuel. Here, we developed a biotechnology method for the rapid screening of microbial ester biosynthesis. Using a colorimetric reaction scheme, esters extracted from fermentation broth were quantitatively converted to a ferric hydroxamate complex with strong absorbance at 520 nm. The assay was validated for ethyl acetate, ethyl butyrate, isoamyl acetate, ethyl hexanoate, and ethyl octanoate, and achieved a z-factor of 0.77. Screening of ethyl acetate production from a combinatorial library of four *Kluyveromyces marxianus* strains on seven carbon sources revealed ethyl acetate biosynthesis from C5, C6, and C12 sugars. This newly adapted method rapidly identified novel properties of *K. marxianus* metabolism and promises to advance high throughput microbial strain engineering for ester biosynthesis.

3.2 Introduction

Microbial strain engineering for chemical synthesis is often limited by the ability to rapidly test and quantify chemical production during fermentation. Advances in genome engineering and combinatorial pathway synthesis have enabled large libraries of mutant and engineered microbial strains,[1-3] but the capacity to effectively screen library function has not kept pace. Often the best method to screen large libraries of engineered strains is to couple the desired phenotype to a colorimetric assay, thus enabling high throughput screening by fluorescence cell sorting, visualization, or plate-based UV-Vis assays. For example, the colorimetric screening of lycopene enabled the rapid identification of high producing *E. coli* strains generated by multiplexed automated genome engineering (MAGE) [1]; and, a colorimetric enzyme assay for L-3,4-dihydoxyphenylalanine (L-DOPA) was recently used to engineer a synthetic pathway for alkaloid biosynthesis in yeast.[4]

Rapid development of engineered microbial strains for ester biosynthesis has, in part, been limited by the analytical methods available to quantify ester biosynthesis. Diesel substitutes in the form of fatty acid methyl and ethyl esters as well as branched chain fatty esters have been produced in *Escherichia coli*,[5, 6] *Saccharomyces cerevisiae*,[7, 8] and *Pichia pastoris*.[5] In each case, ester production was quantified by gas chromatography (GC), GC-MS, or thin layer chromatography (TLC). GC and GC-MS have also been used to quantify medium chain acetate and butyrate flavor and fragrance compounds synthesized in engineered *E. coli*,[9-11] and GC has been used for short chain volatile ester quantification in the yeast *Kluyveromyces marxianus* [12, 13] and *E. coli*.[14]

A potential solution to this problem was developed over 70 years ago. In 1946, Uno Hill of the Inland Steel Company in East Chicago Illinois, published a colorimetric method to identify fatty acids and fatty esters contamination on tin plates.[15] The assay reacted

carboxylate esters with hydroxylamine to produce hydroxamic acid. When ferric iron was reacted with hydroxamic acid an iron complex with strong absorbance in the 500 – 550 nm range is produced (Figure 3.1). Hill's reaction scheme was later modified for the quantitative analysis of fatty acids and fatty esters in blood samples and the analysis of acyl-coenzyme A (CoA) synthetase activity towards ester substrates, thus demonstrating a flexible colorimetric assay for ester quantification.[16, 17]

Here, we adapted and optimized the hydroxylamine/ferric iron reaction scheme to develop a method for the high throughput quantitative analysis of microbial ester biosynthesis. We used an engineered strain of *S. cerevisiae* overexpressing the terminal reaction step of ethyl acetate biosynthesis, alcohol-O-acetyltransferase (Atf),[18] and a low ethyl acetate producing strain of *S. cerevisiae* to validate the assay and test its suitability for high throughput, plated-based screening. To demonstrate the utility of the assay, we rapidly screened four wild type strains of the *K. marxianus* (CBS6556, DSM5422, DSM70106 and DSM5420) to determine their capacity to convert various carbohydrates, including glucose, fructose, galactose, lactose, cellobiose, xylose, and xylitol into the short chain volatile ester, ethyl acetate. *K. marxianus* was selected for screening because a number of different strains have been shown to grow on a variety of different C5 and C6 sugars and overproduction of ethyl acetate has been demonstrated.[13, 19, 20] *K. marxianus* is also an attractive microbial host for volatile ester synthesis as it is thermotolerant to temperatures upwards of 45 °C, genetic engineering tools have been demonstrated for metabolic engineering, and it is a generally regarded as safe (GRAS) organism.[12, 19, 21, 22] To demonstrate the broad utility of the assay we also investigated the extraction and quantification of longer chain esters including ethyl butyrate, isoamyl acetate, ethyl hexanoate and ethyl octanoate.

3.3 Methods and Materials

3.3.1 Strains and culturing conditions

Ethyl acetate biosynthesis were carried out with *S. cerevisiae* BY4742 and *S. cerevsisae* BY4742 with a genome integrated copy of its native alcohol-O-acetyltransferase 1 (ATF1) gene driven by a TDH3 promoter and a TEF2 terminator. Genomic integration at *yprcΔ15* was accomplished by homologous recombination of a linear DNA fragment with 50 base pairs homology up- and downstream of the integration site. Wild-type *S. cerevisiae* and *S. cerevisiae* expressing ATF1 were grown aerobically at 30 °C for 24 h in YM media (0.3% Yeast extract, 0.3% malt extract, 0.5% peptone (DB Difco®)) containing 1% and 10% glucose, respectively. Ethyl acetate biosynthesis was also carried out with four *K. marxianus* strains, including *K. marxianus* CBS6556 from the ATCC culture collection and DSM522, DSM70106, and DSM5420 from DSMX (Deutsche Sammlung von Mikroorganismen und Zellkulturen). *K. marxianus* strains were cultured in 2 mL of YM media with 1% of glucose. Overnight cultures were started from single colonies and were used to inoculate cultures at 37 °C. The optical density (OD) at 600 nm of each culture was measured using a Nanodrop 2000 UV-VIS spectrophotometer (Fisher Scientific). All initial culture ODs were adjusted to 0.05.

To assess the effect of alternative carbon sources on growth and ethyl acetate production in *K. marxianus,* 50 mL cultures were grown in 250 mL baffled flasks containing synthetic defined (SD) media (0.67% yeast nitrogen base without amino acids (DB Difco®), and 0.08% complete supplement mixture (CSM; Sunrise Science Products)) with 1% glucose, fructose, galactose, lactose, cellobiose, xylose or xylitol. Growth was monitored by measuring culture OD as described above. For sugar analysis, samples were collected at

appropriate time points and spun down at 15000 rpm for 10 mins., and supernatant was stored at -20 °C prior to analysis.

3.3.2 Ester Extraction

Produced esters were extracted from culture media by extraction with hexane (ReagentPlus®, ≥99%; Sigma Aldrich). Samples were centrifuged at 5000 rpm for 5 mins., hexane was added to isolated supernatant at volumetric ratios of 1:1, 1:2, or 1:5 and the extraction continued for 10 minutes at room temperature with orbital shaking at 200 rpm. After settling, 60 μL of the hexane layer was removed and used for each assay.

3.3.3 Colorimetric ester quantification assay

The ester quantification assay requires reaction of solvent extracted esters with hydroxylamine followed by reaction with ferric chloride. Hydroxylamine solutions were prepared by mixing 2.5 w/v % of hydroxylamine (Fisher Scientific) in 95% ethanol (Fisher Scientific) with 2.5 w/v % of sodium hydroxide (Fisher Scientific) in 95% ethanol at a 1:1 ratio. Hydroxylamine solutions were prepared fresh for each set of assays and were filtered prior to use. Stock ferric iron(III) solutions were prepared by dissolving 0.4 w/v% iron(III) chloride (Sigma Aldrich) in 50% perchloric acid (Fisher Scientific) and 50% de-ionizied water (MilliQ). Stock ferric solutions were diluted in ethanol (1:20) prior to use. All assays were conducted in flat bottom 96-well plates.

Hydroxymates were produced by combining 20 μL of hydroxylamine stock solution with 60 μL of hexane extracted ester and allowing the reaction to proceed for 5 to 30 minutes. Subsequently, 120 μL of the ferric working solution was added and incubated for 5 to 30 minutes. The production of ferric hydroxamate was measured at 520 nm with a

BioTek Synergy 2, Multi-Mode UV-Vis plate reader. In the optimized assay protocol, the hydroxymate reaction proceeded for 10 mins., while ferric chelation was allowed to proceed for 5 mins. at room temperature.

3.3.4 Ester analysis by headspace gas chromatography (GC)

The colorimetric ester assay was compared to ester quantification by headspace gas chromatography. Two-mL fermentation cultures were centrifuged at 4°C and 5000 rpm for 5 mins. and 1 mL of the supernatant was placed in a 10 mL vial with septum top (Fisher) for headspace GC analysis. Prior to analysis, 1 mL of NaCl was added to the solution to decrease ethyl acetate solubility as previously described.[14] Ethyl acetate concentrations were measured using an Agilent 7890A system equipped with a Restek Rtx® –Wax column and a FID detector.

3.3.5 Sugar analysis

Isolated spent media was analyzed using a Waters Alliance 2695 High Pressure Liquid Chromatography (HPLC) system equipped with a Bio-Rad AminexHPX-87H column and Waters 2414 refractive index (RI) detector. The column was heated to 65 °C and analysis was performed using an eluent of 5 mM sulfuric acid at a flow rate of 0.6 mL min^{-1}. Peaks were analyzed using the Empower 2 software.

3.3.6 Z-factor and statistical analysis.

To analyze the suitability of the ester detection assay for high throughput screening, the assay's z-factor was calculated using the resulting A_{520} values of the positive and negative controls. For ethyl acetate biosynthesis, *S. cerevisiae* BY4742 was used as a

negative control and *S. cerevisiae* BY4742 with a genome integrated copy of ATF1 (*yprcΔ15*::ATF1) was used as the positive control. Ethyl acetate production was increased by using media containing 10% glucose. The arithmetic mean, μ, and the standard deviations, σ, of the positive (p) and negative (n) control data sets were used to calculate the z-factor as follows:

$$Z = \frac{3(\sigma_p+\sigma_n)}{|\mu_p-\mu_n|} \qquad \text{Eqn. (1)}$$

All other measurements were made in triplicate, with the mean and the standard deviation reported. Comparison between GC analysis and ester concentration measurements with the colorimetric assay were accomplished by an unpaired two-tailed T-test with a significant difference at $p< 0.05$.

3.4 Results and Discussion

3.4.1 A colorimetric assay for quantifying acyl and acetyl esters in fermentation broth

Based on the two-step reaction scheme identified by Hill,[15] a ferric hydroxamate-based reporter assay for esters extracted from microbial fermentations was developed (Figure 3.1A). In the first reaction, solvent extracted esters were reacted with an alkaline hydroxylamine solution (95% ethanol in water) to produce hydroxamic acid and a primary alcohol corresponding to the R2 group of the ester. The formed hydroxamic acid, (carrying the R1 group of the ester) was then combined with ferric iron to form ferric hydroxamate, which has strong purple color and absorbance at 520 nm (Figure 3.1B). Hexane was selected as the solvent for ester extraction because it is immiscible in culture broth, has a high capacity to extract short and medium chain esters such as ethyl acetate and isoamyl acetate and is miscible in the assay reaction mixture. An optimized protocol for plate-based assays was achieved by varying the times of the first and second reactions, reactant

concentrations, and reaction volume. Highly repeatable results were achieved by reacting 60 μL of hexane extracted ester with 20 μL of hydroxylamine solution for 10 minutes at room temperature followed by the addition of 120 μL of ferric solution for 5 minutes (Figure S3.1). This protocol produced quantitative measures of ethyl acetate, ethyl butyrate isoamyl acetate, ethyl hexanoate, and ethyl octanoate in the range of 0-200 mg L^{-1} of ester in YM media using a hexane to media ratio of 1:1 by volume (Figure 3.1C). The range of testable ester concentrations can be extended by varying the ratio of culture media to solvent from 1:1 to 1:2 or 1:5 (Figure S3.2). Assay calibrations shown in Figure 3.1C vary across the tested esters potentially due to differences in extraction efficiency and equilibrium constants. [23, 24] In each case, 520 nm was used for quantification resulting in the following coefficients: ethyl acetate $6.09x10^{-4}$ mg^{-1} L, isoamyl acetate $3.55x10^{-4}$ mg^{-1} L, ethyl butyrate $2.82x10^{-4}$ mg^{-1} L, ethyl hexanoate $1.17x10^{-4}$ mg^{-1} L, and ethyl octanoate $1.04x10^{-4}$ mg^{-1} L.

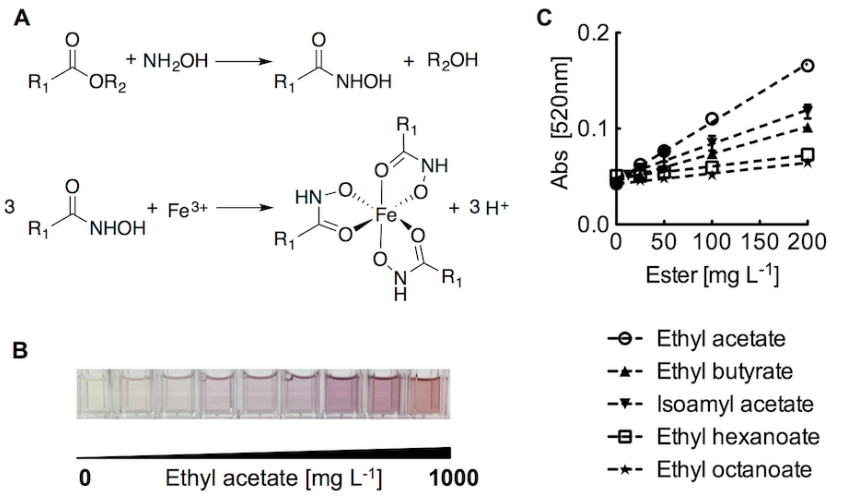


Figure 3.1: A colorimetric assay for rapid screening of microbial ester biosynthesis. (A) Assay reaction scheme. Extracted esters are reacted with hydroxylamine to produce hydroxamic acid and an alcohol. A colored ferric hydroxamate complex is formed by the addition of ferric iron(III). (B) Assay calibrations for ethyl acetate, ethyl butyrate, isoamyl acetate, ethyl hexanoate and ethyl octanoate. (C) Image of the color range produced by the two-step assay from solutions containing 0 to 1000 mg L^{-1} of ethyl acetate.

3.4.2 Z-factor analysis demonstrates suitability of the colorimetric assay for high throughput screening of ester biosynthesis

To be suitable for high throughput screening an assay must produce a statistically significant difference between positive and negative controls as quantified by the Z-factor (Eqn. 1). Z-factor values between 0.5 to 1.0 are suitable for high throughput screening.[25] *S. cerevisiae* BY4742 under aerobic conditions was used as the negative control, while *S. cerevisiae* BY4742 with a genome integrated overexpression cassette of ATF1, the terminal reaction step of ethyl acetate biosynthesis in yeast, was used as the positive control (Figure 3.2A). In wild type *S. cerevisiae*, ATF1 and ester biosynthesis are suppressed in the presence of oxygen,[26] correspondingly the negative control strain produced no ethyl acetate (Figure S3.3; 0 mg L^{-1}, A_{520} = 0.0450±0.0015). The constitutive overexpression of Atf1 in the positive control strain resulted in the formation of 143±9.7 mg L^{-1} of ethyl acetate (A_{520} = 0.176±0.009) and a z-factor value of 0.77.

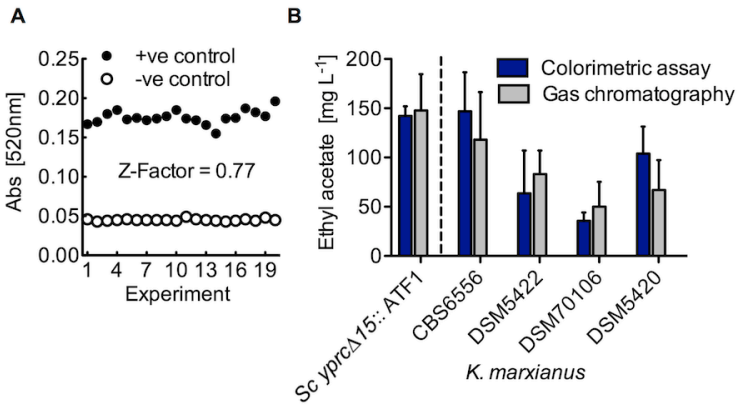


Figure 3.2: Z-factor analysis and validation of a colorimetric assay for quantifying short and medium chain esters in fermentation broth. (A) Z-factor analysis of ethyl acetate microbial biosynthesis. *S. cerevisiae* BY4742 was used as a negative control, while *S. cerevisiae* BY4742 *yprcΔ15::ATF1* was used as a positive control. (B) Ethyl acetate production from *S. cerevisiae* BY4742 *yprcΔ15::ATF1* and *K. marxianus* strains CBS6556, DSM 5422, DSM70106, and DSM5420 quantified by gas chromatography and the ferric hydroxamate colorimetric assay. *K. marxianus* cultures were assayed after 24 h of anaerobic fermentation at 37 °C. *S. cerevisiae* cultures were assayed after 24 h of aerobic growth at 30 °C. Bars represent the arithmetic mean of a minimum of three biological replicates and the error bars represent standard deviations.

In order to validate the assay, ethyl acetate production from the *S. cerevisiae* positive control as well as native ethyl acetate production from four different wild type strains of *K. marxianus* were compared using GC analysis and the colorimetric assay (Figure 3.2B). In each case, there was no statistical difference between the GC results and the ethyl acetate concentrations measured by the colorimetric assay. It is important to note that in its current format the assay does not distinguish between different esters synthesized in a single culture and was developed for the quantification of overproduced esters from native or engineered microbial strains. GC analysis of volatiles produced from *S. cerevisiae* with overexpressed Atf1 revealed that ethyl acetate was the only ester produced in significant quantities under the tested conditions (Figure S3.3). Similar analysis of *K. marxianus* CBS6556 shows that ethyl acetate represents upwards of 93% of the volatiles on synthetic media (Figure S3.4). Hexane extraction selectively partitions ethyl acetate to the organic solvent phase along with any other short and medium chain esters, and it is possible that other esters are quantified by the assay; however, the contributions of other esters are not detected in the comparison of the GC and assay analyses presented in Figure 3.2B.

3.4.3 Combinatorial screening of *K. marxianus* strains and C5, C6, and C12 sugars reveals high ethyl acetate biosynthesis on C6 sugars

Various strains of *K. marxianus* have been shown to grow on C5, C6, and C12 sugars and some strains are known to produce high titers of ethyl acetate.[12, 19, 20, 27] To demonstrate the capabilities of the colorimetric ester assay, we rapidly screened ethyl acetate biosynthesis in four *K. marxianus* strains grown on seven different carbon sources including the C6 sugars glucose, fructose, and galactose, the C5 sugars xylose and xylitol, and the C12 sugars lactose and cellobiose (Figure 3.3).

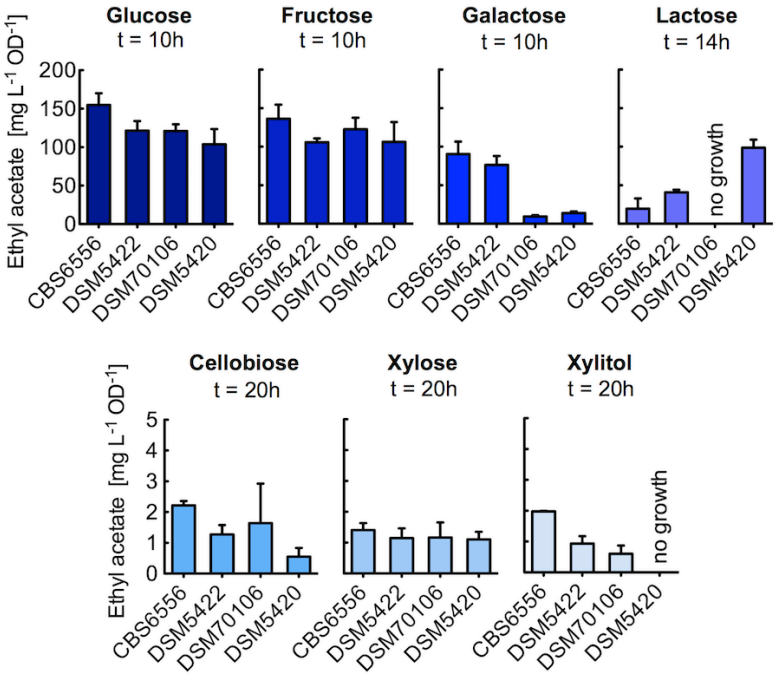


Figure 3.3: Combinatorial screening of *K. marxianus* strains on various C5, C6, and C12 carbon sources. *K. marxianus* strains were cultured at 37°C in SD media containing 1% carbon source. Ethyl acetate production and optical density (OD_{600}) were measured at late exponential phase (10 h for glucose, lactose and galactose, 14 h for lactose, 20 h for cellobiose, xylose and xylitol). Bars represent the arithmetic mean of three biological replicates and the error bars represent the standard deviation.

To determine the point of highest ethyl acetate production, *K. marxianus* strain CBS6556 was grown at 37 °C on different sugars while measuring ethyl acetate production and cell growth at regular time intervals from inoculation to stationary phase. For all sugars, the highest ethyl acetate concentration was achieved as cultures approached stationary phase (Figure S3.5), as such, time points corresponding to this growth phase were selected for screening. Strains grown on glucose, fructose, and galactose were screened at 10 h, strains grown on lactose were analyzed at 14 h, and strains grown on cellobiose, xylose, and xylitol were tested at 20 h. With the exception of two strain/sugar combinations, *K. marxianus* produced measurable quantities of ethyl acetate with each

sugar feed (Figure 3.3). Strains DSM70106 and DSM5420 were the exceptions and did not grow on lactose and xylitol, respectively.

Consistent with previous reports, *K. marxianus* strains CBS6556, DSM5422, DSM70106, and DSM5420 showed measurable growth on C5, C6, and C12 sugars (Figure S3.6).[12, 19, 20] Our combinatorial strain and carbon source screening adds to our understanding of *K. marxianus* sugar metabolism by revealing differences in growth across strains. Growth on glucose and fructose was similar across all tested strains, with strains CBS6556 and DSM5422 resulting in marginally higher OD's than strains DSM70106 and DSM5420 (ODs on glucose: 13.9±0.5, 15.1±0.7, 11.2±0.73 and 12.0±1.5, respectively; ODs on fructose: 13.9±0.6, 15.8±1.0, 10.3±1.1, and 11.1±1.7, respectively). By comparison, growth on galactose, lactose, cellobiose, xylose, and xylitol was highly strain dependent.

The highest specific formation of ethyl acetate was observed with strain CBS6556 on glucose, which produced 155±15.2 mg $L^{-1}OD^{-1}$ (Figure 3.3). In general, ethyl acetate production and titer across all strains was highest on the C6 sugars glucose and fructose (ethyl acetate from glucose, mg $L^{-1}OD^{-1}$: 155±15.2 for CBS6556, 122±12.3 for DSM5422, 121±8.9 for DSM70106, and 104±19.9 for DSM5420; ethyl acetate from fructose, mg $L^{-1}OD^{-1}$: 136±18.4 for CBS6556, 106±5.0 for DSM5422, 123±15.1 for DSM70106, and 106±25.8 for DSM5420; ethyl acetate titers in mg L^{-1} are presented in Figure S3.7). In addition, sugar uptake analysis showed that glucose and fructose were depleted or nearly depleted from culture media as cultures approached stationary phase (Figures S3.8).

The preference of some *K. marxianus* strains towards lactose and galactose as carbon sources has previously been exploited in the development of an ethyl acetate bioprocess using waste whey containing lactose and galactose as a feedstock.[12, 27] Pilot plant studies found that *K. marxianus* DSM5422 with whey permeate as a feedstock could

produce 2.45 g $L^{-1}hr^{-1}$ of ethyl acetate in a 70 L aerobic batch process.[28] An important difference in the pilot plant experiments and our screening experiments is the culture conditions. The pilot plant and bioprocess development by the same research group used aerated cultures with low iron content media, which inhibits enzymes of the electron transport chain and the Krebs cycle to induce fermentation and the formation of ethyl acetate.[29] In our experiments, ethyl acetate production was screened in baffled shake flasks containing media with excess iron and without nutrient limitations. The screen revealed that ethyl acetate production on galactose and lactose was strain dependent. Strains CBS6556 and DSM5422 produced high amounts of ethyl acetate on galactose (90.5±16.1 and 76.4±11.5 mg $L^{-1}OD^{-1}$), but production was reduced on lactose (19.6±13.2 and 50.9±3.3 mg $L^{-1}OD^{-1}$). Galactose uptake in DSM70106 and DSM5420 was limited (Figure S3.8), resulting in poor growth and limited ethyl acetate production in comparison to the other strains (9.6±1.6 mg $L^{-1}OD^{-1}$ and 14.0±2.1 mg $L^{-1}OD^{-1}$ of ethyl acetate for DSM70106 and DSM5420, respectively). DSM5420 grew well on lactose, producing 98.9±10.3 mg $L^{-1}OD^{-1}$ of ethyl acetate, outperforming all other strains in terms of converting lactose to ethyl acetate (Figure 3.3).

Recent studies have focused on developing consolidated bioprocessing (CBP) strains of *K. marxianus* by engineering the heterologous expression of cellulases to release xylose, glucose, and the glucose dimer cellobiose from pretreated biomass. Overexpression of a series of cellulases including a β–glucosidase, an endo-glucanase, and an exo-glucanase resulted in the production of 8 g L^{-1} ethanol on 2% cellobiose.[30] In another study, *K. marxianus* was engineered to produce 43.4 g L^{-1} ethanol from 10% cellobiose by overexpressing a similar series of cellulases.[31] Our results showed that *K. marxianus* is capable of metabolizing cellobiose and synthesizing ethyl acetate at a rate of up to 2.2±0.1

mg $L^{-1}OD^{-1}$ (*K. marxianus* CBS6556) without the addition of heterologous cellulose degrading enzymes (Figures 3.3 and S3.8). Recent transcriptional analysis suggests the presence of a β-glucosidase homolog in *K. marxianus*, which may be responsible for cellobiose degradation.[32] All strains were also able to uptake and grow on xylose, but ODs were significantly reduced in comparison to growth on glucose and ethyl acetate production was low (1.4±0.2 mg $L^{-1}OD^{-1}$ for CBS6556; 1.1±0.3 for DSM5422; 1.2±0.5 for DSM70106; and 1.1±0.2 for DSM5420). CBS6556 showed increased sugar uptake, growth, and ethyl acetate productivity on xylitol in comparison to xylose, but no significant differences were observed with DSM5422 and DSM70106, and DSM5420 did not grow on xylitol (Figure S3.6). A number of reports support the findings that *K. marxianus* has limited capacity to uptake C5 sugars in comparison to its ability to metabolize glucose and other C6 sugars.[20, 33] Genetic engineering of xylose metabolism aimed at eliminating a redox imbalance during xylose catabolism have been demonstrated and result in improved xylose uptake and *K. marxianus* growth.[22, 34, 35]

Taken together, the adaptation of a colorimetric ester assay and ethyl acetate screening from a set of *K. marxianus* strains demonstrates a new method for the rapid quantification of microbial ester biosynthesis. Application to plate-based formats will enable parallel screening of large libraries of wild type and mutant microbial strains that can be integrated with liquid handling robotics to increase throughput. We demonstrated that the assay is applicable to short and medium chain esters and other reports have demonstrated a similar assay for longer chain esters, suggesting broad flexibility and utility for quantifying the overproduction of different fragrance and flavor compounds as well as the biosynthesis of wax esters and diesel- and jet-fuel replacements. In its current format, the assay does not distinguish between different esters and is best applied where one (or

more) specific esters are overproduced. In cases where the identification of all endogenous esters is critical, GC-MS analysis may be more appropriate. Together with the rapid screening of ester biosynthesis enzymes,[36] this assay provides a high throughput platform for screening and engineering microbial ester biosynthesis.

3.5 Concluding Remarks

Here we developed a biotechnology method to rapidly screen microbial ester biosynthesis. The colorimetric assay, which leverages ester reactivity with hydroxylamine to produce a purple colored ferric hydroxamate complex, is suitable for high throughput screening. To demonstrate the capabilities of the assay, we rapidly screened a combinatorial library of four *K. marxianus* strains and seven different carbohydrate carbon sources for the biosynthesis of ethyl acetate, a naturally produced short chain volatile ester. The screen identified two C12 sugars, lactose and cellobiose, as suitable substrates for ethyl acetate fermentation and demonstrated ethyl acetate production on xylose, xylitol, glucose, galactose, and fructose. While the assay is not capable of distinguishing between different esters it is readily applicable for rapid, plate-based screening for ester overproduction in engineered and natural strain libraries. The flexibility of the assay to quantify not only ethyl acetate but also the medium chain esters suggests broad utility and will help match the throughput of strain testing with current technologies that can generate large libraries of engineered microbial strains in the "design-build-test cycle" of synthetic biology.

3.6 Abbreviations: Atf1, alcohol acetyl/acyl transferase; **CBP**, consolidated bioprocessing; **OD**, optical density; **GC**, gas chromatography; **MS**, mass spectrometry; **TLC**, thin layer chromatography.

3.7 Acknowledgement

This work was supported by NSF CBET-1510697.

3.8 Conflict of interest

The authors declare no financial or commercial conflict of interest.

3.9 References

[1] Wang, H. H., Isaacs, F. J., Carr, P. A., Sun, Z. Z., *et al.*, **Programming cells by multiplex genome engineering and accelerated evolution.** *Nature* 2009, *460*, 894-898.

[2] Lynch, M. D., Warnecke, T., Gill, R. T., **SCALEs: multiscale analysis of library enrichment.** *Nat Methods* 2007, *4*, 87-93.

[3] Smanski, M. J., Bhatia, S., Zhao, D. H., Park, Y., *et al.*, **Functional optimization of gene clusters by combinatorial design and assembly.** *Nat Biotechnol* 2014, *32*, 1241-1249.

[4] DeLoache, W. C., Russ, Z. N., Narcross, L., Gonzales, A. M., *et al.*, **An enzyme-coupled biosensor enables (S)-reticuline production in yeast from glucose.** *Nat Chem Biol* 2015, *11*, 465-471.

[5] Tao, H., Guo, D. Y., Zhang, Y. C., Deng, Z. X., Liu, T. G., **Metabolic engineering of microbes for branched-chain biodiesel production with low-temperature property.** *Biotechnol Biofuels* 2015, *8*.

[6] Steen, E. J., Kang, Y. S., Bokinsky, G., Hu, Z. H., *et al.*, **Microbial production of fatty-acid-derived fuels and chemicals from plant biomass.** *Nature* 2010, *463*, 559-U182.

[7] Valle-Rodriguez, J. O., Shi, S. B., Siewers, V., Nielsen, J., **Metabolic engineering of Saccharomyces cerevisiae for production of fatty acid ethyl esters, an advanced biofuel, by eliminating non-essential fatty acid utilization pathways.** *Appl Energ* 2014, *115*, 226-232.

[8] Runguphan, W., Keasling, J. D., **Metabolic engineering of Saccharomyces cerevisiae for production of fatty acid-derived biofuels and chemicals.** *Metab Eng* 2014, *21*, 103-113.

[9] Layton, D. S., Trinh, C. T., **Engineering modular ester fermentative pathways in Escherichia coli.** *Metab Eng* 2014, *26*, 77-88.

[10] Rodriguez, G. M., Tashiro, Y., Atsumi, S., **Expanding ester biosynthesis in Escherichia coli.** *Nat Chem Biol* 2014, *10*, 259-+.

[11] Tai, Y. S., Xiong, M. Y., Zhang, K. C., **Engineered biosynthesis of medium-chain esters in Escherichia coli.** *Metab Eng* 2015, *27*, 20-28.

[12] Urit, T., Loser, C., Wunderlich, M., Bley, T., **Formation of ethyl acetate by Kluyveromyces marxianus on whey: studies of the ester stripping.** *Bioproc Biosyst Eng* 2011, *34*, 547-559.

[13] Loser, C., Urit, T., Nehl, F., Bley, T., **Screening of Kluyveromyces strains for the production of ethyl acetate: Design and evaluation of a cultivation system.** *Eng Life Sci* 2011, *11*, 369-381.

[14] Zhu, J., Lin, J. L., Palomec, L., Wheeldon, I., **Microbial host selection affects intracellular localization and activity of alcohol-O-acetyltransferase.** *Microb Cell Fact* 2015, *14*.

[15] Hill, U. T., **Colorimetric Determination of Fatty Acids and Esters.** *Anal Chem* 1947, *19*, 932-933.

[16] Wofford, N. Q., Beaty, P. S., Mcinerney, M. J., **Preparation of Cell-Free-Extracts and the Enzymes Involved in Fatty-Acid Metabolism in Syntrophomonas-Wolfei.** *J Bacteriol* 1986, *167*, 179-185.

[17] Stern, I., Shapiro, B., **A Rapid and Simple Method for the Determination of Esterified Fatty Acids and for Total Fatty Acids in Blood.** *J Clin Pathol* 1953, *6*, 158-160.

[18] Lin, J. L., Wheeldon, I., **Dual N- and C-Terminal Helices Are Required for Endoplasmic Reticulum and Lipid Droplet Association of Alcohol Acetyltransferases in Saccharomyces cerevisiae.** *PloS one* 2014, *9*.

[19] Fonseca, G. G., de Carvalho, N. M. B., Gombert, A. K., **Growth of the yeast Kluyveromyces marxianus CBS 6556 on different sugar combinations as sole carbon and energy source.** *Appl Microbiol and Biotechnol* 2013, *97*, 5055-5067.

[20] Nonklang, S., Abdel-Banat, B. M. A., Cha-Aim, K., Moonjai, N.*, et al.*, **High-Temperature Ethanol Fermentation and Transformation with Linear DNA in the Thermotolerant Yeast Kluyveromyces marxianus DMKU3-1042.** *Appl Environ Microbiol* 2008, *74*, 7514-7521.

[21] Hoshida, H., Murakami, N., Suzuki, A., Tamura, R.*, et al.*, **Non-homologous end joining-mediated functional marker selection for DNA cloning in the yeast Kluyveromyces marxianus.** *Yeast* 2014, *31*, 29-46.

[22] Zhang, J., Zhang, B., Wang, D. M., Gao, X. L.*, et al.*, **Rapid ethanol production at elevated temperatures by engineered thermotolerant Kluyveromyces marxianus via the NADP(H)-preferring xylose reductase-xylitol dehydrogenase pathway.** *Metab Eng* 2015, *31*, 140-152.

[23] Antonis, A., Platt, D. S., Thorp, J. M., **Automated Method for the Colorimetric Determination of Acyl Esters in Serum.** *J Lipid Res* 1965, *6*, 301-306.

[24] Thompson, A. R., **A Colorimetric Method for the Determination of Esters.** *Aust J Sci Res Ser A* 1950, *3*, 128-135.

[25] Zhang, J. H., Chung, T. D. Y., Oldenburg, K. R., **A simple statistical parameter for use in evaluation and validation of high throughput screening assays.** *J Biomol Screen* 1999, *4*, 67-73.

[26] Fujiwara, D., Kobayashi, O., Yoshimoto, H., Harashima, S., Tamai, Y., **Molecular mechanism of the multiple regulation of the Saccharomyces cerevisiae ATF1 gene encoding alcohol acetyltransferase.** *Yeast* 1999, *15*, 1183-1197.

[27] Kallelmhiri, H., Engasser, J. M., Miclo, A., **Continuous Ethyl-Acetate Production by Kluyveromyces-Fragilis on Whey Permeate.** *Appl Microbioly Biotechnol* 1993, *40*, 201-205.

[28] Loser, C., Urit, T., Stukert, A., Bley, T., **Formation of ethyl acetate from whey by Kluyveromyces marxianus on a pilot scale.** *J Biotechnol* 2013, *163*, 17-23.

[29] Loser, C., Urit, T., Keil, P., Bley, T., **Studies on the mechanism of synthesis of ethyl acetate in Kluyveromyces marxianus DSM 5422.** *Appl Microbiol Biotechnol* 2015, *99*, 1131-1144.

[30] Chang, J. J., Ho, C. Y., Ho, F. J., Tsai, T. Y.*, et al.*, **PGASO: A synthetic biology tool for engineering a cellulolytic yeast.** *Biotechnol Biofuels* 2012, *5*.

[31] Hong, J., Wang, Y., Kumagai, H., Tamaki, H., **Construction of thermotolerant yeast expressing thermostable cellulase genes.** *J Biotechnol* 2007, *130*, 114-123.

[32] Lertwattanasakul, N., Kosaka, T., Hosoyama, A., Suzuki, Y.*, et al.*, **Genetic basis of the highly efficient yeast Kluyveromyces marxianus: complete genome sequence and transcriptome analyses.** *Biotechnol Biofuels* 2015, *8*, 47.

[33] Zhang, M., Shukla, P., Ayyachamy, M., Permaul, K., Singh, S., **Improved bioethanol production through simultaneous saccharification and fermentation of lignocellulosic agricultural wastes by Kluyveromyces marxianus 6556.** *World J Microb Biot* 2010, *26*, 1041-1046.

[34] Zhang, B. A., Li, L. L., Zhang, J., Gao, X. L.*, et al.*, **Improving ethanol and xylitol fermentation at elevated temperature through substitution of xylose reductase in Kluyveromyces marxianus.** *J Ind Microbiol Biotechnol* 2013, *40*, 305-316.

[35] Zhang, B., Zhang, J., Wang, D., Han, R.*, et al.*, **Simultaneous fermentation of glucose and xylose at elevated temperatures co-produces ethanol and xylitol through overexpression of a xylose-specific transporter in engineered Kluyveromyces marxianus.** *Bioresour Technol* 2016, *216*, 227-237.

[36] Lin, J. L., Zhu, J., Wheeldon, I., **Rapid ester biosynthesis screening reveals a high activity alcohol-O-acyltransferase (AATase) from tomato fruit.** *Biotechnol J* 2016, *11*, 700-707.

3.10 Supporting information

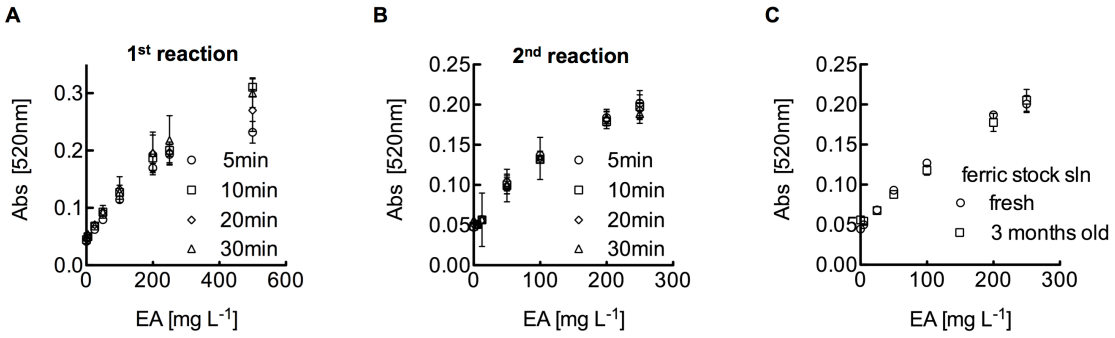


Figure S3.1: Optimization of the colorimetric assay for ester quantification. (A) Varying reaction time of hydroxamic acid formation with ethyl acetate and hydroxylamine (1st reaction). 60 µL of hexane extracted ethyl acetate was incubated with 20 µl of hydroxylamine solution for 5, 10, 20 and 30 minutes at room temperature. Linear trends were achieved with a minimum of 10 minutes of incubation time. (B) Ferric hydroxamate complex formation was optimized by incubating the hydroxamate mixture from the first reaction with 120 µl of ferric perchlorate solution for 5, 10, 20 and 30 minutes (2nd reaction). Five-minutes was determined to be sufficient to reproducibly form the purple ferric complex. (C) A comparison between freshly prepared ferric iron solution and a ferric iron solution stored at 4°C for 3 months.

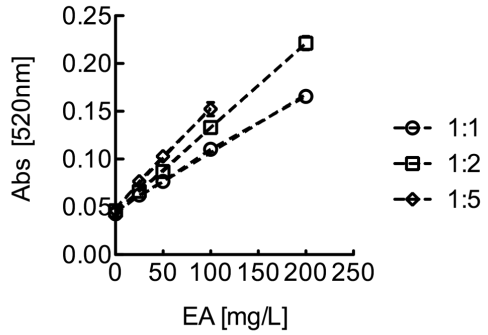
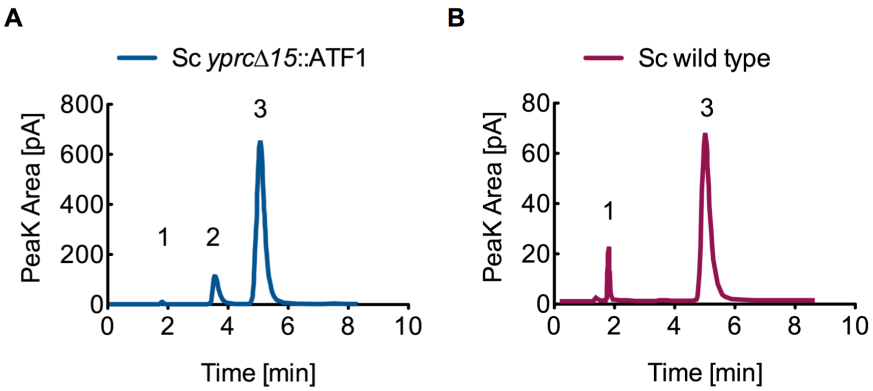


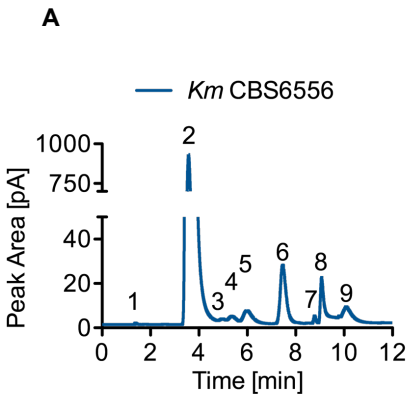
Figure S3.2: Different ratios of ethyl acetate containing YM media to hexane were tested for extraction.



C

Peak #	Retention Time [min]	Peak Area A [pA s]	Peak Area B [pA s]
1	1.8	53.26	94.18
2	3.6	1632.13	-
3	5.0	12397.00	1350.69

Figure S3.3: GC-FID headspace analysis of *S. cerevisiae* aerobic cultures. (A) Chromatogram of volatiles from *S. cerevisiae* BY4742 *yprcΔ15*::ATF1 grown at 30 °C in 2 mL of YM media. Gas analysis was accomplished after 24 h of culture at 30 °C using 1 mL of media with cells removed by centrifugation. (B) Chromatogram of volatiles from *S. cerevisiae* BY4742. Peak 1 corresponds to acetaldehyde, Peak 2 corresponds to ethyl acetate and peak 3 corresponds to ethanol.



B

Peak #	Retention Time [min]	Peak Area [pA s]	% Area
1	1.4	3.5	0.02
2	3.6	14260.5	92.9
3	5.0	42.7	0.28
4	5.4	82.6	0.54
5	6.0	167.1	1.09
6	7.5	446.9	2.91
7	8.8	20.0	0.13
8	9.1	203.2	1.3
9	10.1	127.8	0.83

Figure S3.4: GC-FID headspace analysis of *K. marxianus* CBS6556 aerobic cultures. *K. marxianus* CBS6556 was grown for 10 h in YM media containing 1% glucose. Peak 1 corresponds to acetaldehyde, peak 2 corresponds to ethyl acetate and peak 3 corresponds to ethanol. Ethyl acetate represents ~93% of the total peak area.

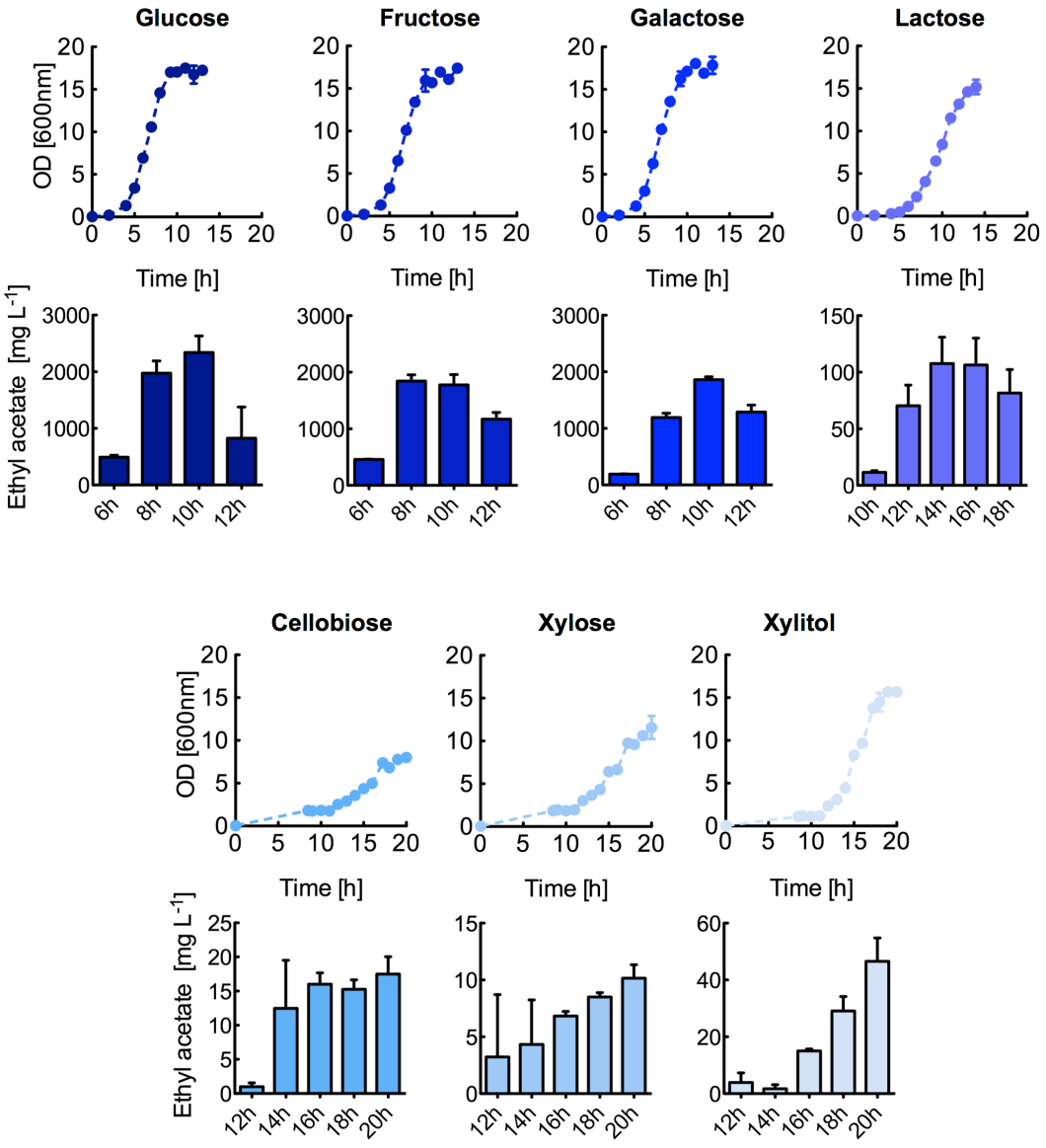


Figure S3.5: Time-dependent ethyl acetate production. To determine the optimal time point for ethyl acetate analysis *K. marxianus* CBS6556 was grown on SD media containing 1% of the appropriate sugars aerobically at 37 °C. The top set of panels shows the optical densities (OD) measured over time, while the bottom set of panels shows ethyl acetate concentration as measured by the developed colorimetric assay. The time point of highest ethyl acetate concentration was determined to be late exponential phase, 10 h for glucose, fructose and galactose; 14h for lactose; and 20h for cellobiose, xylose, and xylitol. Bars represent arithmetic mean of three biological replicates and the error bars represent the standard deviation.

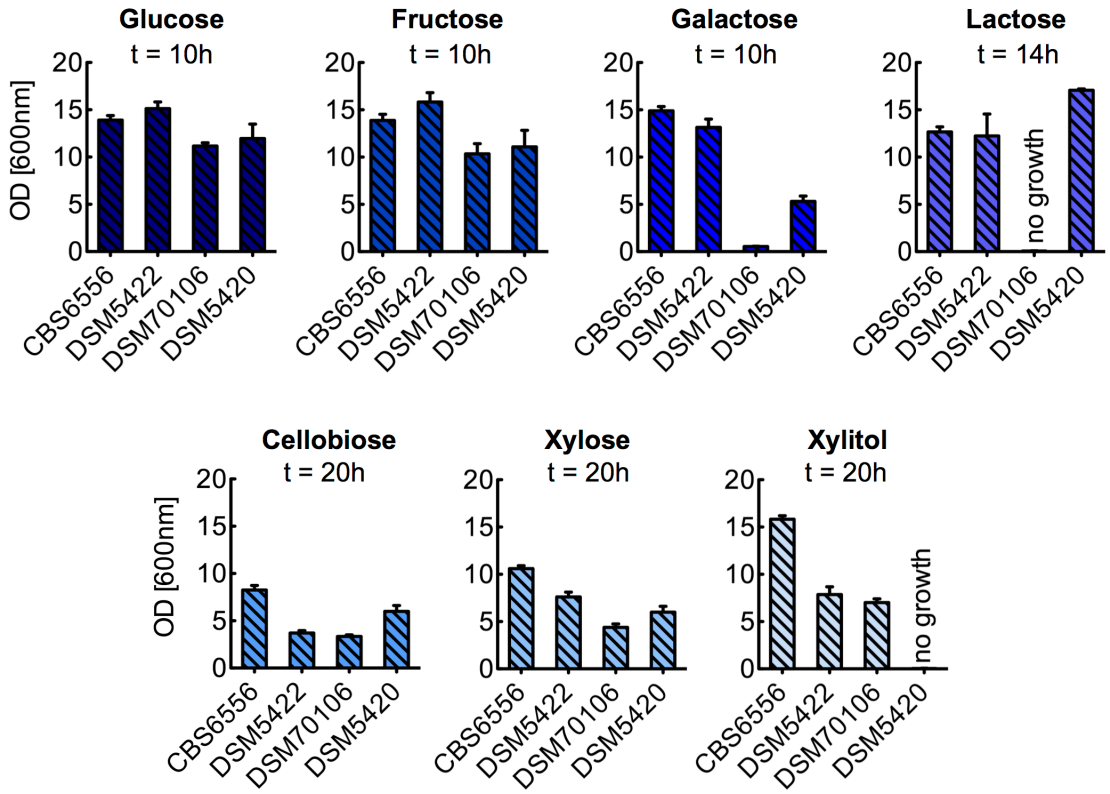


Figure S3.6: Optical density of *K. marxianus* strains at late exponential phase. *K. marxianus* strains were grown for 10 h (glucose, fructose, galactose), 14 h (lactose), and 20 h (cellobiose, xylose, xylitol). Bars represent arithmetic mean of three biological replicates and the error bars represent the standard deviation.

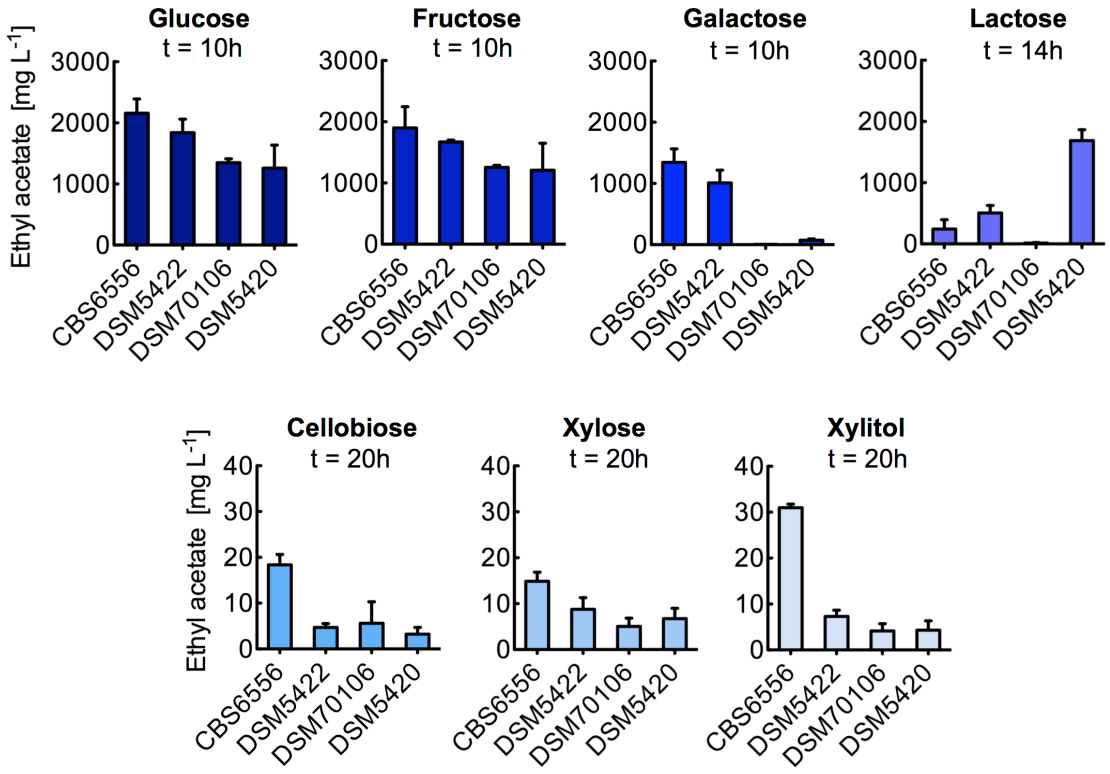


Figure S3.7: Ethyl acetate production by four *K. marxianus* strains at late exponential phase. Four different *K. marxianus* were grown to late exponential phase and ethyl acetate concentration was measured. Glucose, fructose and galactose cultures were grown for 10 h, lactose cultures were grown for 14 h; and cellobiose, xylose and xylitol cultures were grown for 20 h. Bars represent arithmetic mean of three biological replicates and the error bars represent the standard deviation.

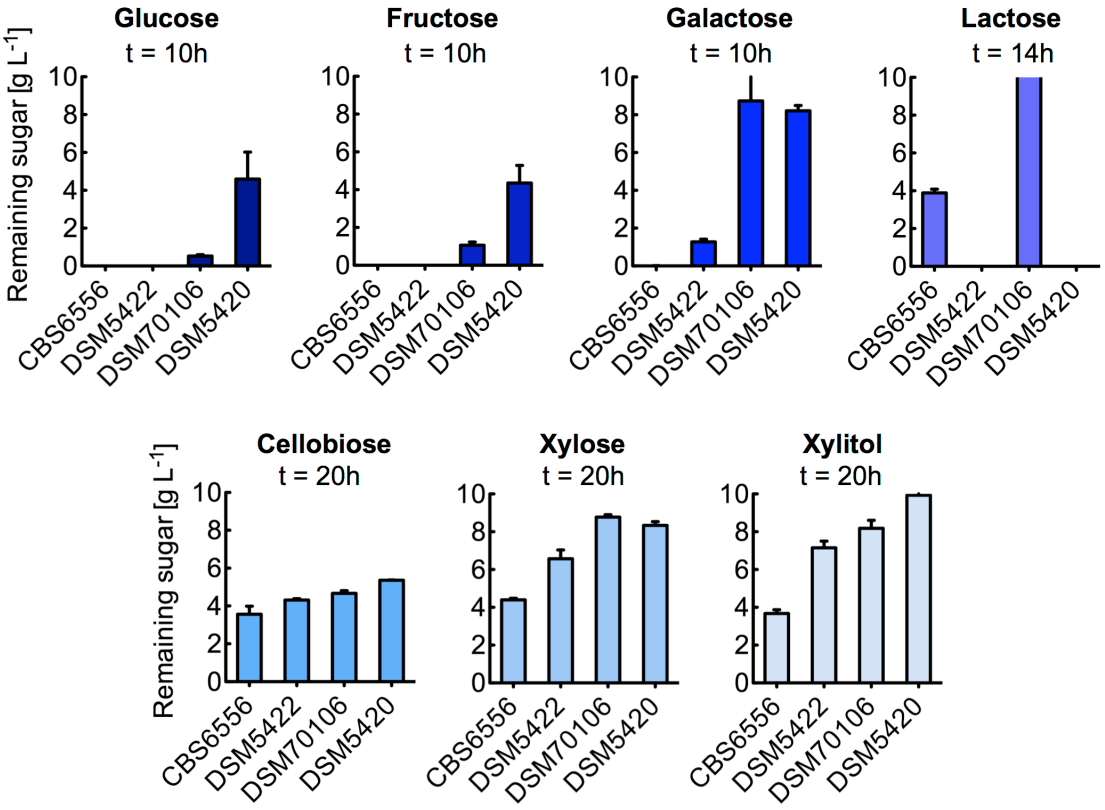


Figure S3.8: Sugar uptake by *K. marxianus* strains CBS6556, DSM5422, DSM70106 at late exponential phase. Strains were grown in media containing 10 g/L of the appropriate carbon source, sugar concentrations at each indicated time point are shown. Bars represent arithmetic mean of three biological replicates and the error bars represent the standard deviation.

Chapter 4: CRISPR-Cas9 enabled genetic disruptions for understanding ethanol and ethyl acetate biosynthesis in *Kluyveromyces marxianus*

4.1 Abstract

4.1.1 Background

The thermotolerant yeast *Kluyveromyces marxianus* shows promise as an industrial host for the biochemical production of fuels and chemicals. Wild type strains are known to ferment high titers of ethanol and can effectively convert a wide range of C_5, C_6, and C_{12} sugars into the volatile short-chain ester ethyl acetate. Strain engineering, however, has been limited due to a lack of advanced genome editing tools and an incomplete understanding of ester and ethanol biosynthesis.

4.1.2 Results

Enabled by the design of hybrid RNA polymerase III promoters, this work adapts the CRISPR-Cas9 system from *Streptococcus pyogenes* for use in *K. marxianus*. The system was used to rapidly create functional disruptions to alcohol dehydrogenase (ADH) and alcohol-O-acetyltransferase (ATF) genes with putative function in ethyl acetate and ethanol biosynthesis. Screening of the *Km*ATF disrupted strain revealed that Atf activity contributes to ethyl acetate biosynthesis, but the knockout reduced ethyl acetate titers by only ~15%. Overexpression experiments revealed that *Km*Adh7 can catalyze the oxidation of hemiacetal to ethyl acetate. Finally, analysis of the *Km*ADH2 disrupted strain showed that the knockout almost completely eliminated ethanol production and resulted in the accumulation of acetaldehyde.

4.1.3 Conclusions

Newly designed RNA polymerase III promoters for sgRNA expression in *K. marxianus* enable a CRISPR-Cas9 genome editing system for the thermotolerant yeast. This system was used to disrupt genes involved in ethyl acetate biosynthesis, specifically *Km*ADH1-7 and *Km*ATF. *Km*Adh2 was found to be critical for aerobic and anaerobic ethanol production. Aerobically produced ethanol supplies the biosynthesis of ethyl acetate catalyzed by *Km*Atf. *Km*Adh7 was found to exhibit activity towards to oxidation of hemiacetal, a possible alternative route for the synthesis of ethyl acetate.

4.2 Background

The yeast *Kluyveromyces marxianus* is a potentially valuable host for industrial biochemical synthesis. Wild type strains are known to have fast growth kinetics, thermotolerance to ~50 °C, the ability to metabolize a range of monomeric and dimeric C_5 and C_6 sugars, and strong fermentation pathways for ethanol production.[1-3] A high capacity for protein expression and secretion is also advantageous and has been exploited in bioprocesses for inulase and galactosidase production.[4, 5] Another industrially relevant phenotype is the native capacity of various strains to synthesize acetate esters. Ethyl acetate and other short-chain volatile esters have use as industrial solvents and as flavor and fragrance compounds; worldwide ethyl acetate demand is ~1.7 million tons per year.[6, 7] This demand is typically met by converting ethanol and acetate into ethyl acetate by Fisher esterification in reactive distillation processes.[6, 8] A recent bioprocessing pilot plant using *K. marxianus* produced ethyl acetate titers of 10.9 g L^{-1} from waste whey feeds with yields of 51.4%, thus providing a single-step alternative to the traditional process that relies on

petroleum feedstocks (for acetate production) and multiple unit operations (ethanol fermentation, acetate production, and Fisher esterification).[8]

Metabolic synthesis of ethyl acetate and other short- and medium-chain esters can occur via a number of pathways.[7] In *Saccharomyces cerevisiae*, ester biosynthesis is primarily catalyzed by alcohol-O-acetyltransferase (Atf or AATase) activity that condenses acetyl-CoA with an alcohol to produce the corresponding ester.[9, 10] Double knockout of ATF1 and ATF2 in *S. cerevisiae* eliminates the synthesis of the medium-chain ester isoamyl acetate and reduces ethyl acetate production by 50%.[11] These enzymes localize to the endoplasmic reticulum and lipid droplets; this has been shown to be critical for high activity.[12, 13] Reverse esterase and alcohol dehydrogenase (Adh) activity may also contribute to acetate ester production. In *Candida utilis*, ethyl acetate is produced by Adh oxidation of a hemiacetal intermediate (the spontaneous product of ethanol and acetaldehyde).[14] Adh1 from *S. cerevisiae* (*Sc*Adh1) and an Adh from *Neurospora crassa* are also known to exhibit hemiacetal oxidation activity,[15, 16] a reaction that is thought to be involved in aldehyde detoxification under conditions of low NADH availability.[7] In *K. marxianus*, both Atf and reverse esterase activities have been identified, and Atf activity is thought to be primarily responsible for ethyl acetate biosynthesis.[17, 18] However, the metabolism of *K. marxianus* is less well understood than that of *S. cerevisiae*, and promiscuous Adh activity toward hemiacetal oxidation and the roles of different Adh enzymes in ester and alcohol metabolism are not yet completely understood.

Seven unique ADH genes have been previously identified in *K. marxianus*. *Km*ADH1, -2, -3, and -4 were identified as paralogs to *K. lactis* ADH genes, and recent transcriptional studies identified 3 additional ADH genes, *Km*ADH5, -6, and -7.[19-23] Protein homology analysis revealed that *Km*Adh5 and -6 are similar to *Sc*Adh4 and -6, respectively.[19, 23] *Km*Adh7

has high protein homology to an Adh from the bacterium *Cupriavidus necator*.[19] Thus far, most studies have focused on *Km*ADH1-4, and the corresponding enzymes have been assigned to the group of zinc-dependent alcohol dehydrogenases that preferentially use NAD(H) over NADP(H).[21, 24] The identified *K. marxianus* ADH genes along with *Km*ATF are summarized in Table 4.1.

Table 4.1. *K. marxianus* alcohol dehydrogenases (Adh) and alcohol-O-acetyltransferase (Atf)

Name	Alternative Names	Putative Function	Homology (Identity, Similarity)	Size	Ref.	Gene Product
*Km*Adh1	-	Ethanol production	*Kl*Adh1 (85.2%, 92.0%) *Kl*Adh2 (86.5%, 92.0%) *Sc*Adh2 (79.4%, 88.0%) *Sc*Adh1 (76.3%, 84.8%)	348aa	[19-24]	BAO40126.1
*Km*Adh2	-	Ethanol production	*Sc*Adh1(86.0%, 91.4%) *Kl*Adh2 (83.7%, 91.1%) *Kl*Adh1 (86.0%, 90.9%) *Sc*Adh2(84.5%, 90.3%)	348aa	[19-24]	BAO40244.1
*Km*Adh3	-	Use of non-fermentable carbon sources	*Kl*Adh3 (91.2%, 94.4%) *Sc*Adh3(77.5%, 87.0%)	375aa	[19-24]	BAO42617.1
*Km*Adh4	ADH4b [19]	Ethanol detoxification	*Kl*Adh4 (90.8%, 93.9%) *Sc*Adh3 (78.9%, 87.6%)	379aa	[19-24]	BAO38616.1
*Km*Adh5	ADHb [20], ADH4a [19]	Unknown	*Sc*Adh3 (65.2%, 78.3%)	418aa	[19, 20, 23]	BAO38463.1
*Km*Adh6		Unknown	*Sc*Adh6 (62.0%, 76.4%) *Sc*Adh7 (58.8%, 74.3%)	366aa	[19, 20, 23]	BAO42650.1
*Km*Adh7	Adha [20], adh [19]	Unknown	Acinetobacter equi adh (67.2%, 81.4%) Snodgrassella alvi adh (67.4%, 80.9%) Cupriavidus necator adh (30.8%, 50.2%)	386aa	[19, 20]	BAO40648.1
*Km*Atf	Atf1 [19]	Unknown	*Kl*Atf (56.0%, 75.8%) *Sc*Atf2 (31.7%, 53.1%) *Sc*Atf1 (35.0%, 52.5%)	515aa	[19, 20]	BAO39498.1

To enable gene disruption studies and identify the roles of *Km*ADH1-7 and *Km*ATF in ethyl acetate and ethanol biosynthesis, we adapted the Type II CRISPR-Cas9 system from *S. pyogenes* for genome editing in *K. marxianus*. *K. marxianus* is known to have a high capacity for DNA repair by non-homologous end joining (NHEJ), which can limit the rapid generation of targeted knockout libraries by traditional genome manipulation techniques.[2, 25] In this work, we used a hybrid promoter strategy to express single-guide RNAs (sgRNAs) that target Cas9 endonuclease to *Km*ADH1-7 and *Km*ATF for functional

disruption in the CBS 6556 strain of *K. marxianus*.[26, 27] Transcriptional and functional analysis of the disruption library revealed the critical role of *Km*ADH2 in ethanol production, and that disruption results in acetaldehyde accumulation. In addition, analysis of the knockout strains coupled with overexpression studies revealed novel hemiacetal activity towards ethyl acetate synthesis from *Km*Adh7 and showed that *Km*Atf has activity towards the condensation of acetyl-CoA and ethanol.

4.3 Results

4.3.1 Thermotolerance and ethyl acetate biosynthesis in *K. marxianus* CBS 6556

Various strains of *K. marxianus* have been shown to have fast growth kinetics at temperatures above 40 °C.[2, 28] For example, the strain CBS 6556 exhibits high growth rates at 45 °C (0.60 ± 0.07 h^{-1}; Figure 4.1A). High growth rates were also observed at lower temperatures and were nearly twice that of the model yeast *S. cerevisiae* BY4742. We have previously identified CBS 6556 as a high producer of ethyl acetate and here demonstrate the growth-associated production of 3.72 ± 0.06 g L^{-1} from the wild type in a controlled aerated bioreactor (Figures 4.1B and C). Selectivity towards ethyl acetate was high, as only a limited amount of ethanol (60 ± 11 mg L^{-1}) was co-produced. Given the fast growth kinetics, thermotolerance to 45 °C, and the high capacity to synthesize ethyl acetate, we selected CBS 6556 as a model *K. marxianus* strain to further understand the roles of *Km*ADHs and *Km*ATF in ethyl acetate and ethanol biosynthesis.

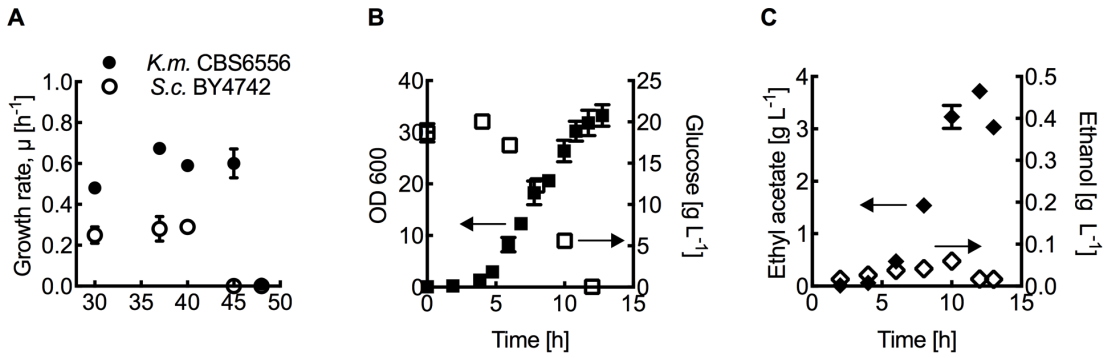


Figure 4.1: Ethyl acetate and ethanol production in thermotolerant *K. marxianus* CBS 6556. (A) Specific growth rates of *K. marxianus* CBS 6556 and *S. cerevisiae* BY4742 were determined by shake flask culturing at temperatures of 30, 37, 40, 45, and 48 °C. (B) Aerated bioreactor growth and glucose consumption of *K. marxianus* CBS 6556 at 60% dissolved oxygen and 37 °C. (C) *K. marxianus* CBS 6556 production of ethyl acetate and ethanol during the bioreactor experiments shown in (B). Data points represent the arithmetic mean of 3 biological replicates and the error bars represent the standard deviation.

4.3.2 Ethyl acetate synthesis activity in *K. marxianus*

Three metabolic pathways are known to produce ethyl acetate: 1) the condensation of ethanol and acetyl-CoA; 2) the reverse esterase synthesis of ethyl acetate from ethanol and acetate; and 3) the oxidation of hemiacetal (Figure 4.2A).[7] Similar to other yeasts, ethanol is synthesized by the decarboxylation of pyruvate and the Adh-mediated reduction of acetaldehyde to ethanol. To identify activities towards ethyl acetate that are present in *K. marxianus*, we conducted a series of cell lysate assays supplemented with appropriate enzyme substrates (Figure 4.2B). The addition of hemiacetal and NAD^+ cofactor to cell lysates synthesized 20.4 ± 3.3 nmol $min^{-1}mg^{-1}$ of ethyl acetate, while the addition of ethanol and acetyl-CoA produced 1.7 ± 0.1 nmol $min^{-1}mg^{-1}$ (rates reported as per mg of lysate protein). A control sample of hemiacetal without NAD^+ produced 7.4 ± 0.8 nmol $min^{-1}mg^{-1}$, likely due to cofactor present in the cell lysate. Reverse esterase activity was significantly limited (0.44 ± 0.02 nmol $min^{-1}mg^{-1}$), and control assays with only ethanol, acetate, and

acetaldehyde did not produce measurable amounts of ethyl acetate. Combined, the assays suggest that ethyl acetate may be produced from one or both of the Atf- and Adh-dependent pathways.

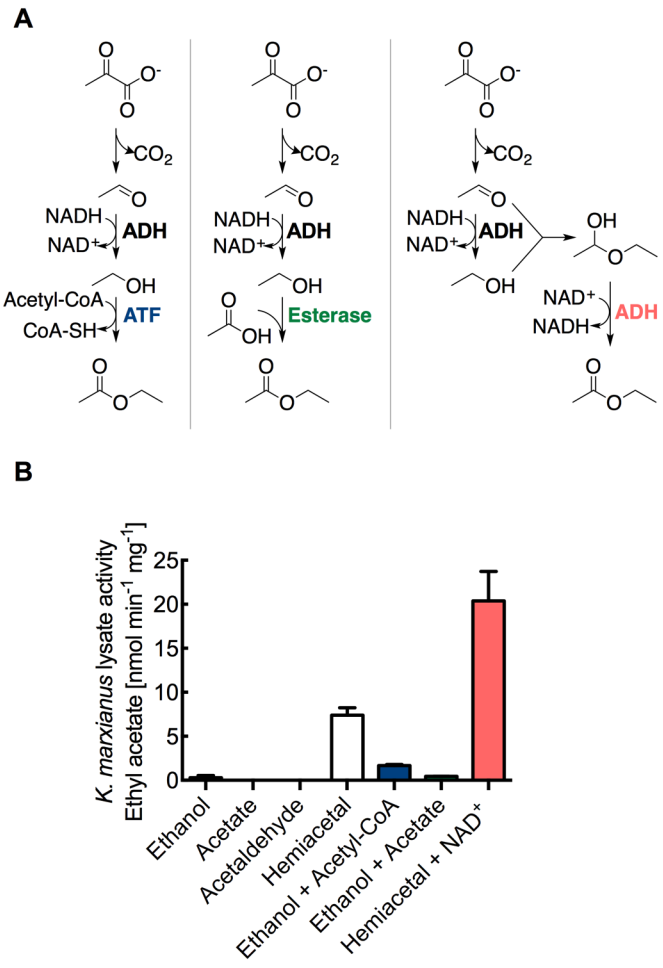


Figure 4.2: Ethyl acetate biosynthetic pathways and synthesis activities in *K. marxianus*. (A) Ethyl acetate biosynthesis via alcohol-O-acetyltransferase (Atf; left), reverse esterase activity (middle), and by alcohol dehydrogenase (Adh) oxidation of hemiacetal (right). (B) *K. marxianus* CBS 6556 lysate assays for ethyl acetate production. Reactions were accomplished with 100 mg of lysate protein buffered to pH 7.2 with 100 mM potassium phosphate and 30 °C. Data points represent the arithmetic mean of 3 biological replicates and the error bars represent the standard deviation.

4.3.3 CRISPR-Cas9 mediated gene disruptions in *K. marxianus*

CRISPR-Cas9 genome editing systems function by targeting the Cas9 endonuclease to specific loci in the genome through complexation with guide RNAs (sgRNAs). sgRNA expression is often limiting to Cas9 function.[27, 29] To address this potential problem, we designed a series of native and hybrid RNA polymerase III promoters for sgRNA expression, including SNR52, $tRNA^{Gly}$, and fusions of SNR52- $tRNA^{Gly}$, SCR1-$tRNA^{Gly}$, and RPR1-$tRNA^{Gly}$ (Figure 4.3A). We targeted xylitol dehydrogenase (XYL2) for promoter testing because successful disruption can be coupled to a phenotype that is easily screened, *i.e.*, the loss of growth on xylitol supplemented agar media plates (Figure 4.3C). XYL2 disruption efficiencies were determined by restreaking a minimum of 30 colonies that were subjected to mutation by CRISPR-Cas9 onto solid media plates with xylitol as the sole carbon source. Disruption efficiency was found to be promoter-dependent (Figure 4.3B). The highest efficiency, 66 ± 8%, was achieved with the RPR1-$tRNA^{Gly}$ promoter, while $tRNA^{Gly}$ achieved 52 ± 15% disruption efficiency. The SNR52, SNR52-$tRNA^{Gly}$, and SCR1-$tRNA^{Gly}$ promoters were less successful, resulting in disruption efficiencies of 10 ± 6, 35 ± 7, and 18 ± 11%, respectively. Gene disruption was also found to be sgRNA sequence-dependent (Table S4.1), and a scrambled sgRNA did not produce a loss of xylitol dehydrogenase function (Figure 4.3B).

With a functional CRISPR-Cas9 system in hand, we created a library of mutant CBS 6556 strains with functional disruptions to *Km*ADH1-7 and *Km*ATF. In each case, sgRNA design was accomplished using a previously published sgRNA scoring algorithm to identify high-scoring sgRNAs upstream of each enzyme's putative active site.[30] The protocol for creating and screening genetic disruptions by CRISPR-Cas9 is schematically described in Figure S4.1. The final steps of the protocol include quantifying the rate of gene disruption,

which was determined by amplifying the gene of interest from the genome, and sequencing the PCR product to identify indels. The indel success rates, which ranged from 10 to 67%, are presented in Table 4.2. Importantly, the mutant strains used in subsequent experiments were cured of the CRISPR-Cas9 plasmid, and the gene of interest was sequenced to identify the specific mutations in each gene (Table 4.2). In each case, the indel created a genetic frameshift mutation that results a premature stop codon, thus functionally disrupting the gene.

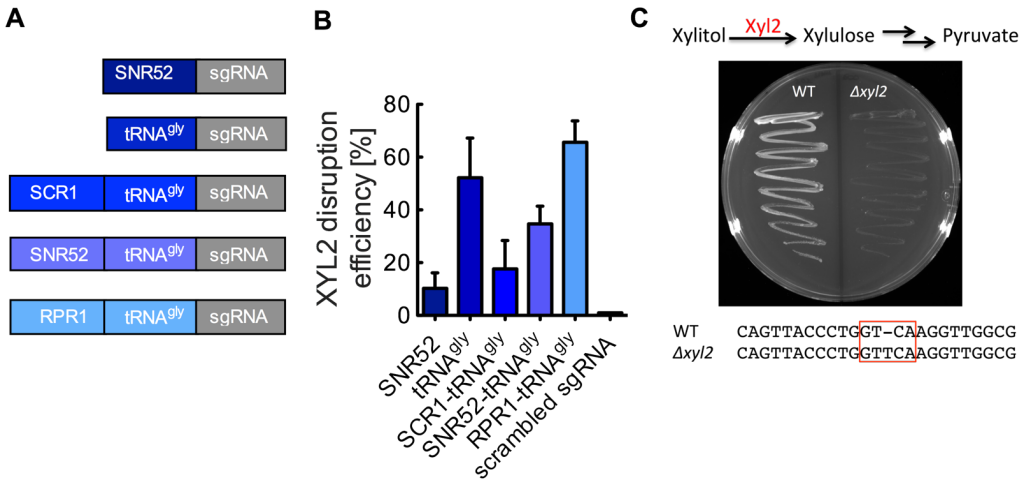


Figure 4.3: CRISPR-Cas9 genome editing in *K. marxianus*. (A) The design of native and hybrid RNA polymerase III promoters for sgRNA expression. The 20 bp sgRNA targeting sequence targets xylitol dehydrogenase (XYL2). (B) CRISPR-Cas9 induced disruption efficiencies of the XYL2 gene. Cultures were transformed with a single vector expressing codon optimized Cas9 and the sgRNA. Culture were grown on selective media for 2 days prior to screening on xylitol supplemented agar media. Data points represent the arithmetic mean of 30 colonies randomly selected from three different transformations. Error bars represent the standard deviation of the samples. (C) XYL2 disrupted strains were restreaked on xylitol containing agar plates and screened for loss of growth. Gene disruption was confirmed by sequencing.

Table 4.2: *Km*ADH and *Km*ATF CRISPR-Cas9 target sequences and knockout efficiencies

Target	sgRNA	Gene Disruption	Indel Success
*Km*ADH1	AGACTTCAAAGCCTTGTACA	WT CCGTGTACAAGGCTTTGAAG KO CCG--------GCTTTGAAG	2/10
*Km*ADH2	GGTACCAGCTGGGATGTGAG	WT AAGCCGCTCACATCCCAGCT KO AAGCCGCT-----CCCAGCT	4/30
*Km*ADH3	GCTATTCCAGAAAAGCAAAA	WT CAGAAAAGCAAAAGGGTGTT KO CAGAAAA-----AGGGTGTT	2/10
*Km*ADH4	GCCATCCCAGAATCCCAAAA	WT TCCCAGAATCCC-AAAAGGG KO TCCCAGAATCCCAAAAAGGG	4/10
*Km*ADH5	ATGGTCTTGAAAGAACACAA	WT GGTCTTGAAAGAA-CACAAG KO GGTCTTGAAAGAACCACAAG	1/10
*Km*ADH6	GTACCACCACCGCAAAGTAG	WT CTACTTTGCGGTGGTGGTAC KO CTACTT-GCGGTGGTGGTAC	2/10
*Km*ADH7	GCTTGAGCTGAGAGATTGAT	WT AAGTCCCATCAA-TCTCTCA KO AAGTCCCATCAAATCTCTCA	3/5
*Km*ATF	ATATAGTCTTCGGCAACACC	WT CGGTGTTGCCGAAGACTATA KO CGGTG--GCCGAAGACTATA	4/10

4.3.4 The effect of *Km*ATF disruption on ethyl acetate biosynthesis

Figures 4.1B and C as well as previous *K. marxianus* studies show that ethyl acetate production is growth-associated, as such *Km*ATF transcript levels were examined at lag, log, and stationary phase.[1, 31] Reverse transcription quantitative PCR (RT-qPCR) analysis showed that under aerobic conditions (the conditions required for high ethyl acetate production) *Km*ATF is upregulated during stationary phase (Figure 4.4A). Anaerobically, *Km*ATF followed a growth-associated expression pattern (Figure S4.2). CRISPR-Cas9 disruption of the gene produced a 15% reduction in ethyl acetate during aerobic growth and a 66% reduction under anaerobic conditions, suggesting that *Km*Atf has activity towards ethyl acetate biosynthesis (Figure 4.4B). The effect of the gene disruption when cultured aerobically on synthetic minimal media was not statistically significant (Figure S4.3). The alcohol-O-acetyltransferase activity of *Km*Atf was confirmed in lysate experiments with overexpression in *S. cerevisiae* (Figure 4.4C); however, *Km*Atf (2.7 ± 0.1 nmol $min^{-1}mg^{-1}$) was found to be less active then *Sc*Atf1 (85.4 ± 3.1 nmol $min^{-1}mg^{-1}$).

Western blot analysis of the *S. cerevisiae* lysates used in the assay confirmed protein expression and suggested that some of the difference in activity was due to reduced enzyme expression (Figure 4.4C).

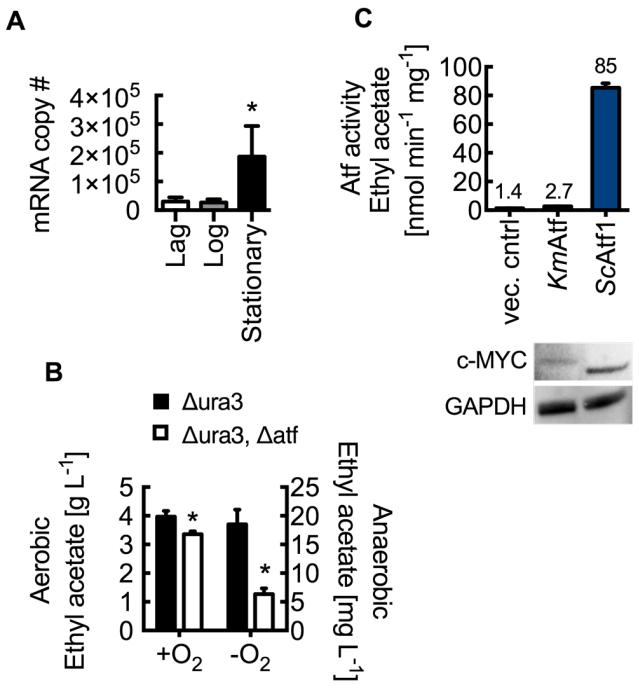


Figure 4.4: Ethyl acetate biosynthesis in *K. marxianus* by alcohol-O-acetyltransferase. (A) RT-qPCR analysis of wild type *K. marxianus* ATF transcript levels at lag (5h), log (10h) and stationary (18h) phases. mRNA copy number represents total number of transcripts for 5 ng of isolated RNA. (B) *Km*ATF disruption reduces ethyl acetate production in both aerobic and anaerobic conditions. Cultures were grown at 37 °C for 10 h (aerobic) and 14 h (anaerobic). (C) Lysate Atf activity from *S. cerevisiae* with overexpressed *Km*ATF and *Sc*ATF1. Rates are reported per mg of lysate protein. Enzyme expression was confirmed by Western blot analysis using anti-c-MYC and anti-GAPDH antibodies. Bars and error bars represent the arithmetic mean and standard deviation of triplicate biological samples. Statistical significance ($p<0.05$) is indicated by "*" and was determined using a T-test.

4.3.5 The effect of *Km*ADH1-7 disruptions on ethyl acetate and ethanol biosynthesis

To determine the effects of *Km*ADH1-7 on ethanol and ethyl acetate biosynthesis, wild-type CBS 6556 and mutant strains were cultured, monitoring growth, gene expression, ethanol, ethyl acetate, and acetaldehyde production. The transcriptional patterns of the *Km*ADHs in wild-type CBS 6556 are shown in Figure 4.5. Transcriptional analysis was conducted aerobically at 5, 10, and 18 hours, corresponding to lag, exponential, and stationary phases, respectively (Figure 4.5A). *Km*ADH1 and -2 were highly expressed after 5 hours of aerobic growth and transcript levels remain consistent after 10 and 18 hours.

Transcript levels of *Km*ADH3 and -4 at 5 hours were significantly lower in comparison with *Km*ADH1 and -2, but were upregulated during stationary phase. Analysis of *Km*ADH5 revealed the lowest initial expression level, with increased expression after 10 hours of aerobic culture. *Km*ADH6 expression remained consistent through each growth phase, and *Km*ADH7 increased expression as cells entered stationary phase, both of which maintained initial levels lower than the highly expressed *Km*ADH1 and -2. For the analysis of anaerobic expression, cultures were grown for 6, 14 and 24 hours to lag, log and stationary phase, respectively (Figure 4.5B). *Km*ADH2 expression was highest during all of the growth stages, indicating its importance in *K. marxianus* metabolism. Similar to *Km*ADH2, *Km*ADH1 was highly expressed throughout culture growth. *Km*ADH3 and -4 showed increased expression at stationary phase, *Km*ADH6 and *Km*ADH7 expression increased upon reaching stationary phase, and *Km*ADH5 expression was insensitive to growth phase.

With respect to cell growth, *Km*ADH1-7 and *Km*ATF disruptions had no effect on cell growth under aerobic conditions (Figure S4.4). Under anaerobic conditions ethanol fermentation pathways are significantly upregulated are necessary for cell redox balance. As expected, *Km*ADH1-7 and *Km*ATF disruptions affected growth (Figure S4.5). More specifically, disruptions of *Km*ADH2 and -4 resulted in growth rates of 0.25 ± 0.05 and 0.23 ± 0.08 h^{-1} compared with 0.36 ± 0.10 h^{-1} of CBS 6556 at 37 °C.

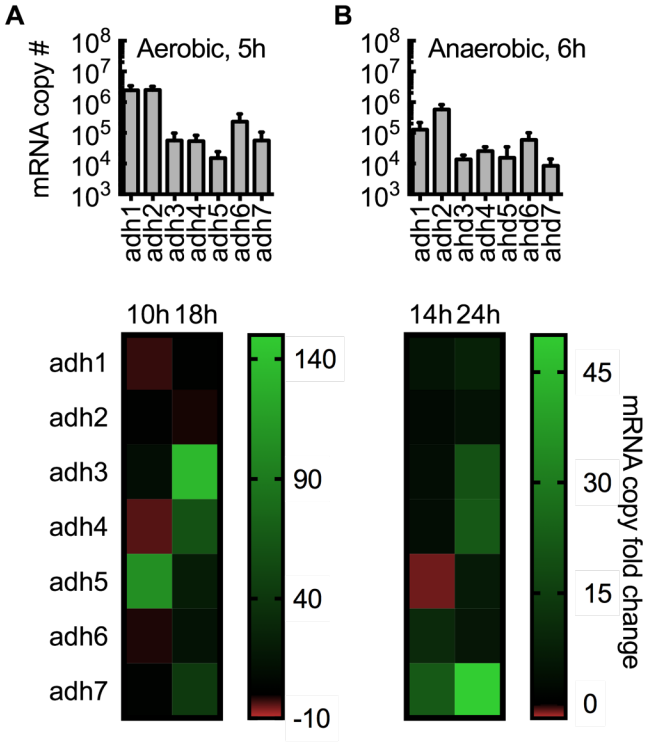


Figure 4.5: Transcriptional analysis of *Km*ADH expression in wild type *K. marxianus* CBS 6556. mRNA copy number at lag, log and stationary phase of aerobic (A) and anaerobic (B) growth was determine by RT-qPCR. The top panels of A and B show the total mRNA copy number of *Km*ADH1-7 expressed in *K. marxianus* strain CBS 6556 at 5 h for aerobic cultures and 6 h for anaerobic cultures. mRNA copy number represents total number of transcripts for 5 ng of isolated RNA. The heat maps below show the mRNA copy fold change of the individual genes at log (10 and 14h) and stationary phases (18 and 24h) compared with the expression at lag phase (5 and 6h). Bars and error bars represent the arithmetic mean and standard deviation of triplicate biological samples.

Metabolite analysis revealed that *Km*ADH1, -2, and -3 mutants exhibited reduced ethyl acetate production compared with the wild type CBS 6556 strain for aerobic cultures (Figure 4.6A). Mutations to *Km*ADH1-3 resulted in specific ethyl acetate titers of 63.3 ± 14.1 mg L^{-1} OD^{-1}, 47.9 ± 9.4 mg L^{-1} OD^{-1} and 38.8 ± 7.4 mg L^{-1} OD^{-1}, compared with wild-type CBS 6556, which produced 157.8 ± 26.9 mg L^{-1} OD^{-1}. Reduced ethanol production accompanied the loss of ethyl acetate biosynthesis (Figure 4.6B). Disruption of *Km*ADH1 produced 7.5 ± 1.4 mg L^{-1} OD^{-1} of ethanol, while functional disruption to *Km*ADH2 and -3 produced only 0.28 ± 0.63 mg L^{-1} OD^{-1} and 0.32 ± 0.01 mg L^{-1} OD^{-1}, respectively. In the case of *Δadh2* an increase in acetaldehyde accumulation was also observed (31.9 ± 10.1 mg L^{-1} OD^{-1}; Figure 4.6C). These results suggest that *Km*ADH2 is critical to the supply of ethanol for aerobic ethyl acetate biosynthesis. Less significant changes in ethyl acetate, ethanol, and

acetaldehyde were observed with disruptions to *Km*ADH4, -5, -6, and -7. It should be noted that the URA3 disrupted strain showed decreased ethyl acetate production of 67.9 ± 38.4 mg L^{-1} OD^{-1} compared with the wild type (Figure S4.6). Figures 4.6D-F present the volatile metabolite analysis for anaerobically grown cultures. While *Km*ADH1-7 disruption did not reduce overall ethanol production, acetaldehyde accumulation (0.05 ± 0.02 mg L^{-1} OD^{-1}) was again observed in the *Δadh2* strain (Figures 4.6E and F). Furthermore, *Km*ADH2 disruption led to a significant increase of anaerobic ethyl acetate production from 1.91 ± 1.16 mg L^{-1} OD^{-1} in the wild type to 5.46 ± 1.23 mg L^{-1} OD^{-1} in the disrupted strain (Figure 4.6D). Note that 1) anaerobic ethanol production was ~50-fold greater than was observed in aerobic cultures, and 2) aerobic ethyl acetate production was ~25-fold greater than was observed anaerobic cultures.

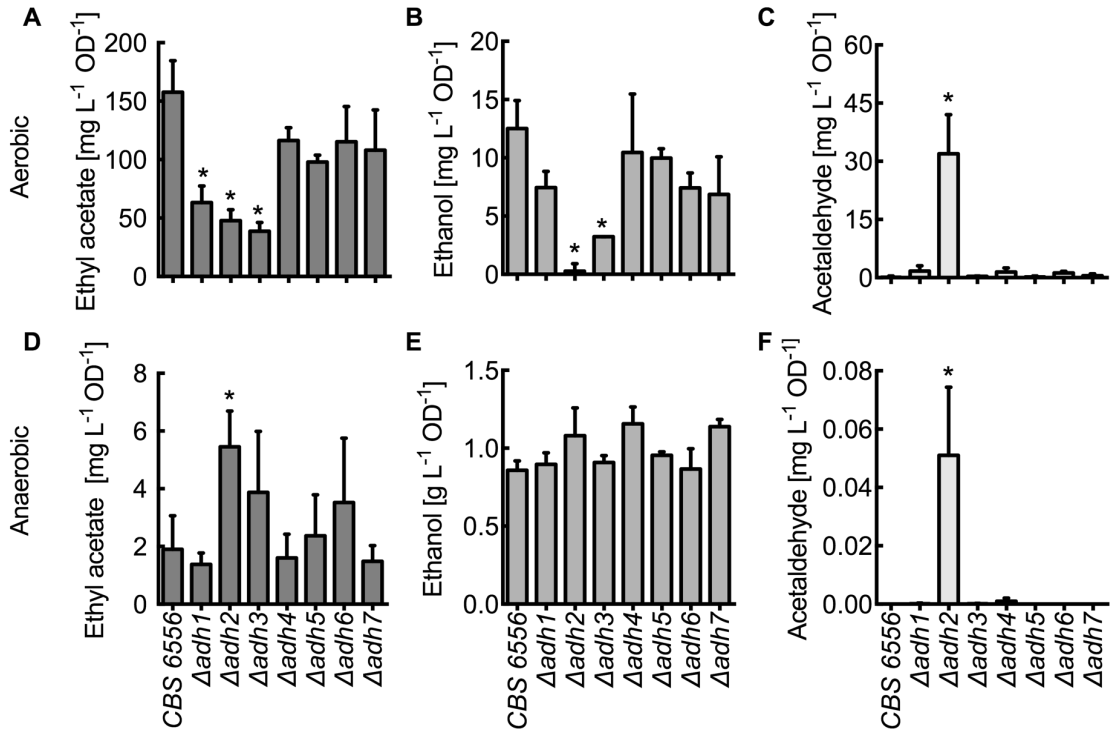


Figure 4.6: Ethyl acetate, ethanol, and acetaldehyde production in *K. marxianus* CBS 6556 and ADH knockout strains. (A-C) Aerobic production of ethyl acetate, ethanol and acetaldehyde production per cell density after 10h of growth in YM media at 37 °C. (D-F) Anaerobic production of ethyl acetate, ethanol and acetaldehyde production per cell density after 14h of growth. Bars and error bars represent the arithmetic mean and standard deviation of triplicate biological samples.

4.3.6 *Km*Adh activity towards ethyl acetate formation from hemiacetal

Because the *Km*ADH1-7 disruption library did not reveal clear gene candidates for ethyl acetate production, each ADH gene was separately overexpressed in *S. cerevisiae* to facilitate lysate enzyme assays. *Sc*Atf1 and -2 are known to be suppressed by oxygen; therefore, the *S. cerevisiae* lysates provided a cell lysate background that lacks the capacity to produce ethyl acetate.[32] Western blot analysis of C-terminal MYC-tag modified enzymes confirmed the overexpression of *Km*Adh1, -2, -5, -6 and -7 (Figure 4.7). *Km*Adh3 and -4 were also expressed but at reduced levels. *S. cerevisiae* lysates overexpressing each

to the enzymes were incubated with hemiacetal and NAD^+ cofactor, and the reactions were allowed to continue for 30 mins. Notably, *Km*Adh7 produced ethyl acetate at a rate of 66.3 ± 2.9 nmol min^{-1} mg^{-1} of lysate protein. No other overexpressed enzymes produced measurable amounts of ethyl acetate. Thus far, the function of *Km*Adh7 has not been described, and the result presented here suggests hemiacetal activity towards the biosynthesis of ethyl acetate.

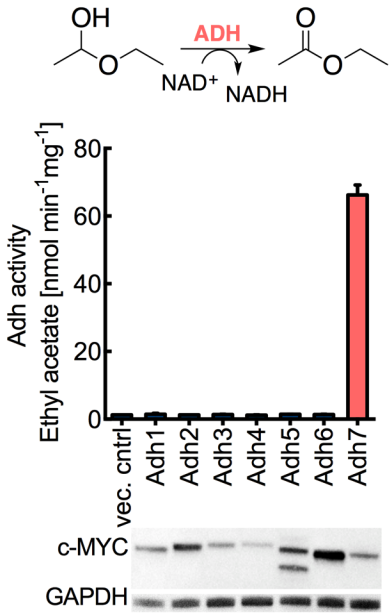


Figure 4.7: Hemiacetal oxidation activity of *K. marxianus* alcohol dehydrogenases. *Km*ADH1-7 were heterologously overexpressed in *S. cerevisiae* BY 4742. Cell lysate assays showed that *Km*Adh7 exhibits activity towards synthesis of ethyl acetate from hemiacetal. Enzyme expression was confirmed by Western blot using anti-c-MYC and anti-GAPDH antibodies. Bars and error bars represent the arithmetic mean and standard deviation of triplicate samples.

4.4 Discussion

The CBS 6556 strain of *K. marxianus* is thermotolerant and exhibits fast growth kinetics at 45 °C (Figure 4.1A). It ferments high titers of ethanol under anaerobic conditions,[2] and aerobically produces significant amounts of the short-chain volatile ester ethyl acetate (Figure 4.1C). CBS 6556 and other strains of *K. marxianus* also have the ability to metabolize various C_5, C_6, and C_{12} carbon sources.[1] Collectively, these characteristics are useful for industrial-scale bioprocessing: high-temperature bioreactors can minimize the need for aseptic conditions, diverse carbon sources allow for the use of the lowest-cost sugars, and the high titer production of volatile compounds can facilitate low-cost product separation through distillation and stripping.[33]

Despite these many advantages, *K. marxianus* strain development has been limited in comparison with the model yeast *S. cerevisiae* because genome editing tools and stable heterologous expression systems developed for *K. marxianus* are limited.[25, 34-36] Enabled by hybrid RNA polymerase III promoters for sgRNA expression, we adopted the CRISPR-Cas9 system from *S. pyogenes* for *K. marxianus* genome editing (Figure 4.3). This system was necessary to create a library of strains with disruptions to genes with suspected function in ethyl acetate and ethanol metabolism including *Km*ADH1-7 and *Km*ATF (Table 4.1). The main observations stemming from our CRISPR-Cas9 experiments are that 1) the RPR1-tRNA^{Gly} hybrid promoter achieved the highest knockout efficiencies (Figure 4.3B), and 2) that gene disruption efficiency was found to be highly dependent on gene target and sgRNA sequence (Table S4.2). We have previously observed a similar result for hybrid sgRNA promoter design in the yeast *Yarrowia lipolytica*, and the sequence dependency of sgRNAs has been described in both human and mouse systems. [27, 30]

The widespread adoption of type II CRISPR-Cas9 technologies for genome editing has made less genetically tractable organisms more accessible.[26, 27] These tools are particularly useful in organisms where non-homologous end joining prevails over DNA repair by homologous recombination and when genomes are diploid.[37, 38] For example, CRISPR-Cas9 genome editing has recently been demonstrated in the yeasts *Y. lipolytica, S. pompe, P. pastoris, K. lactis*, *C. albicans*, and *S. cerevisiae*.[26, 27, 37-42] Many of these examples focus on tool development for genome and metabolic engineering, including standardized and multiplexed methods for heterologous gene integration.[26, 38, 41, 42] The other examples use the genome editing system, as we have done here, to study the effects of genetic disruptions on cell phenotypes and metabolism.[37, 39, 40] To our knowledge, these types of studies have not yet been accomplished in *K. marxianus*.

The Δ*adh1-7* and *Δatf* strains of *K. marxianus* CBS 6556 were created to investigate the effects of these genes on ethyl acetate and ethanol biosynthesis. Our first experiments towards this goal helped identify the targeted knockouts. Specifically, *K. marxianus* lysate assays showed ethyl acetate biosynthesis from both Atf and Adh catalyzed reactions (Figure 4.2). Atf activity was observed when lysates were supplemented with ethanol and acetyl-CoA, while Adh activity was observed when supplementing with hemiacetal and NAD^+ cofactor. These results are in contrast to prior reports that describe Atf and reverse esterase activities as the only reactions responsible for ethyl acetate biosynthesis in *K. marxianus* DSM 5422 and other strains.[17, 18] In the CBS 6556 strain, we found no reverse esterase activity, but did identify Atf and Adh activities as possible routes for ethyl acetate production. On the basis of these results, *Δadh1-7* and *Δatf* strains of *K. marxianus* CBS 6556 were created to investigate the effects of these genes on ethyl acetate and ethanol biosynthesis.

Functional disruption of *Km*ATF resulted in the statistically significant reduction of ethyl acetate biosynthesis under both aerobic and anaerobic conditions (Figure 4.4B). It is important to note here that ethyl acetate production was ~100-fold higher when cells were grown aerobically. Ethyl acetate synthesis through *Km*Atf activity was further confirmed in *S. cerevisiae* lysate assays containing heterologously overexpressed enzyme (Figure 4.4C). Previous studies on *K. lactis* Atf (*Kl*Atf), which is most closely related to *Km*Atf, revealed limited activity towards ethyl acetate in comparison with *Sc*Atf1, a result that was also observed in our studies of *Km*Atf (Figure 4.4C).[13, 43, 44] Taken together, these results suggest that *Km*Atf activity contributes in part to ethyl acetate biosynthesis in *K. marxianus*, but that other alcohol-O-acetyltransferases and/or other metabolic pathways are also responsible.

It is well understood that ethanol production in yeast is accomplished by the Adh-catalyzed reduction of acetaldehyde. Various Adh enzymes have also been shown to catalyze ethyl acetate synthesis through the oxidation of hemiacetal (Figure 4.2A).[7] RT-qPCR analysis of the genes found to be most relevant to ethyl acetate production showed that *Km*ADH2 was constitutively expressed in aerobic growth with glucose as a carbon source, while *Km*ADH3 and *Km*ADH7 expression increased as cells reached stationary phase. These results are in agreement with previously published analyses of the *K. marxianus* transcriptome.[21] The Δ*adh2* strain showed reduced aerobic ethanol and ethyl acetate production, suggesting a large role in ethanol and consequently ethyl acetate production. Disruption of *Km*ADH3 also had a significant effect on ethyl acetate and ethanol formation, but this is likely due to the role of *Km*ADH3 in cellular cofactor balance.[45] Under both aerobic and anaerobic conditions, the disruption of *Km*ADH2 resulted in a large accumulation of acetaldehyde, something that was not observed in the Δ*adh3* strain or the

other knockouts strains, suggesting that *Km*Adh2 is the dominant enzyme in ethanol production when glucose is used as a carbon source.

With respect to *Km*ADH7, the lysate assays conducted with overexpressed *Km*Adh1-7 in *S. cerevisiae* (Figure 4.7) demonstrated that *Km*Adh7 is a source of hemiacetal activity in *K. marxianus* (Figure 4.2). Bioinformatic analysis revealed that *Km*Adh7 has no significant similarity to *Km*Adh1-6. NADH cofactor usage is similar to *Km*Adh1-5, but otherwise *Km*Adh7 appears to be unique among *K. marxianus* Adh enzymes. *Km*Adh7 does, however, show sequence similarity to an Adh from the bacteria *Cupriavidus necator*, but not to other yeast Adh enzymes.[19] The *C. necator* NAD(H)(P)-dependent enzyme exhibits broad substrate specificity including activity towards the oxidation ethanol and 2,3-butanediol, and the reduction of acetaldehyde, acetoin, and diacetyl; however, hemiacetal oxidation has not been demonstrated.[46, 47] Such activity has been reported for Adhs found in both *C. utilis* and *N. crassa*, as well as in *S. cerevisiae*.[14-16] The Δ*adh7* strain of *K. marxianus* CBS 6556 studied here did not result in a reduction in ethyl acetate biosynthesis (Figure 4.6A) and, therefore, the role of *Km*Atf7 in ester biosynthesis remains unclear.

Previous investigations of *K. marxianus* metabolism suggest that Atf activity plays a critical role in ethyl acetate biosynthesis.[17, 18] The studies presented here use new genome editing tools and biochemical assays to confirm the function of *Km*ATF, acetate ester synthesis via the condensation of an alcohol with acetyl-CoA. CRISPR-Cas9 technology adopted for use in *K. marxianus* also allowed for the creation of a library of *Km*ADH disruption strains, identifying *Km*ADH2 as critical for the reduction of acetaldehyde to ethanol (a precursor to ethyl acetate). In addition, *Km*Adh7 was found to exhibit activity towards the oxidation of hemiacetal to ethyl acetate, but the absence of this activity in the *Δadh7* strain of CBS 6556 did not produce a measurable reduction in ethyl acetate

biosynthesis. It was recently postulated that ester biosynthesis in *K. marxianus* may also occur through homologues to the medium-chain acyltransferases Eeb1 and Eht1 from *S. cerevisiae*, the isoamyl acetate-hydrolyzing esterase (Iah1), the N-acetyltransferase Sli1 and/or the alcohol-O-acetyltransferase Eat1.[44, 48] In *S. cerevisiae*, Eeb1 and Eht1 have limited activity towards the acetylation of ethanol and are most active towards medium-chain acyl-coAs and alcohols.[49] If similar activities are found in the *K. marxianus* homologs, then these enzymes are unlikely to contribute to ethyl acetate production. Overexpression of Iah1 in *S. cerevisiae* resulted in lower ester titers due to ester hydrolysis, suggesting that the *K. marxianus* homologue may not contribute to ethyl acetate biosynthesis.[50] A recently published study describes Eat1 as a putative alcohol-O-acetyltransferase capable of producing ethyl acetate.[48] The new genome editing tools created in this work should enable the future study of the role of N-acetyltransferases and other enzymes in ethyl acetate biosynthesis.

4.5 Conclusion

In this work, we developed an efficient CRISPR-Cas9 system for genome editing in *K. marxianus*. This system was used to create a library of single-knockout strains to investigate ethyl acetate and ethanol biosynthetic pathways in *K. marxianus* CBS 6556. Analysis of the knockout strains revealed the importance of *Km*ADH2 in ethanol production in glucose-fed aerobic and anaerobic cultures. With respect to ethyl acetate biosynthesis, *Km*ADH2 is necessary to produce ethanol a substrate for the Atf-catalyzed condensation reaction with acetyl-CoA. Because functional disruption of *Km*ATF did not completely abolish ethyl acetate production, alternative biosynthetic routes are likely present in *K. marxianus*. One

possible pathway is the oxidation of hemiacetal (the spontaneous product of ethanol and acetaldehyde) by Adh activity, an activity that we identify in *Km*Atf7.

4.6 Methods and Materials

4.6.1 Strains and culturing conditions

All strains were purchased from ATCC or DSMZ (Deutsche Sammlung von Microorganismen und Zellkulturen). All materials were purchased from Fisher Scientific unless noted otherwise. All yeast strains used in this study are listed in Table S4.2.

Wild-type *K. marxianus* strain CBS 6556 as well as the *S. cerevisiae* BY4742 strain were grown in shake flasks containing 50 mL YM media (3 g L^{-1} yeast extract, 3 g L^{-1} malt extract, 5 g L^{-1} peptone; DB Difco®, Becton-Dickinson) with 10 g L^{-1} glucose. Overnight cultures were inoculated with isolated single colonies freshly grown on agar YM plates. The overnight cultures were used to inoculated shake flask at an OD of 0.05, which were subsequently cultured at 30, 37, 40, 45, and 48 °C. For *K. marxianus* cell lysate studies strain CBS 6556 was grown for 8h in 1 mL YM media at 37 °C. Then 500 µL was transferred into 50 mL YM media and culture was grown for 16h.

To create knockout strains of *K. marxianus*, cells harboring a CRISPR-Cas9 plasmid were cultured in synthetic defined medium without uracil (SD-U) containing 6.7 g L^{-1} yeast nitrogen base without amino acids DB Difco®; (Becton-Dickinson), 1.92 g L^{-1} yeast synthetic drop-out medium supplements without uracil (Sigma Aldrich), and 20 g L^{-1} glucose or SD-U plates containing 15 g L^{-1} agar. To remove the CRISPR-Cas9 plasmid, cells were grown in YPD medium (5 g L^{-1} Yeast extract, 10 g L^{-1} peptone with 20 g L^{-1} glucose; DB Difco®, Becton-Dickinson) overnight. Screening of XYL2 disruption colonies was achieved on SDX media

containing 6.7 g L^{-1} yeast nitrogen base DB Difco®; (Becton-Dickinson), 0.79 g L^{-1} complete supplement mixture (CSM; Sunrise Science Products) and 20 g L^{-1} xylitol.

K. marxianus alcohol dehydrogenase and alcohol-O-acetyltransferase knockout strains (YS402, YS630, YS671, YS673, YS675, YS679, YS703, YS720 and YS794; see Table S4.2) were cultured in uracil supplemented $SD^{+}U$ media at 37 °C (6.7 g L^{-1} yeast nitrogen base DB Difco®; (Becton-Dickinson), 0.79 g L^{-1} complete supplement mixture (CSM; Sunrise Science Products) and 130 mg L^{-1} uracil (Sunrise Science Products). Overnight cultures were inoculated into 25 mL of media in 250 mL baffled shake flasks (0.05 initial OD) and incubated at 37 °C for the length of the experiment. Initial and final optical cell densities (OD) were measured using the Nanodrop 2000c UV-Vis spectrometer (Thermo Scientific) at 600 nm.

For Adh and Atf overexpression studies, *S. cerevisiae* strains (YAL1-9 and YS202, see Table S4.2) with either an empty vector or an Adh/Atf expression plasmid were grown at 30 °C for 8h in 1 mL SD-U and then transferred to a shake flask with 50 mL SD-U and grown for an additional 16h.

4.6.2 Bioreactor cultures

K. marxianus strain CBS 6556 cultures were grown in a 1 L stirred bioreactor in batch mode (Biosatat®A, Satorius AG). The vessel was equipped with four baffles, two Rushton impellers, gassing tube; exhaust-gas cooler; ports for supplementation, inoculation and sampling; and sensors for dissolved oxygen, temperature, and pH. Cultures were grown in synthetic defined (SD) media (6.7 g L^{-1} yeast nitrogen base without amino acids, DB Difco®; and 0.79 g L^{-1} CSM, Sunrise Science Products) containing 20 g L^{-1} glucose and 1 mL of a 1:1000 dilution of antifoaming agent (Antifoam B Emulsion, Sigma Aldrich) at 37 °C.

Twenty-five-mL overnight cultures were used to inoculate the reactor at an initial cell density of 0.08 OD. Dissolved oxygen (DO) concentration was maintained at 60% saturation by constant aeration with 1000 ccm air and varying stir rate. Media pH of 5 was maintained by titration with 1 M sodium hydroxide. When necessary, liquid was taken through the sampling port. Gas sampling was accomplished by collecting 0.4 L of exhaust gas in a 1 L gas sampling bag connected to the gas sampling port (Supel™Inert Multi-Layer Foil Gas Sampling Bag, Sigma Aldrich). OD was measured and spent media and gas samples were analyzed by GC-FID. Media samples for glucose analysis were spun down to removed cells and the supernatant was stored at -20 °C prior to analysis.

4.6.3 Glucose analysis

The spent media of the bioreactor experiments was analyzed for residual glucose using the Glucose (GO) Assay Kit (GAGO-20; Sigma-Aldrich). The procedure was slightly modified. Briefly, 200 µL of sample was mixed with 400 µL Glucose Assay reagent. After reaction for 30 minutes at 37 °C the reaction was stopped by addition of 400 µL 12N H_2SO_4. 250 µL of the solution mixed was then transferred into a 96-well plate, and absorbance was determined at 540nm using a BioTek Synergy 2 UV-Vis plate reader.

4.6.4 Molecular cloning and plasmids construction

All cloning was accomplished using Phusion polymerase, restriction endonucleases and Gibson assembly master mix purchased from New England BioLabs (NEB). DNA oligos were purchased from Integrated DNA Technologies (IDT). Chemically competent DH5α *Escherichia coli* was used for plasmid propagation. Following transformation, *E. coli* cells

were grown in LB medium containing 100 mg L^{-1} ampicillin. All plasmids and primers used are listed in Table S4.3 and S4.4.

The CRISPR-Cas9 plasmid was constructed using pJSK316-GPD (a kind gift from Dr. Dae-Hyuk Kweon, Sungkyunkwan University) that contains the backbone necessary for plasmid retention in *K. marxianus*.[51] This plasmid was digested with the restriction enzymes KpnI and SacII. The Tef1p-Cas9-Cyct cassette was amplified from p414 (Addgene #43802) using primers P1379/P1525.[52] The structural guide RNA containing a *Sc*SNR52 promoter, an Ade2 target sequence and the structural guide RNA (SNR52-sgRNA-Ade2; Figure S4.7) were designed based on previously described sequences.[52] The SNR52-sgRNA-Ade2 fragment was amplified using primer P1626/P1530. The Cas9 and sgRNA fragments were inserted into the digested pJSK316-GPD by Gibson Assembly.[53] For increased editing efficiency the Cas9 nuclease sequence was codon optimized using Optimizer (http://genomes.urv.es/OPTIMIZER/) and the codon usage table for *K. marxianus*.[54] The resulting sequence was then manually altered, as shown in Figure S4.8, to allow for production of 3 similar sized gBlocks (double stranded fragments from IDT DNA). The CRISPR plasmid (pIW333) was cut with SpeI and XhoI, and the 3 gBlocks were inserted by Gibson Assembly to create a new *Km*CRISPR plasmid (pIW360). Different RNA polymerase III promoters were designed to assess the expression of sgRNAs and CRISPR-Cas9 efficiency.[27] The xylitol dehydrogenase gene (XYL2) was used as reporter gene for Cas9-induced gene disruption. The CRISPR-Cas9 plasmids with varying sgRNA promoters were constructed by inserting different RNA polymerase III promoters into pIW360. *Sc*SNR52 was amplified from pIW360 using P1626/P1789. The *K. marxianus* RNA polymerase III promoters were amplified from isolated *K. marxianus* CBS 6556 genomic DNA. *Km*SNR52 was amplified with P1792/P1806, while P1792/P1793 were used to

generate the SNR52 fragment for *Km*SNR52-tRNA^{Gly}. *Km*SCR1 and *Km*RPR1 were amplified using primers P1795/P1796 and P1798/P1799, respectively. The tRNA^{Gly} sequence was amplified with primers P1839/40 for insertion by itself or with P1790/P1807, P1794/P1807, P1797/P1807, P1800/P1807 for insertion with *Sc*SNR52, *Km*SNR52, *Km*SCR1 and *Km*RPR1, respectively. The sgRNA fragments targeting XYL2 were amplified using primers P1773/P1530 and P1775/P1530. To exchange the sgRNA target sequence, the reverse primer of the promoter fragment and the forward primer of the sgRNA fragment were replaced by appropriate primers containing the target sequence.

For ADH and ATF overexpression in *S. cerevisiae*, the 8 genes were separately cloned into the pRS426 vector containing a PGK1 expression cassette (pIW21).[13] The ADH and ATF genes of interest were identified in the annotated genome of *K. marxianus* DMKU3-1042 as KLMA_40102, KLMA_40220, KLMA_80306, KLMA_20158, KLMA_20005 KLMA_80339 and KLMA_40624 for *Km*ADH1-7, respectively and KLMA_30203 for *Km*ATF.[19] Blast searches confirmed the presences of each gene in the unannotated genome of CBS 6556. All proteins were tagged with a C-terminal Myc tag and cloned into pIW21 at the SacII and SpeI sites using Gibson Assembly. Coding sequences for ADH2-7 and ATF were made using the primers shown in Table S4.4 and cloned into pIW695 that was cut with SacII and AvrII. The resulting plasmids pIW696-702 are listed in Table S4.3.

4.6.5 Transformation of *K. marxianus*

Plasmid and linear DNA transformation were performed using a previously reported protocol with the following modifications.[55] Briefly, 1.5 mL of *K. marxianus* cells were grown to stationary phase and harvested by centrifugation at 5000 rpm for 1 min. After washing with 1 mL of sterile water, cells were suspended in 100 ug carrier DNA (salmon sperm DNA) and 0.2-1 ug of plasmid or linear DNA. 400 mL of transformation mix

(40% polyethylene glycol 3350, 0.1 M lithium acetate, 10 mM Tris-HCl (pH 7.5), 1 mM EDTA and 10 mM DTT) was then added and the solution was incubated at room temperature for 15 min. Subsequently, the transformation mix was heat shocked at 47 °C for 15 min and cells were plated.

4.6.6 Creation of a URA3 auxotrophic strain

To create a URA3 disruption strain, a truncated *K. marxianus* URA3 fragment (missing 160 bp of the coding region) with 500 bp homology upstream and downstream was transformed into *K. marxianus* strain CBS 6556. Transformed cells were recovered overnight and plated on 5-fluororotic acid 5-FOA containing plates. Solid media contained 6.7 g L^{-1} yeast nitrogen base (Becton-Dickinson), 1.92 g L^{-1} yeast synthetic drop-out medium supplements without uracil (Sigma Aldrich), 50 mg L^{-1} uracil (Sunrise Science Products), 1 g L^{-1} (5-FOA; Sigma Aldrich) and 20 g L^{-1} glucose. Colonies were selected based on colony PCR and selected colonies were sequenced (Figure S4.9). To create the URA3 disruption homology donor, overlap extension PCR was performed as previously described.[56] Overlapping fragments of upstream and downstream regions of the sequence targeted for deletion are amplified using primers P1019/P1920 and P1021/P1022, respectively. Resulting fragments are purified and used for overlap PCR with P1019/P1022. The resulting fragment was purified and used for transformation. To confirm disruption, genomic DNA was screened using primers P1072/1073 where a knockout resulted in a 200 bp amplified fragment, with the wild-type gene producing a 360 bp fragment. The knockout design is shown in Figure S4.10.

4.6.7 CRISPR-Cas9 mediated gene disruption

The *K. marxianus* CRISPR-Cas9 system developed in this work was adopted from systems developed for *Y. lipolytica* and *S. cerevisiae*.[27, 52] Cas9 was codon optimized for *K. marxianus* and was expressed from a plasmid using the *S. cerevisiae* Tef1 promoter. For sgRNA expression, RNA polymerase III promoters in *K. marxianus* were identified by blasting the *K. lactis* genes of SNR52 (NC_006042.1), RPR1 (NC_006042.1) and SCR1 (NC_006042.1) against the draft genome of *K. marxianus* strain CBS 6556 (accession number: AKFM00000000).[57] The search yielded *K. marxianus* SNR52, RPR1 and SCR1 that had 86%, 75% and 82% identity to the respective *K. lactis* genes. The promoter regions were identified by searching for conserved A and B-box motifs as previously described.[27, 58] Promoter regions were defined as ~100 bp upstream of the A box until the start of the coding region of the gene. For *Km*SCR1 the boxes were within the transcribed region, which is why the promoter was chosen as the start of the aligned sequence to about 30bp downstream of the identified B-box. The glycine tRNA ($tRNA^{Gly}$) was identified by blasting the annotated tRNA-Gly (AGG) from *K. marxianus* strain DMKU3-1042 (RNA central; URS00003CECDB; [19,50]) against the genome of CBS 6556 as described above.

sgRNA target sequences for xylose dehydrogenase (XYL2), ADH and ATF knockouts were identified using the sgRNA design tool hosted by the Broad Institute (http://www.broadinstitute.org/rnai/public/analysis-tools/sgrna-design).[30] Target sequences were checked for secondary structures using the IDT OligoAnalyzer Tool 3.1 (https://www.idtdna.com/calc/analyzer) and uniqueness within the *K. marxianus* genome using BLAST. All sgRNA sequences are listed in Tables 4.2 and S4.1.

For ADH and ATF disruptions, transformed *K. marxianus* cells were cultured in 2 mL SD-U media or plated on SD-U plates and grown for 2 days at 30 °C. If cultured in liquid

media, 50 mL of cell culture was transferred into new media after 1 day of culturing. After 2 days of growth colonies were screened by amplifying the CRISPR-Cas9 edited region in the genome by colony PCR and subsequent sequencing of the purified PCR fragments. Positively confirmed disruptions colonies were saved at -80 °C after the plasmid was removed.

4.6.7 Adh and Atf protein sequence analysis

Homology of the *K. marxianus* Adh1-7 and Atf proteins to other proteins was analyzed by Pairwise Sequence Alignment using the EMBOSS Needle software form EMBL-EBI. Analyzed sequences are shown in Table S4.5.

4.6.8 Headspace gas chromatography

One-mL of culture supernatant or cell lysate reaction was used for headspace GC analysis in a 10 mL headspace vial containing 1g of NaCl and 20 µL of 5 g L^{-1} 1-pentanol as internal standard. Volatile metabolite concentration was measured using an Agilent 7890A system equipped with an Agilent DB-624UI column and an FID detector. For metabolite separation, the temperature was held at 40 °C for 2 min, then increased 20 °C min^{-1} to 70 °C and 50 °C min^{-1} to 220 °C and held for 2 min. For the bioreactor off gas analysis 1 mL of the off gas was injected from the gas sample bag by manual injection.

4.6.8 Reverse transcription quantitative PCR (RT-qPCR)

Total RNA was extracted using the YeaStar™ RNA Kit (Zymo Research). RNA was DNAse treated (DNAse I, NewEngland Biolabs) for 30 min. and subsequently purified using the RNA Clean & Concentrator™-5 Kit (Zymo Research). RNA was used for the reverses transcription reaction (iScript™ Reverse Transcription Supermix for RT-qPCR, Bio-Rad) and

cDNA was used for SYBR Green qPCR (SsoAdvanced™ Universal SYBR® Green Supermix. Bio-Rad) using the Bio-Rad CFX Connect™. Primers used for the qPCR reaction are listed in Table S4.4. For Figures 4.4A, 4.5A, and 4.5B (top panel) total copy number was calculated from a standard curve with GAPDH as an internal standard. Fold changes in Figures 4.5A and B (bottom panel) were calculated under consideration of the reaction efficiency as previously described.[59] Log and stationary phase expression were compared with lag phase expression using transcript level normalized to GAPDH.

4.6.9 Total cell lysates

Cell lysis was performed as described earlier.[13] In short, cell were harvested, washed and resuspended in equal volumes of wet cell pellets, 425–600 µm acid-washed glass beads (G8772, Sigma-Aldrich), and ice-cold lysis buffer (100 mM potassium phosphate buffer, 2 mM magnesium chloride, 5 mM DTT, and Pierce™ Protease Inhibitor Tablets). The cells were disrupted at 4 °C by vortexing 10 times for 30s with a 30s cooling step between each vortexing. The beads were removed by centrifugation at 500g for 5 min at 4 °C, and the supernatant transferred to a pre-cooled 1.5 mL tube. The protein concentrations of whole cell lysates were determined by Pierce™ 660 nm Protein Assay.

4.6.10 Enzyme activity assay

K. marxianus strain CBS 6556 and *S. cerevisiae* strains containing Adh and Atf expression plasmids are described in Table S4.3. One hundred-µg of total cell lysate isolated from the various strains was used in each reaction. Lysates were incubated in 100 mM potassium phosphate (pH 7.4), 500 mM ethanol and 0.5 mM acetyl-CoA to test for Atf activity and 100 mM potassium phosphate (pH 7.4), 100 mM acetaldehyde, 1 M ethanol and

30 mM NAD^+ (Sigma Aldrich) to test for hemiacetal activity of Adh. To test for esterase activity, cell lysates were incubated in potassium phosphate (pH 7.4), 500 mM ethanol and 500 mM potassium acetate. The reaction mixtures were incubated for 30 min at 30 °C. The samples were analysis by headspace chromatography as described above. To allow for hemiacetal production, reaction mixtures with 10 M ethanol and 1 M acetaldehyde were mixed beforehand and pH was adjusted to 10 using sodium hydroxide as previously described.[15] The chromatograms of the hemiacetal solution and vector control experiments showed no significant peak for ethyl acetate (Figure S4.11).

4.6.11 Western blot analysis

Western blot analysis was done to confirm the expression of Adh and Atf proteins with C-terminal c-Myc tag. 2.5 OD of cells were lysed using 0.1 M NaOH. Samples were loaded onto a 10-well Any kD™ Mini-PROTEAN® TGX™ Precast Protein Gel (BioRad) and run for 1h at 150 V. Samples were then electrophoretically transferred overnight to a PVDF membrane at 25 V. Membranes were blocked with 5% non-fat milk in TBST buffer for 1h at room temperature and incubated with anti-c-Myc mouse antibody (Sc-40, Santa Cruz Biotech) or anti-GAPDH (PA1-987) diluted to 1:2000 and 1:5000 in TBST buffer with 1% non-fat milk. Goat anti-mouse IgG-HRP (31430) diluted to 1:10000 was added as secondary antibody and incubated at room temperature for 30 min. After washing with TBST, HRP substrate (Bio-Rad) was used for signal detection. Blots were imaged using the BioRad ChemiDoc™ MP System with the Image Lab software.

4.6.12 Statistical analysis

Data points represent arithmetic means of at least triplicate biological samples, and error bars represent the standard deviation. Comparisons between two samples were accomplished by an unpaired two-tailed T-test with a significant difference at $p<0.05$. Groups of samples were analyzed by one-way ANOVA with a Tukey post-hoc test and considered significant at $p<0.05$. Statistical analysis and plotting of data points was performed using the GraphPad Prism software.

4.7 List of abbreviations

CRISPR, Clustered regularly interspaced short palindromic repeats; Adh, alcohol dehydrogenase, Atf, alcohol-O-acetyltransferase, Acetyl-CoA, Acetyl coenzyme A; NAD(P)(H), Nicotinamide adenine dinucleotide (phosphate) (reduced); sgRNA, single guide RNA RT-qPCR, reverse transcription quantitative polymerase chain reaction; 5-FOA, 5-fluororotic acid; OD, optical density; DO, dissolved oxygen; EDTA, ethylenediaminetetraacetic acid; DTT, Dithiothreitol; TBST, tris-buffered saline Tween 20; HRP, horseradish peroxidase

4.8 Acknowledgements

We thank Dr. Dae-Hyuk Kweon (Sungkyunkwan University) for kindly providing the pJSKM316GPD *K. marxianus* expression plasmid. This work was supported by NSF CBET-1510697.

This work was supported by NSF CBET-1510697.

4.9 References

[1] Lobs AK, Lin JL, Cook M, Wheeldon I: **High throughput, colorimetric screening of microbial ester biosynthesis reveals high ethyl acetate production from Kluyveromyces marxianus on C5, C6, and C12 carbon sources**. *Biotechnol. J.* 2016, **11**(10):1274-1281.

[2] Nonklang S, Abdel-Banat BMA, Cha-Aim K, Moonjai N, Hoshida H, Limtong S, Yamada M, Akada R: **High-Temperature Ethanol Fermentation and Transformation with Linear DNA in the Thermotolerant Yeast Kluyveromyces marxianus DMKU3-1042**. *Appl. Environ. Microbiol.* 2008, **74**(24):7514-7521.

[3] Fonseca GG, de Carvalho NMB, Gombert AK: **Growth of the yeast Kluyveromyces marxianus CBS 6556 on different sugar combinations as sole carbon and energy source**. *Appl. Microbiol. Biotechnol.* 2013, **97**(11):5055-5067.

[4] Rocha SN, Abrahao-Neto J, Cerdan ME, Gonzalez-Siso MI, Gombert AK: **Heterologous expression of glucose oxidase in the yeast Kluyveromyces marxianus**. *Microb. Cell Fact.* 2010, **9**.

[5] Gombert AK, Madeira JV, Cerdan ME, Gonzalez-Siso MI: **Kluyveromyces marxianus as a host for heterologous protein synthesis**. *Appl. Microbiol. Biotechnol.* 2016, **100**(14):6193-6208.

[6] Loser C, Urit T, Bley T: **Perspectives for the biotechnological production of ethyl acetate by yeasts**. *Appl. Microbiol. Biotechnol.* 2014, **98**(12):5397-5415.

[7] Park YC, Shaffer CE, Bennett GN: **Microbial formation of esters**. *Appl. Microbiol. Biotechnol.* 2009, **85**(1):13-25.

[8] Loser C, Urit T, Stukert A, Bley T: **Formation of ethyl acetate from whey by Kluyveromyces marxianus on a pilot scale**. *J. Biotechnol.* 2013, **163**(1):17-23.

[9] Fujii T, Nagasawa N, Iwamatsu A, Bogaki T, Tamai W, Hamachi M: **Molecular-Cloning, Sequence-Analysis, and Expression of the Yeast Alcohol Acetyltransferase Gene**. *Appl. Microbiol. Biotechnol.* 1994, **60**(8):2786-2792.

[10] Nagasawa N, Bogaki T, Iwamatsu A, Hamachi M, Kumagai C: **Cloning and nucleotide sequence of the alcohol acetyltransferase II gene (ATF2) from Saccharomyces cerevisiae Kyokai No. 7**. *Biosci. Biotech. Bioch.* 1998, **62**(10):1852-1857.

[11] Verstrepen KJ, Van Laere SDM, Vanderhaegen BMP, Derdelinckx G, Dufour JP, Pretorius IS, Winderickx J, Thevelein JM, Delvaux FR: **Expression levels of the yeast alcohol acetyltransferase genes ATF1, Lg-ATF1, and ATF2 control the formation of a broad range of volatile esters**. *Appl. Environ. Microbiol.* 2003, **69**(9):5228-5237.

[12] Lin JL, Wheeldon I: **Dual N- and C-Terminal Helices Are Required for Endoplasmic Reticulum and Lipid Droplet Association of Alcohol Acetyltransferases in Saccharomyces cerevisiae**. *PloS ONE* 2014, **9**(8).

[13] Zhu J, Lin JL, Palomec L, Wheeldon I: **Microbial host selection affects intracellular localization and activity of alcohol-O-acetyltransferase**. *Microb. Cell Fact.* 2015, **14**.

[14] Kusano M, Sakai Y, Kato N, Yoshimoto H, Tamai Y: **A novel hemiacetal dehydrogenase activity involved in ethyl acetate synthesis in Candida utilis**. *J. Biosci. Bioeng.* 1999, **87**(5):690-692.

[15] Park YC, San KY, Bennett GN: **Characterization of alcohol dehydrogenase 1 and 3 from Neurospora crassa FGSC2489**. *Appl. Microbiol. Biotechnol.* 2007, **76**(2):349-356.

[16] Kusano M, Sakai Y, Kato N, Yoshimoto H, Sone H, Tamai Y: **Hemiacetal dehydrogenation activity of alcohol dehydrogenases in Saccharomyces cerevisiae**. *Biosci. Biotech. Bioch.* 1998, **62**(10):1956-1961.

[17] Kallelmhiri H, Miclo A: **Mechanism of Ethyl-Acetate Synthesis by Kluyveromyces Fragilis**. *FEMS Microbiol. Lett.* 1993, **111**(2-3):207-212.

[18] Loser C, Urit T, Keil P, Bley T: **Studies on the mechanism of synthesis of ethyl acetate in Kluyveromyces marxianus DSM 5422**. *Appl. Microbiol. Biotechnol.* 2015, **99**(3):1131-1144.

[19] Lertwattanasakul N, Kosaka T, Hosoyama A, Suzuki Y, Rodrussamee N, Matsutani M, Murata M, Fujimoto N, Suprayogi, Tsuchikane K *et al*: **Genetic basis of the highly efficient yeast Kluyveromyces marxianus: complete genome sequence and transcriptome analyses**. *Biotechnol. Biofuels* 2015, **8**.

[20] Gao JQ, Yuan WJ, Li YM, Xiang RJ, Hou SB, Zhong SJ, Bai FW: **Transcriptional analysis of Kluyveromyces marxianus for ethanol production from inulin using consolidated bioprocessing technology**. *Biotechnol. Biofuels* 2015, **8**.

[21] Lertwattanasakul N, Sootsuwan K, Limtong S, Thanonkeo P, Yamada M: **Comparison of the gene expression patterns of alcohol dehydrogenase isozymes in the thermotolerant yeast Kluyveromyces marxianus and their physiological functions**. *Biosci. Biotech. Bioch.* 2007, **71**(5):1170-1182.

[22] Bozzi A, Saliola M, Falcone C, Bossa F, Martini F: **Structural and biochemical studies of alcohol dehydrogenase isozymes from Kluyveromyces lactis**. *BBA-Protein Struct. M.* 1997, **1339**(1):133-142.

[23] de Smidt O, du Preez JC, Albertyn J: **The alcohol dehydrogenases of Saccharomyces cerevisiae: a comprehensive review**. *FEMS Yeast Res.* 2008, **8**(7):967-978.

[24] Liang JJ, Zhang ML, Ding M, Mai ZM, Wu SX, Du Y, Feng JX: **Alcohol dehydrogenases from Kluyveromyces marxianus: heterologous expression in Escherichia coli and biochemical characterization**. *BMC Biotechnol.* 2014, **14**.

[25] Abdel-Banat BMA, 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* 2010, **27**(1):29-39.

[26] Schwartz C, Shabbir-Hussain M, Frogue K, Blenner M, Wheeldon I: **Standardized Markerless Gene Integration for Pathway Engineering in Yarrowia lipolytica**. *ACS Synth. Biol.* 2017, **6**(3):402-409.

[27] Schwartz CM, Hussain MS, Blenner M, Wheeldon I: **Synthetic RNA Polymerase III Promoters Facilitate High-Efficiency CRISPR-Cas9-Mediated Genome Editing in Yarrowia lipolytica**. *ACS Synth. Biol.* 2016, **5**(4):356-359.

[28] Banat IM, Nigam P, Marchant R: **Isolation of Thermotolerant, Fermentative Yeasts Growing at 52-Degrees-C and Producing Ethanol at 45-Degrees-C and 50-Degrees-C**. *World J. Microb. Biotechnol.* 1992, **8**(3):259-263.

[29] Ryan OW, Skerker JM, Maurer MJ, Li X, Tsai JC, Poddar S, Lee ME, DeLoache W, Dueber JE, Arkin AP *et al*: **Selection of chromosomal DNA libraries using a multiplex CRISPR system**. *Elife* 2014, **3**.

[30] Doench JG, Hartenian E, Graham DB, Tothova Z, Hegde M, Smith I, Sullender M, Ebert BL, Xavier RJ, Root DE: **Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation**. *Nat. Biotechnol.* 2014, **32**(12):1262-1267.

[31] Urit T, Loser C, Wunderlich M, Bley T: **Formation of ethyl acetate by Kluyveromyces marxianus on whey: studies of the ester stripping**. *Bioproc. Biosyst. Eng.* 2011, **34**(5):547-559.

[32] Plata C, Mauricio JC, Millan C, Ortega JM: **Influence of glucose and oxygen on the production of ethyl acetate and isoamyl acetate by a Saccharomyces cerevisiae strain during alcoholic fermentation**. *World J. Microb. Biotechnol.* 2005, **21**(2):115-121.

[33] Wheeldon I, Christopher P, Blanch H: **Integration of heterogeneous and biochemical catalysis for production of fuels and chemicals from biomass**. *Curr. Op. Biotechnol.* 2017, **45**:127-135.

[34] Pecota DC, Rajgarhia V, Da Silva NA: **Sequential gene integration for the engineering of Kluyveromyces marxianus**. *J. Biotechnol.* 2007, **127**(3):408-416.

[35] Ribeiro O, Gombert AK, Teixeira JA, Domingues L: **Application of the Cre-loxP system for multiple gene disruption in the yeast Kluyveromyces marxianus**. *J. Biotechnol.* 2007, **131**(1):20-26.

[36] Heo P, Yang TJ, Chung SC, Cheon Y, Kim JS, Park JB, Koo HM, Cho KM, Seo JH, Park JC *et al*: **Simultaneous integration of multiple genes into the Kluyveromyces marxianus chromosome**. *J. Biotechnol.* 2013, **167**(3):323-325.

[37] Vyas VK, Barrasa MI, Fink GR: **A Candida albicans CRISPR system permits genetic engineering of essential genes and gene families**. *Sci. Adv.* 2015, **1**(3):e1500248.

[38] Wagner JM, Alper HS: **Synthetic biology and molecular genetics in non-conventional yeasts: Current tools and future advances**. *Fungal Genet. Biol.* 2016, **89**:126-136.

[39] Jacobs JZ, Ciccaglione KM, Tournier V, Zaratiegui M: **Implementation of the CRISPR-Cas9 system in fission yeast**. *Nat. Comm.* 2014, **5**.

[40] Min K, Ichikawa Y, Woolford CA, Mitchell AP: **Candida albicans Gene Deletion with a Transient CRISPR-Cas9 System**. *Msphere* 2016, **1**(3).

[41] Weninger A, Hatzl AM, Schmid C, Vogl T, Glieder A: **Combinatorial optimization of CRISPR/Cas9 expression enables precision genome engineering in the methylotrophic yeast Pichia pastoris**. *J. Biotechnol.* 2016, **235**:139-149.

[42] Horwitz AA, Walter JM, Schubert MG, Kung SH, Hawkins K, Platt DM, Hernday AD, Mahatdejkul-Meadows T, Szeto W, Chandran SS *et al*: **Efficient Multiplexed Integration of Synergistic Alleles and Metabolic Pathways in Yeasts via CRISPR-Cas**. *Cell Syst.* 2015, **1**(1):88-96.

[43] Van Laere SDM, Saerens SMG, Verstrepen KJ, Van Dijck P, Thevelein JM, Delvaux FR: **Flavour formation in fungi: characterisation of KlAtf, the Kluyveromyces lactis orthologue of the Saccharomyces cerevisiae alcohol acetyltransferases Atf1 and Atf2**. *Appl. Microbiol. Biotechnol.* 2008, **78**(5):783-792.

[44] Gethins L, Guneser O, Demirkol A, Rea MC, Stanton C, Ross RP, Yuceer Y, Morrissey JP: **Influence of carbon and nitrogen source on production of volatile fragrance and flavour metabolites by the yeast Kluyveromyces marxianus**. *Yeast* 2015, **32**(1):67-76.

[45] Lertwattanasakul N, Shigemoto E, Rodrussamee N, Limtong S, Thanonkeo P, Yamada M: **The crucial role of alcohol dehydrogenase Adh3 in Kluyveromyces marxianus mitochondrial metabolism**. *Biosci. Biotechnol. Biochem.* 2009, **73**(12):2720-2726.

[46] Steinbuchel A, Schlegel HG: **A Multifunctional Fermentative Alcohol-Dehydrogenase from the Strict Aerobe Alcaligenes-Eutrophus - Purification and Properties**. *Eur. J. Biochem.* 1984, **141**(3):555-564.

[47] Jendrossek D, Steinbuchel A, Schlegel HG: **Alcohol-Dehydrogenase Gene from Alcaligenes-Eutrophus - Subcloning, Heterologous Expression in Escherichia-Coli, Sequencing, and Location of Tn5 Insertions**. *J. Bacteriol.* 1988, **170**(11):5248-5256.

[48] Kruis AJ, Levisson M, Mars AE, van der Ploeg M, Garces Daza F, Ellena V, Kengen SWM, van der Oost J, Weusthuis RA: **Ethyl acetate production by the elusive alcohol acetyltransferase from yeast**. *Metab. Eng.* 2017, **41**:92-101.

[49] Lin JL, Zhu J, Wheeldon I: **Rapid ester biosynthesis screening reveals a high activity alcohol-O-acyltransferase (AATase) from tomato fruit**. *Biotechnol. J.* 2016, **11**(5):700-707.

[50] Lilly M, Bauer FF, Lambrechts MG, Swiegers JH, Cozzolino D, Pretorius IS: **The effect of increased yeast alcohol acetyltransferase and esterase activity on the flavour profiles of wine and distillates**. *Yeast* 2006, **23**(9):641-659.

[51] Lee KS, Kim JS, Heo P, Yang TJ, Sung YJ, Cheon Y, Koo HM, Yu BJ, Seo JH, Jin YS *et al*: **Characterization of Saccharomyces cerevisiae promoters for heterologous gene expression in Kluyveromyces marxianus**. *Appl. Microbiol. Biot.* 2013, **97**(5):2029-2041.

[52] DiCarlo JE, Norville JE, Mali P, Rios X, Aach J, Church GM: **Genome engineering in Saccharomyces cerevisiae using CRISPR-Cas systems**. *Nucl. Acids Res.* 2013, **41**(7):4336-4343.

[53] Gibson DG, Young L, Chuang RY, Venter JC, Hutchison CA, 3rd, Smith HO: **Enzymatic assembly of DNA molecules up to several hundred kilobases**. *Nat. Methods* 2009, **6**(5):343-345.

[54] Puigbo P, Guzman E, Romeu A, Garcia-Vallve S: **OPTIMIZER: a web server for optimizing the codon usage of DNA sequences**. *Nucl. Acids Res.* 2007, **35**(Web Server issue):W126-131.

[55] Antunes DF, de Souza CG, de Morais MA: **A simple and rapid method for lithium acetate-mediated transformation of Kluyveromyces marxianus cells**. *World J. Microb. Biotechnol.* 2000, **16**(7):653-654.

[56] Nelson MD, Fitch DHA: **Overlap Extension PCR: An Efficient Method for Transgene Construction**. *Mol. Methods Evol. Gene.* 2011, **772**:459-470.

[57] Jeong H, Lee D, Kim S, Kim H, Lee K, Song J, Kim B, Sung B, Park J, Sohn J *et al*: **Genome sequence of the thermotolerant yeast Kluyveromyces marxianus var. marxianus KCTC 17555**. *Eukaryot. Cell* 2012, **11**:1584 - 1585.

[58] Marck C, Kachouri-Lafond R, Lafontaine I, Westhof E, Dujon B, Grosjean H: **The RNA polymerase III-dependent family of genes in hemiascomycetes: comparative RNomics, decoding strategies, transcription and evolutionary implications**. *Nucl. Acids Res.* 2006, **34**(6):1816-1835.

[59] Pfaffl MW: **A new mathematical model for relative quantification in real-time RT-PCR**. *Nucl. Acids Res.* 2001, **29**(9):e45.

4.10 Supporting Information

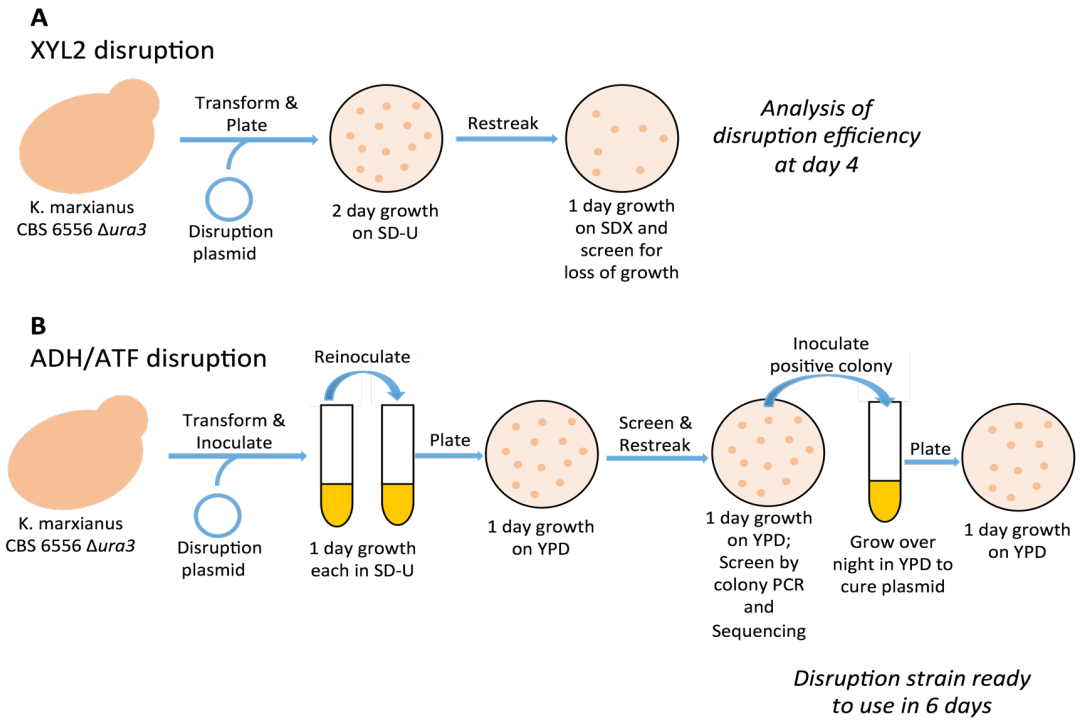


Figure S4.1: Schematic workflow of CRISPR-Cas9 mediated gene disruption and screening. Description of the workflow for XYL2 (A) and ADH/ATF (B) disruptions and screening. (A) To determine the efficiency of the sgRNA expression promoter systems, disruptions were screened by the phenotypic loss of growth on xylitol. After transformation, cells were plated on selective media and grown for 2 days. Colonies were restreaked on rich (YPD) and xylitol (SDX) solid media. At day 4, colonies with growth on rich but no growth on SDX media were considered to have a disruption in the XYL2 gene. Selected colonies were confirmed by sequencing. (B) For ADH and ATF disruptions, cells were transformed and grown in selective media for 1 day. To enrich colonies with the plasmid, cells were reinoculated in new media for another day and subsequently plated on rich media. Random colonies selected for screening were subjected to colony PCR and restreaked on fresh solid media plates. Colony PCR products were sequenced to identify indels. Finally, colonies with successful gene disruptions were grown over night in rich media to cure the CRIPSR-Cas9 plasmid.

Table S4.1: sgRNA and efficiencies used in the study

Target	Target Sequence	Score	Strand	Indel Success	sgRNA promoter
XYL2	ACGATCGCCAACCTTGACCA	0.66	antisense	68/90	**ScSNR52-tRNA^{Gly}**
ADH1	AAAGAACGTCGACTTGGCCG	0.832	sense	0/20	
ADH1 T2	GGCAGCCTGGACAGCGTCAG	0.486	antisense	0/20	
ADH1T3	AGACTTCAAAGCCTTGTACA	0.453	antisense	2/10	
ADH2	GTGACCTTGCCGGTATCAAA	0.686	sense	0/20	
ADH2 T2	GTCACCAGCCTTCATTTCAG	0.697	antisense	0/20	
ADH2 T3	GGTACCAGCTGGGATGTGAG	0.421	antisense	4/30	
ADH3	GCTATTCCAGAAAAGCAAAA	0.826	sense	2/10	
ADH4	GCCATCCCAGAATCCCAAAA	0.825	sense	4/10	
ADH5	ATGGTCTTGAAAGAACACAA	0.716	sense	1/10	
ADH6	GTACCACCACCGCAAAGTAG	0.751	antisense	2/10	
ADH7	GTATTAGGCCATGAAGGTAT	0.709	sense	0/20	
ADH7 T2	TCTCCTTAGCCATAGCCAAA	0.883	antisense	0/20	
ADH7 T3	GCTTGAGCTGAGAGATTGAT	0.695	antisense	3/5	
ATF T1	GCTGAAACAGAGTTTCAGCA	0.958	sense	0/20	
ATF T2	ATATAGTCTTCGGCAACACC	0.551	antisense	4/10	
XYL2 T1	ACGATCGCCAACCTTGACCA	0.622	antisense	58/90	**KmRPR1-tRNA^{Gly}**
XYL2 T2	GCAATTCAAGATAAGTTGGG	0.839	sense	2/30	
scrambled	AGTCCGGTCATTACAACTTA	-	-	0	

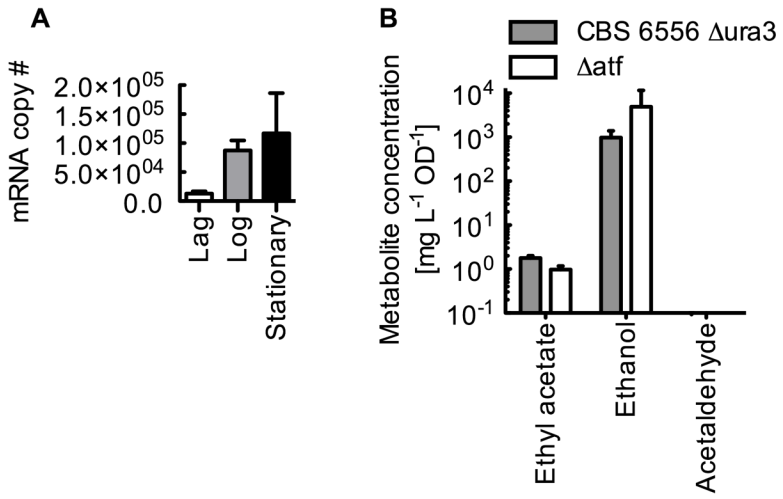


Figure S4.2: Anaerobic ATF expression and impact of *Km*ATF knockout on volatile metabolite production. (A) mRNA copy number of 5 ng input total RNA of wildtype *Km*ATF at different growth stages of anaerobic growth and (B) ethyl acetate, ethanol and acetaldehyde production of the URA3 deficient background strain compared to the *Km*ATF disruption strain. Bars and error bars represent the arithmetic mean and standard deviation of three samples.

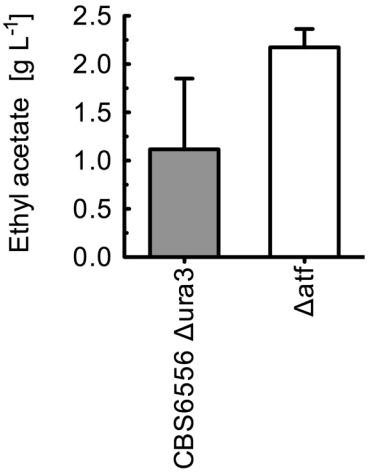


Figure S4.3: Aerobic ethyl acetate production of *Km*ATF knockout on SD media. *Km*ATF disruption does not lead to a significant decrease in ethyl acetate production compared to the URA3 deficient strains when the strains are grown aerobically on SD media. Bars and error bars represent the arithmetic mean and standard deviation of three samples.

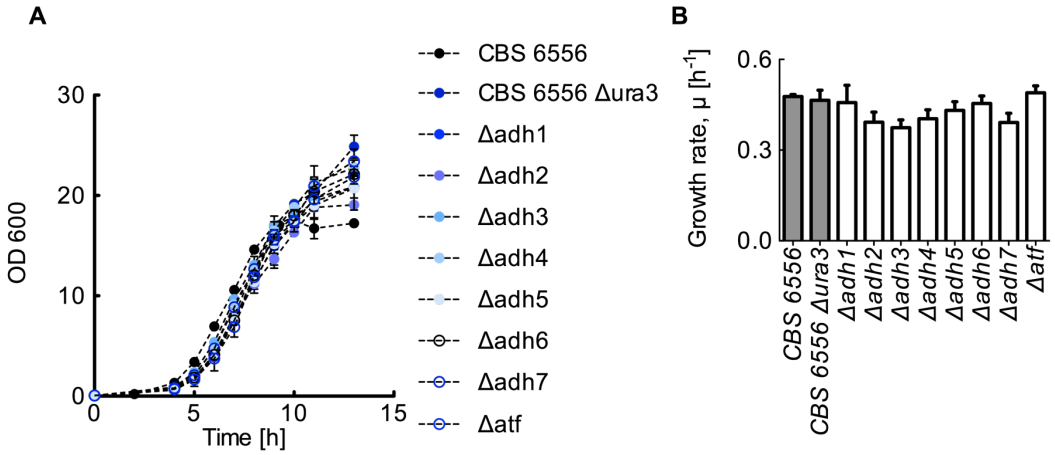


Figure S4.4: Aerobic growth of *Km*ADH and *Km*ATF disruption strains. (A) Growth curves and (B) growth rates of *K. marxianus* wild type, URA3 deficient and *KmADH1-7* and *Km*ATF disruption strains under aerobic conditions. Data points and bars represent the arithmetic mean of 3 biological replicates and error bars represent the standard deviation.

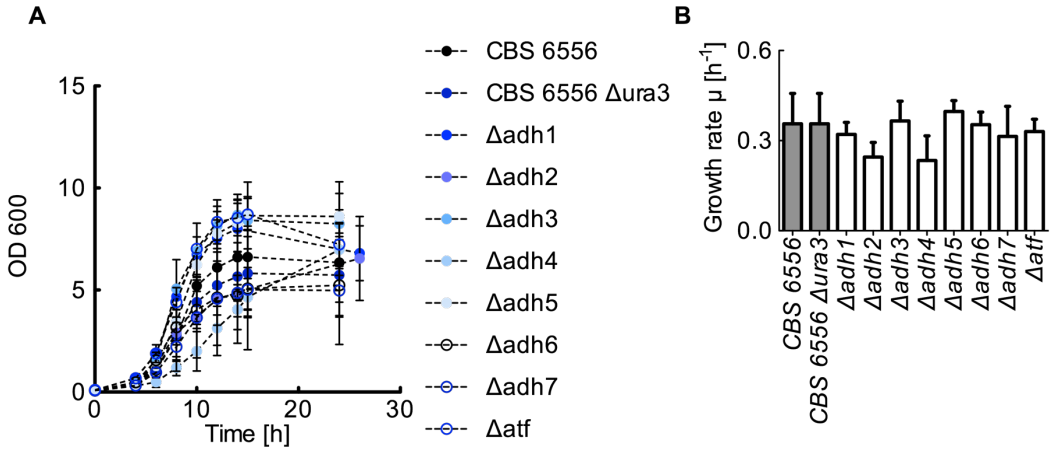


Figure S4.5: Anaerobic growth of *Km*ADH and *Km*ATF disruption strains. (A) Growth curves and (B) growth rates of *K. marxianus* wild type, URA3 deficient and *Km*ADH1-7 and *Km*ATF disruption strains under aerobic conditions. Data points and bars represent the arithmetic mean of triplicate biological replicates and error bars represent the standard deviation

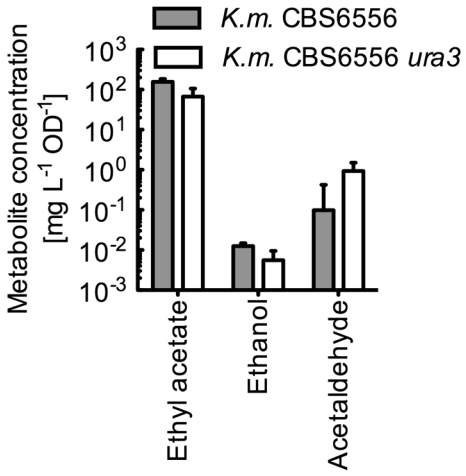


Figure S4.6: Metabolite production of wild type and Ura3 deficient *K. marxianus* CBS6556 strains. Ethyl acetate and ethanol production are significantly reduced in the URA3 disrupted strain. Bars represent the arithmetic mean of biological triplicates and error bars represent the standard deviation.

Table S4.2: Strains used in the study

Strain	Genotype	Source
YS5	*S. cerevisiae* BY4742 MAT*α his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0*	GE Healthcare
YS8	*S .cerevisiae* YS5 + pIW14	Lin et al [1]
YS202	*S .cerevisiae* YS5 + pIW107	Zhu *et al.*[2]
YS302	*K. marxianus* CBS6556	DMSZ
YS402	*K. marxianus* CBS6556 *ura3Δ*	This study
YS630	*K. marxianus* CBS6556 *ura3Δ adh4Δ*	This study
YS671	*K. marxianus* CBS6556 *ura3Δ adh3Δ*	This study
YS673	*K. marxianus* CBS6556 *ura3Δ adh5Δ*	This study
YS675	*K. marxianus* CBS6556 *ura3Δ adh6Δ*	This study
YS679	*K. marxianus* CBS6556 *ura3Δ atfΔ*	This study
YS703	*K. marxianus* CBS6556 *ura3Δ adh2Δ*	This study
YS720	*K. marxianus* CBS6556 *ura3Δ adh7Δ*	This study
YS794	*K. marxianus* CBS6556 *ura3Δ adh1Δ*	This study
YAL1	*S .cerevisiae* YS5 + pIW695	This study
YAL2	*S .cerevisiae* YS5 + pIW696	This study
YAL3	*S .cerevisiae* YS5 + pIW697	This study
YAL4	*S .cerevisiae* YS5 + pIW698	This study
YAL5	*S .cerevisiae* YS5 + pIW699	This study
YAL6	*S .cerevisiae* YS5 + pIW700	This study
YAL7	*S .cerevisiae* YS5 + pIW701	This study
YAL8	*S .cerevisiae* YS5 + pIW702	This study

Table S4.3: Plasmids used in the study

Name	Description	Source
pIW14	pRS426 PGK1p-PGK1t	Lin et al [1]
pIW21	pRS426 PGK1p-ATF1-GFP-PGK1t	Lin et al [1]
pIW107	pRS426 PGK1p-ATF1-C-Myc-PGK1t	Zhu *et al*. [2]
pIW243	p414-TEF1p-Cas9-CYC1t (S.c. expression, *S. pyogenes* human optimized)	Addgene [3]
pIW272	pJSKM316GPD	Lee *et al*.[4]
pIW333	Tef1-Cas9-Cyc1 ScSNR52-Ade Target2-SUP4	This study
pIW360	Tef1-KmCas9-Cyc1 ScSNR52-Ade Target2-SUP4	This study
pIW443	Tef1-KmCas9-SV40-Cyc1 ScSNR52-XDH1-SUP4	This study
pIW444	Tef1-KmCas9-SV40-Cyc1 KmSNR52-XDH1-SUP4	This study
pIW445	Tef1-KmCas9-SV40-Cyc1 KmSNR52p-Gly-tRNA-XDH1-SUP4	This study
pIW446	Tef1-KmCas9-SV40-Cyc1 KmSCR1p-Gly-tRNA-XDH1-SUP4	This study
pIW447	Tef1-KmCas9-SV40-Cyc1 KmRPR1p-Gly-tRNA-XDH1-SUP4	This study
pIW461	Tef1-KmCas9-SV40-Cyc1 ScSNR52-Gly-tRNA-ADH1-SUP4	This study
pIW462	Tef1-KmCas9-SV40-Cyc1 ScSNR52-Gly-tRNA-ADH2-SUP4	This study
pIW463	Tef1-KmCas9-SV40-Cyc1 ScSNR52-Gly-tRNA-ADH3-SUP4	This study
pIW464	Tef1-KmCas9-SV40-Cyc1 ScSNR52-Gly-tRNA-ADH4-SUP4	This study
pIW465	Tef1-KmCas9-SV40-Cyc1 ScSNR52-Gly-tRNA-ADH5-SUP4	This study
pIW466	Tef1-KmCas9P-SV40-Cyc1 ScSNR52-Gly-tRNA-ADH6-SUP4	This study
pIW467	Tef1-KmCas9-SV40-Cyc1 ScSNR52-Gly-tRNA-ADH7-SUP4	This study
pIW468	Tef1-KmCas9-SV40-Cyc1 ScSNR52-Gly-tRNA-ATF-SUP4	This study
pIW492	Tef1-KmCas9-SV40-Cyc1 KmRPR1-Gly-tRNA-Xdh2-SUP4	This study
pIW502	Tef1-GFP-SV40-Cyc1 Gly-tRNA-XDH1-SUP4	This study
pIW503	Tef1-GFP-SV40-Cyc1 KmRPR1-Gly-tRNA-scramble DNA-SUP4	This study
pIW506	Tef1-KmCas9-SV40-Cyc1 ScSNR52-Gly-tRNA-ATF T2-SUP4	This study
pIW508	Tef1-KmCas9-SV40-Cyc1 ScSNR52-Gly-tRNA-ADH1T2-SUP4	This study
pIW509	Tef1-KmCas9-SV40-Cyc1 ScSNR52-Gly-tRNA-ADH2 T2-SUP4	This study
pIW554	Tef1-KmCas9-SV40-Cyc1 ScSNR52-Gly-tRNA-ADH7 T2-SUP4	This study
pIW557	Tef1-KmCas9-SV40-Cyc1 ScSNR52-Gly-tRNA-ADH1 T3-SUP4	This study
pIW558	Tef1-KmCas9-SV40-Cyc1 ScSNR52-Gly-tRNA-ADH2 T3-SUP4	This study
pIW576	Tef1-KmCas9-SV40-Cyc1 ScSNR52-Gly-tRNA-ADH7 T3-SUP4	This study
pIW695	PGK1p-KmAdh1-c-Myc-PGK1t	This study
pIW696	PGK1p-KmAdh2-c-Myc-PGK1t	This study
pIW697	PGK1p-KmAdh3-c-Myc-PGK1t	This study
pIW698	PGK1p-KmAdh4-c-Myc-PGK1t	This study
pIW699	PGK1p-KmAdh5-c-Myc-PGK1t	This study
pIW700	PGK1p-KmAdh6-c-Myc-PGK1t	This study
pIW701	PGK1p-KmAdh7-c-Myc-PGK1t	This study
pIW702	PGK1p-KmAtf-c-Myc-PGK1t	This study

Table S4.4: Primers used in the study

Name	Primer Sequence (5' to 3')
P1019 URA 1F Overl	GAGCATCTTGGTCTTCTGAG
P1020 URA 2R Overl	GCCCGCACCAGTCACACCGTGTGCATTGGTTACATGTGTCTTCAATAG ACAGATA
P1021 URA 3F Overl	CCATATATCTGTCTATTGAAGACACATGTAACCAATGCACACGGTGTG ACTGGTGCGG
P1022 URA 4R Overl	GTATACAATGTGACGCAATGC
P1072 URA3- F	GTC AAA CTT ATG TGC TTC TCT TG
P1073 URA3- R	CGGCAAGCATTAACAACCC
P1379 Cas9 F	CGACTCACTATAGGGCGAATTGGAGCTCCACATAGCTTCAAAATGTTT CTACTCCTTTTT
P1525 Cas9 R	TTATCTTTTCAAAGAcgcGGTACCaagGCAAATTAAAGCCTTCGAGCGT CCCAAAAC
P1530 sgRNA R	ATTAACCCTCACTAAAGGGAACAAAAGCTGctaCCGCGGcagAGACATA AAAAACAAAAAAAGCACC
P1626 SNR52 F	CGCTCGAAGGCTTTAATTTGCcttCCTAGGGCGTCTTTGAAAAGATAAT G
P1684 Adh1 KO Seq F	GCCATTGGTGGGTGGTCACG
P1685 Adh1 KO Seq R	TATCCTTGGTCTTGGTAAAGTC
P1690 Adh4 KO Seq F	TTACCCGGACTTTATCGAATTAG
P1691 Adh4 KO Seq R	TAGTCACCGATCTCGAAGTT
P1692 Adh5 KO Seq F	TCCACAGTTGGTTTTTACATG
P1693 Adh5 KO Seq R	ATATGGTTGTCAATAATAGCCATC
P1694 Adh6 KO Seq F	ACCACTGGTATCAAGGTTGG
P1695 Adh6 KO Seq R	ACGATCAAGTCCAATTTGTCA
P1698 Atf KO Seq F	ATGATGAAACTACAGAAATTGTCG
P1699 Atf KO Seq R	AGGAAGCAGGATCAGTCTGA
P1773 XDH1 sgRNA F	ACGATCGCCAACCTTGACCATTTTAGAGCTAGAAATAGCAAGTTAAA
P1775 XDH2 sgRNA F	GCAATTCAAGATAAGTTGGGTTTTAGAGCTAGAAATAGCAAGTTAAA
P1789 ScSNR52 tRNA R	GAAATTGCGGCCGCGATCATTTATCTTTCACTGCGG
P1790 tRNA ScSNR52 F	AATGATcgcggccgCAATTTCTCTTCTACCACGAACTC
P1792 KmSNR52 F	CGCTCGAAGGCTTTAATTTGCcttCCTAGGgcgGGGAGTGAGTAAAAAA AAGAGAAG
P1793 KmSNR52 tRNA R	AGAAATTGcggccgcGTAAGATTCGAACTGCGGACG
P1794 tRNA KmSNR52 F	ATCTTAcgcggccgCAATTTCTCTTCTACCACGAACTC
P1795 KmSCR1 F	CGCTCGAAGGCTTTAATTTGCcttCCTAGGgcgTGTTATACTTGGATAAG TGGCTC
P1796 KmSCR1 tRNA R	agaaattgcggccgcTGGGAAAATTTGCTAAATCGTTAC
P1797 tRNA SCR1 F	tttcccagcggccgCAATTTCTCTTCTACCACGAACTC
P1798 KmRPR1 F	CGCTCGAAGGCTTTAATTTGCcttCCTAGGgcgTATACTCCAACTTGGTC GAAAG
P1799 KmRPR1 tRNA R	agaaattgcggccgcATCTAAATTCTCTCTTTTTCCTTCAA
P1800 tRNA RPR1 F	tttagatgcggccgCAATTTCTCTTCTACCACGAACTC

Name	Primer Sequence (5' to 3')
P1806 KmSNR52 Xdh1 R	ATGGTCAAGGTTGGCGATCGTCctcgagGGTAAGATTCGAACTGCGGAC G
P1807 GlytRNA Xdh1 R	AAAATGGTCAAGGTTGGCGATCGTCctcgagGTTGACACTGACGGGATT CGA
P1811 ADH1 KO sgRNA F	AAAGAACGTCGACTTGGCCGTTTTAGAGCTAGAAATAGCAAGTTAAA
P1812 ADH1 KO R	CGGCCAAGTCGACGTTCTTTCctcgagGTTGACACTGACGGGATTC
P1813 ADH2 KO sgRNA F	GTGACCTTGCCGGTATCAAATTTTAGAGCTAGAAATAGCAAGTTAAA
P1814 ADH2 KO R	TTTGATACCGGCAAGGTCACCctcgagGTTGACACTGACGGGATTC
P1815 ADH3 KO sgRNA F	GCTATTCCAGAAAAGCAAAATTTTAGAGCTAGAAATAGCAAGTTAAA
P1816 ADH3 KO R	TTTTGCTTTTCTGGAATAGCCctcgagGTTGACACTGACGGGATTC
P1817 ADH4 KO sgRNA F	GCCATCCCAGAATCCCAAAATTTTAGAGCTAGAAATAGCAAGTTAAA
P1818 ADH4 KO R	TTTTGGGATTCTGGGATGGCCctcgagGTTGACACTGACGGGATTC
P1819 ADH5 KO sgRNA F	ATGGTCTTGAAAGAACACAATTTTAGAGCTAGAAATAGCAAGTTAAA
P1820 ADH5 KO R	TTGTGTTCTTTCAAGACCATCctcgagGTTGACACTGACGGGATTC
P1821 ADH6 KO sgRNA F	GTACCACCACCGCAAAGTAGTTTTAGAGCTAGAAATAGCAAGTTAAA
P1822 ADH6 KO R	CTACTTTGCGGTGGTGGTACCctcgagGTTGACACTGACGGGATTC
P1823 ADH7 KO sgRNA F	GTATTAGGCCATGAAGGTATTTTTAGAGCTAGAAATAGCAAGTTAAA
P1824 ADH7 KO R	ATACCTTCATGGCCTAATACCctcgagGTTGACACTGACGGGATTC
P1825 ATF KO sgRNA F	GCTGAAACAGAGTTTCAGCATTTTAGAGCTAGAAATAGCAAGTTAAA
P1826 ATF KO R	TGCTGAAACTCTGTTTCAGCCctcgagGTTGACACTGACGGGATTC
P1827 KmXDH2 R	CCCAACTTATCTTGAATTGCCctcgagGTTGACACTGACGGGATTC
P1833 Adh2 KO Seq F	TTGCCATTGGTCGGTGGTCAC
P1834 Adh3 KO Seq F	GAACTTAGCTCAGTCAAGTCCGAA
P1839 GlytRNA only F	GCTTTAATTTGCcttCCTAGGgcgGCGGCCGCAATTTCTCTT
P1840 GlytRNA only R	cgcCCTAGGaagGCAAATTAAAGCCTTCG
P1841 scramble DNA sgRNA F	AGTCCGGTCATTACAACTTATTTTAGAGCTAGAAATAGCAAGTTAAA
P1842 tRNA scramble DNA R	TAAGTTGTAATGACCGGACTCctcgagGTTGACACTGACGGGATTC
P1850 ATF1 T2 KO sgRNA F	ATATAGTCTTCGGCAACACCTTTTAGAGCTAGAAATAGCAAGTTAAA
P1851 ATF1 T2 KO R	GGTGTTGCCGAAGACTATATCctcgagGTTGACACTGACGGGATTC
P1857 ADH1 T2 KO sgRNA F	CctcgagGGGCAGCCTGGACAGCGTCAGTTTTAGAGCTAGAAATAGCAA GTTAAA
P1858 ADH1 T2 KO R	AAAACTGACGCTGTCCAGGCTGCCCctcgagGTTGACACTGACGGGATTC
P1860 ADH2 T2 KO sgRNA F	CctcgagGGTCACCAGCCTTCATTTCAGTTTTAGAGCTAGAAATAGCAAG TTAAA
P1861 ADH2 T2 KO R	AAAACTGAAATGAAGGCTGGTGACCctcgagGTTGACACTGACGGGATT C
P1868 ADH3 T2 KO sgRNA F	CctcgagGGGACAATTCACAGAATTCACTTTTAGAGCTAGAAATAGCAA GTTAAA
P1869 ADH3 T2 KO R	GCTCTAAAAGTGAATTCTGTGAATTGTCCCctcgagGTTGACACTGACGG GATTC
P1871 ADH3 T3 KO sgRNA F	CctcgagGGTCACCAGCCTTCAAGCCAGTTTTAGAGCTAGAAATAGCAAG TTAAA

Name	Primer Sequence (5' to 3')
P1872 ADH3 T3 KO R	GCTCTAAAACTGGCTTGAAGGCTGGTGACCctcgagGTTGACACTGACGGGATTC
P1874 KmADH3 Seq R	AAGGCTTCGGCACCCAATTG
P1877 KmAct1Q F	CCCAATGAACCCAAAGAATAACAG
P1878 KmAct1Q R	GATAGCATGAGGCAAGGAGAAACC
P1879 KmGapdhQ F	GTCCAGAAAGAACATCGAAGTTGTC
P1880 KmGapdhQ R	GTAGCTGGGTCTCTTTCTTGGAAG
P1932 ADH7 T2 KO sgRNA F	gagGTCTCCTTAGCCATAGCCAAATTTTAGAGCTAGAAATAGCAAGTTAAA
P1933 ADH7 T2 KO R	TAAAATTTGGCTATGGCTAAGGAGACctcgagGTTGACACTGACGGGATTC
P1937 ADH7 KO Seq R	CCACCTTCCTTGAGAACACG
P1938 ADH1 T3 sgRNA F	CctcgagGAGACTTCAAAGCCTTGTACATTTTAGAGCTAGAAATAGCAAGTTAAA
P1939 ADH1 T3 R	GCTCTAAAATGTACAAGGCTTTGAAGTCTCctcgagGTTGACACTGACGGGATTC
P1941 ADH2 T3 sgRNA F	CctcgagGGGTACCAGCTGGGATGTGAGTTTTAGAGCTAGAAATAGCAAGTTAAA
P1942 ADH2 T3 R	GCTCTAAAACTCACATCCCAGCTGGTACCCctcgagGTTGACACTGACGGGATTC
P1946 ADH7 T3 sgRNA F	gagGGCTTGAGCTGAGAGATTGATTTTTAGAGCTAGAAATAGCAAGTTAAA
P1947 ADH7 T3 R	TAAAAATCAATCTCTCAGCTCAAGCCctcgagGTTGACACTGACGGGATTC
P1949 ADH7 KO Seq F	CAACTACATTGAATATAAACATATATATATATCAGC
AL30 ADH2 KO Seq R	GACAATGTCCTTAGACTTGGT
AL87 KmAdh1 qPCR F	CGTTACTGGCTGGGAAATCG
AL88 KmAdh1 qPCR R	GAACCGTCGTGTGTGTAACC
AL89 KmAdh2 qPCR F	GTCATCAAGGCTACCAACGG
AL90 KmAdh2 qPCR R	CATCGGACTTACACTTGGCG
AL91 KmAdh3 qPCR F	GTCCACACGGTGTCATCAAC
AL92 KmAdh3 qPCR R	TAACGACGTGGGAGAAGACC
AL93 KmAdh4 qPCR F	CAGACCAGCATTACCACCAC
AL94 KmAdh4 qPCR R	TCCAACTTACCGCCGTTTTC
AL95 KmAdh5 qPCR F	CTTCCACCTCCATTGACTGC
AL96 KmAdh5 qPCR R	AACTCAACGCCCTTCAAAGC
AL97 KmAdh6 qPCR F	AAGAGATACGGCTGTGGTCC
AL98 KmAdh6 qPCR R	ATGGCGTAAACTTCAGCACC
AL99 KmAdh7 qPCR F	CGGTGTCCATGGAAAGTCTG
AL100 KmAdh7 qPCR R	TGGCAAGCTTTTCGGACTTC
AL101 KmAtf1 qPCR F	CTGTCCCCGTTGATGAATCG
AL102 KmAtf1 qPCR R	TGGTGTCAATGTGGCCTTAC
AL114 FP1	TATACATGGGATCA TAAATC

Name	Primer Sequence (5' to 3')
AL115 RP1	CTTTGTCTTGTATGATATC
AL118 KmAdh1 F	ttttctctttttacagatcaCCgcGGATGGCTATTCCAGAAACTCAA
AL119 KmAdh1 Myc R	AGATAAGTTTTTGTTCacctccgcctagggatccgcctccTTTGGAAGTGTCAACGACAA
AL120 KmAdh1 Myc R Extension	ATCTATCGATTTCAATTCAATTCAATACTAGTTTACAGGTCCTCCTCGGAAATCAGCTTTTGTTCacctc
AL121 KmAdh2 F	ttttctctttttacagatcaCCgcGGATGTCTATTCCAACTACTCAAAAGG
AL122 KmAdh2 Myc R	GCTTTTGTTCacctccgcctagggatccgcctccTTTGGAAGTGTCAACAACGTAT
AL123 KmAdh3 F	ttttctctttttacagatcaCCgcGGATGCTTAGATTAACTAACGCCAG
AL124 KmAdh3 Myc R	GCTTTTGTTCacctccgcctagggatccgcctccTTTTTCAGTGTCGACGACGT
AL125 KmAdh4 F	ttttctctttttacagatcaCCgcGGATGTTCAGACTAGCACGCGC
AL126 KmAdh4 Myc R	GCTTTTGTTCacctccgcctagggatccgcctccTTTGGAAGTGTCAACGACGTAT
AL127 KmAdh5 F	ttttctctttttacagatcaCCgcGGATGTTTCATAGAAGAGCATTGAAG
AL128 KmAdh5 Myc R	GCTTTTGTTCacctccgcctagggatccgcctccGCATTCATAGGCCTGTCTGA
AL129 KmAdh6 F	ttttctctttttacagatcaCCgcGGATGTCCTACCCAGATAGTTTCC
AL130 KmAdh6 Myc R	GCTTTTGTTCacctccgcctagggatccgcctccTTTTTGAGCCTTGAACTCTCC
AL131 KmAdh7 F	ttttctctttttacagatcaCCgcGGATGTTTCGTAAGGTCACATCTG
AL132 KmAdh7 Myc R	TTTTGTTCacctccgcctagggatccgcctccAAAGTTAATAATAAGTTTCATAGCCTTT
AL133 KmAtf F	ttttctctttttacagatcaCCgcGGATGATGAAACTACAGAAATTGTCG
AL134 KmAtf Myc R	GCTTTTGTTCacctccgcctagggatccgcctccCAATGTAGTCAAGTTGTTTTCAAA

TCTTTGAAAAGATAATGTATGATTATGCTTTCACTCATATTTATACAGAAACTTGATGTTTTCTTTCGAGTATATACAAGGTGATTACATGTACGTTTGAAGTACAACTCTAGATTTTGTAGTGCCCTCTTGGGCTAGCGGTAAAGGTGCGCATTTTTTCACACCCTACAATGTTCTGTTCAAAAGATTTTGGTCAAACGCTGTAGAAGTGAAAGTTGGTGCGCATGTTTCGGCGTTCGAAACTTCTCCGCAGTGAAAGATAAATGATCAAAGCATTGATCTGCTTTGCTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAAGAAAGATAAATGATCCTTGAAAAAGTGGCACCGAGTCGGTGGTGCTTTTTTTGTTTTTTATGTCT

Figure S4.7: gblock SNR52-Ade2-sgRNA. Gblock containing the ScSNR52 promoter (grey shaded), a 20bp Ade2 target sequence (red) and the structural guide RNA for Cas9 recruitment.

CTTGCTCATTAGAAAGAAAGCATAGCAATCTAATCTAAGTTTTCTAGAACTAGTGGATCCCCCGGGaaaaATGGACAAGAAGTAC
TCTATCGGTTTGGACATCGGTACCAACTCTGTTGGTTGGGCTGTTATCACCGACGAATACAAGGTTCCATCTAAGAAGTTCAAG
GTTTTGGGTAACACCGACAGACACTCTATtAAGAAGAAtTTGATCGGTGCTTTGTTGTTCGACTCTGGTGAgACCGCTGAAGCTA
CCAGATTGAAGAGAACCGCTcGtAGAAGATACACCAGAAGAAAGAACAGAATCTGTTAtTTGCAAGAAATtTTCTCTAACGAgAT
GGCTAAGGTcGACGAtTCTTTCTTtCACAGATTGGAAGAATCTTTCTTGGTcGAAGAgGAtAAGAAaCACGAAAGACACCCAATCT
TCGGTAACATtGTTGACGAAGTTGCTTACCACGAAAAGTACCCAACCATCTACCAtTTGAGAAAaAAGTTGGTTGAtTCcACCGAC
AAGGCTGACTTGAGATTGATCTACTTGGCcTTGGCcCACATGATCAAGTTCAGAGGTCACTTCTTGATtGAAGGTGACTTGAACC
CAGACAAtTCTGACGTTGACAAGTTGTTCATCCAATTGGTTCAAACCTACAACCAATTGTTCGAgGAAAAtCCAATCAACGCTTCc
GGTGTcGACGCcAAaGCcATtTTGTCcGCtAGATTaTCTAAaTCTcGtAGATTaGAAAACTTGATtGCcCAATTGCCAGGTGAgAAGA
AGAAtGGTTTGTTCGGTAACTTGATCGCcTTGTCcTTGGGTTTaACCCCAAAtTTCAAGTCTAAtTTCGACTTGGCTGAgGAtGCTA
AGTTaCAATTGTCcAAGGACACCTACGACGAtGACTTGGAtAACTTaTTGGCTCAAATtGGTGAtCAATACGCcGACTTaTTCTTGG
CTGCcAAGAAtTTGTCTGAtGCcATCTTGTTGTCcGAtATCTTGAGAGTTAACACCGAAATCACCAAGGCTCCATTGTCcGCTTCTA
TGATCAAGAGATACGAtGAACAtCAtCAAGAtTTGACCTTaTTGAAaGCTTTGGTTAGACAACAgTTGCCAGAAAAtTACAAGGAA
ATtTTCTTtGAtCAATCTAAGAACGGTTAtGCTGGTTACATCGAtGGTGGTGCgTCTCAAGAgGAATTCTACAAGTTCATtAAGCCA
ATCTTGGAgAAGATGGACGGTACCGAAGAATTGTTGGTTAAGTTGAACAGAGAAGAtTTGTTGAGAAAGCAAAGAACCTTCGAC
AACGGTTCTATCCCACACCAAATCCACTTGGGTGAATTGCACGCTATCTTGAGAAGACAAGAAGAtTTCTACCCATTCTTGAAGG
AtAACAGAGAgAAGATCGAAAAGATtTTGACCTTCAGAATCCCATACTACGTTGGTCCATTGGCTAGAGGTAACTCTAGATTCGC
TTGGATGACCAGAAAGTCTGAAGAAACCATCACCCCATGGAACTTCGAAGAAGTcGTTGACAAGGGTGCTTCTGCTCAATCTTTC
ATCGAAAGAATGACCAACTTCGACAAGAACTTGCCAAACGAAAAGGTTTTGCCAAAGCACTCTTTGTTGTACGAATACTTCACC
GTTTACAACGAATTGACCAAGGTTAAGTACGTTACCGAAGGTATGAGAAAGCCAGCTTTCTTGTCTGGTGAACAAAAGAAGGCT
ATCGTTGACTTGTTGTTCAAGACtAACAGAAAGGTTACCGTcAAGCAgTTGAAGGAAGAtTACTTCAAGAAGATCGAATGTTTCG
ACTCTGTcGAAATtTCTGGTGTcGAAGACAGATTCAAtGCTTCTTTGGGTACtTACCACGAtTTGTTGAAaATCATtAAGGAtAAGG
AtTTCTTaGAtAACGAgGAgAACGAAGAtATtTTGGAgGAtATtGTcTTGACtTTGACtTTGTTtGAgGAtAGAGAAATGATCGAgGAA
AGATTGAAGACtTACGCTCAtTTGTTCGAtGAtAAGGTcATGAAGCAATTGAAGAGAAGAAGATACACtGGTTGGGGTAGATTGTC
cAGAAAGTTGATCAACGGTATtAGAGAtAAGCAATCcGGTAAGACCATtTTGGAtTTCTTGAAGTCTGACGGTTTCGCTAAtAGAA
ACTTtATGCAATTGATtCACGACGAtTCTTTGACtTTtAAGGAAGAtATtCAAAAGGCTCAAGTcTCTGGTCAgGGTGAtTCTTTGCA
CGAACACATCGCTAACTTGGCTGGTTCTCCAGCTATCAAaAAaGGTATtTTGCAAACtGTTAAGGTTGTTGACGAATTGGTTAAG
GTTATGGGTAGACACAAGCCAGAgAAtATtGTcATCGAgATGGCTAGAGAAAAtCAAACCACtCAAAAGGGTCAAAAGAACTCcAG
AGAgAGAATGAAaAGAATtGAgGAAGGTATCAAGGAgTTGGGTTCcCAAATtTTGAAGGAACACCCAGTcGAgAACACtCAATTGC
AAAAtGAAAAGTTGTACTTGTAtTACTTGCAAAACGGTAGAGAtATGTACGTTGACCAAGAATTGGACATtAACAGATTGTCTGA
CTACGAtGTTGAtCACATCGTTCCACAATCcTTtTTGAAGGAtGACTCTATCGAtAACAAGGTTTTGACtAGATCcGACAAGAAtAGA
GGTAAGTCTGAtAACGTTCCATCTGAAGAAGTcGTcAAGAAGATGAAGAACTACTGGAGACAgTTGTTGAAtGCTAAGTTGATtAC
tCAAAGAAAGTTCGAtAACTTGACtAAGGCTGAgAGAGGTGGTTTGTCTGAgTTGGAtAAGGCcGGTTTCATtAAGAGACAATTGG
TcGAAACCAGACAAATCACCAAGCACGTTGCTCAAATCTTGGACTCTAGAATGAACACCAAGTACGACGAAAACGACAAaTTGAT
tAGAGAgGTTAAGGTcATCACCTTGAAGTCcAAaTTGGTTTCcGACTTtAGAAAGGAtTTCCAATTCTACAAGGTcAGAGAAATcAA
tAACTACCACCAtGCTCACGACGCcTACTTGAACGCTGTTGTcGGTACtGCTTTGATcAAGAAGTACCCAAAGTTGGAgTCTGAATT
tGTcTACGGTGAtTACAAaGTTTACGACGTTAGAAAGATGATtGCTAAGTCTGAgCAAGAAATCGGTAAGGCTACtGCTAAaTACT
TCTTCTACTCcAACATtATGAAtTTCTTCAAGACCGAgATtACCTTGGCcAACGGTGAAATCAGAAAGAGACCATTGATtGAAACtA
ACGGTGAAACCGGTGAgATtGTcTGGGACAAGGGTAGAGAtTTCGCTACtGTTAGAAAGGTcTTGTCTATGCCACAAGTcAACATtG
TcAAGAAGACCGAAGTcCAgACtGGTGGTTTCTCcAAGGAATCTATtTTGCCAAAGAGAAACTCTGAtAAGTTGATtGCcAGAAAGA
AaGACTGGGAtCCAAAGAAGTACGGTGGTTTCGACTCTCCAACCGTTGCTTACTCcGTTTTGGTTGTcGCTAAGGTTGAAAAGGG
TAAGTCcAAGAAGTTGAAGTCcGTcAAGGAgTTGTTGGGTATtACCATCATGGAAAGATCTTCcTTCGAAAAGAACCCAATtGAtT
TCTTGGAgGCTAAGGGTTACAAGGAAGTcAAGAAGGACTTGATCATtAAGTTGCCAAAGTACTCTTTGTTCGAATTGGAgAACGG
TAGAAAGAGAATGTTGGCTTCcGCTGGTGAATTGCAAAAGGGTAACGAATTGGCTTTGCCATCTAAGTACGTTAACTTCTTGTA
CTTGGCTTCcCACTACGAAAAGTTGAAGGGTTCTCCAGAAGAtAACGAACAAAAGCAATTGTTCGTTGAACAACAtAAGCACTAC
TTGGACGAAATCATCGAACAAATCTCTGAATTCTCTAAGAGAGTTATCTTGGCTGACGCTAACTTGGACAAGGTTTTGTCTGCT
TACAACAAGCACAGAGACAAGCCAATCAGAGAgCAAGCTGAAAACATCATCCACTTGTTCACtTTGACCAACTTGGGTGCTCCAG
CTGCTTTCAAGTACTTCGACACCACCATCGACAGAAAGAGATACACCTCTACCAAGGAAGTTTTGGACGCTACCTTGATCCACCA
ATCTATCACCGGTTTGTACGAAACCAGAATCGACTTGTCTCAATTGGGTGGTGACTCTAGAGCTGACCCAAAGAAGAAGAGAAA
GGTTTGATCTCTTCTCGAGTCATGTAATTAGTTATGTCACGCTTACATTC

Figure S4.8: Codon optimized Cas9 with overlap sequences to plasmid backbone. The for *K. marxianus* codon optimized version of SpCas9 was manually changed to avoid highly repetitive sequences to allow for production of the gBlocks. The gray sequences represent overlaps to the plasmid backbone.

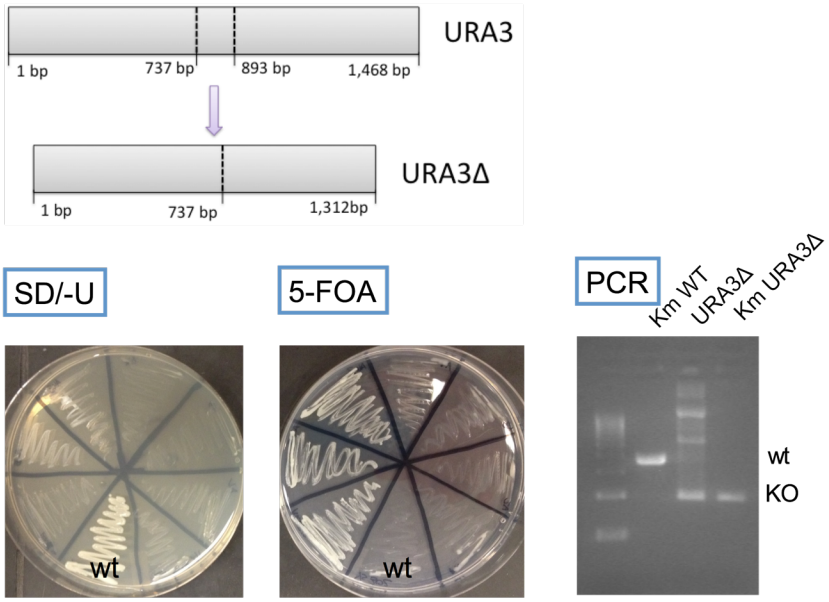


Figure S4.9: Schematic of URA3 knockout and verification by growth and PCR. The URA3 knockout was achieved by homologous recombination of a truncated URA3 fragment into the native gene. This leads to a deletion of about 160 bp of the coding region of URA3. Selective pressure for homologous recombination is applied by growing the colonies transformed with the truncated URA3 on agar plates containing 5-fluoroorotic acid (5-FOA). Ability to grow on 5-FOA and inability of growth on SD/-U media indicated URA3 knockout. The knockout was confirmed using primers that bind 100 bp upstream and downstream of the region to be excised so that there is a band of 200 bp for a knockout and 360 bp for the wild type The lanes of the gel show a 100 bp ladder, the wild type URA3 amplicon, as well as the URA3Δ fragment and the knocked out URA3 gene.
>gi|313024|emb|Z21934.1| K.marxianus URA3 gene

GAGCATCTTGGTCTTCTGAGCTCATTATACCTCAATCAAAACTGAAATTAGGTGCCTGTCACGGCTCTTTTTTTACTGT
ACCTGTGACTTCCTTTCTTATTTCCAAGGATGCTCATCACAATACGCTTCTAGATCTATTATGCATTATAATTAATAGTT
GTAGCTACAAAAGGTAAAAGAAAGTCCGGGGCAGGCAACAATAGAAATCGGCAAAAAAAAACTACAGAAATACTAAGAGC
TTCTTCCCATTCAGTCATCGCATTTCGAAACAAGAGGGGAATGGCTCTGGCTAGGGAACTAACCACCATCGCCTGACTCT
ATGCACTAACCACGTGACTACATATATGTGATCGTTTTAACATTTTCAAAGGCTGTGTGTCTGGCTGTTTCCATTAATTT
TCACTGATTAAGCAGTCATATTGAATCTGAGCTCATCACCAACAAGAAATACTACCGTAAAAGTGTAAAAGTTCGTTTA
AATCATTTGTAAACTGGAACAGCAAGAGGAAGTATCATCAGCTAGCCCCATAAACTAATCAAAGGAGGATGTCGACTAA
GAGTTACTCGGAAAGAGCAGCTGCTCATAGAAGTCCAGTTGCTGCCAAGCTTTTAAACTTGATGGAAGAGAAGAAGTCA
AACTTATGTGCTTCTCTTGATGTTCGTAAAACAGCAGAGTTGTTAAGATTAGTTGAGGTTTTGGGTCCATA**TATCTGTC**
TATTGAAGACACATGTAGATATCTTGGAGGATTTCAGCTTTGAGAATACCATTGTGCCGTTGAAGCAATTAGCAGAGA
AACACAAGTTTTTGATATTTGAAGACAGGAAGTTTGCCGACATTGGGAACACTGTTAAATTACAATACACGTCTGGTGT
ATACCGTATCGCCGAATGGTCTGATATC**ACCAATGCACACGGTGTGACTGGTGCGG**GCATTGTTGCTGGTTTGAAGCA
AGGTGCCGAGGAAGTTACGAAAGAACCTAGAGGGTTGTTAATGCTTG**CCG**AGTTATCGTCCAAGGGGTCTCTAGCGCAC
GGTGAATACACTCGTGGGACCGTGGAAATTGCCAAGAGTGATAAGGACTTTGTTATTGGATTTATTGCTCAAAACGATA
TGGGTGGAAGAGAAGAGGGCTACGATTGGTTGATCATGACGCCAGGTGTTGGTCTTGATGACAAAGGTGATGCTTTGGG
ACAACAATACAGAACTGTGGATGAAGTTGTTGCCGGTGGATCAGACATCATTATTGTTGGTAGAGGTCTTTTCGCAAAG
GGAAGAGATCCTGTAGTGGAAGGTGAGAGATACAGAAAGGCGGGATGGGACGCTTACTTGAAGAGAGTAGGCAGATCCG
CTTAAGAGTTCTCCGAGAACATGCAGAGGTTCGAGTGTACTCGGATCAGAAGTTACAAGTTGATCGTTTATATATAAAC
TATACAGAGATGTTAGAGTGTAATG**GCATTGCGTCACATTGTATAC**

Figure S4.10: URA3 knockout fragment Nucleotide sequence of the *K. marxianus* Ura3 gene. The open reading frame of the URA3 protein is shown in blue. The gray shaded area shows the 168bp that are being deleted after overlap PCR. Primers binding sites to create the overlap fragments are shown in bold. Screening primers P1072/1073 are displayed in Table S5.4

Table S4.5: Alcohol dehydrogenases and Alcohol-O-acetyltransferases analyzed for homology to *K. marxianus* proteins

Protein	Uniprot Reference
ScAdh1	P00330
ScAdh2	P00331
ScAdh3	P07246
ScAdh4	P10127
ScAdh5	P38113
ScAdh6	Q04894
ScAdh7	P25377
Cupriavidus necator adh	P14940
Snodgrassella alvi adh	WP_025331133
Acinetobacter equi adh	WP_054580671
ScAtf1	P40353
ScAtf2	P53296
KlAtf	Q6CJX7
KlAdh1	P20369
KlAdh2	P49383
KlAdh3	P49384
KlAdh4	P49385

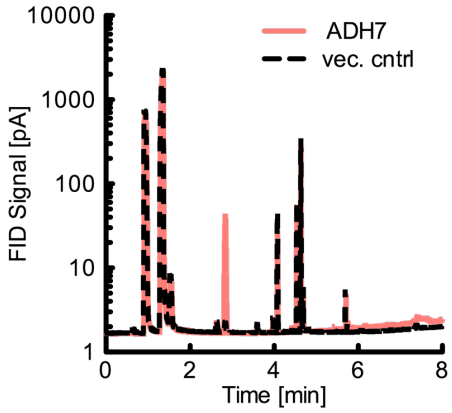


Figure S4.11: Comparison of GC chromatograms of the hemiacetal reaction. The GC samples for *Km*Adh7 overexpression and vector control are shown to confirm the absence of an ethyl acetate peak when the hemiacetal mixture is incubated with protein lacking an alcohol dehydrogenase that is active towards hemiacetal.

Supporting references

[1] Lin JL, Wheeldon I: **Dual N- and C-Terminal Helices Are Required for Endoplasmic Reticulum and Lipid Droplet Association of Alcohol Acetyltransferases in Saccharomyces cerevisiae**. *PloS one* 2014, **9**(8).

[2] Zhu J, Lin JL, Palomec L, Wheeldon I: **Microbial host selection affects intracellular localization and activity of alcohol-O-acetyltransferase**. *Microbial cell factories* 2015, **14**.

[3] DiCarlo JE, Norville JE, Mali P, Rios X, Aach J, Church GM: **Genome engineering in Saccharomyces cerevisiae using CRISPR-Cas systems**. *Nucleic acids research* 2013, **41**(7):4336-4343.

[4] Lee KS, Kim JS, Heo P, Yang TJ, Sung YJ, Cheon Y, Koo HM, Yu BJ, Seo JH, Jin YS *et al*: **Characterization of Saccharomyces cerevisiae promoters for heterologous gene expression in Kluyveromyces marxianus**. *Applied microbiology and biotechnology* 2013, **97**(5):2029-2041.

Chapter 5: Enhancing the high native capacity of the thermotolerant *Kluyveromyces marxianus* to produce ethyl acetate

5.1 Abstract

Recent advances in synthetic biology have enabled the development of non-host organisms for chemicals biosynthesis. These new tools have enabled the ability to manipulate desired traits in non-model organisms. The yeast *Kluyveromyces marxianus* is a promising candidate for chemicals biosynthesis. Its natural capacity to produce short and medium chain volatile esters at high rates, along with rapid growth kinetics at temperatures upwards of 45 °C make it especially interesting as platform for biotechnological ester production. The recent development of an efficient CRISPR-Cas9 genome editing tool in *K. marxianus* has allowed for genomic disruptions and integrations and was applied to elucidate and improve the ethyl acetate production pathway in *K. marxianus*. Disruptions of *K. marxianus* alcohol acetyl/acyltransferases revealed that in contrast to prior claims, Atf only marginally contribute to bulk ethyl acetate formation and that the newly found Eat1 is responsible for the formation of bulk ethyl acetate in *K. marxianus*. Eat1 was found to be localized to the mitochondria, allowing for high ethyl acetate production through utilization of the mitochondrial acetyl-CoA pool that is fostered by rapid growth kinetics. Overexpression of Eat1 increased ester production significantly suggesting that this reaction step is a bottleneck to high flux. To further increase ethyl acetate production in *K. marxianus* we developed a CRISPR interference (CRISPRi) system to knock down electron transport chain and/or TCA cycle enzyme expression to divert carbon flux from TCA cycle towards increase ethyl acetate production.

5.2 Background

A rising interest in bioprocessing along with the advent of sophisticated synthetic biology tools have enabled the development of non-model organisms for chemicals biosynthesis. One approach to build efficient microbial production systems is the choice of production organism based on a desired phenotype. The non-model yeast *Kluyveromyces marxianus* has gained increasing interest based on multiple characteristics. Of particular interest are its capability to produce high amounts of ethyl acetate from various carbon sources as well as its multi-stress tolerance and fast growth characteristics.[1-4] Wild type strains have been known to produce upwards of 2 g $L^{-1} h^{-1}$ in aerated bioreactors, but thus far ester biosynthesis pathways have not been well understood.[5,6, 7]

Ester production pathways in *S. cerevisiae* have been studied and ethyl acetate biosynthesis has been attributed to alcohol acetyltransferase Atf1 and 2 activities, where Atf1 is the main contributor to ethyl acetate formation. Atf1 and 2, as well as other acyltransferases are located to the endoplasmic reticulum (ER) and lipid droplets and localization has been shown to be essential for activity.[8] Double knockouts of Atf1 and 2 lead to a reduction of ethyl acetate by 50% hinting to the presence of alternative production routes.[9] Recent work led to the discovery of a new alcohol acetyltransferase named ethanol acetyltransferase Eat1 in the yeast *Wickerhamomyces anomalus*.[10] Expression of this Eat1 along with homologs from other yeast species led to significant ethyl acetate production, and disruption of Eat1 in *S. cerevisiae* and *K. lactis* lead to a decrease in ethyl acetate by 50 and 80%, respectively.[10]

Studies on ester biosynthesis in *K. marxianus* are limited and while prior research suggests the importance of Atf in ester biosynthesis, we previously showed that Atf only marginally contributes to ester production in *K. marxianus*.[6, 7] Alternatively, Eat1 from *K.*

marxianus showed the highest ethyl acetate production among the tested newly discovered Eat1's when expressed in *S. cerevisiae*, thus presenting itself as a possible candidate to contribute to ethyl acetate production in *K. marxianus*.[10]

Eukaryotic metabolite production is oftentimes compartmentalized to different organelles within the cell. To be able to maximize production yields, research has been focused on co-localizing whole pathways. Significant increases have been achieved, for example by co-localizing the isobutanol pathway to the mitochondria.[11] In another example Lin *et al*. achieved a 1.7-fold increase in ethyl acetate production in *S. cerevisiae* by co-localizing ALD6 and ACS1 to the lipid droplets.[12] While ester production in *K. marxianus* has been thought to take place in the cytosol no studies have been conducted that confirm this hypothesis and further studies are needed to guide metabolic engineering approaches.

K. marxianus has been know to produce great amounts of esters, especially under trace metal limitation. The positive impact of trace metal limitation on ethyl acetate production has been thought to be caused by a decrease of flux through the tricarboxylic acid cycle (TCA cycle) caused by a lack of iron or copper that are essential for TCA cycle and electron transport chain (ETC) enzymes.[7] The decreased flux was suggested to free up acetyl-coA for ester synthesis. [7] This effect has been observed in other yeasts like *Candida utilis* and *C. pseudotropicalis*.[13, 14]

The type II bacterial CRISPR system has been applied in a myriad of yeast to allow for efficient genome editing.[6, 15, 16] Two point mutations in the Cas9 nuclease lead to a catalytically dead mutant (dCas9) that, when targeted to the promoter region of a gene, can act as a transcriptional repressor.[17] Targeting dCas9 to the transcription start site (TSS) or within 200bp upstream has shown to be most effective.[17] Development of a CRISPR

interference system will allow us to survey the impact of decreased TCA cycle and ETC gene activity on ethyl acetate.

In this chapter we elucidated the function of different acyltransferases on acetate ester production and found that Eat1 is critical for ethyl acetate biosynthesis, as well as the production of other acetate esters. To efficiently engineer ester production in *K. marxianus* we analyzed the localization of Eat1 and found that Eat1 and thus ester production is localized to the mitochondria and that localization of Eat1 is essential for high ethyl acetate production yields. Overexpression of Eat1 was able to increase ethyl acetate biosynthesis by 1.8–fold from 126.5 to 223.0 mg L^{-1} OD^{-1}. CRISPRi mediated knockdowns were used to shuttle acetyl-coA towards ethyl acetate production, thus increasing ethyl acetate production by 3.8–fold compared to the scrambled sgRNA controls.

5.3 Methods and Materials

5.3.1 Strains and culturing conditions

All strains were purchased from ATCC or DSMZ (Deutsche Sammlung von Microorganismen und Zellkulturen). All materials were purchased from Fisher Scientific unless noted otherwise. All yeast strains used in this study are listed in Table S5.1.

K. marxianus strain CBS 6556 Δura3 (YS402) was used as a wild type strain in this work.[6] *S. cerevisiae* strain BY4742 (YS5) with Ds-Red tagged OM45 was used to study mitochondrial localization.

To create knockout strains of *K. marxianus,* cells harboring a CRISPR-Cas9 plasmid were grown on (SD-U) containing 6.7 g L^{-1} yeast nitrogen base without amino acids DB Difco®; (Becton-Dickinson), 1.92 g L^{-1} yeast synthetic drop-out medium supplements without uracil (Sigma Aldrich), and 20 g L^{-1} glucose or SD-U plates containing 15 g L^{-1} agar.

To remove the CRISPR-Cas9 plasmid, cells were grown in YPD medium (5 g L^{-1} Yeast extract, 10 g L^{-1} peptone with 20 g L^{-1} glucose; DB Difco®, Becton-Dickinson) overnight.

K. marxianus CRISPRi cultures were grown in 50 mL SD-U media in 250 mL baffled (0.2 initial OD) shake flasks at 250rpm and 37°C for 14h. Initial and final optical cell densities (OD) were measured using the Nanodrop 2000c UV-Vis spectrometer (Thermo Scientific) at 600 nm.

K. marxianus disruption and Eat1 integration strains (see Table S5.1) were cultured in YPD media (5 g L^{-1} Yeast extract, 10 g L^{-1} peptone with 20 g L^{-1} glucose; DB Difco®, Becton-Dickinson) at 37 °C. Overnight cultures were inoculated into 25 mL of media in 250 mL

5.3.2 Molecular cloning and plasmids construction

All cloning was accomplished using Q5 polymerase, restriction endonucleases and Hifi assembly master mix purchased from New England BioLabs (NEB). DNA oligos were purchased from Integrated DNA Technologies (IDT). Chemically competent DH5α *Escherichia coli* was used for plasmid propagation. Following transformation, *E. coli* cells were grown in LB medium containing 100 mg L^{-1} ampicillin. All plasmids and primers used are listed in Table S5.2 and S5.3.

To construct a version of the Cas9 nuclease lacking nuclease activity two point mutations were introduced by digestion of pIW601 with XmaI/AvrII and assembly with two dCas9 fragments containing the two mutations necessary (Figure S5.1) that were amplified from pIW601 with P1853/54 and P1855/56, respectively.

The CRISPR-Cas9 and –dCas9 plasmids were constructed using pIW601 (Addgene ID 98907) and pIW602, respectively. For integration of the 20bp target sequence, pIW601

or pIW602 were digested with PspXI and assembled with a 60bp cassette containing 20 upstream and downstream homology as well as the target sequence. The cassette was constructed by annealing two complementary 60bp primers as described before.[18] Target sequences were checked for secondary structures using the IDT OligoAnalyzer Tool 3.1 (https://www.idtdna.com/calc/analyzer) and uniqueness within the *K. marxianus* genome using BLAST. All sgRNA sequences and primers are listed in Tables S5.3 and S5.4

S. cerevisiae Eat expression plasmids were constructed by amplifying Eat1 and the truncations with AL276, AL280-283 and AL277 (Table S5.3) and cloned into pRS426 vector containing a PGK1 expression cassette (pIW21). The resulting plasmids are listed in Table S5.2.

For genomic integration into the URA3 locus a homology donor plasmid was constructed. The marker of pIW272 was exchanged for *S. cerevisiae* His3 by assembling the amplified backbone and the His3 marker, amplified from plasmid pIW8. The resulting plasmid pIW578 was digested to insert 1kb upstream and downstream homology donors, respectively. Expression cassettes for genomic integration (Ampliefied from genomic DNA using AL47/116 AL49/117) were inserted into the cut pIW578 to generate PIW634 (Table S5.2 and S5.3)

To clone CRISPRi plasmids expressing sgRNAs targeting multiple genes, plasmids targeting a single gene were digested with AvrII or NsiI. Cassettes expressing sgRNAs additional were amplified via PCR (for example, with Cr_1539 and Cr_1540) and cloned into the digested backbone using Gibson Assembly as has been previously described.[19]

5.3.3 Transformation of *K. marxianus* and *S. cerevisiae*

Plasmid and linear DNA transformation were performed as described previously.[6]

K. marxianus cells were grown to stationary phase, washed with sterile water, and were suspended in 100 ug carrier DNA (salmon sperm DNA) and 0.2-1 ug of plasmid or linear DNA. 400 mL of transformation mix (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 and the solution was incubated at room temperature for 15 min and subsequently heat shocked at 47°C. The transformed cells were plated on solid SD-U agar plates. *S. cerevisiae* cells were transformed using the Zymo Frozen-EZ Yeast Transformation II Kit and plated on solid SD-U agar plates.

5.3.5 Strain construction

For gene disruptions, transformed *K. marxianus* cells were plated on SD-U plates and grown for 2 days at 30 °C. After 2 days of growth colonies were screened by amplifying the CRISPR-Cas9-edited region in the genome by colony PCR and subsequent sequencing of the purified PCR fragments.

The native OM45 gene of *S. cerevisiae* BY4742 (YS5) was fused to DsRed by homologous recombination to create a mitochondrial marker strain (YS578)

For genomic integrations, *K. marxianus* CBS6556 Δura3 was transformed with the CRISPR plasmids (pIW538) and a homology donor plasmid and plated on SD-U plates. After 2 days of growth at 30°C, colonies were picked and restreaked on YPD plates. The resulting colonies were subjected to colony PCR using a three primer approach as described before.[20] In brief, 2 primers binding outside of the homology region (P1928 /29) and the promoter (AL275) were used for PCR, where a 2kb band represents no integration through homologous recombination and a 1.2kb band corresponds to genomic integration of the expression cassette at the URA3 site.

Positively confirmed disruptions or integrations were grown over night in YPD liquid cultures for plasmid removal and saved at -80 °C.

5.3.6 Eat1 protein sequence analysis

The Eat1 protein sequenced was analyzed for secondary structures using the CFSSP software. Truncations were made when secondary structure changed at residue 6, 14, 18 and 35.[21] Mitochondrial localization was assessed using the MitoFates software.[22]

5.3.7 Headspace gas chromatography

Volatile metabolite concentration was measured using an Agilent 7890A system equipped with a Phenomenex ZB-624 column and an FID detector. For metabolite separation, the temperature was held at 40 °C for 2 min, then increased 20 °C min^{-1} to 70 °C and 50 °C min^{-1} to 220 °C and held for 2 min.

5.3.8. Reverse transcription quantitative PCR (RT-qPCR)

Total RNA was extracted as described before.[6] In short, RNA was extracted from 5 OD of cells using the YeaStar™ RNA Kit (Zymo Research). Following DNAse treatment (DNAse I, NewEngland Biolabs), RNA was purified using the RNA Clean & Concentrator™-5 Kit (Zymo Research). RNA was subjected to reverses transcription reaction (iScript™ Reverse Transcription Supermix for RT-qPCR, Bio-Rad) and cDNA was used for SYBR Green qPCR (SsoAdvanced™ Universal SYBR® Green Supermix. Bio-Rad) using the Bio-Rad CFX Connect™. Primers used for the qPCR reaction are listed in Table S5.3. Fold changes were calculated using the Pfaffl method under consideration of the reaction efficiency as previously described and normalized to Act1 expression.[23]

5.3.8 MitoTracker staining

For Eat1 localization studies *K. marxianus* cultures with an integrated copy of Eat1-GFP and its truncations (see Table S5.1) were grown over night in YPD media and transferred to fresh media to OD 0.1. At different time points, 500ul of cells were harvested, resuspended in fresh media with 100nM MitoTracker® Red CMXRos (Cell Singaling Technology) and incubated at 37°C for 15 mins. Subsequently, the cells were washed twice and resuspended in media and imaged immediately.

5.3.9 Fluorescence microscopy

S. cerevisiae and stained *K. marxianus* cells (see 5.3.8) were imaged with an Olympus BX51 microscope (UPlanFL 100X 1.30 oil-immersion objective lens, mercury lamp) and fluorescent images were captured by a Q-Imaging Retiga Exi CCD camera. Images were processed using the CellSens Dimension 1.7 software (Olympus).

5.3.10 Western blot analysis

Western blot analysis was done to quantify protein translation of the Eat1-GFP fusion protein and its truncations in *K. marxianus*. 2.5 OD of cells were lysed as described earlier and loaded onto a 10-well Any kD™ Mini-PROTEAN® TGX™ Precast Protein Gel (BioRad) and run for 1h at 150 V.[6] Samples were then electrophoretically transferred overnight to a PVDF membrane at 25 V. Membranes were blocked with 5% non-fat milk in TBST buffer (Tris-buffered saline with Tween20) for 1h at room temperature and incubated with rabbit anti-GFP antibody (NB600-303, Novus Biologicals) or mouse anti-GAPDH (PA1-987) diluted to 1:2000 and 1:5000 in TBST buffer with 1% non-fat milk. Goat anti-rabbit (65-6120, Invitrogen) and anti-mouse IgG-HRP (31430) diluted to 1:10000 were added as

secondary antibodies and incubated at room temperature for 30 min. After washing with TBST, the Immobilon™ Western Chemiluminescent HRP substrate (Millipore) was used for signal detection. Blots were imaged and protein levels were quantified using the BioRad ChemiDoc™ MP System with the Image Lab software.

5.3.11 Statistical analysis

Data points represent arithmetic means of at least triplicate biological samples, and error bars represent the standard deviation. Groups of samples were analyzed by one-way ANOVA with a Tukey post-hoc test and considered significant at $p<0.05$. Statistical analysis and plotting of data points was performed using the GraphPad Prism software.

5.4 Results

5.4.1 Eat1 disruption abolishes acetate ester biosynthesis

K. marxianus has been known to produce great amounts of ethyl acetate, however the production pathway remains to be elucidated. Previously performed cell lysate activity studies suggest strong activity of one or more alcohol dehydrogenases (Adh) and limited activity of one or more alcohol acetyltransferasese (AATases) towards ethyl acetate production.[6] Although disruption and overexpression of all known Adh's and AATases showed activity of KmAdh7 towards ethyl acetate biosynthesis, this enzyme seems to have no importance in *in vivo* ethyl acetate production in *K. marxianus*. Recently a new alcohol acetyltransferase, named ethanol acetyltransferase (Eat1) was identified in *W. anomalus*.[10] Overexpression of the *K. marxianus* homologue of Eat1 showed significant ethyl acetate production when expressed in *S. cerevisiae*. We therefore aimed to elucidate the role of all known AATases in *K. marxianus*. CRISPR-Cas9 mediated disruption of ATF,

EAT1, EHT1 and EEB1 revealed that EAT1 disruption almost completely abolishes ethyl acetate production from 180.08 mg L^{-1} OD^{-1} in the wild type of K. marxianus CBS 6556 to 0.80 ± 0.13 mg L^{-1} OD^{-1} (Figure 5.1). Analogously, loss of Eat1 activity led to accumulation of ethanol and acetaldehyde from 58.46 ± 45.64 and 0.78 ± 0.50 mg L^{-1} OD^{-1} to 205.91 ± 16.03 and 4.50 ± 0.71 mg L^{-1} OD^{-1}, respectively. Similarly, EAT1 disruption significantly decreased production of other acetate esters such as isoamyl acetate and 2-phenylethyl acetate from 2.52 ± 0.93 and 1.71 ± 0.48 mg L^{-1} OD^{-1} in the wild type to from 0.15 ± 0.01 and 0.19 ± 0.06 mg L^{-1} OD^{-1}, respectively (Figure S5.2).

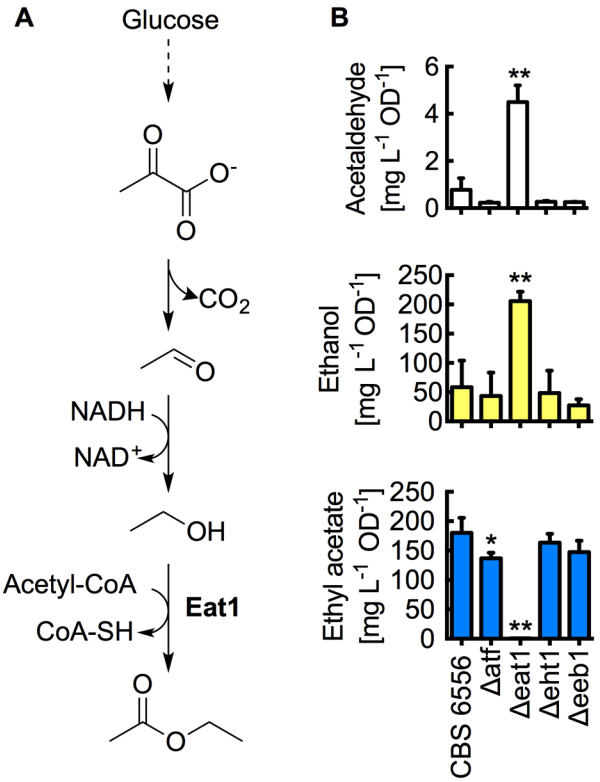


Figure 5.1: Eat1 is responsible for ethyl acetate production in *K. marxianus*. (A) Ethyl acetate biosynthesis pathway in K. marxianus. (B) Strain CBS6556 Δura3 with a series of AATase disruptions were screened for acetaldehyde, ethanol and ethyl acetate production. Cultures were grown in YPD for 10h. Bars represent arithmetic mean and error bars represent standard deviation

5.4.2 Mitochondrial localization of Eat1 is important for ethyl acetate production in *K. marxianus*

Metabolic pathways in eukaryotic are highly organized and compartmentalized. Therefore, to understand and engineer such pathways in yeast a deeper understanding about the pathway and its localization in question is essential in guiding metabolic engineering approaches. Ester production through Atf1 in *S. cerevisiae* have been known to be localized to either ER or lipid droplets.[8] Limited knowledge about ester production in *K. marxianus,* much less localization of the pathway, has kept researchers from studying the localization of ester production in *K. marxianus*. Based on the disruption results we expressed an integrated copy an Eat1-Gfp fusion protein in a Δeat1 strain of *K. marxianus* and found that Eat1 is localized to the mitochondria (Figure 5.2A). Similarly, when expressed in *S. cerevisiae* expressing a mitochondrial marker (OM45-DsRed), the signal of Eat1-Gfp overlaps with the mitochondrial signal. Different N-terminal truncations show that 14 amino acids need to be truncated to partially delocalize Eat1 from the mitochondria and a 19 aa truncation is required for full localization to the cytosol (Figure S5.3). These two truncations were chosen for further study in *K. marxianus*. A time course study revealed that Eat1 localization is independent of growth stage and that the localization pattern of the two truncated Eat1 proteins are similar to what has been observed in *S. cerevisiae* (Figure S5.3). Further examination of the Eat1 sequence predicted a 19aa preprotein sequence and the presence of two cut sites (MPP, Icp55) that are characteristic for mitochondrial matrix proteins (Figure 5.2B).[22, 24] Based on these predictions, it is suspected that the preprotein is cut at the MPP cut site upon transportation to the mitochondria and subsequentially cleaved at the Icp55 site for stabilization. With these predictions in mind, protein expression of the native Eat1 and the three truncations was analyzed on the level of

transcription and translation (Figure 5.2C and D). The results showed that neither mRNA nor protein levels differed significantly in either of the tested Eat1's. To test whether Eat1 localization to the mitochondria is required for ethyl acetate production, *K. marxianus* Δeat1 strains expressing wild type Eat1 and two truncations (Δ1-14Eat1 and Δ1-19 Eat1) were screened for ethyl acetate production. GC analysis revealed that stepwise truncation and thus localization of Eat1 to the cytosol leads to a stepwise decrease in ethyl acetate from 150.5 mg L^{-1} OD^{-1} in when the native Eat1 is overexpressed to 99.6 mg L^{-1} OD^{-1} for Δ1-14Eat1 and 45.1 mg L^{-1} OD^{-1} for Δ1-19 Eat1 (Figure 5.2E). Analogously, ethanol production increased in a similar fashion from 92.3 to 102.8 and 154.4 mg L^{-1} OD^{-1} for Eat1, Δ1-14Eat1 and Δ1-19 Eat1, respectively (Figure S5.4). It is noteworthy that fusion of a C-terminal Gfp decreases ethyl acetate production from all the Eat1 proteins (Figure S5).

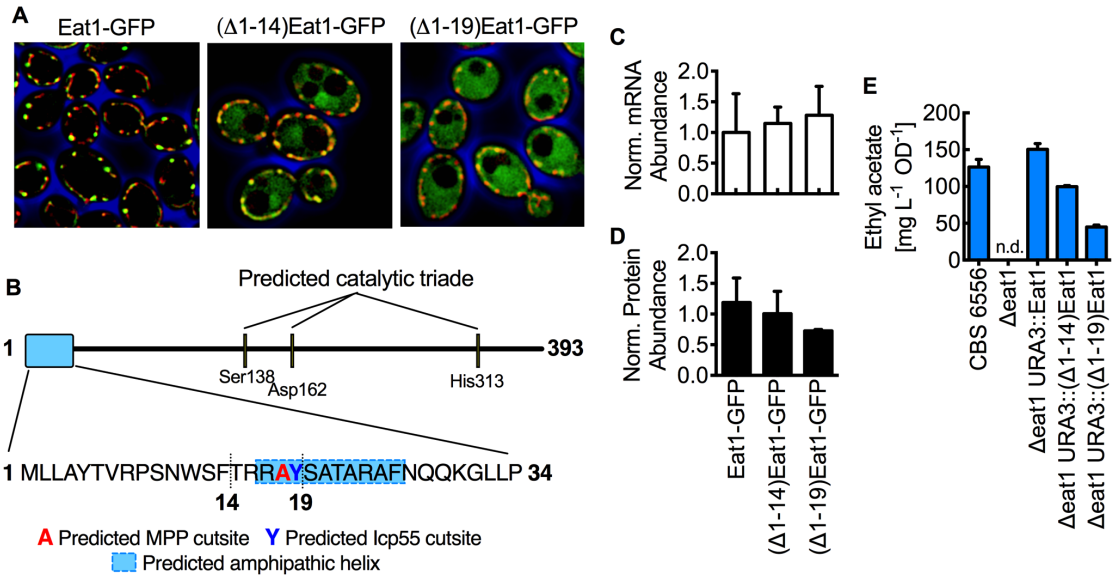


Figure 5.2: Eat1 is localized to the mitochondria and localization is important for ethyl acetate production. *K. marxianus* strain CBS 6556 Δura3 Δeat1 expressing wild type Eat1 and 2 truncations. (A) Fluorescence microscopy of *K. marxianus* expressing Eat1-Gfp variants reveals mitochondrial localization of the native and stepwise cytosolic localization of the truncated proteins. (B) Sequence analysis predicted a preprotein sequence with two cut sites characteristic for mitochondrial matrix proteins. Transcriptional (C) and translational analysis (D) revealed no impact of the truncation on protein expression. Fermentation results (E) indicated negative impact of truncation on ethyl acetate yields. All bars represent arithmetic means of at least n=3 replicates and error bars represent the standard deviation.

5.4.3 Eat1 overexpression increases ethyl acetate production

To increase product formation in an organism that is already producing high amounts of the desired compound, it is crucial to assess bottlenecks and limitations in the pathway. Because the wild type of *K. marxianus* still produces ethanol, we hypothesize that ethanol is not limiting. Given the fast growth of *K. marxianus* and the mitochondrial localization of Eat1 there is a hypothetical excess of acetyl-coA present for ester production. However, Eat1 is competing for acetyl-coA with citrate synthase that converts acetyl-coA to citrate. Therefore, Eat1 overexpression could potentially increase ester production. To determine if Eat1 presents a bottleneck in the production of ethyl acetate we compared ester production between the wild type and different Eat1 overexpressing strains.

The results showed that Eat1 overexpression increases ester production from 126.5 mg L^{-1} OD^{-1} in the wild type strain to 150.6 mg L^{-1} OD^{-1} when Eat1 is overexpressed and the native Eat1 copy is disrupted and 223.0 mg L^{-1} OD^{-1} when Eat1 is overexpressed in addition to the native Eat1 expression (Figure 5.3). This trend suggests Eat1 is the limiting enzyme for ethyl acetate synthesis. The fact that ethyl acetate is not significantly different in the wild type compared to the overexpressing strain with native Eat1 disrupted, indicates a high native expression level of Eat1. The results also indicate that ethanol production is coupled to ethyl acetate biosynthesis, where higher ethyl acetate production, catalyzed by increased Eat1 expression, leads to a decrease in ethanol production (Figure 5.3).

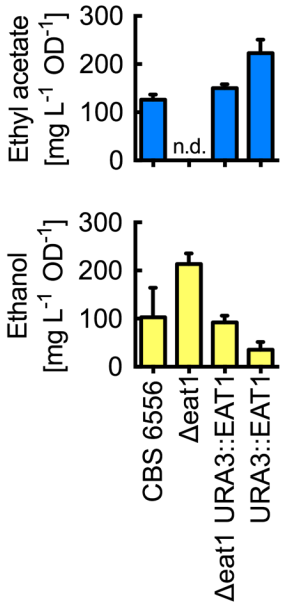


Figure 5.3: Eat1 overexpression significantly decreases ethyl acetate biosynthesis. Comparison of ethyl acetate and ethanol production in wild type, Δeat1 and two EAT1 overexpressing strains. Cultures were grown in YPD for 10h. Bars represent arithmetic mean and error bars represent standard deviation.

5.4.5 CRISPRi- mediated knockdowns enhance ethyl acetate production by decreasing flux through the TCA cycle

K. marxianus has a highly respiratory metabolism, characterized by fast growth and rapid TCA cycle flux. Previous research stated an increase in ethyl acetate production, when iron or copper was deprived from the media.[25] It has been suggested that the lack of metal ions in the media inhibits enzymes containing copper or iron in the TCA cycle or electron transport chain.[7] This hypothesis was confirmed by mimicking the effect of limited metal ions by addition of electron transport chain inhibitors.[7] The study also showed that too much inhibition of the electron transport chain is detrimental for growth and ethyl acetate production. Therefore, we decided to knockdown enzymes of the electron transport chain and TCA cycle to test this hypothesis and increase ethyl acetate production by freeing up mitochondrial acetyl-coA through CRISPRi-mediated knockdowns. Fpur TCA cycle enzymes and 5 electron transport chain enzymes were targeted. The TCA cycle

enzymes targeted are the iron cluster containing aconitase (Aco2b), the flavoprotein subunit of succinate dehydrogenase (Sdh1) and malate dehydrogenase (Mdh1). Similarly, the electron transport chain targets comprise the iron-sulfur containing subunit of succinate dehydrogenase (Sdh2), Rieske iron-sulfur protein (Rip1) and heme-containing Cytochrome c1 (Cyt1) of the mitochondrial cytochrome bc1 complex, and the transcriptional activator (Mss51) and subunit 7 (COX7) of the cytochrome c oxidase (Figure 4A). In *S. cerevisiae,* homologs to all of these targets are essential for respiratory growth.[26]

For maximum knockdown efficiency two sgRNA's were targeted to the TATA box and transcription start site of the gene of interest. Knockdown efficiencies were determined by qRT-PCR and compared to a scrambled DNA control. Results showed that with the exception of CYT1, all CRISPRi targeting resulted in a significant decrease in mRNA transcription (Figure 5.4B). Knockdown efficiencies ranged from a 35.5% knockdown of RIP1 to 81.5% for MDH1. The results also showed that the ACO2b, SDH2, RIP1 and MSS51 knockdowns significantly increased ethyl acetate production from 48.3 mg L^{-1} OD^{-1} to 77.2, 67.1, 73.3 and 68.7 mg L^{-1} OD^{-1}, respectively (Figure 5.4C). Ethanol production remains unchanged (Figure S5.6A).

To investigate the effect each sgRNA has on the knockdown efficiency we tested the knockdown efficiency of the individual sgRNA's for these 4 genes. The results show that ACO2b sgRNA 1 and SDH2 sgRNA 2 performed better than the two sgRNAs combined and that both sgRNAs were required for efficient RIP1 and MSS51 knockdown (Figure S5.7). To determine the best combination of the single ACO2b and SDH2 sgRNAs and the double sgRNAs for RIP1 and MSS51, the different targets were multiplexed (Figure 5.4D). Multiplexing targets SDH2 with MSS51 or RIP1, ACO2b, MSS51 and Rip1, ACO2b, SDH2 and

MSS51 or all 4 targets significantly increased ester production to 94.6, 91.9, 92.9, 120.2 and 161.0 mg L^{-1} OD^{-1}, respectively. Here the best performing strain is the ARSM strain that targets all 4 genes led to a 3.8-fold increase. Only multiplexed knockdown of all 4 targets ARSM impacted ethanol production where a knockdown led to a decrease from 0.54 g L^{-1} OD^{-1} to 0.19 L^{-1} OD^{-1} (Figure S5.6B).

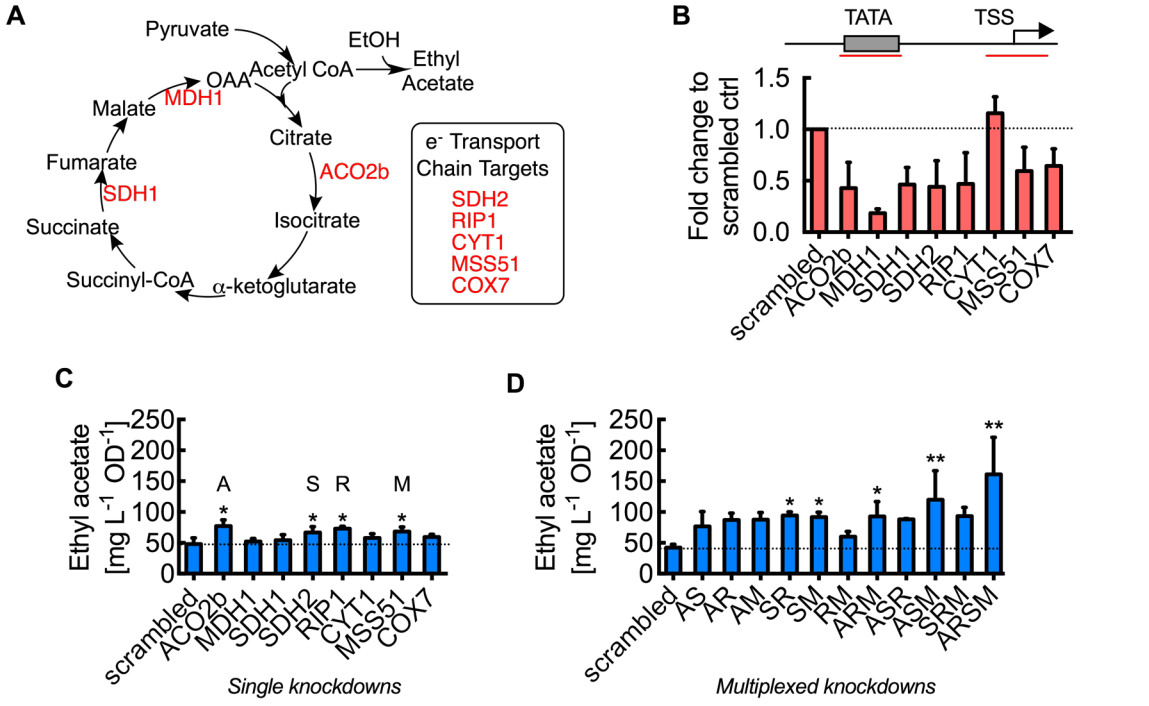


Figure 5.4: CRISPRi mediated knockdown of TCA cycle and electron transport chain induces ethyl acetate production. (A) Different enzymes of the TCA cycle and electron transport chain were chosen as targets for knockdowns using CRISPR interference. (B) qRT-PCR was performed to test if sgRNA's targeted to TATA box and transcription start site (TSS) were able to knockdown expression of 8 target genes. (C) Ethyl acetate production was measured for the strains with the 8 knockdown plasmids. (D) The 4 single knockdowns that increased ethyl acetate production significantly were multiplexed to further increase ester production. Bars represent arithmetic means of at least n=3 replicates and error bars represent the standard deviation.

5.5 Discussion

Recent advances in genome editing, tools such as CRISPR-Cas9, have enabled the development of non-model organisms for biotechnological chemicals production. Here, microorganisms with a specific desirable trait are chosen to build and improve on using metabolic engineering approaches. For example the oleaginous yeast *Y. lipolytica* is chosen as a host organism for production of lipids and acetyl-CoA derivatives because it efficiently shuttles acetyl-CoA from the mitochondria to the cytosol where it can be converted into a variety of products.[2] Similarly, the capability of the yeast *S. stipitis* to efficiently ferment xylose is used for small molecules production.[2]

The non-model yeast *Kluyveromyces marxianus* produces high amounts of the ester ethyl acetate, while side product formation such as ethanol is low.[6] Along with its fast growth kinetics, thermo- and low pH tolerance this yeast presents itself as an ideal candidate for ester biosynthesis. While other yeasts such as *P. anomala* or *C. utilis* are able to produce similar amounts of esters, fermentations are usually much slower than in *K. marxianus*.[27] In addition to this, *K. marxianus*' thermotolerance enables efficient *in situ* product removal. For an economically feasible process high yields and production rates, as well as simple product recovery are desired. Therefore, *K. marxianus* is a most promising candidate host for industrial ethyl acetate production.[27]

Ester production pathways have been studied intensively and ethyl acetate production in *S. cerevisiae* has been attributed to Atf1 and 2 activities.[9] So far, ester production pathways in *K. marxianus* have remained cryptic. Out of the three proposed pathways for ester productions that are catalyzed through alcohol acetyl/acyltransferase (AATase), esterase or alcohol dehydrogenase (Adh), the AATase pathway was suggested to be the likely pathway for ester production in *K. marxianus*.[6, 7] However, disruption of Atf

in *K. marxianus* as the sole known alcohol acetyltransferase only reduced ester production by 15% and disruptions of other acyltransferases did not impact ester production (Figure 5.1). Recently, Kruis *et. al* discovered a new class of AATases called ethanol acetyltransferases (Eat) in *Wickerhamomyces anomalus* (or *Pichia anolama*) and showed high activity of several yeast Eat's including a homolog from *K. marxianus* when expressed in *S. cerevisiae*.[10] Disruption of this homolog in *K. marxianus* decreased ethyl acetate, isoamyl acetate as well as 2-phenylethyl acetate by 99.6, 93.9 and 88.7%, respectively (Figure 5.1 and S5.2). These results suggest that Eat1 is the essential enzyme for bulk acetate ester formation.

Because ester production in *S. cerevisiae* by Atf1 and 2 is localized to the ER or lipid droplets, localization studies of *K. marxianus* Eat1 are essential to guide metabolic engineering approaches. Our studies found that Eat1 is localized to the mitochondria and that mitochondrial localization is essential for high flux ester biosynthesis. The mitochondrial localization of Eat1 allows access to the mitochondrial acetyl-CoA pool that is fostered by fast growth kinetics. Conventional production of acetyl-CoA derivatives in organisms that do not have an efficient pyruvate dehydrogenase bypass or acetyl-coA shuttle system is usually hampered by low cytosolic acetyl-CoA availability. Metabolic engineering for increased ester production has, so far, relied on engineering the pathway in the cytosol or at the lipid droplets in *S. cerevisiae*.[12] However, intensive engineering is usually required to increase acetyl-CoA concentrations for cytosolic ester production.[12, 28] The mitochondrial localization of Eat1 and the nature of its alcoholic substrates and ester products allows for transport across the mitochondrial membranes and thus high ester production. Ethyl acetate in *K. marxianus* has stoichiometrically and experimentally been proven to be dependent on oxygen availability.[27, 29] The mitochondrial localization

of Eat1 may be part of the reason for this observation. It has been reported that in most organisms the pyruvate dehydrogenase complex (that produces mitochondrial acetyl-coA from pyruvate) is only active under aerobic conditions.[30, 31] Therefore, a lack of PDH activity could explain a decrease in ester yields by *K. marxianus* under anaerobic conditions.

Overexpression of Eat1 with a strong promoter in the Δeat1 background does not change ethyl acetate production significantly compared to the wild type, suggesting high native expression of the protein. This is in accordance with expression data obtained in *W. anomalus*.[10] Overexpression of Eat1 on top of the native copy significantly increases ethyl acetate production and reduces ethanol production indicating availability of mitochondrial acetyl-coA and enzyme saturation when only one the native copy is present.

Previously published work suggests increased ethyl acetate production under iron or copper limited conditions.[3, 7] The electron transport chain complexes, as well as aconitase and succinate dehydrogenase in the TCA cycle, depend on the presence of metal clusters for electron transport. We mimicked the effect by selectively knocking down different ETC and TCA cycle enzymes. CRISPR-Cas9 mediated knockdowns were able to significantly increase ester production (Figure 5.4C and D). Our results indicated a positive effect of Aco2b, Sdh2, Rip1 and Mss51 knockdowns. The iron-sulfur cluster containing Aco2b that catalyzes the second committed step of the TCA cycle from citrate to isocitrate is the only target directly involved in the TCA cycle where a knockdown had an effect on ester production. Sdh2, Rip1 and Mss51 are all involved in electron transfer of the ETC. Sdh2 is responsible for transferring electrons from succinate to ubiquinone, Rip1 is essential to complex III activity by oxidizing ubiquinol and transferring electrons to cytochromes c1 and b6G, and Mss51 is essential for assembly and translational activation of the cytochrome oxidase subunit 1 (COX1), which is one of the three core subunits of complex IV.[32-34]

Interestingly, with the exception of ARSM, none of the CRISPRi constructs reduced ethanol production (Figure S5.6). These results do not allow for a clear determination if ETC and TCA cycle knockdowns increase acetyl-CoA or ethanol availability, or impact ethyl acetate synthesis in a different way such as cofactor availability etc.

Previous studies suggest that a lack of trace metals in the media inhibits NADH oxidation and thus slow down flux through the TCA cycle freeing up acetyl-CoA.[7] A similar effect has been observed in *C. utilis*, where iron limitation induced ethyl acetate production.[35, 36] The same phenotype was recovered by the addition of ECT inhibitors to a moderate concentration. While three tested inhibitor increased ethyl acetate production, they seem to impact ester synthesis in different ways. Antimycin A and cyanide were used to inhibit electron transfer of ETC complex III and IV, respectively. Here, moderate addition of the inhibitor increased ethanol and ethyl acetate production and slowed down growth. Higher concentrations of the inhibitors completely disrupted ethyl acetate synthesis and it was hypothesized that ECT activity, and thus the ability to oxidize NADH, is essential for ester production.[7] A different inhibitor, carboxin interacts with the ubiquinone binding site of succinate dehydrogenase and prevents transfer of electrons to ubiquinone. NADH oxidation, and therefore ETC activity, however can still be maintained by the internal and external NADH dehydrogenases (Nde1 and Ndi1). Carboxin induced ethyl acetate synthesis in a dose dependent manner without accumulation of ethanol and was previously found to decrease succinate oxidation and thus reducing flux of acetyl-coA into the TCA cycle.[7, 37] To determine the impact that the TCA cycle and ETC knockdowns have on *K. marxianus* metabolism and especially ester production further studies such as transcriptome and metabolome analysis are required.

5.6 Conclusion

In this work we identified Eat1 as the crucial enzyme for acetate ester production in the non-model yeast *Kluyveromyces marxianus*. Localizations studies revealed that Eat1 is localized to the mitochondria and that this localization is critical for high ethyl acetate production. Overexpression of Eat1 significantly increased ethyl acetate production while reducing ethanol, indicating this step to be the bottleneck for the production of ethyl acetate. To further increase ester production we developed a CRISPR interference system that was able to efficiently knockdown electron transfer chain and TCA cycle enzymes thus slowing down flux through the TCA cycle and allowing for increased ester production.

5.7 References

[1] Lobs AK, Lin JL, Cook M, Wheeldon I: **High throughput, colorimetric screening of microbial ester biosynthesis reveals high ethyl acetate production from Kluyveromyces marxianus on C5, C6, and C12 carbon sources**. *Biotechnology journal* 2016, **11**(10):1274-1281.

[2] Lobs AK, Schwartz C, Wheeldon I: **Genome and metabolic engineering in non-conventional yeasts: Current advances and applications**. *Synth Syst Biotechnol* 2017, **2**(3):198-207.

[3] Urit T, Loser C, Wunderlich M, Bley T: **Formation of ethyl acetate by Kluyveromyces marxianus on whey: studies of the ester stripping**. *Bioproc Biosyst Eng* 2011, **34**(5):547-559.

[4] Fonseca GG, Heinzle E, Wittmann C, Gombert AK: **The yeast Kluyveromyces marxianus and its biotechnological potential**. *Applied microbiology and biotechnology* 2008, **79**(3):339-354.

[5] Loser C, Urit T, Stukert A, Bley T: **Formation of ethyl acetate from whey by Kluyveromyces marxianus on a pilot scale**. *Journal of biotechnology* 2013, **163**(1):17-23.

[6] Lobs AK, Engel R, Schwartz C, Flores A, Wheeldon I: **CRISPR-Cas9-enabled genetic disruptions for understanding ethanol and ethyl acetate biosynthesis in Kluyveromyces marxianus**. *Biotechnology for biofuels* 2017, **10**:164.

[7] Loser C, Urit T, Keil P, Bley T: **Studies on the mechanism of synthesis of ethyl acetate in Kluyveromyces marxianus DSM 5422**. *Applied microbiology and biotechnology* 2015, **99**(3):1131-1144.

[8] Lin JL, Wheeldon I: **Dual N- and C-Terminal Helices Are Required for Endoplasmic Reticulum and Lipid Droplet Association of Alcohol Acetyltransferases in Saccharomyces cerevisiae**. *PloS one* 2014, **9**(8).

[9] Verstrepen KJ, Van Laere SDM, Vanderhaegen BMP, Derdelinckx G, Dufour JP, Pretorius IS, Winderickx J, Thevelein JM, Delvaux FR: **Expression levels of the yeast alcohol acetyltransferase genes ATF1, Lg-ATF1, and ATF2 control the formation of a broad range of volatile esters**. *Applied and environmental microbiology* 2003, **69**(9):5228-5237.

[10] Kruis AJ, Levisson M, Mars AE, van der Ploeg M, Garces Daza F, Ellena V, Kengen SWM, van der Oost J, Weusthuis RA: **Ethyl acetate production by the elusive alcohol acetyltransferase from yeast**. *Metab Eng* 2017, **41**:92-101.

[11] Avalos JL, Fink GR, Stephanopoulos G: **Compartmentalization of metabolic pathways in yeast mitochondria improves the production of branched-chain alcohols**. *Nature biotechnology* 2013, **31**(4):335-+.

[12] Lin JL, Zhu J, Wheeldon I: **Synthetic Protein Scaffolds for Biosynthetic Pathway Colocalization on Lipid Droplet Membranes**. *ACS synthetic biology* 2017, **6**(8):1534-1544.

[13] Willetts A: **Ester formation from ethanol by Candida pseudotropicalis**. *Antonie Van Leeuwenhoek* 1989, **56**(2):175-180.

[14] Corzo G, Revah S, Christen P: **Effect of oxygen on the ethyl acetate production from continuous ethanol stream by Candida utilis in submerged cultures**. 1995.

[15] Schwartz CM, Hussain MS, Blenner M, Wheeldon I: **Synthetic RNA Polymerase III Promoters Facilitate High-Efficiency CRISPR-Cas9-Mediated Genome Editing in Yarrowia lipolytica**. *ACS synthetic biology* 2016, **5**(4):356-359.

[16] DiCarlo JE, Norville JE, Mali P, Rios X, Aach J, Church GM: **Genome engineering in Saccharomyces cerevisiae using CRISPR-Cas systems**. *Nucleic acids research* 2013, **41**(7):4336-4343.

[17] Smith JD, Suresh S, Schlecht U, Wu M, Wagih O, Peltz G, Davis RW, Steinmetz LM, Parts L, St Onge RP: **Quantitative CRISPR interference screens in yeast identify chemical-genetic interactions and new rules for guide RNA design**. *Genome Biol* 2016, **17**:45.

[18] Schwartz C, Frogue K, Ramesh A, Misa J, Wheeldon I: **CRISPRi repression of nonhomologous end-joining for enhanced genome engineering via homologous recombination in Yarrowia lipolytica**. *Biotechnol Bioeng* 2017, **114**(12):2896-2906.

[19] Schwartz C, Shabbir-Hussain M, Frogue K, Blenner M, Wheeldon I: **Standardized Markerless Gene Integration for Pathway Engineering in Yarrowia lipolytica**. *ACS synthetic biology* 2017, **6**(3):402-409.

[20] Kumar TA: **CFSSP: Chou and Fasman Secondary Structure Prediction server**. *Wide Spectrum* 2013, **1**(9):15 - 19.

[21] Fukasawa Y, Tsuji J, Fu SC, Tomii K, Horton P, Imai K: **MitoFates: Improved Prediction of Mitochondrial Targeting Sequences and Their Cleavage Sites**. *Molecular & Cellular Proteomics* 2015, **14**(4):1113-1126.

[22] Pfaffl MW: **A new mathematical model for relative quantification in real-time RT-PCR**. *Nucleic acids research* 2001, **29**(9):e45.

[23] Schmidt O, Pfanner N, Meisinger C: **Mitochondrial protein import: from proteomics to functional mechanisms**. *Nat Rev Mol Cell Bio* 2010, **11**(9):655-667.

[24] Urit T, Stukert A, Bley T, Loser C: **Formation of ethyl acetate by Kluyveromyces marxianus on whey during aerobic batch cultivation at specific trace element limitation**. *Applied microbiology and biotechnology* 2012, **96**(5):1313-1323.

[25] Merz S, Westermann B: **Genome-wide deletion mutant analysis reveals genes required for respiratory growth, mitochondrial genome maintenance and mitochondrial protein synthesis in Saccharomyces cerevisiae**. *Genome Biology* 2009, **10**(9).

[26] Loser C, Urit T, Bley T: **Perspectives for the biotechnological production of ethyl acetate by yeasts**. *Applied microbiology and biotechnology* 2014, **98**(12):5397-5415.

[27] van Rossum HM, Kozak BU, Pronk JT, van Maris AJA: **Engineering cytosolic acetyl-coenzyme A supply in Saccharomyces cerevisiae: Pathway stoichiometry, free-energy conservation and redox-cofactor balancing**. *Metab Eng* 2016, **36**:99-115.

[28] Kallelmhiri H, Miclo A: **Mechanism of Ethyl-Acetate Synthesis by Kluyveromyces-Fragilis**. *Fems Microbiol Lett* 1993, **111**(2-3):207-212.

[29] Snoep JL, Degraef MR, Westphal AH, Dekok A, Demattos MJT, Neijssel OM: **Differences in Sensitivity to Nadh of Purified Pyruvate-Dehydrogenase Complexes of Enterococcus-Faecalis, Lactococcus-Lactis, Azotobacter-Vinelandii and Escherichia-Coli - Implications for Their Activity in-Vivo**. *Fems Microbiol Lett* 1993, **114**(3):279-283.

[30] Canelas AB, van Gulik WM, Heijnen JJ: **Determination of the cytosolic free NAD/NADH ratio in Saccharomyces cerevisiae under steady-state and highly dynamic conditions**. *Biotechnol Bioeng* 2008, **100**(4):734-743.

[31] Gatti DL, Meinhardt SW, Ohnishi T, Tzagoloff A: **Structure and Function of the Mitochondrial Bc1 Complex - a Mutational Analysis of the Yeast Rieske Iron Sulfur Protein**. *J Mol Biol* 1989, **205**(2):421-435.

[32] Siep M, van Oosterum K, Neufeglise H, van der Spek H, Grivell LA: **Mss51p, a putative translational activator of cytochrome c oxidase subunit-1 (COX1) mRNA, is required for synthesis of Cox1p in Saccharomyces cerevisiae**. *Curr Genet* 2000, **37**(4):213-220.

[33] Oyedotun KS, Lemire BD: **The quaternary structure of the Saccharomyces cerevisiae succinate dehydrogenase - Homology modeling, cofactor docking, and molecular dynamics simulation studies**. *Journal of Biological Chemistry* 2004, **279**(10):9424-9431.

[34] Thomas KC, Dawson PS: **Relationship between iron-limited growth and energy limitation during phased cultivation of Candida utilis**. *Can J Microbiol* 1978, **24**(4):440-447.

[35] Armstrong DW, Yamazaki H: **Effect of Iron and Edta on Ethyl-Acetate Accumulation in Candida-Utilis**. *Biotechnology letters* 1984, **6**(12):819-824.

[36] Opekarova M, Sigler K: **Separation of Transport and Metabolic Steps in the Uptake of Succinate by Kluyveromyces-Fragilis**. *Folia Microbiol* 1987, **32**(3):200-205.

5.8 Supporting information

Table S5.1 Yeast strains

Name	Genotype	Reference
YS5	*S. cerevisiae* BY4742 MAT*α his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0*	GE Healthcare
YS402	*K. marxianus* CBS 6556 Δura3	[1]
YS578	*S. cerevisiae* BY4742 OM45-DsRed	This work
YS679	*K. marxianus* CBS 6556 Δura3 Δatf	[1]
YS964	*K. marxianus* strain CBS 6556 Δura3 Δeht1	This work
YS965	*K. marxianus* strain CBS 6556 Δura3 Δeeb	This work
YS969	*K. marxianus* strain CBS 6556 Δura3 Δeat1	This work
YS1052	*K. marxianus* CBS 6556 Δura3 URA3::GPDp-Eat1-Cyc1t	This work
YS1053	*K. marxianus* CBS 6556 Δura3 URA3::GPDp-Eat1-GFP-Cyc1t	This work
YS1064	*K. marxianus* CBS 6556 Δura3 Δeat1 URA3::GPDp-Eat1-Cyc1t	This work
YS1066	*K. marxianus* CBS 6556 Δura3 Δeat1URA3::GPDp-Eat1-GFP-Cyc1t	This work
YS1087	*K. marxianus* CBS 6556 Δura3 Δeat1 URA3::GPDp-(Δ1-42)Eat1-Cyc1t	This work
YS1089	*K. marxianus* CBS 6556 Δura3 Δeat1 URA3::GPDp-(Δ1-52)Eat1-Cyc1t	This work
YS1091	*K. marxianus* CBS 6556 Δura3 Δeat1 URA3::GPDp-(Δ1-42)Eat1-GFP- Cyc1t	This work
YS1093	*K. marxianus* CBS 6556 Δura3 Δeat1 URA3::GPDp-(Δ1-52)Eat1-GFP-Cyc1t	This work

Table S5.2 Plasmids

Name	Sequence	Reference
pIW8	pRS423	Gifted by Nancy Da Silva
pIW21	pRS426 PGK1p-ATF1-GFP-PGK1t	Lin *et al.* [2]
pIW272	pJSKM316-GDP	Lee *et al.*[3]
pIW512	Tef1-Km-dCas9-SV40-Cyc1 KmPRP1-Gly-tRNA-KmMDH1 T1-SUP4	This work
pIW513	Tef1-Km-dCas9-SV40-Cyc1 KmPRP1-Gly-tRNA-KmSDH1 T1-SUP4	This work
pIW514	Tef1-Km-dCas9-SV40-Cyc1 KmPRP1-Gly-tRNA-KmSDH2 T1-SUP4	This work
pIW578	pIW272 His3 marker	This work
pIW601	Tef1-Km-Cas9-SV40-Cyc1 KmPRP1-Gly-tRNA-PspXI-SUP4	This work
pIW602	Tef1-Km-dCas9-SV40-Cyc1 KmPRP1-Gly-tRNA-PspXI4-SUP4	This work
pIW630	Tef1-Km-dCas9-SV40-Cyc1 KmPRP1-Gly-tRNA-scramble DNA-SUP4	This work
pIW634	Homology donor Ura3 GPDp-Cyc1t (derived from pIW578)	This work
pIW653	Tef1-Km-dCas9-SV40-Cyc1 KmPRP1-Gly-tRNA-Aco2b T1-SUP4	This work
pIW654	Tef1-Km-dCas9-SV40-Cyc1 KmPRP1-Gly-tRNA- Aco2b T2-SUP4	This work
pIW778	Tef1-Km-dCas9-SV40-Cyc1 KmPRP1-Gly-tRNA-KmPDB1 T2-SUP4	This work
pIW779	Tef1-Km-dCas9-SV40-Cyc1 KmPRP1-Gly-tRNA-KmPDB1 2x-SUP4	This work
pIW780	Tef1-Km-dCas9-SV40-Cyc1 KmPRP1-Gly-tRNA-KmMDH1 T2-SUP4	This work
pIW781	Tef1-Km-dCas9-SV40-Cyc1 KmPRP1-Gly-tRNA-KmMDH1 2x-SUP4	This work
pIW782	Tef1-Km-dCas9-SV40-Cyc1 KmPRP1-Gly-tRNA-KmSDH1 T2-SUP4	This work
pIW783	Tef1-Km-dCas9-SV40-Cyc1 KmPRP1-Gly-tRNA-KmSDH1 2x-SUP4	This work
pIW784	Tef1-Km-dCas9-SV40-Cyc1 KmPRP1-Gly-tRNA-KmSDH2 T2-SUP4	This work
pIW785	Tef1-Km-dCas9-SV40-Cyc1 KmPRP1-Gly-tRNA-KmSDH2 2x-SUP4	This work
pIW786	Tef1-Km-dCas9-SV40-Cyc1 KmPRP1-Gly-tRNA-KmACO2b 2x-SUP4	This work
pIW821	Tef1-Km-dCas9-SV40-Cyc1 KmPRP1-Gly-tRNA-KmMss51 T1-SUP4	This work
pIW822	Tef1-Km-dCas9-SV40-Cyc1 KmPRP1-Gly-tRNA-KmMss51 T2-SUP4	This work
pIW823	Tef1-Km-dCas9-SV40-Cyc1 KmPRP1-Gly-tRNA-KmCox7 T1-SUP4	This work
pIW824	Tef1-Km-dCas9-SV40-Cyc1 KmPRP1-Gly-tRNA-KmCox7 T2-SUP4	This work
pIW825	Tef1-Km-dCas9-SV40-Cyc1 KmPRP1-Gly-tRNA-KmRip1 T1-SUP4	This work
pIW826	Tef1-Km-dCas9-SV40-Cyc1 KmPRP1-Gly-tRNA-KmRip1 T2-SUP4	This work
pIW827	Tef1-Km-dCas9-SV40-Cyc1 KmPRP1-Gly-tRNA-KmCyt1 T1-SUP4	This work
pIW828	Tef1-Km-dCas9-SV40-Cyc1 KmPRP1-Gly-tRNA-KmCyt1 T2-SUP4	This work
pIW829	Tef1-Km-dCas9-SV40-Cyc1 KmPRP1-Gly-tRNA-KmMss51 2x-SUP4	This work
pIW830	Tef1-Km-dCas9-SV40-Cyc1 KmPRP1-Gly-tRNA-KmCox7 2x-SUP4	This work
pIW967	URA3 T1 HD:: GPDp-Eat1-Cyc1t	This work
pIW968	URA3 T1 HD:: GPDp-Eat1-GFP-Cyc1t	This work
pIW972	PGK1p-KmEat1-GFP-PGK1t	This work
pIW973	PGK1p-KmEat1(delta1-18)-GFP-PGK1t	This work
pIW977	PGK1p-KmEat1(delta1-42)-GFP-PGK1t	This work
pIW978	PGK1p-KmEat1(delta1-57)-GFP-PGK1t	This work
pIW998	PGK1p-KmEat1(delta1-105)-GFP-PGK1t	This work
pIW1005	URA3 T1 HD:: GPDp-Eat1(delta1-42)-Cyc1t	This work
pIW1006	URA3 T1 HD:: GPDp-Eat1(delta1-57)-Cyc1t	This work
pIW1007	URA3 T1 HD:: GPDp-Eat1(delta1-42)-GFP-Cyc1t	This work
pIW1008	URA3 T1 HD:: GPDp-Eat1(delta1-57)-GFP-Cyc1t	This work
pIW1023	Tef1-Km-dCas9-SV40-Cyc1 Mss52 2x Rip1 2x	This work
pIW1024	Tef1-Km-dCas9-SV40-Cyc1 Mss52 2x Sdh 2	This work
pIW1025	Tef1-Km-dCas9-SV40-Cyc1 Mss52 2x Aco2b 1	This work
pIW1026	Tef1-Km-dCas9-SV40-Cyc1 Rip1 2x Sdh 2	This work
pIW1027	Tef1-Km-dCas9-SV40-Cyc1 Rip1 2x Aco2b 1	This work
pIW1028	Tef1-Km-dCas9-SV40-Cyc1 Aco2b 1 Sdh 2	This work

Name	Sequence	Reference
pIW1029	Tef1-Km-dCas9-SV40-Cyc1 Mss52 2x Rip1 2x Sdh 2	This work
pIW1030	Tef1-Km-dCas9-SV40-Cyc1 Mss52 2x Rip1 2x Aco2b 1	This work
pIW1031	Tef1-Km-dCas9-SV40-Cyc1 Mss52 2x Rip1 2x Sdh 2 Aco2b 1	This work
pIW1063	Tef1-Km-dCas9-SV40-Cyc1 Rip1 2x Sdh 2 Aco2b 1	This work
pIW1064	Tef1-Km-dCas9-SV40-Cyc1 Mss52 2x Sdh 2 Aco2b 1	This work

Table S5.3 Primers

Name	Sequence (5'-> 3')
P1853 dCas9 1F	AAGTTTTCTAGAACTAGTGGATCCCCCGGGAAAAATGGACAAGAAGTACTCTATCGGTTTGGCTATCGGTACCAACTCTGTTGG
P1854 dCas9 1 R	TTGTGGAACGATAGCATCAACATCGTAGTCAGACAATC
P1855 dCas9 2 F	TACGATGTTGATGCTATCGTTCCACAATCCTTTTTG
P1856 dCas9 2 R	ACCAAGTTGGAGTATACGCCCTAGGAAGGCAAATTAAAGCCTTCG
P1877 KmAct1Q F	CCCAATGAACCCAAAGAATAACAG
P1878 KmAct1Q R	GATAGCATGAGGCAAGGAGAAACC
P1894 KmMDH1 KD sgRNA F	CCTCGAGGATTATCATACCAAGAAGTAGTTTTAGAGCTAGAAATAGCAAGTTAAA
P1895 KmMDH1 KD R	TAAAACTACTTCTTGGTATGATAATCCTCGAGGTTGACACTGACGGGATTC
P1897 KmSDH1 KD sgRNA F	CCTCGAGGTTAGTATTATGTGATTGGGGTTTTAGAGCTAGAAATAGCAAGTTAAA
P1898 KmSDH1 KD R	CTAAAACCCCAATCACATAATACTAACCTCGAGGTTGACACTGACGGGATTC
P1900 KmSDH2 KD sgRNA F	CCTCGAGGTCTTTGAAGTCGCAAAAAAATTTTAGAGCTAGAAATAGCAAGTTAAA
P1901 KmSDH2 KD R	CTAAAATTTTTTTGCGACTTCAAAGACCTCGAGGTTGACACTGACGGGATTC
P1928 URA3 HDR Screen F	TTGTATCATCACTCCCAGTCAA
P1929 URA3 HDR Screen R	ACTGAGTATAGTGAGTAGTAGTGGTTGAT
AL20 pIW272 His3 F	CCAAAGGTGTTCTTATGTAGGACGTCAAAACTGTATTATAAGTAAATGCATGTATA
AL21 pIW272 His3 R	AAAAATATAGAGTGTACTAGCCTAGGAAGCGTGGTGCACTCTCAGT
AL22 His3 F	ACTGAGAGTGCACCACGCTTCCTAGGCTAGTACACTCTATATTTTTTTATGCCT
AL23 His3 R	TTTACTTATAATACAGTTTTGACGTCCTACATAAGAACACCTTTGGTGG
AL45 dCas9 scrambled F	TCGAATCCCGTCAGTGTCAAATTCGGATGAACGTAGGAATTTTTAGAGCTAGAAATAGCA
AL46 dCas9 scrambled R	TGCTATTTCTAGCTCTAAAAATTCCTACGTTCATCCGAATTTGACACTGACGGGATTCGA
AL47 URA3 T1 HD1 F	CGAATTGGAGCTCCACCGCGGTGGCGTCAACAGTACTCGAATATAATGC
AL49 URA3 T2 HD1 F	CGCTCGAAGGCTTTAATTTGCGCGCGGGTCTCTAGCGCACGGTGA
AL55 MDH1 QPCR F	CCACTTCCATCGCCAAGAAC
AL56 MDH QPCR R	TTTTGCTTCAAGACCTCGGC
AL57 SDH1 QPCR F	CGACAAGACCTTCGCTGATG
AL58 SDH1 QPCR R	TTTGGCGGTTTGAGTAGCAC
AL59 SDH2 QPCR F	CTTGTGTGCTTGTTGTTCCAC
AL60 SDH2 QPCR R	GATACCCGAGGCCTCATCTC
AL75 Aco2b_1 F	TCGAATCCCGTCAGTGTCAAAATATATAGACGGAGCCAATTTTTAGAGCTAGAAATAGCA
AL76 Aco2b_1 R	TGCTATTTCTAGCTCTAAAAATTGGCTCCGTCTATATATTTTGACACTGACGGGATTCGA
AL77 Aco2b_2 F	TCGAATCCCGTCAGTGTCAAAACATTATCAACAGATGATATTTTAGAGCTAGAAATAGCA
AL78 Aco2b_2 R	TGCTATTTCTAGCTCTAAAATATCATCTGTTGATAATGTTTTGACACTGACGGGATTCGA
AL116 URA3 HD1 R	GTATTGATAATGATAAACTGCGGCCCCATGGGCATTAACAACCCTCTAGGTT
AL117 URA3 HD2 R	AAGGGAACAAAAGCTGGTACCGGCCGTGGTTCAAAGCAGCGTTTTGC
AL139 KmRPR1 2nd F	TTTTTTTGTTTTTTATGTCTCTGCCGCGGTATACTCCAACTTGGTCGAAAG
AL140 KmsgRNA 2nd R	ACTAAAGGGAACAAAAGCTGGTACCGCTCGAGGAGACATAAAAAACAAAAAAAGCACC
AL145 MDH1 2 F	TCGAATCCCGTCAGTGTCAATAGTAGTATATGTATGTACTTTTTAGAGCTAGAAATAGCA
AL146 MDH1 2 R	TGCTATTTCTAGCTCTAAAAAGTACATACATATACTACTATTGACACTGACGGGATTCGA
AL147 SDH1 2 F	TCGAATCCCGTCAGTGTCAAATGCTCAGTAGATCAGTGAGTTTTAGAGCTAGAAATAGCA
AL148 SDH1 2 R	TGCTATTTCTAGCTCTAAAACTCACTGATCTACTGAGCATTTGACACTGACGGGATTCGA
AL149 SDH2 2 F	TCGAATCCCGTCAGTGTCAACAACTTACAAATCTTTTTTTTTTTTAGAGCTAGAAATAGCA
AL165 MSS51 1 F	TCGAATCCCGTCAGTGTCAAGCACATTTATCCCTATCATTTTTTAGAGCTAGAAATAGCA
AL166 MSS51 1 R	TGCTATTTCTAGCTCTAAAAAATGATAGGGATAAATGTGCTTGACACTGACGGGATTCGA
AL167 MSS51 2 F	TCGAATCCCGTCAGTGTCAAAGTTGGAAGTATAGATACTTTTTTAGAGCTAGAAATAGCA
AL168 MSS51 2 R	TGCTATTTCTAGCTCTAAAAAAGTATCTATACTTCCAACTTTGACACTGACGGGATTCGA
AL169 COX7 1 F	TCGAATCCCGTCAGTGTCAAAAAGAGTTGTGATTAGAGAGTTTTAGAGCTAGAAATAGCA

Name	Sequence (5'-> 3')
AL170 COX7 1 R	TGCTATTTCTAGCTCTAAAACTCTCTAATCACAACTCTTTTTGACACTGACGGGATTCGA
AL171 COX7 2 F	TCGAATCCCGTCAGTGTCAAATAATAAAAAAAAATTTATATTTTAGAGCTAGAAATAGCA
AL172 COX72 R	TGCTATTTCTAGCTCTAAAATATAAATTTTTTTTTTATTATTTGACACTGACGGGATTCGA
AL173 RIP1 1 F	TCGAATCCCGTCAGTGTCAATATTAATGCTTTCTTTTGAATTTTAGAGCTAGAAATAGCA
AL174 RIP1 1 R	TGCTATTTCTAGCTCTAAAATTCAAAAGAAAGCATTAATATTGACACTGACGGGATTCGA
AL175 RIP2 1 F	TCGAATCCCGTCAGTGTCAATCGGTTCTGATTTGATTTTGTTTTAGAGCTAGAAATAGCA
AL176 RIP1 2 R	TGCTATTTCTAGCTCTAAAACAAAATCAAATCAGAACCGATTGACACTGACGGGATTCGA
AL177 CYT1 1 F	TCGAATCCCGTCAGTGTCAACTTAAATTTATATTCTTTGTTTTTAGAGCTAGAAATAGCA
AL178 CYT1 1 R	TGCTATTTCTAGCTCTAAAAACAAAGAATATAAATTTAAGTTGACACTGACGGGATTCGA
AL179 CYT1 2 F	TCGAATCCCGTCAGTGTCAATTGAGTGGTTTTTTTATAGCTTTTAGAGCTAGAAATAGCA
AL180 CYT1 2 R	TGCTATTTCTAGCTCTAAAAGCTATAAAAAAACCACTCAATTGACACTGACGGGATTCGA
AL181 Eht1 cPCR F	TTGGTATACCCCTTTACCGAAAAGCA
AL182 Eht1 cPCR R	TGTCAGACAATTCAAACAACTCTCTCCC
AL183 Eeb1 cPCR F	GATTAGAAGAATTGAGATTTGAAGATAGTATCATGAAAAACGGG
AL197 Aco2b qPCR F	AACATGGGTGCTGAAATCGG
AL198 Aco2b qPCR R	TCATCGGCAACCAACAAGTC
AL199 Mss51 qPCR F	AAGAGTACTGGATGGGCGAC
AL200 Mss51 qPCR R	CGTACTGGTCCTGCACAATG
AL201 Cox7 qPCR F	TGCCATCCCCACAAAGAATC
AL202 Cox7 qPCR R	AGCGAACAAACCCCAGAATG
AL203 Rip1 qPCR F	CCTCAAGCCGATTCTGACAG
AL204 Rip1 qPCR F	GAACCAACCACCGAAGTCAC
AL205 Cyt1 qPCR F	TGTTTCTCACACCAACGCTG
AL206 Cyt1 qPCR R	TCTGGCAGCTTGTTCGTTTG
AL216 Eht1 T3	TCGAATCCCGTCAGTGTCAAAGAAATTCAATGTCTTTTATTTTTAGAGCTAGAAATAGCA
AL218 Eeb1 T3	TCGAATCCCGTCAGTGTCAAACTATCCTGATGGTGGTGAATTTTAGAGCTAGAAATAGCA
AL225 Eat1 T1	CGTCAGTGTCAACCTCGAGGGAGCCTTCAACCAACAAAAATTTTAGAGCTAGAAATAGCA
AL227 Eat1 cPCR F	GGCATATTTGCACACAATTATTAGC
AL228 Eat1 cPCR R	CCACAAGGTTAACCTTACCCAA
AL250 Eeb1 cPCR T3 R	AGTGGATTCGAACTGTTCATCAGGA
AL254 Eat1 T1 R	TGCTATTTCTAGCTCTAAAATTTTTGTTGGTTGAAGGCTCCCTCGAGGTTGACACTGACG
AL270 Eat1 URA3 T1 HD F	TTCTAGAACTAGTGGATCCCCCGGGATGCTTCTCGCTTACACCGT
AL271 Eat1 URA3 T1 HD R	GACATAACTAATTACATGACTCGAGTTTAATCTCTAGCACTTTTGAGAGATTCAG
AL272 Eat1-GFP URA3 T1 HD R	CACCAGTCATGCTAGCCATCCCTCCGCCTAGGGATCCGCCTCCATCTCTAGCACTTTTGAG AGATTCAG
AL275 HDR Screen GPDp 2R	AAATGGCGAGTATTGATAATGATAAACTG
AL276 KmEat1 PGKp F	TTTCTCTTTTTTACAGATCACCGCGGATGCTTCTCGCTTACACCGT
AL277 GFP PGKt R	ATCTATCGATTTCAATTCAATTCAATACTAGTTCATTTGTATAGTTCATCCATGCCAT
AL280 delta 1-18 Eat1 F	TTCTCTTTTTTACAGATCACCGCGGATGGTCAGACCATCTAACTGGTCTTT
AL281 delta 1-42 Eat1 F	TTCTCTTTTTTACAGATCACCGCGGATGACCAGAAGAGCATACTCTGC
AL282 delta 1-57 Eat1 F	TTCTCTTTTTTACAGATCACCGCGGATGTCTGCTACTGCCAGAGCCTT
AL283 delta 1-105 Eat1 F	TTCTCTTTTTTACAGATCACCGCGGATGCCAATCAAGGAAACTGTCGATATGG
AL285 delta 1-42 Eat1 F Km int URA3	TTCTAGAACTAGTGGATCCCCCGGGATGACCAGAAGAGCATACTCTGC
AL286 delta 1-57 Eat1 Km int URA3	TTCTAGAACTAGTGGATCCCCCGGGATGTCTGCTACTGCCAGAGCCTT
AL287 Ecoli Eat1 F	GTTTAACTTTAAGAAGGAGATATACCATGCTTCTCGCTTACACCGT
AL288 Ecoli Eat1 R	AGTGGTGGTGGTGGTGGTGCTCGAGATCTCTAGCACTTTTGAGAGATTCAG
AL289 Ecoli Eat1(1-	GTTTAACTTTAAGAAGGAGATATACCATGACCAGAAGAGCATACTCTGC

Name	Sequence (5'-> 3')
42) F	
AL290 Ecoli Eat1(1-57) F	TTCTAGAACTAGTGGATCCCCCGGGATGACCAGAAGAGCATACTCTGC
AL291 qPCR GFP F	TCGGGATCTGTACGACGATG
AL292 qPCR GFP R	TTCACCCTCTCCACTGACAG
Cr_1539	GTGGTGCTTTTTTTGTTTTTTATGTCTctgCCGCTATACTCCAACTTGGTCGAAAGATG
Cr_1540	CAATCAAGCTCGGATTACGGTGTTCACTCCAGACATAAAAAACAAAAAAAGCACCACC
Cr_1541	GGAGTGAACACCGTAATCCGAGCTTGATTGTATACTCCAACTTGGTCGAAAGATG
Cr_1542	CATCTTTCGACCAAGTTGGAGTATACCGCAGACATAAAAAACAAAAAAAGCACCACC
Cr_1547	CACTAAAGGGAACAAAAGCTGgtaCCGCAGACATAAAAAACAAAAAAAGCACCACC
Cr_1548	CGCTCGAAGGCTTTAATTTGCcttCCTAGTATACTCCAACTTGGTCGAAAGATG
Cr_1549	CTTTCGACCAAGTTGGAGTATAcgcCCTAGAGACATAAAAAACAAAAAAAGCACCACC
Cr_1551	GTCCTGCTGTTACTCACGGTCTGTAATCCAAGACATAAAAAACAAAAAAAGCACCACC
Cr_1552	TGGATTACAGACCGTGAGTAACAGCAGGACTATACTCCAACTTGGTCGAAAGATG
Cr_1598	GACGCTCGAAGGCTTTAATTTGCcttCCTAGGTATACTCCAACTTGGTCGAAAGATG
Cr_1599	CTTTCGACCAAGTTGGAGTATAcgcCCTAGAGACATAAAAAACAAAAAAAGCACCACC
Cr_1645	GCCAGCAAAACTAAAAAACTGTATTATAAGTAAATGCAGGCTTTAATTTGCCTTCCTAGG
Cr_1646	ATTGAAGCTCTAATTTGTGAGTTTAGTATACATGCACCTCACTAAAGGGAACAAAAGC

Table S5.4 sgRNA target sequences

Name	20 Sequence and PAM (5'-> 3')
Atf1 KO	ATATAGTCTTCGGCAACACCGGG
Eat1 KO	GAGCCTTCAACCAACAAAAAGGG
Eeb1 KO	ACTATCCTGATGGTGGTGAAGGG
Eht1 KO	AGAAATTCAATGTCTTTTATGGG
Ura3 for HR	TTGCCGAGTTATCGTCCAAGGGG
Scrambled sgRNA	ATCCGGATGAACGTAGGAAT (no PAM)
Aco2b sgRNA 1	AATATATAGACGGAGCCAATCGG (TATA, antisense)
Aco2b sgRNA 2	AACATTATCAACAGATGATATGG (TSS, antisense)
Mdh1 sgRNA 1	ATTATCATACCAAGAAGTAGAGG (TATA, antisense)
Mdh1 sgRNA 2	TAGTAGTATATGTATGTACTTGG (TSS, antisense)
Sdh1 sgRNA 1	TTAGTATTATGTGATTGGGGGGG (TATA, antisense)
Sdh1 sgRNA 2	ATGCTCAGTAGATCAGTGAGAGG (TSS, sense)
Sdh2 sgRNA 1	TCTTTGAAGTCGCAAAAAAATGG (TATA, antisense)
Sdh2 sgRNA 2	CAACTTACAAATCTTTTTTTTGG (TSS, sense)
Rip1 sgRNA 1	TATTAATGCTTTCTTTTGAATGG (TSS, antisense)
Rip1 sgRNA 2	TCGGTTCTGATTTGATTTTGTGG (TATA, sense)
Cyt1 sgRNA 1	CTTAAATTTATATTCTTTGTTGG (TSS, antisense)
Cyt1 sgRNA 2	TTGAGTGGTTTTTTTATAGCAGG (TATA, sense)
Mss51 sgRNA 1	GCACATTTATCCCTATCATTTGG (TSS, antisense)
Mss51 sgRNA 2	AGTTGGAAGTATAGATACTTTGG (TATA, sense)
Cox7 sgRNA 1	AAAGAGTTGTGATTAGAGAGAGG (TSS, sense)
Cox7 sgRNA 2	ATAATAAAAAAAAATTTATATGG (TATA, sense)

> K. marxianus codon-optimized dead S. pyogenes Cas9 (dCas)
MDKKYSIGL**A**IGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYT
RRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDST
DKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSR
RLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAA
KNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPENYKEIFFDQSKNGYAGYIDGGA
SQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIE
KILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL
LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDR
FNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTG
WGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSP
AIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENT
QLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVD**A**IVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVV
KKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEND
KLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRK
MIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNI
VKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITI
MERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY
EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLT
NLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD

Figure S5.1: Km codon-optimized version of dCas9 from *S. pyogenes*. Mutation of D10A and H840A renders Cas9 unable to cut double strand DNA (bold red amino acids)

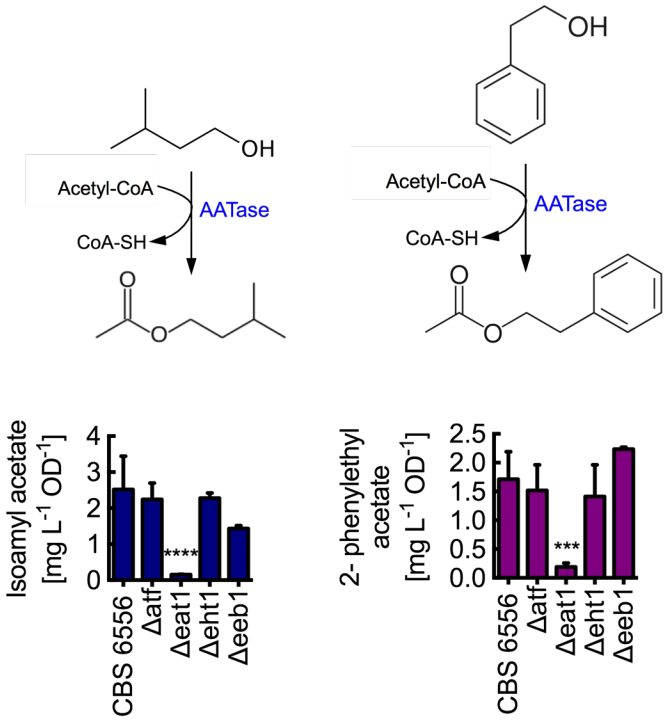


Figure S5.2: Eat1 is responsible for acetate ester formation in *K. marxianus.* Disruption of Eat1 decreases production of isoamyl and phenylethyl acetate. Bars represent arithmetic mean and error bars represent standard deviation.

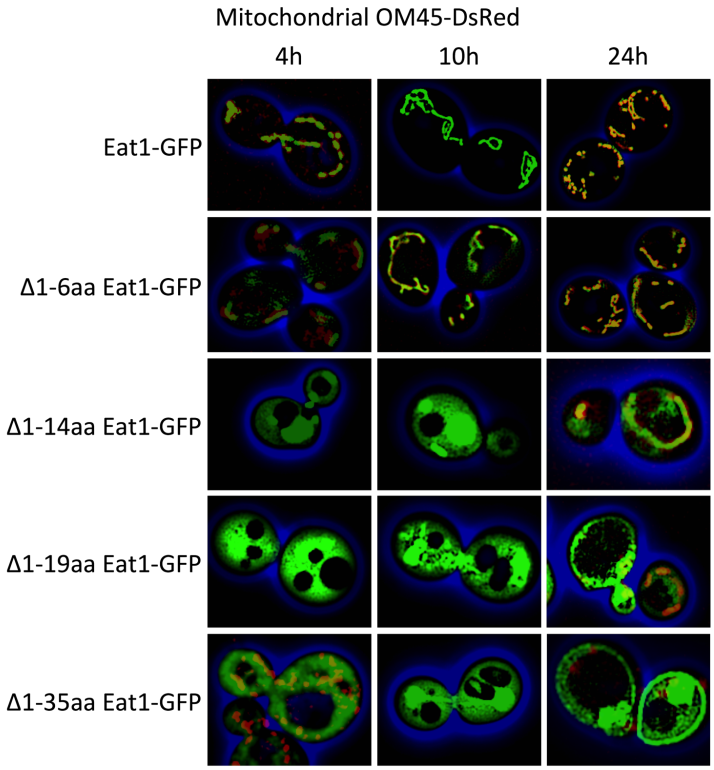


Figure S5.3: Localization of *K. marxianus* Eat1 and truncations when expressed in *S. cerevisiae.* Microscopic analysis revealed time independent localization of *Km*Eat1 to the mitochondria of *S. cerevisiae* with a mitochondrial marker (OM45-DsRed). Stepwise truncation leads to a stepwise localization of the protein to the cytosol, where at least 14aa need to be truncated from the N-terminus.

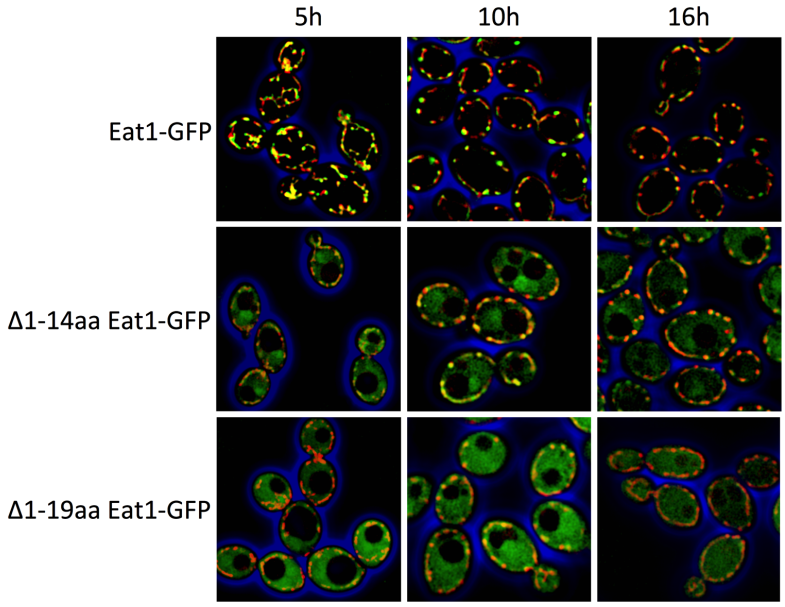


Figure S5.4: Time course of Eat1-GFP expression and N-terminal truncations in *K. marxianus*. Time course studies time independent localization of Eat1 to the mitochondria of *K. marxianus*. Stained with a mitotracker stain. N-terminal truncations of 14 and 19 aa lead to partial and complete delocalization of the protein to the cytosol.

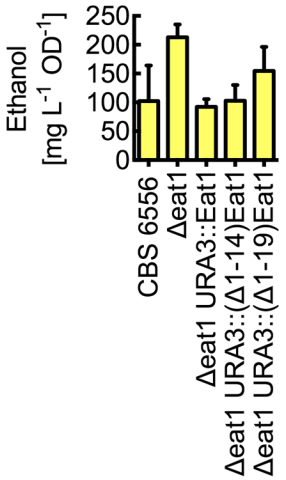


Figure S5.5: Eat1 truncation leads to accumulation of ethanol. Expression of wild type and N-terminally truncated Eat1 proteins lead to increasing ethanol production with disruption of mitochondrial localization. Bars represent arithmetic means of at least n=3 replicates and bars represent the standard deviation.

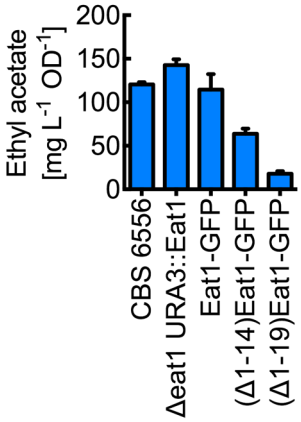


Figure S5.6: GFP fusion to Eat1 reduces ethyl acetate production. A C- terminal fusion of GFP to Eat1 and its truncations lead to a slight decrease in ethyl acetate production. Bars represent arithmetic means of at least n=3 replicates and bars represent the standard deviation.

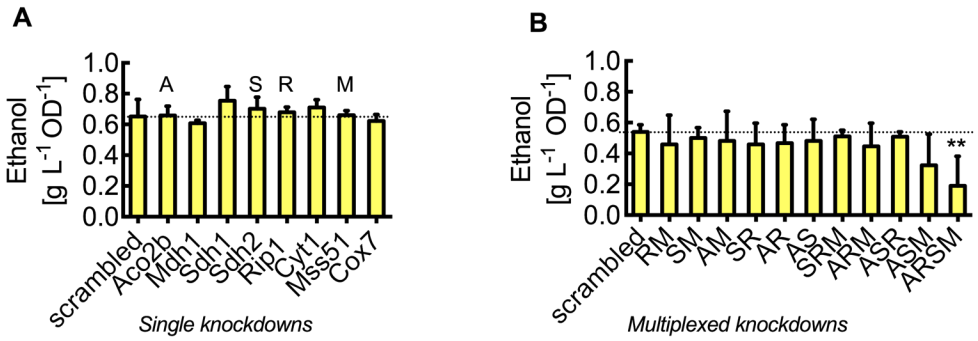


Figure S5.6: Impact of CRISPRi-meditated single target and multiplexed knockdowns on ethanol production. (A) Single gene knockdowns did not impact ethanol production. (B) Multiplexed knockdowns only affected ethanol production for the quadruple knockdown of ARSM. Bars represent arithmetic means of at least n=3 replicates and bars represent the standard deviation.

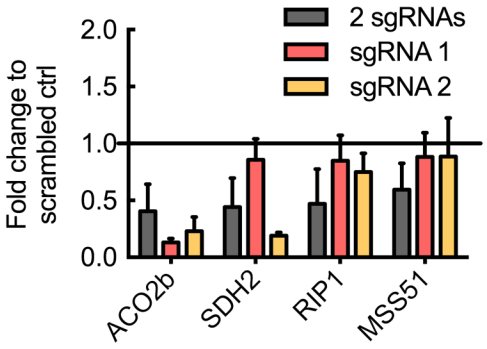


Figure S5.7: qRT-PCR to determine knockdown efficiencies of individual sgRNA's compared to the double target system. Efficient knockdown of Rip1 and Mss51 require both sgRNA's to be expressed while increased or similar knockdown efficiency can be achieved for Aco2b and Sdh2 when only sgRNA1 and sgRNA2 are expressed, respectively. Bars represent arithmetic means of at least n=3 replicates and bars represent the standard deviation.

Supporting References

[1] Lobs AK, Engel R, Schwartz C, Flores A, Wheeldon I: **CRISPR-Cas9-enabled genetic disruptions for understanding ethanol and ethyl acetate biosynthesis in Kluyveromyces marxianus**. *Biotechnology for biofuels* 2017, **10**:164.

[2] Lin JL, Wheeldon I: **Dual N- and C-Terminal Helices Are Required for Endoplasmic Reticulum and Lipid Droplet Association of Alcohol Acetyltransferases in Saccharomyces cerevisiae**. *PloS one* 2014, **9**(8).

[3] Lee KS, Kim JS, Heo P, Yang TJ, Sung YJ, Cheon Y, Koo HM, Yu BJ, Seo JH, Jin YS *et al*: **Characterization of Saccharomyces cerevisiae promoters for heterologous gene expression in Kluyveromyces marxianus**. *Applied microbiology and biotechnology* 2013, **97**(5):2029-2041.

Chapter 6: Summary and Conclusion

This work advances the concept of phenotype-based host selection of non-conventional yeasts for biotechnological chemical and protein production. Based on a high native capacity for ester production we chose the yeast *Kluyveromyces marxianus* as a host for ethyl acetate biosynthesis. In the context of this thesis we developed a high-throughput colorimetric assay that allowed us to rapidly screen for ester production by a variety of *K. marxianus* strains on different carbon sources to choose the best host for further work. Because previously established genome engineering tools were sparse, inefficient and tedious, we developed a CRISPR-Cas9 system for efficient genome editing. We applied this system to study volatile metabolite biosynthesis in *K. marxianus* strain CBS 6556 and found that Adh1, 2, and 3 activities are important for ethanol and ethyl acetate production. Furthermore, we identified a novel activity of Adh7 towards the oxidation of hemiacetal to ethyl acetate. In contrast to previous hypotheses, we found that Atf only marginally contributes to ethyl acetate biosynthesis in *K. marxianus*. Subsequent disruption studies revealed that Eat1 is the critical enzyme for acetate ester biosynthesis in *K. marxianus* and that mitochondrial localization of Eat1 is crucial for high ethyl acetate production. Overexpression of Eat1 was increased ethyl acetate biosynthesis by 1.8–fold. To increase carbon flux towards ethyl acetate production we knocked down TCA cycle and electron transport chain enzyme expression using CRISPR interference, which led to an increase in ethyl acetate by 3.8–fold compared to the scrambled sgRNA control.

As discussed in chapter 2, non-conventional yeasts bear an array of advantages over model organisms like *S. cerevisiae*. The emergence of CRISPR as an efficient genome editing tool has opened the door for engineering these yeasts. In chapter 3, we describe an efficient

CRISPR-Cas9 genome editing system that allows for efficient genomic disruptions. The system was used to create a library of alcohol dehydrogenase and alcohol acetyltransferase disruptions. For the system to be efficient an effective sgRNA is essential. While there are a variety of scoring algorithms for sgRNA efficiency in model organisms and mammalian cells, such a tool has not been developed for *K. marxianus*. Therefore, *K. marxianus* genome editing would greatly benefit from a genome wide survey that assesses sgRNA efficiency and off-target effects.

CRISPR–Cas9 systems from *Streptococcus pyogenes* have been developed for many non-conventional yeasts of interest. Alternative systems such as CRISPR-Cpf1 and xCas9 that recognizes a wide array of PAM sequences may be adapted to increase the coverage of accessible DNA for editing, especially in genomic regions or organisms rich in AT. In addition, researchers have worked on increasing CRISPR-Cas9 efficiencies by fine-tuning sgRNA expression using different promoter systems such as the bacterial T7 system and tuned the fidelity of Cas9 (Cas9-HF) to avoid off-target effects. With the increasing application of genome wide CRISPR libraries the whole genome can now be surveyed for comprehensive assessment of gene function and to achieve specific genome manipulation with a beneficial phenotype.

The work in this thesis has mostly focused on the production of the volatile ester ethyl acetate. So far, analysis of ester production has mostly relied on slow chromatography such as gas chromatography. Conventional ester analysis by chromatography such as GC or HPLC is time consuming and low-throughput. We developed a colorimetric screen that allows for high-throughput analysis of ethyl acetate production in a multiwell plates format. The system is intended for analysis of organisms that produce or overproduce mainly one single ester. Cross-reactivity of different esters in the reaction can be minimized by choice

of the extraction solvent but not completely eliminated. Therefore, the assay is not suitable for analysis of strains producing high amounts of different esters as observed in a variety of beer and wine yeast strains. The assay, as described in chapter 3, is used for end-point analysis of ester production. Improvements to this method could be achieved by choosing a solvent that is suitable for *in situ* product removal. This would allow for overlay cultures in multiwell culturing plates and therefore minimize labor, evaporation and cytotoxicity of the ester. This assay can further be streamlined by using robotics for culturing as well as performing the assay.

With the ever-faster development of synthetic biology tools, gene synthesis and next-generation sequencing, the bottleneck of the design-build-test cycle in strain engineering has become the testing of the strains. Conventional methods for small molecule detection such as GC or HPLC do not match the throughput required to screen genetic libraries of clones generated by high-throughput genome editing tools. To screen large libraries transduction of target molecule concentrations into an easily measureable signal, such as luminescence, fluorescence or absorbance is sought. Such signals can easily be detected in multiwell plates but to survey large libraries single cell-based methods such as fluorescence activated cell sorting (FACS) or droplet-based microfluidic sorting are desired to match throughput.

This thesis describes the use of *K. marxianus* as host for ethyl acetate production. We chose *K. marxianus* strain CBS 6556 as a host for elucidating and engineering ester pathways because of its high ethyl acetate production on glucose, fast growth characteristics up to 45°C and the availability of a shotgun sequence. Because the family of *K. marxianus* contains a genetically and physiologically diverse group of strains, different strains might be chosen for processes other than ethyl acetate production. The results of the

ADH disruption library show that Adh1, 2, and 3 activities are essential for ethanol and ethyl acetate production. Under the conditions tested the cytosolic Adh2 appears to be the predominant enzyme for ethanol synthesis from acetaldehyde. Both Adh1 and Adh2 have been indicated in ethanol synthesis, explaining decreased ethanol synthesis in the disruption strains. The mitochondrial Adh3 however, has been thought to be important for the use of non-fermentable carbon sources and ethanol degradation in the mitochondria. Therefore, a decrease in ethanol and ethyl acetate in the ADH3 disruption strain was not anticipated. Another study suggests the importance of Adh3 for mitochondrial NADH oxidation and the ethanol-acetaldehyde shuttle. In this study, Adh3 is suggested to produce ethanol from acetaldehyde through NADH oxidation, thus reducing flux through the electron transport chain and subsequent ROS production. These two contradicting studies and the impact of Adh3 disruption on ethanol and ethyl acetate biosynthesis highlight the importance of further studies on Adh3 and its role in ethyl acetate synthesis and redox balance.

Disruption of the EAT1 in *K. marxianus* almost completely abolishes ethyl acetate biosynthesis while accumulating ethanol and acetaldehyde. The enzyme, belonging to a new class of alcohol acetyltransferases, has recently been discovered and only limited work has been done to characterize this class of enzymes. Disruption of a *S. cerevisiae* homolog led to the reduction of ethyl acetate by 50%, suggesting that Eat1 is also important for ethyl acetate production *in S. cerevisiae*. Not much is known about Eat1's substrate specificity, expression patterns or the role in yeast metabolism. Our findings suggest that Eat1 has broad substrate specificity towards a wide range of alcohol including isoamyl and phenyl ethyl alcohol. Thus Eat1 may be interesting for engineering longer chain ester biosynthesis pathways. In contrast to previous beliefs on ethyl acetate biosynthesis, we found that Eat1

and thus ethyl acetate biosynthesis is localized to the mitochondria, which allows for high ethyl acetate production based on high mitochondrial acetyl-CoA levels. Previous acetate ester engineering efforts (targeted to the cytosol or lipid droplets) have been hampered by low cytosolic acetyl-CoA availability. The mitochondrial localization of Eat1 enables access to the acetyl-CoA pool thus alleviating this limitation. The increase in ethyl acetate production when Eat1 is overexpressed suggests that this step is the bottleneck for ethyl acetate production in the wild type (Chapter 5). With increasing Eat1 expression levels and activity, acetyl-CoA will most likely be limiting and prohibit further improvements in titer.

In an attempt to increase acetyl-CoA and/or ethanol availability for ester synthesis we knocked down different electron transport chain and TCA cycle enzymes using CRISPR interference. The results showed that knockdown of one TCA cycle enzyme and three ETC enzymes significantly impact ester production. Multiplexing to knock down all four targets led to an even bigger increase in ethyl acetate production. Interestingly, only the quadruple knockout significantly deceased ethanol production, possibly indicating an increase of acetyl-CoA availability. In general, increased ester production could be caused by an increase in ethanol, an increase in acetyl-CoA availability or another cause such as cofactor balance. Previous research has suggested increased ethyl acetate production along with an increase in ethanol production when the electron transport chain is inhibited. Here ethanol production is possibly induced by excess in NADH. In contrast, inhibition of the TCA cycle flux only increased ethyl acetate production indicating a backup in the TCA cycle to free up acetyl-CoA. Additional studies such as metabolite profiling and transcriptomics would help elucidating the effect of ETC and TCA cycle inhibition and guide further engineering approaches.

Because ethyl acetate is directly derived from the central carbon metabolism, an in-depth knowledge of the *K. marxianus* metabolism is essential for successful engineering. Especially the redox state of the cell is of interest as ethyl acetate production is dependent on oxygen availability. Because high ethyl acetate production is native to the wild type of *K. marxianus*, the pathway is most likely subjected to stringent regulation. To achieve maximum ester yields in *K. marxianus* it is therefore essential to decipher and uncouple regulatory networks in the cell.

This work has laid the foundation for developing *K. marxianus* as biotechnological host for ester and possibly other chemicals production. Furthermore, the elucidation of the ethyl acetate production pathway has shed light on the so far elusive metabolism of *K. marxianus* and has enabled metabolic engineering of ester production in this yeast.