Engineering of metabolism and membrane transport in *Saccharomyces cerevisiae* for improved industrial performance

JASMINE M. BRACHER

Engineering of metabolism and membrane transport in *Saccharomyces cerevisiae* for improved industrial performance

Dissertation
for the purpose of obtaining the degree of doctor
at Delft University of Technology
by the authority of the Rector Magnificus, prof.dr.ir. T.H.J.J. van der Hagen,
chair of the Board for Doctorates
to be defended publicly on
Thursday 17 January 2019 at 10:00 o'clock

by

Jasmine Melanie BRACHER

Master of Science in Molecular Biology, University of Basel, Switzerland
born in Basel, Switzerland

This dissertation has been approved by the promotors.

Composition of the doctoral committee:

Rector Magnificus	Chairperson
Prof.dr. J.T. Pronk	Delft University of Technology, promotor
Prof.dr.ir. A.J.A. van Maris	KTH Royal Institute of Technology, promotor

Independent members:

Prof.dr. P.A.S. Daran-Lapujade	Delft University of Technology
Prof.dr. A.J.M. Driessen	Rijksuniversiteit Groningen
Prof.dr. J.H. de Winde	Universiteit Leiden
Dr. G.J. Smits	Universiteit van Amsterdam
Prof.dr. M.F. Gorwa-Grauslund	Lund University

Substitute member:

Prof.dr.ir. M.C.M. van Loosdrecht	Delft University of Technology

The research presented in this thesis was performed at the Industrial Microbiology Section, Department of Biotechnology, Faculty of Applied Sciences, Delft University of Technology, The Netherlands. This work was supported by the BE-Basic R&D Program, which was granted an FES subsidy from the Dutch Ministry of Economic Affairs, Agriculture and Innovation (EL&I). The BE-Basic project 'Omniyeast', within which this research was performed, received financial support from DSM.

Cover	Jasmine M. Bracher
Layout	Renate Siebes | Proefschrift.nu
Printed by	ProefschriftMaken
ISBN	978-94-6380-178-2

© 2018 Jasmine M. Bracher

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronically, mechanically, by photo-copying, recording or otherwise, without the prior written permission of the author.

Table of contents

SUMMARY 5

SAMENVATTING 11

CHAPTER 1 17
INTRODUCTION: Engineering the yeast *Saccharomyces cerevisiae* for 1st - and 2nd generation bioethanol production

CHAPTER 2 55
The *Penicillium chrysogenum* transporter *Pc*AraT enables high affinity, glucose insensitive L-arabinose transport in *Saccharomyces cerevisiae*

CHAPTER 3 87
Laboratory evolution of a glucose-phosphorylation-deficient, arabinose-fermenting *S. cerevisiae* strain reveals mutations in *GAL2* that enable glucose-insensitive L-arabinose uptake

CHAPTER 4 125
Reassessment of requirements for anaerobic xylose fermentation by engineered, non-evolved *Saccharomyces cerevisiae* strains

CHAPTER 5 155
Laboratory evolution of a biotin-requiring *Saccharomyces cerevisiae* strain for full biotin prototrophy and identification of causal mutations

OUTLOOK 187

ACKNOWLEDGEMENTS 191

CURRICULUM VITAE 195

LIST OF PUBLICATIONS 197

Summary

Nearly 200 parties signed and committed to the Paris agreement in 2017, which comprises the long-term goal to keep the average global temperature increase well below 2 degrees above pre-industrial levels. Virtually all possible scenarios drafted to reach this goal include a strongly increased use of biofuels for transport by land, sea and air. Bioethanol, whose production could, in principle and in contrast to fossil fuel production, involve a closed carbon cycle, is generated by microbial fermentation of sugars from plant-derived starch or agricultural waste. This liquid transport fuel provides a readily implementable alternative to fossil fuels as it combines the advantages of sustainable fuel production and compatibility with existing combustion engine technologies, without a requirement for time-consuming and expensive changes in our current infrastructure.

To date, bioethanol is the largest volume product of industrial biotechnology. 99% of this ethanol is generated via 1st generation processes, largely derived by fermentation of hydrolysed sugar cane or corn starch by bakers' yeast (*Saccharomyces cerevisiae*). So-called '2nd generation' bioethanol, for which the first commercial-scale plants are now starting up, is made by fermentation of sugars present in lignocellulosic biomass, typically harvested from agricultural waste streams, such as wheat straw or sugar beet pulp. Whilst such feedstocks enable a "food *and* fuel" scenario, their industrial implementation brings along additional challenges for yeasts and biotechnologists. Hydrolysis of lignocellulosic biomass, in particular the cellulose and hemicellulose fractions, releases a mixture of different sugars as well as inhibiting compounds that impair growth and viability of *S. cerevisiae.* Whilst glucose is the most abundant fermentable carbon source, the pentose sugar D-xylose can cover up to 30% of the total sugar content. The fraction of the pentose L-arabinose typically varies between 2 – 20%, depending on the feedstock used. Although

pentose sugars cannot be fermented to ethanol by wild-type *S. cerevisiae* strains, international research efforts over the past two decades yielded metabolic engineering strategies to enable anaerobic conversion of D-xylose and L-arabinose to ethanol by *S. cerevisiae.*

The expression of heterologous, D-xylose- and L-arabinose-isomerase based pathways from fungi or bacteria, together with over-expression of genes of the non-oxidative pentose phosphate pathway (PPP) and deletion of the unspecific aldose-reductase gene *GRE3* within *S. cerevisiae* allows this yeast to aerobically metabolize both sugars. The recent advances in metabolic engineering tools, such as CRISPR-Cas9-assisted genome editing, greatly advanced the construction and characterization of metabolically engineered *S. cerevisiae* strains with improved yields, kinetics and robustness in 2nd generation ethanol production processes.

CHAPTER 1 of this thesis summarizes the history and recent advances in metabolic engineering of *S. cerevisiae* for the production of bioethanol from an academic and industrial perspective.

Cellular uptake of the native substrate glucose by *S. cerevisiae* is highly efficient and involves complex regulation of the expression levels and *in vivo* activities of 20 hexose transporters, which together offer an adaptable range of transport affinities and capacities suited for various extracellular glucose levels. A similar system for the uptake of pentose sugars has not evolved within *S. cerevisiae.* Whilst D-xylose can enter the cells via multiple hexose transporters, L-arabinose transport depends on the single galactose transporter Gal2, which however exhibits a low affinity for L-arabinose and whose expression is strongly repressed by glucose, which is abundantly present in lignocellulosic hydrolysates. Hence, glucose is a potent competitive inhibitor of L-arabinose transport via Gal2. The combination of both characteristics delays the total fermentation time of a mix of glucose and L-arabinose due to sequential uptake of those sugars and causes a prominent “tailing effect”, reflected by very slowly decreasing L-arabinose concentrations towards the end of the fermentation. A low ethanol productivity, impairs inhibitor tolerance of the cells and negatively affects process economics.

The objective of **CHAPTER 2** of this thesis was to increase the affinity of L-arabinose transport in engineered, L-arabinose fermenting *S. cerevisiae* strains. To this end, the analysis of chemostat-based transcriptome data of L-arabinose-grown *Penicillium chrysogenum* cultures yielded a selection of candidate L-arabinose transporter genes. Their separate expression in an engineered *S. cerevisiae* strain lacking *GAL2,* and growth tests on L-arabinose led to the identification of one transporter that enabled growth within the described setting. In contrast to Gal2, this protein, termed *Pc*AraT, was shown to allow

L-arabinose transport even in presence of glucose, theoretically enabling co-fermentation of those two hydrolysate-borne carbon sources. Uptake experiments with radiolabelled sugars revealed a very high L-arabinose affinity of this proton-symporter and the absence of transport activity for both, glucose and xylose. Four hundred fifty-fold lower residual L-arabinose concentrations in chemostat cultures with engineered strains expressing *Pc*AraT, as compared to similar strains expressing Gal2, corroborated those results and confirmed the value of *Pc*AraT as a promising addition to engineered, L-arabinose fermenting *S. cerevisiae* strains.

*Pc*AraT conferred glucose-insensitive, high-affinity L-arabinose uptake which is specifically relevant for "mopping up" remaining L-arabinose towards the end of a fermentation. However, the transport capacity of this transporter alone did not allow fast consumption of high quantities of L-arabinose. The endogenous galactose transporter Gal2, confers L-arabinose transport with a satisfying transport capacity, but suffers from severe glucose repression.

The aim of **CHAPTER 3** of this thesis was to improve L-arabinose transport capacity of engineered, L-arabinose consuming *S. cerevisiae* in glucose-xylose-arabinose media in order to allow co-fermentation and ultimately reduce total fermentation times of mixtures of those sugars. To this end, a glucose-phosphorylation-negative, *Pc*AraT expressing, *S. cerevisiae* strain that could transport but not metabolize glucose, was generated and subjected to adaptive laboratory evolution in sequential anaerobic batch reactors. The latter contained glucose-xylose-arabinose media, thereby generating a selective advantage for cells capable of taking-up L-arabinose in presence of high concentrations of competing sugars. Whole- genome sequencing identified duplications of *GAL2* as well as mutations within this gene in multiple independently evolved single-colony isolates. The relevance of individual mutations was assessed both via reversion to wild-type alleles in evolved strains and by introduction of mutations into the native parental strain, as well as by quantitative transport assays using radiolabeled sugars. Whilst the mutated *GAL2* alleles revealed novel Gal2 characteristics in favour of L-arabinose transport, duplication of *GAL2* alone was shown to, in itself, not confer instantaneous anaerobic growth on L-arabinose. However, transport assays with $Gal2^{N376T,T89I}$ revealed glucose-insensitive L-arabinose transport characteristics and showed that this transporter completely lost its ability to transport glucose. The transporter encoded by a second evolved allele, $GAL2^{N376T}$, showed high L-arabinose transport capacities at the cost of reduced glucose transport abilities. Lastly, $Gal2^{T89I}$ showed a decreased glucose affinity while increasing overall L-arabinose affinity at the cost of transport capacity of both sugars. Restoring this allele to wild-type within an evolved strain nearly doubled the total fermentation time of 20 g L^{-1} L-arabinose within a glucose-arabinose-xylose mixture from 25 to 45 h. Deletion

of *Pc*AraT within the same evolved strain reduced overall growth rate from 0.12 h^{-1} to 0.1 h^{-1} and consequently increased the overall fermentation time. With the discovery and functional expression of *Pc*AraT in *S. cerevisiae,* and the selection for and characterization of mutated Gal2 alleles with improved L-arabinose transport characteristics in presence of glucose, valuable elements for improved L-arabinose fermentation were provided for the generation of novel yeast strains for 2nd generation bioethanol production.

With D-xylose accounting for up to 30% of the total sugar content of lignocellulosic biomass, anaerobic fermentation of D-xylose is paramount to ensure an economically viable process. Not surprisingly, a significant part of the research effort to engineer *S. cerevisiae* for biofuel production has been dedicated to anaerobic D-xylose fermentation. Remarkably, reports on metabolic engineering of D-xylose-fermenting, xylose-isomerase-based *S. cerevisiae* strains disagree as to whether extensive laboratory evolution is necessary to enable anaerobic growth on D-xylose or not. In particular, two previous publications from the TU Delft's industrial microbiology group, using the same laboratory strain and a similar metabolic engineering strategy, in one case reported instantaneous anaerobic xylose fermentation after targeted metabolic engineering while, in the second case, anaerobic growth on this pentose required anaerobic evolution and an additional mutation.

The objective of **CHAPTER 4** of this thesis was to systematically reassess the genetic requirements for anaerobic growth on D-xylose of engineered, CEN.PK-derived *S. cerevisiae* strains. To this end, reconstruction of D-xylose-metabolizing strains allowed for the identification of subtle differences in cultivation procedures that strongly affected anaerobic growth. Inoculum density was found to be a particularly important factor. Cultures of a particular engineered strain inoculated with 0.2 g biomass L^{-1} readily picked up anaerobic growth on D-xylose upon inoculation of the anaerobic bioreactor while, at an inoculum density of 0.02 g biomass L^{-1}, the same strain required an anaerobic lag phase of 7-8 d prior to initiation of anaerobic growth. Mimicking the higher initial CO_2 concentrations in high-inoculum density cultures by sparging low-inoculum-density cultures with CO_2-enriched nitrogen gas or by using L-aspartate instead of ammonium as nitrogen source, reproduced the phenotype of the former. Finally, omission of over-expression cassettes of the PPP paralogs *NQM1* and *TKL2* allowed the (re-)construction of a solely metabolically engineered strain that instantaneously fermented xylose when inoculated at low inoculum densities and sparged with pure nitrogen gas. Together, these observations resolved apparent contradictions in the literature on metabolic engineering strategies required for anaerobic growth of *S. cerevisiae* on D-xylose and expanded our knowledge on the potential relevance of CO_2 availability and anaplerotic carboxylation reactions on anaerobic D-xylose fermentation by metabolically engineered *S. cerevisiae* strains.

To ensure standardization and reproducibility of scientific results, yeast growth media in academia as well as industry are routinely supplemented with vitamins to prevent potential vitamin limitations from affecting the studied phenotype. Careful investigation of the exact vitamin requirements and prototrophies (i.e the ability to synthesize certain vitamins intracellularly), has the potential to significantly reduce media costs. Previously, whole genome sequencing of the laboratory strain *S. cerevisiae* CEN.PK113-7D revealed the presence of all biosynthesis genes assumed to be required to synthesize biotin (vitamin H/B7). Nevertheless, this strain showed only barely detectable growth in biotin-free media.

CHAPTER 5 of this thesis was inspired by the academic interest in unravelling the as yet unknown aspects of the biotin biosynthesis pathway in *S. cerevisiae*, and by the attractivity of biotin prototrophic *S. cerevisiae* strains for industrial and academic use. To this end, CEN.PK113-7D was selected in media lacking this expensive supplement. Using parallel sequential batch cultures and accelerostat regimes yielded biotin prototrophic strains, some of which grew as fast in the absence of biotin as in its presence. Whole-genome sequencing and reverse-engineering studies identified a massive amplification of *BIO1* to be predominantly responsible for improved growth in biotin-free media. The additional deletion of the genes encoding the transporters Tpo1 and/or Pdr12, which were found to be inactivated in evolved prototrophic strains, further improved the growth rates of a *BIO1*-overexpressing strain. The identification of metabolic engineering targets for obtaining biotin prototrophic yeasts in **CHAPTER 5** has the potential to improve process economics of a possibly large number of biotechnological products and to inspire future researchers to expand this method to other expensive growth factors.

Samenvatting

In 2017 tekenden bijna 200 partijen het klimaatakkoord in Parijs waarmee zij zich committeerden aan het doel om de gemiddelde wereldwijde temperatuurstijging onder de 2 graden te houden. Bijna alle mogelijke scenario's die gebaseerd zijn op dit doel omvatten een sterke toename in het gebruik van biobrandstof voor transport ter land, ter zee en in lucht. In tegenstelling tot fossiele brandstof maken biobrandstoffen, in principe, een gesloten koolstofcyclus mogelijk. De biobrandstof ethanol wordt gevormd door microbiële vergisting van plantaardig zetmeel, rietsuiker of reststromen uit de landbouw. Deze vloeibare brandstof is een direct implementeerbaar alternatief voor fossiele brandstof, omdat het de voordelen van duurzame brandstofproductie combineert met toepasbaarheid in bestaande verbrandingsmotoren zonder dat hiervoor dure modificaties van de huidige brandstofinfrastructuur nodig zijn.

Momenteel is bioethanol, binnen de industriële biotechnologie, het product dat met het grootste volume geproduceerd wordt. 99% van deze ethanol wordt geproduceerd in zogenaamde "eerste generatie"-processen waarin bakkersgist (*Saccharomyces cerevisiae*) de koolhydraten uit suikerriet of mais fermenteert. "Tweede generatie" bioethanol, waarvoor de eerste fabrieken op industriële schaal nu worden gestart, wordt gemaakt door vergisting van suikers die afkomstig zijn van lignocellulose-bevattende plantaardige biomassa. Deze biomassa is afkomstig van landbouwafval zoals tarwestro en suikerbietenpulp en maakt een "voedsel *en* brandstof" scenario mogelijk. De industriële implementatie van deze processen brengt echter wel uitdagingen met zich mee voor de gist en voor biotechnologen. Hydrolyse van lignocellulose-bevattende biomassa levert een mengsel van suikers afkomstig van cellulose en hemicellulose, maar ook verbindingen die de groei en levensvatbaarheid van *S. cerevisiae* negatief beïnvloeden.

Qua suikersamenstelling bevat de gehydrolyseerde lignocellulose, naast glucose, ook 30% D-xylose en 2 – 20% L-arabinose. Wildtype stammen van *S. cerevisiae* zijn niet in staat de pentoses D-xylose en L-arabinose te fermenteren. Daarom heeft internationaal onderzoek zich in de afgelopen twee decennia gericht op strategieën waarbij de stofwisseling van *S. cerevisiae* zo wordt aangepast dat de anaërobe omzetting van D-xylose en L-arabinose naar ethanol mogelijk wordt.

Een combinatie van genetische aanpassingen van *S. cerevisiae* geeft deze gist het vermogen de genoemde pentoses om te zetten. Hiervoor moeten stofwisselingsroutes gebaseerd op D-xylose- en L-arabinose-isomerase, afkomstig uit andere schimmels en bacteriën, gecombineerd worden met het tot overexpressie brengen van het niet-oxidatieve gedeelte van de pentosefosfaatroute en deletie van het niet-specifieke aldose-reductasegen *GRE3*. Ontwikkelingen in methoden om de stofwisseling van gist aan te passen, zoals het aanpassen van het genoom met behulp van CRISPR-Cas9, hebben de constructie en het karakteriseren van *S. cerevisiae* stammen waarin de kinetiek, robuustheid en ethanolopbrengst in tweede-generatie processen zijn verbeterd, mogelijk gemaakt.

HOOFSTUK 1 van dit proefschrift vat de geschiedenis en recente vooruitgang in het aanpassen van de stofwisseling van *S. cerevisiae* van productie van ethanol samen vanuit een academisch en industrieel perspectief.

Opname van glucose door *S. cerevisiae* is efficiënt en omvat een complexe regulatie van genen en de *in vivo* activiteit van 20 hexose-transporteiwitten. Deze transporteiwitten hebben een brede reikwijdte in affiniteit voor glucose en het vermogen om op verschillende concentraties van glucose te reageren. Hoewel een vergelijkbaar systeem voor de opname van pentoses niet is geëvolueerd in *S. cerevisiae*, hebben verschillende hexosetransport-eiwitten ook affiniteit voor D-xylose en maken transport van deze pentose mogelijk. L-arabinose transport is daarentegen vooral afhankelijk van één galactose-transport eiwit, Gal2. Dit transporteiwit heeft een lage affiniteit voor L-arabinose en de expressie van *GAL2* wordt sterk onderdrukt door glucose, dat veel voorkomt in gehydrolyseerd lignocellulose. Glucose is hierdoor een sterke en industrieel relevante remmer van L-arabinose opname door Gal2. De combinatie van lage affiniteit en repressie door glucose leidt tot een langere fermentatietijd, omdat hierdoor eerst glucose en dan pas L-arabinose opgenomen wordt. Dit wordt met name aan het einde van fermentatieprocessen duidelijk door een “naijleffect” dat zichtbaar een langzame afname van de L-arabinoseconcentratie. Dit naijleffect heeft negatieve gevolgen voor de productiviteit van de ethanol productie.

De doelstelling van **HOOFDSTUK 2** van dit proefschrift was het verbeteren van de affiniteit van L-arabinose transport in genetisch aangepaste, L-arabinosefermenterende

S. cerevisiae -stammen. Door het analyseren van transcriptoomdata van *Penicillium chrysogenum*, gekweekt in een chemostaat met L-arabinose als limiterend substraat, konden kandidaat-transporteiwitten voor deze suiker geïdentificeerd worden. De op deze manier geïdentificeerde genen werden elk tot expressie gebracht in een *S. cerevisiae*-stam waarin *GAL2* was verwijderd, waardoor onderzocht kon worden welke van de genen groei op L-arabinose mogelijk maakten. In tegenstelling tot Gal2 gaf het *Penicillium*-transporteiwit *Pc*AraT, de mogelijkheid om L-arabinose op te nemen in de aanwezigheid van glucose. Hierdoor konden, in theorie, deze twee suikers gelijktijdig gefermenteerd worden. Opname-experimenten met radioactief gemerkte suikers toonden aan dat *Pc*AraT een zeer hoge affiniteit voor L-arabinose had en dat dit eiwit geen glucose of D-xylose transporteerde. De residuele L-arabinoseconcentratie in chemostaatcultures was vierhonderdvijftig keer lager voor stammen die *Pc*AraT tot expressie brachten dan voor vergelijkbare stammen die L-arabinose alleen konden transporteren via Gal2. Deze waarneming ondersteunde de experimenten met radioactieve suikers en bevestigde dat *Pc*AraT een bijdrage kan leveren aan de constructie van nieuwe L-arabinose-fermenterende *S. cerevisiae* stammen.

De introductie van *Pc*AraT in *S. cerevisiae* zorgde voor een glucose-ongevoelige L-arabinose-opname met een hoge affiniteit, die vooral relevant is om de residuele L-arabinose aan het einde van een fermentatie op te nemen. Echter, expressie van *Pc*AraT maakte geen snel transport van L-arabinose mogelijk. Het endogene galactosetransporteiwit Gal2 kan L-arabinose wel snel transporteren, maar de expressie van Gal2 wordt onderdrukt door glucose.

Het doel van **HOOFDSTUK 3** van dit proefschrift was het verbeteren van de capaciteit van genetisch aangepaste, L-arabinose consumerende, *S. cerevisiae* om L-arabinose te transporteren. Hierbij werd gebruik gemaakt van glucose-xylose-arabinose medium om gelijktijdige fermentatie van deze suikers en, uiteindelijk, een afname van de totale fermentatietijd mogelijk te maken. Hiervoor werd *S. cerevisiae* stam gemaakt die *Pc*AraT tot expressie bracht maar geen glucose kon fosforyleren, waardoor deze stam glucose kon opnemen maar niet verder kon omzetten. Door evolutie in het laboratorium, waarbij gebruik werd gemaakt van een sequentiële anaërobe batchcultures op glucose-xylose-arabinose medium, kon geselecteerd worden op cellen die L-arabinose konden opnemen in de aanwezigheid van een hoge concentratie van de andere suikers. Na 'whole genome sequencing' werden duplicaties van het gen voor de *GAL2* transporteiwit en verschillende mutaties binnen dit gen aangetroffen in onafhankelijk geëvolueerde cellijnen. De relevantie van individuele mutaties werd onderzocht door deze ongedaan te maken in de verschillende geëvolueerde stammen en door deze mutaties in de wildtype stam te introduceren. Daarnaast werd transport kwantitatief bestudeerd met radioactief gemerkte suikers. Alleen Duplicatie van Gal2 was niet voldoende om directe anaërobe groei op L-arabinose

mogelijk te maken, maar een aantal mutaties resulteerde in Gal2 varianten met verbeterd L-arabinosetransport. Het allel *GAL2*$^{N371T,\ T89I}$ codeerde voor een glucose-onafhankelijk L-arabinose transport en leidde tot verlies van glucosetransport. Een tweede geëvolueerd allel, *GAL2*^{N376T}, had een verhoogde L-arabinosetransportcapaciteit, die ten koste ging van de capaciteit om glucose te transporteren. Een derde allel, *GAL2*^{T89I}, had een verlaagde affiniteit voor glucose terwijl de algehele affiniteit voor L-arabinose groter geworden was. Deze veranderingen gingen ten koste van de capaciteit om beide suikers te transporteren. Toen dit allel werd vervangen door een wildtype allel, verdubbelde de fermentatietijd van 20 g L^{-1} L-arabinose in het glucose-arabinose-xylose medium (van 25 naar 45 uur). Deletie van *Pc*AraT in dezelfde stam verlaagde de groeisnelheid van 0.12 uur^{-1} tot 0.1 uur^{-1}, waardoor de totale fermentatietijd toenam. Met de ontdekking en functionele expressie van *Pc*AraT in *S. cerevisiae* en de selectie voor en karakterisering van de gemuteerde Gal2-allelen met –verbeterde L-arabinose– transport in de aanwezigheid van glucose, zijn belangrijke elementen geleverd voor een verbeterde fermentatie van L-arabinose en de productie van nieuwe giststammen voor 2de generatie bioethanolproductie.

Omdat de suikers uit lignocellulose biomassa voor 30% bestaan uit D-xylose, is de anaërobe vergisting van deze suiker essentieel voor een productief economisch proces. Het is daarom niet verrassend dat een significant gedeelte van het onderzoek naar het gebruik van *S. cerevisiae* in de bioethanolproductie zich heeft gericht op anaërobe fermentatie van D-xylose. Publicaties over D-xylosefermenterende *S. cerevisiae*-stammen die xylose-isomerase tot expressie brengen, komen tot tegenstrijdige conclusies over de noodzaak van evolutie in het laboratorium om anaërobe groei op D-xylose mogelijk te maken. Twee publicaties van de Industriële Microbiologie-groep van de TU Delft, die dezelfde stam en vergelijkbare genetische modificaties gebruikten, rapporteerde in een studie directe anaërobe groei op xylose terwijl in de andere studie evolutie en een additionele mutatie nodig waren om anaërobe groei op deze pentose mogelijk te maken.

In **HOOFDSTUK 4** van dit proefschrift zijn de vereisten voor anaërobe groei op D-xylose van gemodificeerde *S. cerevisiae*-stammen met een CEN.PK achtergrond, opnieuw onder de loep genomen. De reconstructie van de stammen met het vermogen tot vergisting van D-xylose resulteerde in de identificatie van subtiele verschillen in de gebruikte kweekmethoden, die grote gevolgen hadden voor anaërobe groei. Vooral de dichtheid waarmee een culture geënt bleek hierbij van groot belang. Als een entdichtheid van 0.2 g biomassa L^{-1} gebruikt werd, begon de cultuur direct te groeien terwijl het 7–8 dagen durende voordat een cultuur met een entdichtheid van 0.02 g biomassa L^{-1} begon te groeien. Door het doorleiden van met CO_2 verrijkt stikstofgas door cultures met een lage entdichtheid culture kon de initiële, hogere, CO_2^- concentratie in cultures met een hoge entdichtheid gesimuleerd worden. In deze cultures, en ook in cultures waarin L-aspartaat in plaats van

ammonium als stikstofbron werd gebruikt, begon anaërobe groei vrijwel meteen Ook kon door het weglaten van overexpressiecassettes voor *NQM1* en *TKL2*, twee paralogen van pentosefosfaatroutegenen, kon een genetisch aangepaste stam geconstrueerd worden die direct D-xylose begon te fermenteren bij een lage entdichtheid, zelfs wanneer de reactor met puur stikstofgas doorborreld werd. Deze waarnemingen losten de schijnbare tegenstellingen in de literatuur op over strategieën om met genetische aanpassingen *S. cerevisiae* anaeroob D-xylose te laten fermenteren. Daarnaast vergrootten deze resultaten de kennis over de mogelijke relevantie van de beschikbaarheid van CO_2 en anaplerotische carboxyleringsreacties op anaerobe D-xylose-fermentatie door genetisch gemodificeerde *S. cerevisiae*-stammen.

Om de standaardisering en reproduceerbaarheid van wetenschappelijke resultaten te bewaken, worden in zowel academische als industriële context, vitamines aan kweekmedia toegevoegd. Hiermee wordt limitatie van deze componenten tijdens de kweek van gist voorkomen. Onderzoek naar de precieze behoefte aan vitamines en prototrophieën (het vermogen van de cel bepaalde vitamines te maken), kan de kosten van medium mogelijk significant verminderen. In het genoom van de laboratoriumstam *S. cerevisiae* CEN.PK113-7D kon de complete biosynthetische route voor de vitamine biotine synthese geïdentificeerd worden. Desondanks groeide deze stam heel slecht, met een bijna onmeetbaar lage snelheid, in medium zonder biotine.

HOOFDSTUK 5 van dit proefschrift was geïnspireerd door een academische interesse om de nog onbekende aspecten van biotinesynthese in *S. cerevisiae* te ontrafelen. Daarnaast is een biotine-prototrofe stam vanuit zowel industrieel als academisch perspectief interessant. Door gebruik van parallelle sequentiële batchcultures en een accelerostaatregime in media zonder dit dure supplement, konden variantie van de CEN.PK113-7D-stam geselecteerd worden die snel groeiden in de afwezigheid van biotine. In de geëvolueerde stammen werd met behulp van ‘whole-genome-sequencing’ en ‘reverse metabolic engineering’ een amplificatie van het gen *BIO1* geïdentificeerd, die de verbeterde groei in medium zonder biotine grotendeels verklaarde. Deletie van genen die codeerden voor de transporteiwitten Tpo1 en/of Pdr12, die geïnactiveerd waren in geëvolueerde stammen, zorgde voor een verdere verbetering van de groei van stammen die *BIO1* tot overexpressie brachten. De identificatie van kandidaat genen voor het maken van biotineprototrofe stammen in **HOOFDSTUK 5** kan de kosten van veel biotechnologische producten positief beïnvloeden en toekomstige onderzoekers inspireren de gebruikte methoden in te zetten voor andere dure mediumsupplementen.

1.

INTRODUCTION

Engineering the yeast *Saccharomyces cerevisiae* for 1st - and 2nd generation bioethanol production

Adapted from the publication entitled
"*Saccharomyces cerevisiae* strains for second-generation ethanol production: from academic exploration to industrial implementation" by Mickel L.A. Jansen, Jasmine M. Bracher, Ioannis Papapetridis, Maarten D. Verhoeven, Hans de Bruijn, Paul P. de Waal, Antonius J.A. van Maris, Paul Klaassen and Jack T. Pronk in
FEMS Yeast Research, Volume 17, Issue 5, August 2017
(https://doi.org/10.1093/femsyr/fox044)

General introduction

Alcoholic fermentation is a key catabolic process in most yeasts and in many fermentative bacteria, which concentrates the heat of combustion of carbohydrates into two thirds of their carbon atoms ($(CH_2O)_n \rightarrow \frac{1}{3}n\ C_2H_6O + \frac{1}{3}n\ CO_2$). Its product, ethanol, has been used as an automotive fuel for over a century (Bernton *et al.* 1982).

With an estimated global production of 100 Mton (Renewable Fuels Association 2017), ethanol is the largest-volume product in industrial biotechnology. Its production is, currently, mainly based on fermentation of cane sugar or hydrolysed corn starch with the yeast *Saccharomyces cerevisiae.* Such 'first generation' bioethanol processes are characterized by high ethanol yields on fermentable sugars (> 90% of the theoretical maximum yield of 0.51 g ethanol·(g hexose sugar)$^{-1}$), ethanol titers of up to 21% (w/w) and volumetric productivities of 2 to 3 kg·m^{-3}·h^{-1} (Thomas and Ingledew 1992, Della-Bianca *et al.* 2013, Lopes *et al.* 2016).

Over the past two decades, a large international effort, involving researchers in academia, research institutes and industry, aimed to access abundantly available agricultural and forestry residues, as well as fast-growing energy crops, as alternative feedstocks for fuel ethanol production (Rude and Schirmer 2009). Incentives for this effort, whose relative impact depends on geographical location and varies over time, include reduction of the carbon footprint of ethanol production (Otero *et al.* 2007), prevention of competition with food production for arable land (Nordhoff 2007, Tenenbaum 2008), energy security in fossil-fuel importing countries (Farrell *et al.* 2006) and development of rural economies (Kleinschmidt 2007). Techno-economic forecasts of low-carbon scenarios for global energy supply almost invariably include liquid biofuels as a significant contributor (Yan *et al.* 2010). Moreover, successful implementation of economically and environmentally sustainable 'second generation' bioethanol processes can pave the way for similar processes to produce other biofuels and commodity chemicals (Pereira *et al.* 2015).

In contrast to starch, a plant storage carbohydrate that can be easily hydrolysed, the major carbohydrate polymers in lignocellulosic plant biomass (cellulose, hemicellulose and, in some cases, pectin) contribute to the structure and durability of stalks, leaves and roots (Hahn-Hägerdal *et al.* 2006). Consistent with these natural functions and with their chemical diversity and complexity, mobilization of these polymers by naturally occurring cellulose-degrading microorganisms requires complex arrays of hydrolytic enzymes (Lynd *et al.* 2002, Van den Brink and de Vries 2011).

The second-generation ethanol processes that are now coming on line at demonstration and full commercial scale (**Table 1**) are mostly based on fermentation of lignocellulosic biomass hydrolysates by engineered strains of *S. cerevisiae*. While this yeast has a strong

Table 1. Overview of operational commercial-scale (demonstration) plants for second-generation bioethanol production. Data for US and Canada reflect status in May 2017 (source: (UNCTAD 2016, Ethanol Producer Magazine 2017), data for other countries (source: (UNCTAD 2016, Ethanol Producer Magazine 2017) reflect status in 2016.

Company/Plant	Country	Feedstock	Capacity $ML \cdot y^{-1}$
DuPont Cellulosic Ethanol LLC - Nevada	USA (IA)	Corn stover	113.6
Poet-DSM Advanced Biofuels LLC - Project Liberty[1]	USA (IA)	Corn cobs/corn stover	75.7
Quad County Cellulosic Ethanol Plant	USA (IA)	Corn fiber	7.6
Fiberight Demonstration Plant	US (VA)	Waste stream	1.9
ICM Inc. Pilot integrated Cellulosic Biorefinery	US (MO)	Biomass crops	1.2
American Process Inc. – Thomaston Biorefinery	USA (GA)	Other	1.1
ZeaChem Inc. – Demonstration plant	US (OR)	Biomass crops	1.0
Enerkem Alberta Biofuels LP	Canada (AB)	Sorted municipal solid waste	38
Enerkem Inc.-Westbury	Canada (QC)	Woody biomass	5.0
Iogen Corporation	Canada (ON)	Crop residue	2.0
Woodlands Biofuels Inc. – Demonstration plant	Canada (ON)	Woody biomass	2.0
GranBio	Brazil	Bagasse	82.4
Raizen	Brazil	Sugarcane bagasse/ straw	40.3
Longlive Bio-technology Co. Ltd. – Commercial demo	China	Corn cobs	63.4
Mussi Chemtex / Beta Renewables	Italy	Arundo donax, rice & wheat straw	75
Borregaard Industries AS – ChemCell Ethanol	Norway	Wood pulping residues	20

[1] With expansion capacity to 94.6 ML per year

track record in first-generation bioethanol production and its amenability to genetic modification is excellent, *S. cerevisiae* cannot hydrolyse cellulose or hemicellulose. Therefore, in conventional process configurations for second-generation bioethanol production, the fermentation step is preceded by chemical/physical pretreatment and enzyme-catalysed hydrolysis by cocktails of fungal hydrolases, which can either be produced on- or off site (**Figure 1**, (Sims-Borre 2010). Alternative process configurations, including simultaneous saccharification and fermentation (SSF) and consolidated bioprocessing (CBP) by yeast cells expressing heterologous hydrolases are intensively investigated (Olson *et al.* 2012, Den Haan *et al.* 2015). However, the high temperature optima of fungal enzymes and low productivity of heterologously expressed hydrolases in *S. cerevisiae* have so far precluded large-scale implementation of these alternative strategies for lignocellulosic ethanol production (Vohra *et al.* 2014, Den Haan *et al.* 2015).

Over the past decade, academic and industrial research efforts have enabled the development of metabolic engineering strategies for fermentation of lignocellulosic hydrolysates

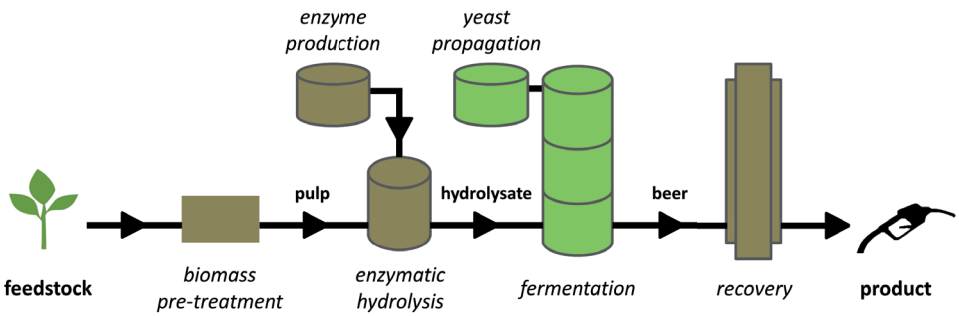


Figure 1. Schematic process-flow diagram for ethanol production from lignocellulose, based on physically separated processes for pretreatment, hydrolysis and fermentation, combined with on-site cultivation of filamentous fungi for production of cellulolytic enzymes and on-site propagation of engineered pentose-fermenting yeast strains.

with engineered *S. cerevisiae* strains and in implementing these in advanced industrial strain platforms (**Table 1**). This chapter discusses challenges and progress in metabolic engineering of *S. cerevisiae* for fermentation of lignocellulosic biomass, as well as developments in the industrial implementation of such 'second generation' ethanol production processes. Furthermore, the concept and potential of engineered, vitamin prototrophic yeast strains to improve biotechnological process economics is discussed.

Fermenting lignocellulosic hydrolysates: challenges for yeast strain development

A wide range of agricultural and forestry residues, as well as energy crops, are being considered as feedstocks for bioethanol production (Khoo 2015). Full-scale and demonstration plants using raw materials such as corn stover, sugar-cane bagasse, wheat straw and switchgrass are now in operation (**Table 1**). These lignocellulosic feedstocks have different chemical compositions, which further depend on factors such as seasonal variation, weather and climate, crop maturity and storage conditions (Kenney *et al.* 2013). Despite this variability, common features of feedstock composition and biomass-deconstruction methods generate several generic challenges that have to be addressed in the development of yeast strains for second-generation bioethanol production.

Pentose fermentation

For large-volume products such as ethanol, maximizing the product yield on feedstock and, therefore, efficient conversion of all potentially available substrate molecules in the feedstock is of paramount economic importance (Lin and Tanaka 2006). In addition to readily fermentable hexoses such as glucose and mannose, lignocellulosic biomass

contains substantial amounts of D-xylose and L-arabinose. These pentoses, derived from hemicellulose and pectin polymers in plant biomass, cannot be fermented by wild-type *S. cerevisiae* strains. D-xylose and L-arabinose typically account for 10 to 25% and 2 to 3%, respectively, of the carbohydrate content of lignocellulosic feedstocks (Lynd 1996). However, in some feedstocks, such as corn fiber hydrolysates and sugar beet pulp, the L-arabinose content can be up to ten-fold higher (Grohmann and Bothast 1994, Grohmann and Bothast 1997). Early studies already identified metabolic engineering of *S. cerevisiae* for efficient, complete pentose fermentation as key prerequisite for its application in second-generation ethanol production (Bruinenberg *et al.* 1983, Kötter *et al.* 1990, Hahn-Hägerdal *et al.* 2001, Sedlak and Ho 2001).

Acetic acid inhibition

Since hemicellulose is acetylated (Van Hazendonk *et al.* 1996), its complete hydrolysis inevitably results in the release of acetic acid. Bacterial contamination during biomass storage, pretreatment and/or fermentation may further increase the acetic acid concentrations to which yeasts are exposed in the fermentation process. First-generation bioethanol processes are typically run at pH values of 4 to 5 to counter contamination with lactic acid bacteria (Beckner *et al.* 2011). At these low pH values, undissociated acetic acid (pKa = 4.76) easily diffuses across the yeast plasma membrane. In the near-neutral pH environment of the yeast cytosol, the acid readily dissociates and releases a proton, which forces cells to expend ATP for proton export via the plasma-membrane ATPase to prevent cytosolic acidification (Verduyn *et al.* 1992, Axe and Bailey 1995, Pampulha and Loureiro-Dias 2000). The accompanying accumulation of the acetate anion in the cytosol can cause additional toxicity effects (Russel 1992, Palmqvist and Hahn-Hägerdal 2000b, Ullah *et al.* 2013). Acetic acid concentrations in some lignocellulosic hydrolysates exceed 5 g L^{-1}, which can cause strong inhibition of anaerobic growth and sugar fermentation by *S. cerevisiae* (Taherzadeh *et al.* 1997). Acetic acid tolerance at low culture pH is therefore a key target in yeast strain development for second-generation ethanol production.

Inhibitors formed during biomass deconstruction

In biomass deconstruction, a trade-off exists between the key objective to release all fermentable sugars at minimal process costs and the need to minimize generation and release of compounds that compromise yeast performance. Biomass deconstruction generally encompasses three steps: (i) size reduction to increase surface area and reduce degree of polymerization, (ii) thermal pretreatment, often at low pH and high pressure, to disrupt the crystalline structure of cellulose while already (partly) solubilizing hemicellulose and/or lignin and (iii) hydrolysis with cocktails of fungal cellulases and hemicellulases to

release fermentable sugars (Hendriks and Zeeman 2009, Silveira *et al.* 2015, Narron *et al.* 2016). Several inhibitors of yeast performance are generated in chemical reactions that occur during biomass deconstruction and, especially, in high-temperature pretreatment. 5-Hydroxymethyl-2-furaldehyde (HMF) and 2-furaldehyde (furfural) are formed when hexoses and pentoses, respectively, are exposed to high temperature and low pH (Dunlop 1948, Ulbricht *et al.* 1984, Palmqvist and Hahn-Hägerdal 2000b). These furan derivatives inhibit yeast glycolysis, alcoholic fermentation and TCA cycle (Banerjee *et al.* 1981, Modig *et al.* 2002, Sárvári Horváth *et al.* 2003) while, additionally, depleting intracellular pools of NAD(P)H and ATP (Almeida *et al.* 2007). Their further degradation, during biomass deconstruction, yields formic acid and levulinic acid (Dunlop 1948, Ulbricht *et al.* 1984), whose inhibitory effects overlap with those of acetic acid (Palmqvist and Hahn-Hägerdal 2000b).

Inhibitor profiles of hydrolysates depend on biomass structure and composition as well as on the type and intensity of the biomass deconstruction method used (Almeida *et al.* 2007, Kumar *et al.* 2009). During pressurized pretreatment at temperatures above 160 °C, phenolic inhibitors are generated by partial degradation of lignin. This diverse class of inhibitors includes aldehydes, ketones, alcohols and aromatic acids (Almeida *et al.* 2007). Ferulic acid, a phenolic compound that is an integral part of the lignin fraction of herbaceous plants (Lawther *et al.* 1996, Klinke *et al.* 2002) is a potent inhibitor of *S. cerevisiae* fermentations (Larsson *et al.* 2000). The impact of phenolic inhibitors on membrane integrity and other cellular functions depends on the identity and position of functional groups and carbon-carbon double bonds (Adeboye *et al.* 2014).

Concentrations of inorganic salts in hydrolysates vary depending on the feedstock used (Klinke *et al.* 2004). Moreover, high salt concentrations in hydrolysates can originate from pH adjustments during pretreatment (Jönsson *et al.* 2013). Salt- and osmotolerance can therefore be important additional requirements in yeast strain development (Casey *et al.* 2013).

The inhibitors in lignocellulosic hydrolysates do not always act independently but can exhibit complex synergistic effects, both with each other and with ethanol (Taherzadeh *et al.* 1999, Palmqvist and Hahn-Hägerdal 2000b, Liu *et al.* 2004), while their impact can also be modulated by the presence of water-insoluble solids (Koppram *et al.* 2016). Furthermore, their absolute and relative impact can change over time due to variations in feedstock composition, process modifications, or malfunctions in biomass deconstruction. While process adaptations to detoxify hydrolysates have been intensively studied (Sivers *et al.* 1994, Palmqvist and Hahn-Hägerdal 2000a, Canilha *et al.* 2012, Jönsson *et al.* 2013), the required additional unit operations typically result in a loss of fermentable sugar and are generally considered to be too expensive and complicated. Therefore, as research on

optimization of biomass deconstruction processes continues, tolerance to the chemical environments generated by current methods is a key design criterion for yeast strain development.

Yeast strain development for second-generation ethanol production: key concepts

For almost three decades, yeast metabolic engineers have vigorously explored strategies to address the challenges outlined above. This quest benefited from rapid technological development in genomics, genome editing, evolutionary engineering and protein engineering. **Box 1** lists key technologies and examples of their application in research on yeast strain development for second-generation ethanol production.

Xylose fermentation

Efficiently linking D-xylose metabolism to glycolysis requires two key modifications of the *S. cerevisiae* metabolic network (**Figure 2**) (Jeffries and Jin 2004, Van Maris *et al.* 2007): introduction of a heterologous pathway that converts D-xylose into D-xylulose and, simultaneously, alleviation of the limited capacity of the native *S. cerevisiae* xylulokinase and non-oxidative pentose-phosphate pathway (PPP). Two strategies for converting D-xylose into D-xylulose have been implemented in *S. cerevisiae*: (i) simultaneous expression of

Box 1. Overview of key technologies used for development of *Saccharomyces cerevisiae* strains for second-generation bioethanol production (left column) and examples of their application (right column).

Metabolic engineering Application of recombinant-DNA techniques for the improvement of catalytic and regulatory processes in living cells, to improve and extend their applications in industry (Bailey 1991).	Metabolic engineering of pentose-fermenting strains commenced with the functional expression of pathways for XR/XDH- (Kötter and Ciriacy 1993, Tantirungkij *et al.* 1993) or XI-based (Kuyper *et al.* 2003) xylose utilization and pathways for isomerase-based arabinose utilization (Becker and Boles 2003, Wisselink *et al.* 2007). Further research focused on improvement of pathway capacity (Kuyper *et al.* 2005a, Wiedemann and Boles 2008), engineering of sugar transport (Fonseca *et al.* 2011, Subtil and Boles 2011, Nijland *et al.* 2014, Nijland *et al.* 2016), redox engineering to decrease byproduct formation and increase ethanol yield (Roca *et al.* 2003, Sonderegger and Sauer 2003, Watanabe *et al.* 2005, Guadalupe-Medina *et al.* 2010, Yu *et al.* 2010, Wei *et al.* 2013, Henningsen *et al.* 2015, Papapetridis *et al.* 2016, Zhang *et al.* 2016a) and expression of alternative pathway enzymes (Brat *et al.* 2009, Ota *et al.* 2013). Expression of heterologous hydrolases provided the first steps towards consolidated bioprocessing (Ha *et al.* 2011a, Ilmén *et al.* 2011, Sadie *et al.* 2011, Den Haan *et al.* 2015).

Box 1 continues on next page.

Box 1 – *Continued*

Evolutionary engineering Application of laboratory evolution to select for industrially relevant traits (Sauer 2001). Also known as adaptive laboratory evolution (ALE).	Evolutionary engineering in repeated-batch and chemostat cultures has been intensively utilized to improve growth and fermentation kinetics on pentoses (e.g., (Sonderegger and Sauer 2003, Kuyper *et al.* 2005b, Wisselink *et al.* 2009, Garcia Sanchez *et al.* 2010, Zhou *et al.* 2012, Demeke *et al.* 2013a, Kim *et al.* 2013, Lee *et al.* 2014) and inhibitor tolerance (Wright *et al.* 2011, Koppram *et al.* 2012, Almario *et al.* 2013, Smith *et al.* 2014, González-Ramos *et al.* 2016).
Whole genome (re)sequencing Determination of the entire DNA sequence of an organism.	Availability of a high-quality reference genome sequence is essential for experimental design in metabolic engineering. When genomes of strains that have been obtained by non-targeted approaches (e.g. evolutionary engineering or mutagenesis) are (re)sequenced, the relevance of identified mutations can subsequently be tested by their reintroduction in naïve strains, non-evolved strains and/or by classical genetics (reverse engineering; (Oud *et al.* 2012)). This approach has been successfully applied to identify mutations contributing to fast pentose fermentation (Nijland *et al.* 2014, dos Santos *et al.* 2016, Hou *et al.* 2016a) and inhibitor tolerance (e.g., (Pinel *et al.* 2015, González-Ramos *et al.* 2016).
Quantitative trait loci (QTL) analysis QTL identifies alleles that contribute to (complex) phenotypes based on their meiotic co-segregation with a trait of interest (Liti and Louis 2012, Wilkening *et al.* 2014). In contrast to whole-genome sequencing alone, QTL analysis can identify epistatic interactions.	QTL analysis currently enables resolution to gene or even nucleotide level (Swinnen *et al.* 2012). QTL analysis has been used to identify alleles contributing to high-temperature tolerance (Sinha *et al.* 2006), ethanol tolerance (Swinnen *et al.* 2012) and improved ethanol-to-glycerol product ratios (Hubmann *et al.* 2013). The requirement of QTL analysis for mating limits its applicability in aneuploidy and/or poorly sporulating industrial *S. cerevisiae* strains.
Protein engineering Modification of the amino acid sequences of proteins with the aim to improve their catalytic properties, regulation and/or stability in industrial contexts (Marcheschi *et al.* 2013).	Protein engineering has been used to improve the pentose-uptake kinetics, reduce the glucose sensitivity and improve the stability of yeast hexose transporters (e.g., (Farwick *et al.* 2014, Wang *et al.* 2015, Li *et al.* 2016b, Young *et al.* 2014, Reznicek *et al.* 2015, Shin *et al.* 2015, Nijland *et al.* 2016)). The approach has been utilized to improve the redox cofactor specificity of XR and/or XDH to decrease xylitol formation (Petschacher *et al.* 2005, Watanabe *et al.* 2005, Watanabe *et al.* 2007, Petschacher and Nidetzky 2008, Krahulec *et al.* 2009). Directed evolution of xylose isomerase yielded XI variants with increased enzymatic activity (Lee *et al.* 2012). Directed evolution of native yeast dehydrogenases has yielded strains with increased HMF tolerance (Moon and Liu 2012).
Genome editing Where 'classical' genetic engineering encompass iterative, one-by-one introduction of genetic modifications, genome editing techniques enable simultaneous introduction of multiple (types of) modifications at different genomic loci (Sander and Joung 2014).	The combination of CRISPR-Cas9-based genome editing (DiCarlo *et al.* 2013, Mans *et al.* 2015) with *in vivo* assembly of DNA fragments has enabled the one-step introduction of all genetic modifications needed to enable *S. cerevisiae* to ferment xylose (Tsai *et al.* 2015, Shi *et al.* 2016, Verhoeven *et al.* 2017). Recent developments have enabled the application of the system in industrial backgrounds (Stovicek *et al.* 2015). CRISPR-Cas9 has been used in reverse engineering studies to rapidly introduce multiple single-nucleotide mutations observed in evolutionary engineering experiments in naïve strains (e.g., (van Rossum *et al.* 2016)).

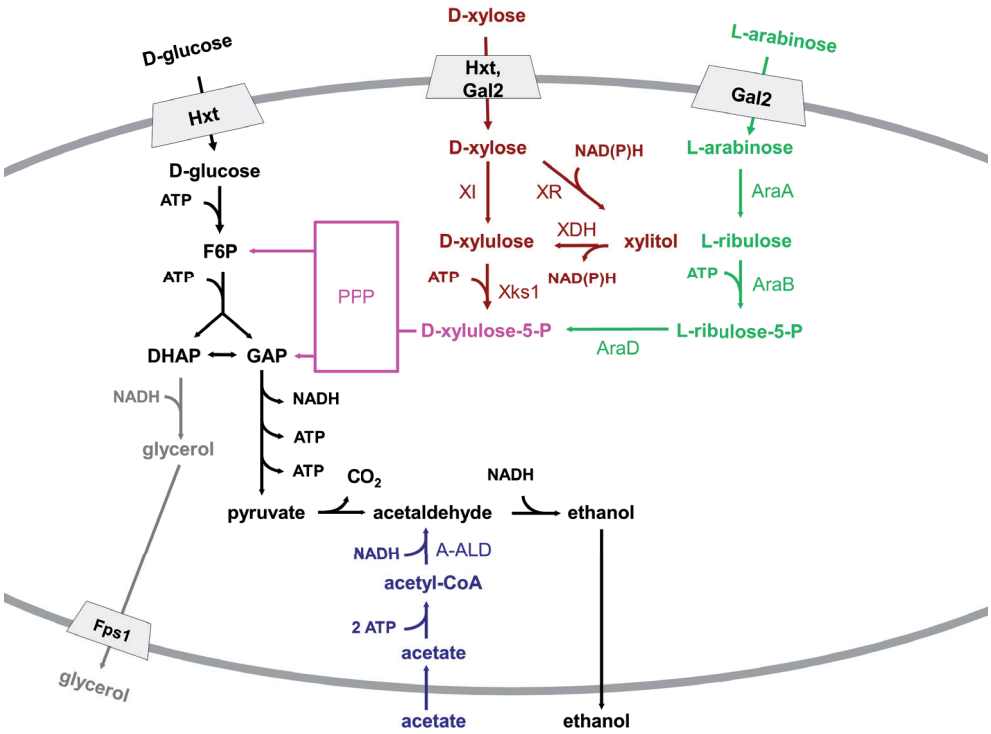


Figure 2. Key strategies for engineering carbon and redox metabolism in *S. cerevisiae* strains for alcoholic fermentation of lignocellulosic feedstocks. Colours indicate the following pathways and processes: **Black**: native *S. cerevisiae* enzymes of glycolysis and alcoholic fermentation; **Magenta**: native enzymes of the non-oxidative pentose-phosphate pathway (PPP), overexpressed in pentose-fermenting strains; **Red**: conversion of D-xylose into D-xylulose-5-phosphate by heterologous expression of a xylose isomerase (XI) or combined expression of heterologous xylose reductase (XR) and xylitol dehydrogenase (XDH), together with the overexpression of (native) xylulokinase (Xks1); **Green**: conversion of L-arabinose into D-xylulose-5-phosphate by heterologous expression of a bacteria AraA/AraB/AraD pathway; **Blue**: expression of a heterologous acetylating acetaldehyde dehydrogenase (A-ALD) for reduction of acetic acid to ethanol; **Grey**: native glycerol pathway.

heterologous xylose reductase (XR) and xylitol dehydrogenase (XDH) and (ii) expression of a heterologous xylose isomerase (XI).

The first *S. cerevisiae* strains engineered for xylose utilization were based on expression of XR and XDH from the xylose-metabolising yeast *Scheffersomyces stipitis* (Kötter and Ciriacy 1993). Due to the non-matching redox-cofactor preferences of these enzymes, these strains produced large amounts of the by-product D-xylitol (Kötter and Ciriacy 1993, Hahn-Hägerdal *et al.* 2001, Jeffries 2006). Modification of these cofactor preferences by protein engineering resulted in reduced xylitol formation under laboratory conditions (Watanabe *et al.* 2007, Runquist *et al.* 2010a). A much lower xylitol formation by XR/XDH-based strains in lignocellulosic hydrolysates was attributed to NADH-dependent reduction of furfural, which may contribute to *in situ* detoxification of this inhibitor (Moniruzzaman *et*

al. 1997, Wahlbom and Hahn-Hägerdal 2002, Sedlak and Ho 2003, Katahira *et al.* 2006, Karhumaa *et al.* 2007). A potential drawback of XR/XDH-based strains for application in large-scale anaerobic processes is that, even after prolonged laboratory evolution, their anaerobic growth rates are very low (Sonderegger and Sauer 2003).

Combined expression of a fungal XI (Harhangi *et al.* 2003) and overexpression of the native *S. cerevisiae* genes encoding xylulokinase and non-oxidative PPP enzymes enabled anaerobic growth of a laboratory strain on D-xylose. In anaerobic cultures of this strain, in which the aldose-reductase encoding *GRE3* gene was deleted to eliminate xylitol formation, ethanol yields on D-xylose were the same as on glucose (Kuyper *et al.* 2005a). This metabolic engineering strategy has been successfully applied in different *S. cerevisiae* genetic backgrounds and/or with different XI genes enabling aerobic xylose consumption (Brat *et al.* 2009, Madhavan *et al.* 2009, Ha *et al.* 2011b, Dun 2012, Hector *et al.* 2013, Hou *et al.* 2016b).

Most studies, except for Kuyper *et al.* (2005), using a strategy similar to the above-mentioned one, report that anaerobic growth on xylose requires subsequent evolutionary engineering under anaerobic or oxygen-limited conditions (Kuyper *et al.* 2005a, Brat *et al.* 2009, Parachin *et al.* 2011, Zhou *et al.* 2012, Demeke *et al.* 2013b, Hector *et al.* 2013, Parreiras *et al.* 2014, Verhoeven *et al.* 2017). A reassessment of the requirements for anaerobic D-xylose fermentation of CEN.PK-based *S. cerevisiae* strains concerning strain design and cultivation conditions, is presented in **Chapter 4** of this thesis.

Laboratory evolution (**Box 1**) for faster D-xylose fermentation and analysis of evolved strains identified high-level expression of XI as a major contributing factor (Zhou *et al.* 2012, Demeke *et al.* 2015, Hou *et al.* 2016a). Multi-copy introduction of XI expression cassettes, optimization of their codon usage, and mutagenesis of their coding sequences have contributed to higher D-xylose fermentation rates (Brat *et al.* 2009, Lee *et al.* 2012, Crook *et al.* 2016). Whole-genome sequencing of evolved D-xylose-fast-fermenting strains expressing *Piromyces* XI identified mutations affecting intracellular homeostasis of Mn^{2+}, a preferred metal ion for this XI (Verhoeven *et al.* 2017). Other mutations affected stress-response regulators and, thereby, increased expression of yeast chaperonins that assisted functional expression of XI (Hou *et al.* 2016a). Consistent with this observation, co-expression of the *Escherichia coli* GroEL and GroES chaperonins enabled *in vivo* activity of *E. coli* XI in *S. cerevisiae* (Xia *et al.* 2016). A positive effect of mutations in the *PHO13* phosphatase gene on xylose fermentation rates in XI- and XR/XDH-based strains has been attributed to transcriptional upregulation of PPP-related genes by an as yet unknown mechanism (Ni *et al.* 2007, Van Vleet *et al.* 2008, Bamba *et al.* 2016, Xu *et al.* 2016). Additionally, Pho13 has been implicated in dephosphorylation of the PPP intermediate sedoheptulose-7-phosphate (Xu *et al.* 2016). For other mutations in evolved strains, e.g.

in genes involved in iron-sulfur cluster assembly and in the MAP-kinase signaling pathway (dos Santos *et al.* 2016, Sato *et al.* 2016), the mechanisms by which they affect D-xylose metabolism remain to be identified.

Arabinose fermentation

The metabolic engineering strategy for constructing L-arabinose-fermenting *S. cerevisiae* is based on heterologous expression of a bacterial pathway for conversion of L-arabinose into xylulose-5-phosphate, involving L-arabinose isomerase (AraA), L-ribulokinase (AraB) and L-ribulose-5-phosphate-4-epimerase (AraD) (Lee *et al.* 1986). Together with the non-oxidative PPP and glycolysis, these reactions enable redox-cofactor-balanced alcoholic fermentation of L-arabinose (**Figure 2**).

Combined expression of *Bacillus subtilis* or *B. licheniformis araA* and *E. coli araBD* (Becker and Boles 2003, Bettiga *et al.* 2008, Wiedemann and Boles 2008) allowed aerobic growth of *S. cerevisiae* on L-arabinose. Anaerobic growth of *S. cerevisiae* on arabinose was first achieved by expressing the *Lactobacillus plantarum araA, B* and *D* genes in an XI-based xylose-fermenting strain that already overexpressed the enzymes of the non-oxidative PPP (**Figure 2**), followed by evolutionary engineering under anaerobic conditions (Wisselink *et al.* 2007). Increased expression levels of *GAL2*, which encodes a galactose transporter that also transports L-arabinose (Kou *et al.* 1970), was essential for L-arabinose fermentation (Becker and Boles 2003, Wisselink *et al.* 2010, Subtil and Boles 2011, Subtil and Boles 2012). Increased expression of the transaldolase and transketolase isoenzymes Nqm1 and Tkl2 contributed to an increased rate of arabinose fermentation in strains evolved for fast arabinose fermentation (Wisselink *et al.* 2010). The set of arabinose isomerase genes that can be functionally expressed in *S. cerevisiae* was recently expanded by coexpression of *E. coli araA* with the *groEL* and *groES* chaperonins (Xia *et al.* 2016).

Engineering of sugar transport and mixed-substrate fermentation

In early *S. cerevisiae* strains engineered for pentose fermentation, uptake of D-xylose and L-arabinose exclusively relied on their native hexose transporters. While several of the 18 *S. cerevisiae* Hxt transporters (Hxt1-17 and Gal2) transport D-xylose, their K_m values for this pentose are one to two orders of magnitude higher than for glucose (Reifenberger *et al.* 1997, Hamacher *et al.* 2002, Lee *et al.* 2002, Saloheimo *et al.* 2007, Farwick *et al.* 2014). High-affinity glucose transporters, which are only expressed at low glucose concentrations (Diderich *et al.* 1999), display a lower K_m for D-xylose than low-affinity glucose transporters (Hamacher *et al.* 2002, Lee *et al.* 2002).

The galactose transporter Gal2, which also catalyses high-affinity glucose transport (Reifenberger *et al.* 1997) also has a much higher K_m for L-arabinose than for glucose (Subtil and Boles 2011, Subtil and Boles 2012). Consequently, strains depending on Gal2 for L-arabinose import fail to grow at low L-arabinose concentrations (Subtil and Boles 2011). Furthermore, kinetic competition in presence of glucose has been shown to persist in strains in which glucose-induced transcriptional repression of *GAL2* has been prevented (Horak and Wolf 1997, Özcan and Johnston 1999, Horak *et al.* 2002).

The higher affinities of Hxt transporters for glucose, combined with the transcriptional repression of Gal2 (Horak and Wolf 1997, Horak *et al.* 2002) and other high-affinity Hxt transporters (Diderich *et al.* 1999, Sedlak and Ho 2004) at high glucose concentrations, contribute to a sequential use of glucose and pentoses during mixed-substrate cultivation of engineered strains that depend on Hxt-mediated pentose uptake. Furthermore, the high K_m values of Hxt transporters for pentoses cause a deceleration of sugar fermentation during the pentose-fermentation phase. This 'tailing' effect is augmented by accumulation of ethanol and by the reduced inhibitor tolerance of *S. cerevisiae* at low sugar fermentation rates (Bellissimi *et al.* 2009, Ask *et al.* 2013, Demeke *et al.* 2013b). Intensive efforts have been made to generate yeast strains that can either co-consume hexoses and pentose sugars or sequentially consume all sugars in hydrolysates in an economically acceptable time frame (Kim *et al.* 2012, Moysés *et al.* 2016).

Recently constructed glucose-phosphorylation-negative, pentose-fermenting *S. cerevisiae* strains enabled evolutionary engineering experiments for *in vivo* directed evolution of Hxt variants that supported growth on D-xylose or L-arabinose in the presence of high glucose concentrations (Farwick *et al.* 2014, Nijland *et al.* 2014, Shin *et al.* 2015, Wisselink *et al.* 2015). Several of the evolved *HXT* alleles were confirmed to encode transporters whose D-xylose-transport kinetics were substantially less sensitive to glucose inhibition (Farwick *et al.* 2014, Nijland *et al.* 2014, Shin *et al.* 2015, Wisselink *et al.* 2015). Remarkably, independent evolutionary engineering studies aimed at selecting glucose-insensitive D-xylose and L-arabinose Hxt transporters yielded single-amino-acid substitutions at the exact corresponding positions in Hxt7(N370), Gal2 (N376), and in a chimera of Hxt3 and Hxt6 (N367) (Farwick *et al.* 2014, Nijland *et al.* 2014, Wisselink *et al.* 2015).

Additionally, replacing N-terminal lysine residues with arginine residues in low-affinity hexose transporters, such as Hxt1 and Hxt36, to prevent ubiquitination-driven degradation at low levels or absence of glucose, improved xylose fermentation when glucose levels decline in mixed sugar fermentations (Nijland *et al.* 2016). Additional Hxt- and Gal2 variants with improved relative affinities for pentoses and glucose were obtained by *in vitro* directed evolution, *HXT*-gene shuffling, and knowledge-based protein engineering (Farwick *et al.* 2014, Reznicek *et al.* 2015, Wang *et al.* 2017, Nijland *et al.* 2018) (**Box 1**).

Chapter 3 of this thesis discusses adaptive laboratory evolution of a glucose-phosphorylation negative, L-arabinose consuming *S. cerevisiae* strain for improved L-arabinose uptake under anaerobic conditions in presence of glucose and D-xylose, and the investigation and characterization of the resulting genetic adaptations within Gal2.

So far, only few low-, moderate- and high-affinity pentose transporters from pentose-metabolizing filamentous fungi or non-*Saccharomyces* yeasts, have been functionally expressed in *S. cerevisiae* (Weierstall *et al.* 1999, Leandro *et al.* 2006, Katahira *et al.* 2008, Du *et al.* 2010, Runquist *et al.* 2010b, Subtil and Boles 2011, Young *et al.* 2012, Ferreira *et al.* 2013, Colabardini *et al.* 2014, Knoshaug *et al.* 2015, Li *et al.* 2015, Reis *et al.* 2016). Expression and/or activity of several of these transporters were further improved by directed evolution (Young *et al.* 2012, Li *et al.* 2015, Li *et al.* 2016b) or evolutionary engineering (Moysés *et al.* 2016, Wang *et al.* 2016). Such high-affinity transporters may be suited to 'mop up' low concentrations of pentoses towards the end of a fermentation process. Since high-affinity sugar transporters are typically proton symporters, care should be taken to avoid scenarios in which their simultaneous expression with Hxt-like transporters, which mediate facilitated diffusion, causes futile cycles and negatively affects inhibitor tolerance.

Chapter 2 of this thesis presents the identification of a *Penicillium chrysogenum* L-arabinose transporter that had been functionally expressed and characterized in *S. cerevisiae.* This high-affinity L-arabinose specific transporter added a valuable accessory to the *S. cerevisiae* engineering toolbox for improved pentose fermentation (Bracher *et al.* 2018).

Evolutionary engineering experiments played a major role in accelerating mixed-sugar utilization by engineered pentose-fermenting strains (Sonderegger and Sauer 2003, Kuyper *et al.* 2005b, Wisselink *et al.* 2009, Sanchez *et al.* 2010, Zhou *et al.* 2012). Repeated batch cultivation on a sugar mixture can favour selection of mutants that rapidly ferment one of the sugars, while showing deteriorated fermentation kinetics with other sugars in the mixture. In practice, such trade-off scenarios can increase rather than decrease the time required for complete conversion of sugar mixtures (Wisselink *et al.* 2009). A modified strategy for repeated batch cultivation, designed to equally distribute the number of generations of selective growth on each of the individual substrates in a mixture, enabled acceleration of the anaerobic conversion of glucose-xylose-arabinose mixtures by an engineered *S. cerevisiae* strain (Wisselink *et al.* 2009). A recent study explored an interesting alternative to using a single, glucose-xylose-arabinose co-fermenting *S. cerevisiae* strain, by analyzing growth of a consortium of three specialized strains, each with the ability to ferment only one of the three sugars (Caballero and Ramos 2017, Verhoeven *et al.* 2018). While this study indicated clear advantages in terms of strain stability during long-term cultivation on sugar mixtures, it remains to be seen whether this approach can contribute to faster overall conversion of lignocellulosic hydrolysates.

Recently, an elegant evolutionary engineering experiment with an XI-based strain harbouring a forced glucose-xylose co-consumption pathway, via deletion of *PGI1* and *RPE1* and altering the cofactor specificity of 6-phosphogluconate dehydrogenase, enabled the identification and characterization of relevant mutations that, when reverse-engineered in a strain without the restriction for a forced co-consumption, led to a significantly improved co-consumption phenotype which dramatically reduced total fermentation time of a mix of 20 g L^{-1} glucose and 10 g L^{-1} xylose (Papapetridis *et al.* 2018).

Inhibitor tolerance

Yeast enzymes involved in detoxification of specific inhibitors provide logical targets for metabolic engineering. For example, overexpression of native $NAD(P)^+$-dependent alcohol dehydrogenases stimulates conversion of furfural and HMF to the less toxic alcohols furanmethanol and furan-2,5-dimethanol, respectively (Jeppsson *et al.* 2003, Lewis Liu *et al.* 2008, Almeida *et al.* 2009). Similarly, combined overexpression of the aldehyde dehydrogenase Ald5, the decarboxylase Pad1 and the alcohol acetyltransferases Atf1 and Atf2 increased resistance to several phenolic inhibitors (Adeboye *et al.* 2017).

Genome-wide expression studies have revealed intricate, strain- and context-dependent stress-response networks as major key contributors to inhibitor tolerance (Abbott *et al.* 2007, Almeida *et al.* 2007, Li and Yuan 2010, Mira *et al.* 2010, Liu 2011, Ullah *et al.* 2013, Guo and Olsson 2014). An in-depth transcriptome analysis implicated *SFP1* and *ACE2*, which encode transcriptional regulators involved in ribosomal biogenesis and septum destruction after cytokinesis, respectively, in the phenotype of an acetic-acid and furfural-tolerant strain. Indeed, overexpression of these transcriptional regulators significantly enhanced ethanol productivity in the presence of these inhibitors (Chen *et al.* 2016).

Whole-genome resequencing of tolerant strains derived from evolutionary engineering, mutagenesis and/or genome shuffling has yielded strains with increased tolerance whose causal mutations could be identified (Almario *et al.* 2013, Demeke *et al.* 2013a, Pinel *et al.* 2015, González-Ramos *et al.* 2016, Thompson *et al.* 2016). Physiological and evolutionary engineering experiments demonstrated the importance of high sugar fermentation rates for acetic acid tolerance (Bellissimi *et al.* 2009, Wright *et al.* 2011). When the acetic-acid concentration in anaerobic, xylose-grown continuous cultures was continually increased over time, evolving cultures acquired the ability to grow at acetic-acid concentrations that prevented growth of the non-evolved *S. cerevisiae* strain. However, after growth in the absence of acetic acid, full expression of their increased tolerance required pre-exposure to a lower acetic-acid concentration. This observation indicated that the acquired tolerance was inducible rather than constitutive (Wright *et al.* 2011). Constitutive tolerance

to acetic acid was shown to reflect the fraction of yeast populations able to initiate growth upon exposure to acetic acid stress (Swinnen *et al.* 2014). Based on this observation, an evolutionary engineering strategy that involved alternating batch cultivation cycles in the presence and absence of acetic acid was successfully applied to select for constitutive acetic acid tolerance (González-Ramos *et al.* 2016).

Exploration of the natural diversity of inhibitor tolerance among *S. cerevisiae* strains (Favaro *et al.* 2013, Wimalasena *et al.* 2014, Field *et al.* 2015) is increasingly used to identify genes and alleles that contribute to tolerance. In particular, combination of whole genome sequencing and classical genetics is a powerful approach to identify relevant genomic loci, genes and even nucleotides (Liti and Louis 2012) (Quantitative Trait Loci (QTL) analysis, see **Box 1**). For example, Meijnen *et al.* (2016) used whole-genome sequencing of pooled tolerant and sensitive segregants from crosses between a highly acetic-acid tolerant *S. cerevisiae* strain and a reference strain to identify mutations in five genes that contributed to tolerance.

Reduction of acetic acid to ethanol: converting an inhibitor into a co-substrate

Even small improvements of the product yield on feedstock can substantially improve the economics of biotechnological processes for manufacturing large-volume products such as ethanol (Van Maris *et al.* 2006, Nielsen *et al.* 2013). In industrial, anaerobic ethanol production processes, a significant amount of sugar is converted into the byproduct glycerol (Nissen *et al.* 2000). Glycerol formation, catalyzed by the two isoforms of glycerol-3-phosphate dehydrogenase (Gpd1 and Gpd2) and of glycerol-3-phosphate phosphatase (Gpp1 and Gpp2), is required during anaerobic growth of *S. cerevisiae* for reoxidation of NADH generated in biosynthetic reactions (Van Dijken and Scheffers 1986, Björkqvist *et al.* 1997). Metabolic engineering strategies to diminish glycerol formation focused on modification of intracellular redox reactions (Nissen *et al.* 2000, Guo *et al.* 2011) or modulation of *GPD1* and *GPD2* expression (Hubmann *et al.* 2011). Replacement of *GPD1* and *GPD2* with a heterologous gene encoding an acetylating acetaldehyde dehydrogenase (A-ALD) and supplementation of acetic acid eliminated glycerol formation in anaerobic *S. cerevisiae* cultures (Guadalupe-Medina *et al.* 2010). By enabling NADH-dependent reduction of acetic acid to ethanol (**Figure 2**), this strategy resulted in a significant increase in the final ethanol yield, while consuming acetic acid. This engineering strategy has recently been extended by altering the redox-cofactor specificities of alcohol dehydrogenase (Henningsen *et al.* 2015) and 6-phosphogluconate dehydrogenase (Papapetridis *et al.* 2016). These further interventions increased the availability of cytosolic NADH for acetate reduction and should, upon implementation in industrial strains, further improve *in situ* detoxification of acetic acid. The A-ALD strategy was also shown to decrease xylitol formation in XR/XDH-based

xylose-fermenting engineered strains by reoxidation of excess NADH formed in the XDH reaction (Wei *et al.* 2013, Zhang *et al.* 2016a).

Development of industrial yeast strains and processes

Much of the research discussed in the preceding paragraphs was based on laboratory yeast strains, grown in synthetic media whose composition can be different from that of industrial lignocellulosic hydrolysates. **Table 2** provides examples of ethanol yields and biomass-specific conversion rates that have been obtained with engineered *S. cerevisiae* strains in synthetic media.

While data on the performance of current industrial strains on industrial feedstocks are proprietary, many scientific publications describe the fermentation of hydrolysates by D-xylose-fermenting strains (either XI or XR-XDH-based, but so far without arabinose pathways). These studies cover a wide variety of feedstocks, biomass deconstruction and fermentation strategies (batch, fed-batch, SSF), aeration regimes and nutritional supplementations (e.g. yeast extract, peptone, low-cost industrial supplements, trace elements, nitrogen sources). However, with few exceptions, these data are restricted to final ethanol yields and titers, and do not include quantitative information of the biomass-specific conversion rates (q_{xylose}, $q_{ethanol}$, expressed in $g \cdot (g\ biomass)^{-1} \cdot h^{-1}$) that are essential for strain comparison and process design. **Table 3** summarizes results of studies on fermentation of biomass hydrolysates that include or enable calculation of biomass-specific conversion rates and ethanol yields.

Despite the heterogeneity of the studies included in **Tables 2 and 3**, the available data clearly illustrate that, while even 'academic' strain platforms can exhibit high ethanol yields in hydrolysates, conversion rates under these conditions are much lower than in synthetic media. Improving kinetics and robustness in industrial hydrolysates is therefore the single most important objective in industrial yeast strain development platforms.

Aspects such as spatial and temporal heterogeneity, hydrostatic pressure and CO_2 concentrations, which are highly important for down-scaling aerobic industrial fermentation processes (Noorman 2011), do not represent substantial challenges in down-scaling second-generation ethanol processes (Jansen *et al.* 2017). Provided that anaerobic conditions can be maintained, strain performance can therefore be adequately assessed in small-scale systems. Access to hydrolysates whose composition and concentration are fully representative for the target industrial substrate(s) may be necessary for strain development. This requirement is not a trivial one due to feedstock variability, the plethora of pretreatment options and the limited scalability and continuous innovation in biomass deconstruction (Knoll *et al.* 2013, Li *et al.* 2016a).

Table 2. Ethanol yields ($Y_{E/S}$, g ethanol·(g sugar)$^{-1}$) and biomass-specific rates of xylose and/or arabinose consumption and ethanol production (q_{xylose}, $q_{arabinose}$ and $q_{ethanol}$, respectively, g·(g biomass)$^{-1}$·h^{-1}) in cultures of *S. cerevisiae* strains engineered for pentose fermentation, grown in synthetic media. Asterisks (') indicate values estimated from graphs in the cited reference.

S. cerevisiae strain	Pentose fermentation strategy	Key genetic modifications	Fermentation conditions	$Y_{E/S}$ g·g^{-1}	$q_{ethanol}$ g·g^{-1}·h^{-1}	q_{xylose} g·g·h^{-1}	$q_{arabinose}$ g·g·h^{-1}	Reference
TMB3400	XR/XDH (*S. stipitis XYL1, XYL2*)	*SsXYL1, SsXYL2* + *XKS1*↑, random mutagenesis	Anaerobic batch (bioreactor), 5% xylose	0.33	0.04	0.13	-	(Karhumaa *et al.* 2007)
GLBRCY87	XR/XDH (*S. stipitis XYL1, XYL2*)	*SsXYL1, SsXYL2, SsXYL3,* evolved on xylose and hydrolysate inhibitors	Semi-anaerobic batch (flask) 5% glucose and 5% xylose	0.34*	0.036*	0.13	-	(Sato *et al.* 2016)
SR8	XR/XDH (*S. stipitis XYL1, XYL2*)	*SsXYL1,Ss XYL2, Ss XYL3, ald6*Δ, evolved on xylose	Anaerobic batch (reactor), 4% xylose	0.39	0.25	0.64	-	(Wei *et al.* 2013)
TMB3421	XR/XDH (*S. stipitis XYL1, XYL2*)	*S. stipitis XYL1*$^{N272D/P275Q}$, *XYL2* + *XKS1*↑ *TAL1*↑ *TKL1*↑ *RPE1*↑ *RKI1*↑ *gre3*Δ, evolved on xylose	Anaerobic batch (reactor), 6% xylose	0.35	0.20	0.57	-	(Runquist *et al.* 2010a)
RWB 217	XI (*Piromyces XylA*)	*Piromyces XylA* + *XKS1*↑ *TAL1*↑ *TKL1*↑ *RPE1*↑ *RKI1*↑ , *gre3*Δ	Anaerobic batch (reactor), 2% xylose	0.43	0.46	1.06	-	(Kuyper *et al.* 2005a)
RWB 218	XI (*Piromyces XylA*)	Derived from RWB 217 after evolution on glucose/xylose mixtures	Anaerobic batch (reactor), 2% xylose	0.41	0.49	1.2	-	(Kuyper *et al.* 2005b)
H131-A3-AL^{CS}	XI (*Piromyces XylA*)	*XylA*, *Xyl3*, *XKS1*↑ *TAL1*↑ *TKL1*↑ *RPE1*↑ *RKI1*↑ , *gre3*Δ, evolved on xylose	Anaerobic batch (reactor), 4% xylose	0.43	0.76	1.9	-	(Zhou *et al.* 2012)
IMS0010	XI/AraABD (*Piromyces XylA, L. plantarum AraA, B,D*	*XylA*; *XKS1*↑ *TAL1*↑ *TKL1*↑ *RPE1*↑ *RKI1*↑ *AraT, AraA*, *AraB*, *AraD*, evolved on glucose, xylose, arabinose mixtures	Anaerobic batch (reactor), 3% glucose, 1.5% xylose and 1.5% arabinose	0.43	-	0.35	0.53	(Wisselink *et al.* 2009)
GS1.11-26	XI/AraABD (*Piromyces XylA, L. plantarum AraA, B,D, K. lactis ARAT).*	*XylA*, *XKS1*↑ *TAL1*↑ *TKL1*↑ *RPE1*↑ *RKI1*↑ *XylA HXT7*↑ *KlAraT, AraA*, *AraB*, *AraD*, *TAL2*↑ *TKL2*↑, several rounds of mutagenesis and evolution on xylose	Semi-anaerobic batch (flask), synthetic medium, 3.5% xylose	0.46	0.48	1.1	-	(Demeke *et al.* 2013a)

Table 3. Ethanol yields on consumed sugar ($Y_{E/S}$, g ethanol·(g sugar)$^{-1}$) and biomass-specific rates of glucose and xylose consumption and ethanol production ($q_{glucose}$, q_{xylose} and $q_{ethanol}$, respectively, g·(g biomass)$^{-1}$·h^{-1}) in cultures of *S. cerevisiae* strains engineered for pentose fermentation, grown in lignocellulosic hydrolysates. Asterisks (*) indicate specific conversion rates estimated from graphs in the cited reference; daggers (†) indicate crude estimates of biomass-specific rates calculated based on the assumption that biomass concentrations did not change after inoculation, these estimates probably overestimate actual biomass-specific conversion rates. [1]Abbreviations of supplements: YE, yeast extract; YP, yeast extract and peptone; YNB, Yeast Nitrogen Base.

***S. cerevisiae* strain**	**Description**	**Feedstock, pretreatment conditions, hydrolysate sugar composition[3]**	**Fermentation conditions, added nutrients[1]**	$Y_{E/S}$ g·g^{-1}	$q_{glucose}$ g·g·h^{-1}	$q_{ethanol}$ g·g·h^{-1}	q_{xylose} g·g·h^{-1}	**Reference**
TMB3400	XR/XDH *S. stipitis XYL1* and XYL2; XKS1↑	Spruce, two-step dilute acid hydrolysis, 1.6% glucose, 0.4% xylose, 1% mannose, 1% galactose.	Anaerobic batch (flasks), $(NH_4)_2HPO_4$ NaH_2PO_4 $MgSO_4$, YE	0.41	0.021	0.005	0.005	(Karhumaa *et al.* 2007)
GLBRCY87	XR/XDH *S. stipitis XYL1, XYL2* and *XYL3* evolved on xylose and hydrolysate inhibitors	Corn Stover, ammonia fiber expansion, 8% glucose, 3.8% xylose.	Semi-anaerobic batch (flasks), pH 5.5, Urea, YNB	0.28	1.4*	0.27*	0.04	(Sato *et al.* 2016)
GLBRCY87	XR/XDH *S. stipitis XYL1, XYL2* and *XYL3* evolved on xylose and hydrolysate inhibitors	Switchgrass, ammonia fiber expansion, 6.1% glucose, 3.9% xylose.	Semi-anaerobic batch (flasks), Urea, YNB	0.35	1.65*	0.28*	0.07	(Sato *et al.* 2016)
MEC1122	XR/XDH, industrial host strain *S. stipitis XYL1*$^{(N272D/P275Q)}$ and *XYL2, XKS1*↑ *TAL1*↑	Corn cobs, autohydrolysis (202 °C), liquid fraction acid-treated. 0.3% glucose, 2.6% xylose.	Oxygen limited batch (flasks), cheese whey, urea, YE, $K_2O_5S_2$	0.3	-	0.12†,*	0.25†	(Costa *et al.* 2017)

S. cerevisiae strain	Description	Feedstock, pretreatment conditions, hydrolysate sugar composition[3]	Fermentation conditions, added nutrients[1]	$Y_{E/S}$ $g \cdot g^{-1}$	$q_{glucose}$ $g \cdot g \cdot h^{-1}$	$q_{ethanol}$ $g \cdot g \cdot h^{-1}$	q_{xylose} $g \cdot g \cdot h^{-1}$	Reference
RWB 218	XI *Piromyces XylA*, *XKS1*↑ *TAL1*↑ *TKL1*↑ *RPE1*↑ *RKI1*↑, *gre3*Δ, evolved on glucose/ xylose mixed substrate	Wheat straw hydrolysate, steam explosion, 5% glucose, 2% xylose	Anaerobic batch (reactor), $(NH_4)_2PO_4$	0.47	1.58†	1.0†	0.32†	(Van Maris *et al.* 2007)
GS1.11-26	XI, AraABD *Piromyces XylA*, *XKS1*↑ *TAL1*↑ *TKL1*↑ *RPE1*↑ *RKI1*↑ *HXT7*↑*AraT, AraA*, *AraB*, *AraD*, *TAL2*↑ *TKL2*↑, several rounds of mutagenesis and evolution on xylose	Spruce (no hydrolysis), acid pre-treated, 6.2% glucose, 1.8% xylose, 1% mannose	Semi-anaerobic batch (flasks), YNB, $(NH_4)_2SO_4$, amino acids added	0.43	2.46†	0.3†	0.11†	(Demeke *et al.* 2013a)
XH7	Multiple integrations of *RuXylA*; *XKS1*↑ *TAL1*↑ *TKL1*↑ *RPE1*↑ *RKI1*↑ *pho13*Δ *gre3*Δ, evolved on xylose	Corn stover, steam explosion, 6.2% glucose, 1.8% xylose	Semi-anaerobic batch (flasks), urea	0.39	0.14	0.080	0.096	(Li *et al.* 2016c)
LF1	Selection mutant of XH7 further evolved on xylose and hydrolysates with MGT transporter introduced	Corn stover, steam explosion, 8.7% glucose, 3.9% xylose	Semi-anaerobic batch (flasks), urea	0.41	0.57	0.34	0.23	(Li *et al.* 2016c)

Due to the complex, multigene nature of inhibitor tolerance, screening of natural and industrial *S. cerevisiae* strains is a logical first step in the development of industrial strain platforms. The power of this approach is illustrated by the Brazilian first-generation bioethanol strain PE-2. Stable maintenance of this strain in non-aseptically operated industrial reactors, over many production campaigns (Basso *et al.* 2008), was attributed to its innate tolerance to the sulfuric-acid washing steps that are employed between fermentation cycles to combat bacterial contamination (Della-Bianca *et al.* 2014). In contrast to most laboratory strains, robust industrial strains of *S. cerevisiae* are heterozygous diploids or polyploids which, additionally, are prone to whole-chromosome or segmental aneuploidy (Zhang *et al.* 2016b, Gorter De Vries *et al.* 2017).

Acquiring high-quality, well annotated genome sequences (**Box 1**) of these complex genomes is an important prerequisite for interpreting the results of strain improvement campaigns and for targeted genetic modification.

Episomal expression vectors carrying auxotrophic marker genes, which are commonly used in academic research, do not allow for stable replication and selection, respectively, in complex industrial media (Pronk 2002, Hahn-Hägerdal *et al.* 2007, Karim *et al.* 2013). Instead, industrial strain development requires chromosomal integration of expression cassettes. Even basic academic designs of xylose- and arabinose-fermenting strains encompass the introduction of 10-12 different expression cassettes (Wisselink *et al.* 2007, Wisselink *et al.* 2010), some of which need to be present in multiple copies (e.g. for high-level expression of XI genes (Verhoeven *et al.* 2017, Zhou *et al.* 2012, Wang *et al.* 2014, Demeke *et al.* 2015)). Additional genetic modifications, on multiple chromosomes in the case of diploid or polyploid strains, are required to reduce by-product formation, improve inhibitor tolerance and/or improve product yields.

Genetic modification of complex industrial yeast genomes has now been strongly accelerated by novel, CRISPR-based genome editing tools (**Box 1**). Non-targeted strategies for strain improvement (**Box 1**) including mutagenesis with chemical mutagens or irradiation, evolutionary engineering, recursive breeding and/or genome shuffling remain essential for industrial strain improvement. Down-scaling, automation and integration with high-throughput screening of the resulting strains in hydrolysates strongly increases the success rates of these approaches (e.g. for ethanol tolerance (Snoek *et al.* 2015)). In non-targeted strain improvement campaigns, it is important to maintain selective pressure on all relevant aspects of strain performance, to avoid trade-offs between, for example, fermentation kinetics with different sugars (glucose, xylose and arabinose), and/or inhibitor tolerance (Wisselink *et al.* 2009, Demeke *et al.* 2013a, Smith *et al.* 2014).

Even when kinetics of yeast growth and fermentation in hydrolysates are suboptimal (**Table 2**) due to the impact of inhibitors and/or strain characteristics, industrial fermentation processes need to achieve complete sugar conversion within acceptable time limits (typically 72 h or less). This can be accomplished by increasing the initial yeast biomass densities, which, in second generation processes, are typically 2- to 8-fold higher than the initial concentrations of 0.125-0.25 $g \cdot L^{-1}$ that are used in first-generation processes without biomass recycling (Jacques *et al.* 2003). Several second-generation bioethanol plants therefore include on-site bioreactors for cost-effective generation of the required yeast biomass. Precultivation in the presence of mild concentrations of inhibitors can prime yeast cells for improved performance upon exposure to stressful conditions (Alkasrawi *et al.* 2006, Sànchez i Nogué *et al.* 2013, Nielsen *et al.* 2015). Especially when biomass propagation uses non-lignocellulosic feedstocks (Steiner 2008, Narendranath and Lewis 2013) and/or is operated aerobically to maximize biomass yields, yeast strain development must take into account the need to maintain pentose- fermentation kinetics and inhibitor tolerance during biomass propagation.

The potential of vitamin prototrophic strains for optimization of industrial process economics

The harsh pretreatment procedures in lignocellulosic hydrolysates as well as feedstock variability can potentially impact the concentrations of key growth factors during fermentation. To ensure standardization and reproducibility of scientific results, the use of synthetic media is widespread in academic and industrial research (Verduyn *et al.* 1992). Furthermore, media compositions for the production of high-value compounds need to be standardized to improve the predictability of the production process and for successful scale-up studies.

Many quantitative studies on yeast physiology use a vitamin solution developed by Verduyn *et al.* (1992). Inclusion of this or a similar vitamin solution is meant to create an environment with an excess of vitamins to prevent limitations that would interfere with cell physiology and perturb experimental data analysis and interpretation (Verduyn *et al.* 1992). Reductions in vitamin concentrations or the complete omission of potentially non-essential vitamins from media ultimately results in a reduction of media costs for research and production and may contribute to the stability of yeast-based processes that use feedstocks with a variable composition.

While the latter could, alternatively, be supplemented with Corn Steep Liquor (CSL), a cheap byproduct of corn starch processing and a source of amino acids, vitamins, trace elements, and minerals (Amartey and Jeffries 1994), such complex medium ingredients may increase the risk of contamination with undesirable microbes. It is therefore of academic

and industrial interest to re-assess the minimal vitamin requirements of laboratory- and industrial yeast strains. Most yeast strains, including *S. cerevisiae* are auxotrophic for several essential vitamins due to incomplete genetic information for synthesis of certain vitamins (Burkholder *et al.* 1944). In nature, vitamin uptake from the environment versus *de novo* synthesis provides an energetic advantage, which probably contributed to the widespread occurrence of vitamin auxotrophies amongst yeast strains that reside in vitamin-rich environments, such as fruits (Helliwell *et al.* 2013). The study of vitamin synthesis in *S. cerevisiae* is (i) essential to discover potential existing prototrophies resulting in immediate cost reductions by adaptation of medium recipes, (ii) of academic interest, as some vitamin biosynthetic routes and precursor origins are not fully unravelled yet, and (iii) has the potential to, once deciphered, serve as a straight-forward metabolic engineering strategy to engineer prototrophic strains for widespread applications in academia and industry.

It has recently been shown that the widely used laboratory *S. cerevisiae* strain CEN.PK113-7D harbours the genes required to synthesise vitamin B7, also known as vitamin H or biotin (Nijkamp *et al.* 2012). CEN.PK-derived strains have also been used to construct the pentose fermenting strain lineage discussed above, starting with strain RWB217 and ultimately resulting in strain IMS0010 which ferments glucose, xylose and arabinose (Kuyper *et al.* 2003, Kuyper *et al.* 2004, Kuyper *et al.* 2005a, Kuyper *et al.* 2005b, Wisselink *et al.* 2007, Wisselink *et al.* 2009, Wisselink *et al.* 2010). However, attempts to grow CEN.PK113-7D in synthetic media lacking biotin yielded extremely slow growing cultures ($\mu \approx 0.01\ h^{-1}$). Interestingly, biotin biosynthesis in yeast and the origin of its pathway precursor, pimelic acid, are not fully understood yet. Attempts to engineer the biotin auxotrophic yeast *Pichia pastoris* for biotin prototrophy by heterologous expression of genes of the biotin biosynthesis pathway resulted in a reduction of medium costs in fed-batch based production processes, even though the resulting strain grew slower than the reference strain (Gasser *et al.* 2010). Furthermore, enhancing biotin synthesis in solventogenic Clostridia has been shown to improve production titres of acetone, butanol, and ethanol, predominantly by increasing cellular viability and performance (Yang *et al.* 2016). Indeed, exponential feeding of extra vitamins and biotin in aerobic fed-batch cultures of *S. cerevisiae* has been shown to increase ethanol productivities and -tolerance (Alfenore *et al.* 2002). **Chapter 5** of this thesis presents an adaptive laboratory evolution study for full biotin prototrophy of CEN.PK113-7D and the identification of targets for engineering biotin prototrophy for laboratory and industrial yeast strains (Bracher *et al.* 2017).

Full-scale implementation of cellulosic ethanol production

Vigorous lab-scale optimization of each of the unit operations in yeast-based ethanol production from lignocellulosic feedstocks enabled the design, construction and operation of processes at pilot scale. Recently, several industrial parties started or announced the first commercial-scale cellulosic ethanol plants, most of which rely on yeast for the fermentation step (**Table 1**). Actual cellulosic ethanol production volumes in the United States of America, derived from registered RIN (Renewable Identification Numbers) credits (United States Environmental Protection Agency 2017), indicate an increase in recent years (**Figure 3**). However, based on these numbers and estimates for plants elsewhere in the world, the global production volume of cellulosic ethanol is still below 1% of that of first-generation processes. This places actual production volumes years behind earlier projections (Lane 2015) and indicates that currently installed commercial-scale plants still operate below their nominal capacity.

Second-generation bioethanol plants are complex, multi-step biorefineries for conversion of crude and variable feedstocks. Just as high-efficiency petrochemical refineries did not appear overnight, optimizing the performance of the current frontrunner plants requires significant process engineering efforts. As remaining challenges in biomass processing and deconstruction are conquered, yeast-based processes for second-generation biofuels should soon leave the demonstration phase, become fully economically viable, and expand production volume. Such an expansion will generate new incentives for improving conversion yields, while reducing carbon footprints and overall costs.

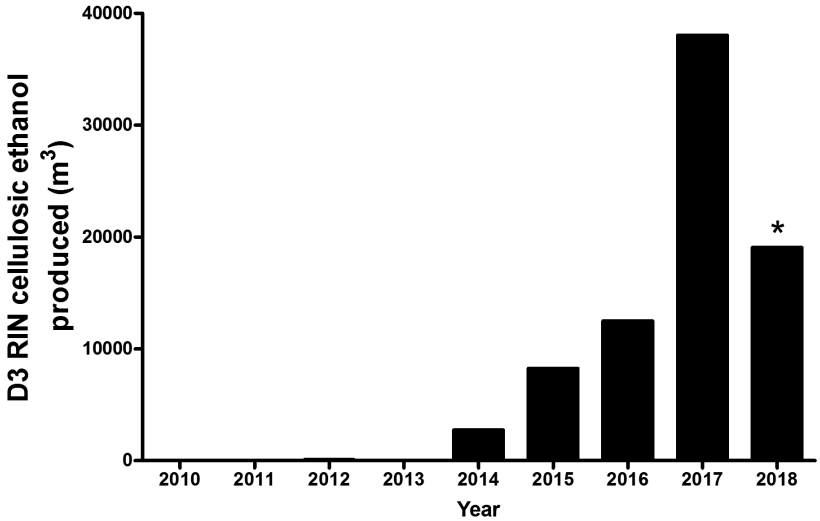


Figure 3. Annual production volumes of cellulosic ethanol in the USA from 2010 until April 2018. Numbers are based on RIN D code 3 RIN (Renewable Identification Number) credits generated (accounted as cellulosic ethanol, (United States Environmental Protection Agency 2017). *Production volume from January – April 2018.

The yeast technology developed for conversion of second-generation feedstocks can also be applied to improve ethanol yields of first-generation bioethanol production processes and plants. For example, in current first-generation ethanol processes, corn fiber is separated from whole stillage and eventually sold as animal feed (Jacques *et al.* 2003, Kim *et al.* 2008). Processes that enable conversion of this fiber-based side stream, which is more easily hydrolysed than other cellulosic feedstocks, in the context of existing first-generation bioethanol facilities, are referred to as 'Gen 1.5' technology. Several Gen 1.5 processes are currently being implemented commercially and have the potential to increase the ethanol yield per bushel of corn by approximately 10% (ICM 2012, Lane 2016, D3MAX 2017).

Scope of this thesis

The main intent of this thesis was to test metabolic engineering strategies for developing *S. cerevisiae* strains for the fermentation of D-xylose and L-arabinose, present in lignocellulosic biomass, to ultimately improve the economic viability of cellulosic ethanol production.

Whilst metabolic engineering of *S. cerevisiae* successfully enables the conversion of L-arabinose to ethanol, the transport of this non-native substrate poses a remaining bottleneck. L-arabinose transport via the endogenous galactose transporter Gal2 suffers from severe glucose inhibition and low L-arabinose transport affinity. The objective of **Chapter 2** of this thesis was to increase the affinity of L-arabinose transport in engineered, L-arabinose fermenting *S. cerevisiae* strains to ameliorate the time-consuming "tailing-effect" observed when, during a batch fermentation, L-arabinose concentrations become low. Candidate transporter genes in the filamentous fungus *Penicillium chrysogenum* were identified by chemostat-based transcriptome analysis. Subsequently, these candidate genes were functionally analysed in an engineered *S. cerevisiae* strain lacking *GAL2.* The most promising transporter, called *Pc*AraT, was further studied by uptake experiments with radiolabelled sugars and by analysing residual L-arabinose concentrations in chemostat cultures of L-arabinose fermenting *S. cerevisiae* strains expressing either *Pc*AraT or Gal2.

Gal2 has the ability to transport L-arabinose with a satisfying transport capacity but suffers from severe glucose repression. Consequently, sequential transport and conversion of glucose and arabinose extend the total fermentation time of lignocellulosic hydrolysates, affecting process economics. The goal of **Chaper 3** of this thesis was to modify the Gal2 protein such to enable uninhibited L-arabinose transport in presence of glucose whilst maintaining high transport capacities to allow co-fermentation of both substrates to ultimately reduce total fermentation times of glucose-arabinose mixtures. To this end, a glucose-phosphorylation-negative *S. cerevisiae* strain was exposed to adaptive laboratory

evolution in sequential anaerobic batch cultures containing glucose-xylose-arabinose media posing selective pressure for L-arabinose transport. Evolved strains were subjected to whole genome sequencing and the impact of the identified mutations was studied by reverse-engineering studies in non-evolved backgrounds and by analysing transport kinetics with radiolabelled sugars.

With the discovery and functional expression of *Pc*AraT in *S. cereivisiae,* and the selection and characterization of mutated Gal2 alleles with improved L-arabinose transport characteristics in presence of glucose, valuable elements for improved pentose fermentation were added to the engineering "plug-and-play" toolbox for the generation of yeast strains for 2nd generation bioethanol production.

Anaerobic fermentation of D-xylose is a key element of lignocellulosic ethanol production. Yet, reports on metabolic engineering of D-xylose-fermenting, XI-based *S. cerevisiae* strains differ where it concerns the necessity for extensive laboratory evolution to enable anaerobic growth on D-xylose. **Chapter 4** of this thesis was inspired by the initial inability to reproduce work from our own group on this matter. Its objective was to reassess the genetic requirements for anaerobic growth on D-xylose of engineered *S. cerevisiae* strains. To this end, the impact of subtle differences in strain construction and cultivation conditions between different published studies were quantitatively analysed.

The laboratory strain CEN.PK113-7D was previously found to harbour all biotin biosynthesis genes known to be required to synthesise this vitamin. Nevertheless, in academia and industry, this strain is cultured in media containing this costly vitamin. The biotin biosynthesis pathway in yeast has not yet been fully elucidated and vitamin prototrophies are an attractive trait for industrial strains from an economic perspective. The goal of **Chapter 5** of this thesis was to gain more insight into the molecular basis of biotin prototrophy in *S. cerevisiae* and to provide a basis for rational reduction of biotin requirements in this yeast. To this end, fast-growing, biotin prototrophic strains were selected via adaptive laboratory evolution. After whole genome sequencing, acquired mutations were reverse engineered in a non-evolved strain background to identify potentially interesting engineering targets for the generation of prototrophic strains.

References

Abbott DA, Knijnenburg TA, De Poorter LMI *et al.* Generic and specific transcriptional responses to different weak organic acids in anaerobic chemostat cultures of *Saccharomyces cerevisiae*. *FEMS Yeast Res* 2007;**7**:819-33.

Adeboye PT, Bettiga M, Olsson L. The chemical nature of phenolic compounds determines their toxicity and induces distinct physiological responses in *Saccharomyces cerevisiae* in lignocellulose hydrolysates. *AMB Express* 2014;**4**:46.

Adeboye PT, Bettiga M, Olsson L. *ALD5*, *PAD1*, *ATF1* and *ATF2* facilitate the catabolism of coniferyl aldehyde, ferulic acid and p-coumaric acid in *Saccharomyces cerevisiae*. *Sci Rep* 2017;**7**:42635.

Alfenore S, Molina-Jouve C, Guillouet S *et al.* Improving ethanol production and viability of *Saccharomyces cerevisiae* by a vitamin feeding strategy during fed-batch process. *Appl Microbiol Biotechnol* 2002;**60**:67-72.

Alkasrawi M, Rudolf A, Lidén G *et al.* Influence of strain and cultivation procedure on the performance of simultaneous saccharification and fermentation of steam pretreated spruce. *Enzyme Microb Technol* 2006;**38**:279-86.

Almario MP, Reyes LH, Kao KC. Evolutionary engineering of *Saccharomyces cerevisiae* for enhanced tolerance to hydrolysates of lignocellulosic biomass. *Biotechnol Bioeng* 2013;**110**:2616-23.

Almeida JRM, Bertilsson M, Hahn-Hägerdal B *et al.* Carbon fluxes of xylose-consuming *Saccharomyces cerevisiae* strains are affected differently by NADH and NADPH usage in HMF reduction. *Appl Microbiol Biotechnol* 2009;**84**:751-61.

Almeida JRM, Modig T, Petersson A *et al.* Increased tolerance and conversion of inhibitors in lignocellulosic hydrolysates by *Saccharomyces cerevisiae*. *Journal of Chemical Technology & Biotechnology* 2007;**82**:340-9.

Amartey S, Jeffries TW. Comparison of corn steep liquor with other nutrients in the fermentation of D-xylose by *Pichia stipitis* CBS 6054. *Biotechnol Lett* 1994;**16**:211-4.

Ask M, Bettiga M, Duraiswamy VR *et al.* Pulsed addition of HMF and furfural to batch-grown xylose-utilizing *Saccharomyces cerevisiae* results in different physiological responses in glucose and xylose consumption phase. *Biotechnol Biofuels* 2013;**6**:181.

Axe DD, Bailey JE. Transport of lactate and acetate through the energized cytoplasmic membrane of *Escherichia coli*. *Biotechnol Bioeng* 1995;**47**:8-19.

Bailey J. Toward a science of metabolic engineering. *Science* 1991;**252**:1668-75.

Bamba T, Hasunuma T, Kondo A. Disruption of *PHO13* improves ethanol production via the xylose isomerase pathway. *Amb Express* 2016;**6**:4.

Banerjee N, Bhatnagar R, Viswanathan L. Inhibition of glycolysis by furfural in *Saccharomyces cerevisiae*. *Appl Microbiol Biotechnol* 1981;**11**:226-8.

Basso LC, De Amorim HV, De Oliveira AJ *et al.* Yeast selection for fuel ethanol production in Brazil. *FEMS Yeast Res* 2008;**8**:1155-63.

Becker J, Boles E. A modified *Saccharomyces cerevisiae* strain that consumes L-arabinose and produces ethanol. *Appl Environ Microbiol* 2003;**69**:4144-50.

Beckner M, Ivey ML, Phister TG. Microbial contamination of fuel ethanol fermentations. *Lett Appl Microbiol* 2011;**53**:387-94.

Bellissimi E, Van Dijken JP, Pronk JT *et al.* Effects of acetic acid on the kinetics of xylose fermentation by an engineered, xylose-isomerase-based *Saccharomyces cerevisiae* strain. *FEMS Yeast Res* 2009;**9**:358-64.

Bernton H, Kovarik B, Sklar S. The forbidden fuel: Power alcohol in the twentieth century. New York: *Caroline House Pubns,* 1982.

Bettiga M, Hahn-Hägerdal B, Gorwa-Grauslund MF. Comparing the xylose reductase/xylitol dehydrogenase and xylose isomerase pathways in arabinose and xylose fermenting *Saccharomyces cerevisiae* strains. *Biotechnol Biofuels* 2008;**1**:16.

Björkqvist S, Ansell R, Adler L *et al.* Physiological response to anaerobicity of glycerol-3-phosphate dehydrogenase mutants of *Saccharomyces cerevisiae*. *Appl Environ Microbiol* 1997;**63**:128-32.

Bracher JM, de Hulster E, Koster CC *et al.* Laboratory evolution of a biotin-requiring *Saccharomyces cerevisiae* strain for full biotin prototrophy and identification of causal mutations. *Appl Environ Microbiol* 2017.

Bracher JM, Verhoeven MD, Wisselink HW *et al.* The *Penicillium chrysogenum* transporter *Pc*AraT enables high-affinity, glucose-insensitive L-arabinose transport in *Saccharomyces cerevisiae*. *Biotechnol Biofuels* 2018;**11**:63.

Brat D, Boles E, Wiedemann B. Functional expression of a bacterial xylose isomerase in *Saccharomyces cerevisiae*. *Appl Environ Microb* 2009;**75**:2304-11.

Bruinenberg PM, Bot PH, Dijken JP *et al.* The role of redox balances in the anaerobic fermentation of xylose by yeasts. *Appl Microbiol Biotechnol* 1983;**18**:287-92.

Burkholder PR, McVeigh I, Moyer D. Studies on some growth factors of yeasts. *J Bacteriol* 1944;**48**:385.

Caballero A, Ramos JL. Enhancing ethanol yields through D-xylose and L-arabinose co-fermentation after construction of a novel high efficient L-arabinose fermenting *Saccharomyces cerevisiae strain*. *Microbiology* 2017.

Canilha L, Chandel AK, Suzane Dos Santos Milessi T *et al.* Bioconversion of sugarcane biomass into ethanol: An overview about composition, pretreatment methods, detoxification of hydrolysates, enzymatic saccharification, and ethanol fermentation. *J Biomed Biotechnol* 2012.

Casey E, Mosier NS, Adamec J *et al.* Effect of salts on the co-fermentation of glucose and xylose by a genetically engineered strain of *Saccharomyces cerevisiae*. *Biotechnol Biofuels* 2013;**6**:83.

Chen Y, Sheng J, Jiang T *et al.* Transcriptional profiling reveals molecular basis and novel genetic targets for improved resistance to multiple fermentation inhibitors in *Saccharomyces cerevisiae*. *Biotechnol Biofuels* 2016;**9**.

Colabardini AC, Ries LNA, Brown NA *et al.* Functional characterization of a xylose transporter in *Aspergillus nidulans*. *Biotechnol Biofuels* 2014;**7**:1.

Costa CE, Romaní A, Cunha JT *et al.* Integrated approach for selecting efficient *Saccharomyces cerevisiae* for industrial lignocellulosic fermentations: Importance of yeast chassis linked to process conditions. *Bioresour Technol* 2017;**227**:24-34.

Crook N, Abatemarco J, Sun J *et al. In vivo* continuous evolution of genes and pathways in yeast. *Nat Commun* 2016;**7**.

D3MAX. Advantages of D3MAX. 2017. https://www.d3maxllc.com/ (24 March 2017, date last accessed).

Della-Bianca BE, Basso TO, Stambuk BU *et al.* What do we know about the yeast strains from the Brazilian fuel ethanol industry? *Appl Microbiol Biotechnol* 2013;**97**:979-91.

Della-Bianca BE, de Hulster E, Pronk JT *et al.* Physiology of the fuel ethanol strain *Saccharomyces cerevisiae* PE-2 at low pH indicates a context-dependent performance relevant for industrial applications. *FEMS Yeast Res* 2014;**14**:1196-205.

Demeke MM, Dietz H, Li Y *et al.* Development of a D-xylose fermenting and inhibitor tolerant industrial *Saccharomyces cerevisiae* strain with high performance in lignocellulose hydrolysates using metabolic and evolutionary engineering. *Biotechnol Biofuels* 2013a;**6**:89.

Demeke MM, Dumortier F, Li Y *et al.* Combining inhibitor tolerance and D-xylose fermentation in industrial *Saccharomyces cerevisiae* for efficient lignocellulose-based bioethanol production. *Biotechnol Biofuels* 2013b;**6**:120.

Demeke MM, Foulquié-Moreno MR, Dumortier F *et al.* Rapid evolution of recombinant *Saccharomyces cerevisiae* for xylose fermentation through formation of extra-chromosomal circular DNA. *PLoS Genet* 2015;**11**:e1005010.

Den Haan R, Van Rensburg E, Rose SH *et al.* Progress and challenges in the engineering of non-cellulolytic microorganisms for consolidated bioprocessing. *Curr Opin Biotechnol* 2015;**33**:32-8.

DiCarlo JE, Norville JE, Mali P *et al.* Genome engineering in *Saccharomyces cerevisiae* using CRISPR-Cas systems. *Nucleic Acids Res* 2013:gkt135.

Diderich JA, Schepper M, van Hoek P *et al.* Glucose uptake kinetics and transcription of *HXT* genes in chemostat cultures of *Saccharomyces cerevisiae*. *J Biol Chem* 1999;**274**:15350-9.

dos Santos LV, Carazzolle MF, Nagamatsu ST *et al.* Unraveling the genetic basis of xylose consumption in engineered *Saccharomyces cerevisiae* strains. *Sci Rep* 2016;**6**.

Du J, Li S, Zhao H. Discovery and characterization of novel D-xylose-specific transporters from *Neurospora crassa* and *Pichia stipitis*. *Mol Biosyst* 2010;**6**:2150-6.

Dun B, Wang, Z., Ye, K., Zhang, B., Li, G., Lu, M. Functional expression of *Arabidopsis thaliana* xylose isomerase in *Saccharomyces cerevisiae*. *Xinjiang Agric Sci* 2012;**49**:681-6.

Dunlop AP. Furfural Formation and Behavior. *Ind Eng Chem* 1948;**40**:204-9.

Ethanol Producer Magazine. U.S. ethanol plants. 2017. http://www.ethanolproducer.com/plants/listplants/US/All/Cellulosic/ (14 February 2017, date last accessed).

Farrell AE, Plevin RJ, Turner BT *et al.* Ethanol can contribute to energy and environmental goals. *Science* 2006;**311**:506.

Farwick A, Bruder S, Schadeweg V *et al.* Engineering of yeast hexose transporters to transport D-xylose without inhibition by D-glucose. *Proc Natl Acad Sci USA* 2014;**111**:5159-64.

Favaro L, Basaglia M, Trento A *et al.* Exploring grape marc as trove for new thermotolerant and inhibitor-tolerant *Saccharomyces cerevisiae* strains for second-generation bioethanol production. *Biotechnol Biofuels* 2013;**6**:168.

Ferreira D, Nobre A, Silva ML *et al. XYLH* encodes a xylose/H+ symporter from the highly related yeast species *Debaryomyces fabryi* and *Debaryomyces hansenii*. *FEMS Yeast Res* 2013;**13**:585-96.

Field SJ, Ryden P, Wilson D *et al.* Identification of furfural resistant strains of *Saccharomyces cerevisiae* and *Saccharomyces paradoxus* from a collection of environmental and industrial isolates. *Biotechnol Biofuels* 2015;**8**.

Fonseca C, Olofsson K, Ferreira C *et al.* The glucose/xylose facilitator Gxf1 from *Candida intermedia* expressed in a xylose-fermenting industrial strain of *Saccharomyces cerevisiae* increases xylose uptake in SSCF of wheat straw. *Enzyme Microb Technol* 2011;**48**:518-25.

Garcia Sanchez R, Karhumaa K, Fonseca C *et al.* Improved xylose and arabinose utilization by an industrial recombinant *Saccharomyces cerevisiae* strain using evolutionary engineering. *Biotechnol Biofuels* 2010;**3**:13.

Gasser B, Dragosits M, Mattanovich D. Engineering of biotin-prototrophy in *Pichia pastoris* for robust production processes. *Metab Eng* 2010;**12**:573-80.

González-Ramos D, Gorter de Vries AR, Grijseels SS *et al.* A new laboratory evolution approach to select for constitutive acetic acid tolerance in *Saccharomyces cerevisiae* and identification of causal mutations. *Biotechnol Biofuels* 2016;**9**:1.

Gorter De Vries AR, Pronk JT, Daran JM. Industrial relevance of chromosomal copy number variation in *Saccharomyces* yeasts. *Appl Environ Microbiol* 2017**:**AEM. 03206-16.

Grohmann K, Bothast R. Pectin-rich residues generated by processing of citrus fruits, apples, and sugar beets.*Enzymatic Conversion of Biomass for Fuels Production*,566: ACS Publications, 1994,pp 372-90.

Grohmann K, Bothast RJ. Saccharification of corn fibre by combined treatment with dilute sulphuric acid and enzymes. *Process Biochem* 1997;**32**:405-15.

Guadalupe-Medina V, Almering MJH, van Maris AJA *et al.* Elimination of glycerol production in anaerobic cultures of a *Saccharomyces cerevisiae* strain engineered to use acetic acid as an electron acceptor. *Appl Environ Microbiol* 2010;**76**:190-5.

Guo Z, Olsson L. Physiological response of *Saccharomyces cerevisiae* to weak acids present in lignocellulosic hydrolysate. *FEMS Yeast Res* 2014;**14**:1234-48.

Guo Z, Zhang L, Ding Z *et al.* Minimization of glycerol synthesis in industrial ethanol yeast without influencing its fermentation performance. *Metab Eng* 2011;**13**:49-59.

Ha S, Galazka JM, Rin Kim S *et al.* Engineered *Saccharomyces cerevisiae* capable of simultaneous cellobiose and xylose fermentation. *Proc Natl Acad Sci USA* 2011a;**108**:504-9.

Ha SJ, Kim SR, Choi JH *et al.* Xylitol does not inhibit xylose fermentation by engineered *Saccharomyces cerevisiae* expressing *xylA* as severely as it inhibits xylose isomerase reaction in vitro. *Appl Microbiol Biotechnol* 2011b;**92**:77-84.

Hahn-Hägerdal B, Galbe M, Gorwa-Grauslund MF *et al.* Bio-ethanol – the fuel of tomorrow from the residues of today. *Trends Biotechnol* 2006;**24**:549-56.

Hahn-Hägerdal B, Karhumaa K, Fonseca C *et al.* Towards industrial pentose-fermenting yeast strains. *Appl Microbiol Biotechnol* 2007;**74**:937-53.

Hahn-Hägerdal B, Wahlbom CF, Gárdonyi M *et al.* Metabolic engineering of *Saccharomyces cerevisiae* for xylose utilization.*Metab Eng*,73: Springer, 2001,53-84.

Hamacher T, Becker J, Gárdonyi M *et al.* Characterization of the xylose-transporting properties of yeast hexose transporters and their influence on xylose utilization. *Microbiology* 2002;**148**:2783-8.

Harhangi HR, Akhmanova AS, Emmens R *et al.* Xylose metabolism in the anaerobic fungus *Piromyces sp*. strain E2 follows the bacterial pathway. *Arch Microbiol* 2003;**180**:134-41.

Hector RE, Dien BS, Cotta MA *et al.* Growth and fermentation of D-xylose by *Saccharomyces cerevisiae* expressing a novel D-xylose isomerase originating from the bacterium *Prevotella ruminicola* TC2-24. *Biotechnol Biofuels* 2013;**6**:84.

Helliwell KE, Wheeler GL, Smith AG. Widespread decay of vitamin-related pathways: coincidence or consequence? *Trends Genet* 2013;**29**:469-78.

Hendriks ATWM, Zeeman G. Pretreatments to enhance the digestibility of lignocellulosic biomass. *Bioresour Technol* 2009;**100**:10-8.

Henningsen BM, Hon S, Covalla SF *et al.* Increasing anaerobic acetate consumption and ethanol yields in *Saccharomyces cerevisiae* with NADPH-specific alcohol dehydrogenase. *Appl Environ Microbiol* 2015;**81**:8108-17.

Horak J, Regelmann J, Wolf DH. Two distinct proteolytic systems responsible for glucose-induced degradation of fructose-1,6-bisphosphatase and the Gal2p transporter in the yeast *Saccharomyces cerevisiae* share the same protein components of the glucose signaling pathway. *J Biol Chem* 2002;**277**:8248-54.

Horak J, Wolf DH. Catabolite inactivation of the galactose transporter in the yeast *Saccharomyces cerevisiae*: ubiquitination, endocytosis, and degradation in the vacuole. *J Bacteriol* 1997;**179**:1541-9.

Hou J, Jiao C, Peng B *et al.* Mutation of a regulator Ask10p improves xylose isomerase activity through up-regulation of molecular chaperones in *Saccharomyces cerevisiae*. *Metab Eng* 2016a.

Hou J, Shen Y, Jiao C *et al.* Characterization and evolution of xylose isomerase screened from the bovine rumen metagenome in *Saccharomyces cerevisiae*. *J Biosci Bioeng* 2016b;**121**:160-5.

Hubmann G, Guillouet S, Nevoigt E. Gpd1 and Gpd2 fine-tuning for sustainable reduction of glycerol formation in *Saccharomyces cerevisiae*. *Appl Environ Microbiol* 2011;**77**:5857-67.

Hubmann G, Mathé L, Foulquié-Moreno MR *et al.* Identification of multiple interacting alleles conferring low glycerol and high ethanol yield in *Saccharomyces cerevisiae* ethanolic fermentation. *Biotechnol Biofuels* 2013;**6**:87.

ICM. Generation 1.5: Grain Fiber to Cellulosic Ethanol Technology. http://icminc.com/products/generation-1-5.html (24 March 2017, date last accessed).

Ilmén M, den Haan R, Brevnova E *et al.* High level secretion of cellobiohydrolases by *Saccharomyces cerevisiae*. *Biotechnol Biofuels* 2011;**4**:30.

Jacques KA, Lyons TP, Kelsall DR. The alcohol textbook: a reference for the beverage, fuel and industrial alcohol industries: *Nottingham University Press,* 2003.

Jansen ML, Bracher JM, Papapetridis I *et al. Saccharomyces cerevisiae* strains for second-generation ethanol production: from academic exploration to industrial implementation. *FEMS Yeast Res* 2017.

Jeffries TW. Engineering yeasts for xylose metabolism. *Curr Opin Biotechnol* 2006;**17**:320-6.

Jeffries TW, Jin Y. Metabolic engineering for improved fermentation of pentoses by yeasts. *Appl Microbiol Biotechnol* 2004;**63**:495-509.

Jeppsson M, Johansson B, Jensen PR *et al.* The level of glucose-6-phosphate dehydrogenase activity strongly influences xylose fermentation and inhibitor sensitivity in recombinant *Saccharomyces cerevisiae* strains. *Yeast* 2003;**20**:1263-72.

Jönsson LJ, Alriksson B, Nilvebrant NO. Bioconversion of lignocellulose: inhibitors and detoxification. *Biotechnol Biofuels* 2013;**6**:16.

Karhumaa K, Sanchez RG, Hahn-Hägerdal B *et al.* Comparison of the xylose reductase-xylitol dehydrogenase and the xylose isomerase pathways for xylose fermentation by recombinant *Saccharomyces cerevisiae*. *Microb Cell Fact* 2007;**6**:5.

Karim AS, Curran KA, Alper HS. Characterization of plasmid burden and copy number in *Saccharomyces cerevisiae* for optimization of metabolic engineering applications. *FEMS Yeast Res* 2013;**13**:107-16.

Katahira S, Ito M, Takema H *et al.* Improvement of ethanol productivity during xylose and glucose co-fermentation by xylose-assimilating *S. cerevisiae* via expression of glucose transporter Sut1. *Enzyme Microb Technol* 2008;**43**:115-9.

Katahira S, Mizuike A, Fukuda H *et al.* Ethanol fermentation from lignocellulosic hydrolysate by a recombinant xylose-and cellooligosaccharide-assimilating yeast strain. *Appl Microbiol Biotechnol* 2006;**72**:1136-43.

Kenney KL, Smith WA, Gresham GL *et al.* Understanding biomass feedstock variability. *Biofuels* 2013;**4**:111-27.

Khoo HH. Review of bio-conversion pathways of lignocellulose-to-ethanol: Sustainability assessment based on land footprint projections. *Renewable and Sustainable Energy Reviews* 2015;**46**:100-19.

Kim SR, Ha S-J, Wei N *et al.* Simultaneous co-fermentation of mixed sugars: a promising strategy for producing cellulosic ethanol. *Trends Biotechnol* 2012;**30**:274-82.

Kim SR, Skerker JM, Kang W *et al.* Rational and evolutionary engineering approaches uncover a small set of genetic changes efficient for rapid xylose fermentation in *Saccharomyces cerevisiae*. *PLOS ONE* 2013;**8**:e57048.

Kim Y, Mosier NS, Hendrickson R *et al.* Composition of corn dry-grind ethanol by-products: DDGS, wet cake, and thin stillage. *Bioresour Technol* 2008;**99**:5165-76.

Kleinschmidt J. Biofueling rural development: making the case for linking biofuel production to rural revitalization date last accessed).

Klinke HB, Ahring BK, Schmidt AS *et al.* Characterization of degradation products from alkaline wet oxidation of wheat straw. *Bioresour Technol* 2002;**82**:15-26.

Klinke HB, Thomsen AB, Ahring BK. Inhibition of ethanol-producing yeast and bacteria by degradation products produced during pre-treatment of biomass. *Appl Microbiol Biotechnol* 2004;**66**:10-26.

Knoll JE, Anderson WF, Richard EP *et al.* Harvest date effects on biomass quality and ethanol yield of new energycane (*Saccharum* hyb.) genotypes in the Southeast USA. *Biomass Bioenergy* 2013;**56**:147-56.

Knoshaug EP, Vidgren V, Magalhães F *et al.* Novel transporters from *Kluyveromyces marxianus* and *Pichia guilliermondii* expressed in *Saccharomyces cerevisiae* enable growth on L-arabinose and D-xylose. *Yeast* 2015;**32**:615-28.

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

Koppram R, Mapelli V, Albers E *et al.* The presence of pretreated lignocellulosic solids from birch during *Saccharomyces cerevisiae* fermentations leads to increased tolerance to inhibitors – a proteomic study of the effects. *PLOS ONE* 2016;**11**:e0148635.

Kötter P, Amore R, Hollenberg CP *et al.* Isolation and characterization of the *Pichia stipitis* xylitol dehydrogenase gene, *XYL2*, and construction of a xylose-utilizing *Saccharomyces cerevisiae* transformant. *Curr Genet* 1990;**18**:493-500.

Kötter P, Ciriacy M. Xylose fermentation by *Saccharomyces cerevisiae*. *Appl Microbiol Biotechnol* 1993;**38**:776-83.

Kou S, Christensen MS, Cirillo VP. Galactose transport in *Saccharomyces cerevisiae* II. Characteristics of galactose uptake and exchange in galactokinaseless cells. *J Bacteriol* 1970;**103**:671-8.

Krahulec S, Klimacek M, Nidetzky B. Engineering of a matched pair of xylose reductase and xylitol dehydrogenase for xylose fermentation by *Saccharomyces cerevisiae*. *Biotechnol J* 2009;**4**:684-94.

Kumar R, Mago G, Balan V *et al.* Physical and chemical characterizations of corn stover and poplar solids resulting from leading pretreatment technologies. *Bioresour Technol* 2009;**100**:3948-62.

Kuyper M, Harhangi HR, Stave AK *et al.* High-level functional expression of a fungal xylose isomerase: the key to efficient ethanolic fermentation of xylose by *Saccharomyces cerevisiae*? *FEMS Yeast Res* 2003;**4**:69-78.

Kuyper M, Hartog MM, Toirkens MJ *et al.* Metabolic engineering of a xylose-isomerase-expressing *Saccharomyces cerevisiae* strain for rapid anaerobic xylose fermentation. *FEMS Yeast Res* 2005a;**5**:399-409.

Kuyper M, Toirkens MJ, Diderich JA *et al.* Evolutionary engineering of mixed-sugar utilization by a xylose-fermenting *Saccharomyces cerevisiae* strain. *FEMS Yeast Res* 2005b;**5**:925-34.

Kuyper M, Winkler AA, Dijken JP *et al.* Minimal metabolic engineering of *Saccharomyces cerevisiae* for efficient anaerobic xylose fermentation: a proof of principle. *FEMS Yeast Res* 2004;**4**:655-64.

Lane J . Cellulosic ethanol: Where are the gallons? Answers for your questions. Biofuels Digest 2015. http://www.biofuelsdigest.com/bdigest/2015/07/01/cellulosic-ethanol-where-are-the-gallons-answers-for-your-questions/ (24 March 2017, date last accessed).

Lane J. What's after Gen 1? The Digest's 2016 Multi-Slide Guide to ICM's Gen 1.5 cellulosic technology. Biofuels Digest 2016a. http://www.biofuelsdigest.com/bdigest/2016/12/14/whats-after-gen-1-the-digests-2016-multi-slide-guide-to-icms-gen-1-5-cellulosic-technology/ (24 March 2017, date last accessed).

Larsson S, Quintana-Sainz A, Reimann A *et al.* Influence of lignocellulose-derived aromatic compounds on oxygen-limited growth and ethanolic fermentation by *Saccharomyces cerevisiae*. *Appl Biochem Biotechnol* 2000;**84-86**:617-32.

Lawther JM, Sun R, Banks WB. Fractional characterization of alkali-labile lignin and alkali-insoluble lignin from wheat straw. *Ind Crops Prod* 1996;**5**:291-300.

Leandro MJ, Gonçalves P, Spencer-Martins I. Two glucose/xylose transporter genes from the yeast *Candida intermedia*: first molecular characterization of a yeast xylose–H+ symporter. *Biochem J* 2006;**395**:543-9.

Lee N, Gielow W, Martin R *et al.* The organization of the *araBAD* operon of *Escherichia coli*. *Gene* 1986;**47**:231-44.

Lee SM, Jellison T, Alper HS. Directed evolution of xylose isomerase for improved xylose catabolism and fermentation in the yeast *Saccharomyces cerevisiae*. *Appl Environ Microbiol* 2012;**78**:5708-16.

Lee SM, Jellison T, Alper HS. Systematic and evolutionary engineering of a xylose isomerase-based pathway in *Saccharomyces cerevisiae* for efficient conversion yields. *Biotechnol Biofuels* 2014;**7**:122.

Lee W, Kim M, Ryu Y *et al.* Kinetic studies on glucose and xylose transport in *Saccharomyces cerevisiae*. *Appl Microbiol Biotechnol* 2002;**60**:186-91.

Lewis Liu Z, Moon J, Andersh BJ *et al.* Multiple gene-mediated NAD(P)H-dependent aldehyde reduction is a mechanism of *in situ* detoxification of furfural and 5-hydroxymethylfurfural by *Saccharomyces cerevisiae*. *Appl Microbiol Biotechnol* 2008;**81**:743-53.

Li BZ, Yuan YJ. Transcriptome shifts in response to furfural and acetic acid in *Saccharomyces cerevisiae*. *Appl Microbiol Biotechnol* 2010;**86**:1915-24.

Li C, Aston JE, Lacey JA *et al.* Impact of feedstock quality and variation on biochemical and thermochemical conversion. *Renewable and Sustainable Energy Reviews* 2016a;**65**:525-36.

Li H, Schmitz O, Alper HS. Enabling glucose/xylose co-transport in yeast through the directed evolution of a sugar transporter. *Appl Microbiol Biotechnol* 2016b;**100**:10215-23.

Li H, Shen Y, Wu M *et al.* Engineering a wild-type diploid *Saccharomyces cerevisiae* strain for second-generation bioethanol production. *Bioresour Bioprocess* 2016c;**3**:51.

Li J, Xu J, Cai P *et al.* Functional analysis of two L-arabinose transporters from filamentous fungi reveals promising characteristics for improved pentose utilization in *Saccharomyces cerevisiae*. *Appl Environ Microbiol* 2015;**81**:4062-70.

Lin Y, Tanaka S. Ethanol fermentation from biomass resources: current state and prospects. *Appl Microbiol Biotechnol* 2006;**69**:627-42.

Liti G, Louis EJ. Advances in quantitative trait analysis in yeast. *PLoS Genet* 2012;**8**:e1002912.

Liu ZL. Molecular mechanisms of yeast tolerance and *in situ* detoxification of lignocellulose hydrolysates. *Appl Microbiol Biotechnol* 2011;**90**:809-25.

Liu ZL, Slininger PJ, Dien BS *et al.* Adaptive response of yeasts to furfural and 5-hydroxymethylfurfural and new chemical evidence for HMF conversion to 2,5-bis-hydroxymethylfuran. *J Ind Microbiol Biotechnol* 2004;**31**:345-52.

Lopes ML, de Lima Paulillo SCdL, Godoy A *et al.* Ethanol production in Brazil: a bridge between science and industry. *Braz J Microbiol* 2016;**47**:64-76.

Lynd LR. Overview and evaluation of fuel ethanol from cellulosic biomass: technology, economics, the environment, and policy. *Annu Rev Energy Env* 1996;**21**:403-65.
Lynd LR, Weimer PJ, van Zyl WH *et al.* Microbial cellulose utilization: fundamentals and biotechnology. *Microbiol Mol Biol Rev* 2002;**66**:506-77.
Madhavan A, Tamalampudi S, Ushida K *et al.* Xylose isomerase from polycentric fungus *Orpinomyces*: gene sequencing, cloning, and expression in *Saccharomyces cerevisiae* for bioconversion of xylose to ethanol. *Appl Microbiol Biotechnol* 2009;**82**:1067-78.
Mans R, van Rossum HM, Wijsman M *et al.* CRISPR/Cas9: a molecular Swiss army knife for simultaneous introduction of multiple genetic modifications in *Saccharomyces cerevisiae*. *FEMS Yeast Res* 2015;**15**::fov004.
Marcheschi RJ, Gronenberg LS, Liao JC. Protein engineering for metabolic engineering: current and next-generation tools. *Biotechnol J* 2013;**8**:545-55.
Mira NP, Palma M, Guerreiro JF *et al.* Genome-wide identification of *Saccharomyces cerevisiae* genes required for tolerance to acetic acid. *Microb Cell Fact* 2010;**9**:79.
Modig T, Lidén G, Taherzadeh MJ. Inhibition effects of furfural on alcohol dehydrogenase, aldehyde dehydrogenase and pyruvate dehydrogenase. *Biochem J* 2002;**363**:769-76.
Moniruzzaman M, Dien B, Skory C *et al.* Fermentation of corn fibre sugars by an engineered xylose utilizing *Saccharomyces* yeast strain. *World J Microbiol Biotechnol* 1997;**13**:341-6.
Moon J, Liu ZL. Engineered NADH-dependent *GRE2* from *Saccharomyces cerevisiae* by directed enzyme evolution enhances HMF reduction using additional cofactor NADPH. *Enzyme Microb Technol* 2012;**50**:115-20.
Moysés DN, Reis VCB, de Almeida JRMM, Lidia Maria Pepe de *et al.* Xylose fermentation by *Saccharomyces cerevisiae*: challenges and prospects. *Int J Mol Sci* 2016;**17**:207.
Narendranath NV, Lewis SM. Systems and methods for yeast propagation. 2013.
Narron RH, Kim H, Chang HM *et al.* Biomass pretreatments capable of enabling lignin valorization in a biorefinery process. *Curr Opin Biotechnol* 2016;**38**:39-46.
Ni H, Laplaza JM, Jeffries TW. Transposon mutagenesis to improve the growth of recombinant *Saccharomyces cerevisiae* on D-xylose. *Appl Environ Microbiol* 2007;**73**:2061-6.
Nielsen F, Tomás-Pejó E, Olsson L *et al.* Short-term adaptation during propagation improves the performance of xylose-fermenting *Saccharomyces cerevisiae* in simultaneous saccharification and co-fermentation. *Biotechnol Biofuels* 2015;**8**:1.
Nielsen J, Larsson C, van Maris A *et al.* Metabolic engineering of yeast for production of fuels and chemicals. *Curr Opin Biotechnol* 2013;**24**:398-404.
Nijkamp JF, van den Broek M, Datema E *et al. De novo* sequencing, assembly and analysis of the genome of the laboratory strain *Saccharomyces cerevisiae* CEN. PK113-7D, a model for modern industrial biotechnology. *Microb Cell Fact* 2012;**11**:36.
Nijland J, Shin H, de Waal P *et al.* Increased xylose affinity of Hxt2 through gene shuffling of hexose transporters in *Saccharomyces cerevisiae*. *J Appl Microbiol* 2018;**124**:503-10.
Nijland JG, Shin HY, de Jong RM *et al.* Engineering of an endogenous hexose transporter into a specific D-xylose transporter facilitates glucose-xylose co-consumption in *Saccharomyces cerevisiae*. *Biotechnol Biofuels* 2014;**7**:168.
Nijland JG, Vos E, Shin HY *et al.* Improving pentose fermentation by preventing ubiquitination of hexose transporters in *Saccharomyces cerevisiae*. *Biotechnol Biofuels* 2016;**9**:158.
Nissen TL, Kielland-Brandt MC, Nielsen J *et al.* Optimization of ethanol production in *Saccharomyces cerevisiae* by metabolic engineering of the ammonium assimilation. *Metab Eng* 2000;**2**:69-77.
Noorman H. An industrial perspective on bioreactor scale-down: What we can learn from combined large-scale bioprocess and model fluid studies. *Biotechnol J* 2011;**6**:934-43.
Nordhoff S. Editorial: Food vs fuel – the role of biotechnology. *Biotechnol J* 2007;**2**:1451.
Olson DG, McBride JE, Joe Shaw A *et al.* Recent progress in consolidated bioprocessing. *Curr Opin Biotechnol* 2012;**23**:396-405.
Ota M, Sakuragi H, Morisaka H *et al.* Display of *Clostridium cellulovorans* xylose isomerase on the cell surface of *Saccharomyces cerevisiae* and its direct application to xylose fermentation. *Biotechnol Prog* 2013;**29**:346-51.
Otero JM, Panagiotou G, Olsson L. Fueling industrial biotechnology growth with bioethanol. In: Olsson L (ed.) *Biofuels*, Berlin, Heidelberg: Springer Berlin Heidelberg, 2007,1-40.

Oud B, van Maris AJ, Daran J-M *et al.* Genome-wide analytical approaches for reverse metabolic engineering of industrially relevant phenotypes in yeast. *FEMS Yeast Res* 2012;**12**:183-96.
Özcan S, Johnston M. Function and regulation of yeast hexose transporters. *Microbiol Mol Biol Rev* 1999;**63**:554-69.
Palmqvist E, Hahn-Hägerdal B. Fermentation of lignocellulosic hydrolysates. I: inhibition and detoxification. *Bioresour Technol* 2000a;**74**:17-24.
Palmqvist E, Hahn-Hägerdal B. Fermentation of lignocellulosic hydrolysates. II: inhibitors and mechanisms of inhibition. *Bioresour Technol* 2000b;**74**:25-33.
Pampulha ME, Loureiro-Dias MC. Energetics of the effect of acetic acid on growth of *Saccharomyces cerevisiae*. *FEMS Microbiol Lett* 2000;**184**:69-72.
Papapetridis I, van Dijk M, Dobbe AP *et al.* Improving ethanol yield in acetate-reducing *Saccharomyces cerevisiae* by cofactor engineering of 6-phosphogluconate dehydrogenase and deletion of *ALD6*. *Microb Cell Fact* 2016;**15**:67.
Papapetridis I, Verhoeven MD, Wiersma SJ *et al.* Laboratory evolution for forced glucose-xylose co-consumption enables identification of mutations that improve mixed-sugar fermentation by xylose-fermenting *Saccharomyces cerevisiae*. *FEMS Yeast Res* 2018.
Parachin NS, Bergdahl B, van Niel EW *et al.* Kinetic modelling reveals current limitations in the production of ethanol from xylose by recombinant *Saccharomyces cerevisiae*. *Metab Eng* 2011;**13**:508-17.
Parreiras LS, Breuer RJ, Narasimhan RA *et al.* Engineering and two-stage evolution of a lignocellulosic hydrolysate-tolerant *Saccharomyces cerevisiae* strain for anaerobic fermentation of xylose from AFEX pretreated corn stover. *PloS one* 2014;**9**:e107499.
Pereira LG, Dias MOS, Mariano AP *et al.* Economic and environmental assessment of n-butanol production in an integrated first and second generation sugarcane biorefinery: Fermentative versus catalytic routes. *Appl Energy* 2015;**160**:120-31.
Petschacher B, Leitgeb S, Kavanagh Kathryn L *et al.* The coenzyme specificity of *Candida tenuis* xylose reductase (*AKR2B5*) explored by site-directed mutagenesis and X-ray crystallography. *Biochem J* 2005;**385**:75-83.
Petschacher B, Nidetzky B. Altering the coenzyme preference of xylose reductase to favor utilization of NADH enhances ethanol yield from xylose in a metabolically engineered strain of *Saccharomyces cerevisiae*. *Microb Cell Fact* 2008;**7**:9.
Pinel D, Colatriano D, Jiang H *et al.* Deconstructing the genetic basis of spent sulphite liquor tolerance using deep sequencing of genome-shuffled yeast. *Biotechnol Biofuels* 2015;**8**:53.
Pronk JT. Auxotrophic yeast strains in fundamental and applied research. *Appl Environ Microbiol* 2002;**68**:2095-100.
Reifenberger E, Boles E, Ciriacy M. Kinetic characterization of individual hexose transporters of *Saccharomyces cerevisiae* and their relation to the triggering mechanisms of glucose repression. *Eur J Biochem* 1997;**245**:324-33.
Reis TF, Lima PBA, Parachin NS *et al.* Identification and characterization of putative xylose and cellobiose transporters in *Aspergillus nidulans*. *Biotechnol Biofuels* 2016;**9**:204.
Renewable Fuels Association. World fuel ethanol production. http://ethanolrfa.org/resources/industry/statistics/ (24 March 2017, date last accessed).
Reznicek O, Facey SJ, de Waal PP *et al.* Improved xylose uptake in *Saccharomyces cerevisiae* due to directed evolution of galactose permease Gal2 for sugar co-consumption. *J Appl Microbiol* 201a;**119**:99-111.
Roca C, Nielsen J, Olsson L. Metabolic engineering of ammonium assimilation in xylose-fermenting *Saccharomyces cerevisiae* improves ethanol production. *Appl Environ Microbiol* 2003;**69**:4732-6.
Rude MA, Schirmer A. New microbial fuels: a biotech perspective. *Curr Opin Microbiol* 2009;**12**:274-81.
Runquist D, Hahn-Hägerdal B, Bettiga M. Increased ethanol productivity in xylose-utilizing *Saccharomyces cerevisiae* via a randomly mutagenized xylose reductase. *Appl Environ Microbiol* 2010a;**76**:7796-802.
Runquist D, Hahn-Hägerdal B, Rådström P. Comparison of heterologous xylose transporters in recombinant *Saccharomyces cerevisiae*. *Biotechnol Biofuels* 2010b;**3**:1.

Russel JB. Another explanation for the toxicity of fermentation acids at low pH: anion accumulation versus uncoupling. *J Appl Microbiol* 1992;**73**:363-70.
Sadie CJ, Rose SH, den Haan R *et al.* Co-expression of a cellobiose phosphorylase and lactose permease enables intracellular cellobiose utilisation by *Saccharomyces cerevisiae*. *Appl Microbiol Biotechnol* 2011;**90**:1373-80.
Saloheimo A, Rauta J, Stasyk V *et al.* Xylose transport studies with xylose-utilizing *Saccharomyces cerevisiae* strains expressing heterologous and homologous permeases. *Appl Microbiol Biotechnol* 2007;**74**:1041-52.
Sànchez i Nogué V, Narayanan V, Gorwa-Grauslund MF. Short-term adaptation improves the fermentation performance of *Saccharomyces cerevisiae* in the presence of acetic acid at low pH. *Appl Microbiol Biotechnol* 2013;**97**:7517-25.
Sanchez RG, Karhumaa K, Fonseca C *et al.* Improved xylose and arabinose utilization by an industrial recombinant *Saccharomyces cerevisiae* strain using evolutionary engineering. *Biotechnol Biofuels* 2010;**3**:1.
Sander JD, Joung JK. CRISPR-Cas systems for editing, regulating and targeting genomes. *Nat Biotech* 2014;**32**:347-55.
Sárvári Horváth I, Franzén CJ, Taherzadeh MJ *et al.* Effects of furfural on the respiratory metabolism of *Saccharomyces cerevisiae* in glucose-limited chemostats. *Appl Environ Microbiol* 2003;**69**:4076-86.
Sato TK, Tremaine M, Parreiras LS *et al.* Directed evolution reveals unexpected epistatic interactions that alter metabolic regulation and enable anaerobic xylose use by *Saccharomyces cerevisiae*. *PLoS Genet* 2016;**12**:e1006372.
Sauer U. Evolutionary engineering of industrially important microbial phenotypes. *Metabolic Engineering*, Springer Berlin Heidelberg, 2001,129-69.
Sedlak M, Ho NW. Expression of E. coli *araBAD* operon encoding enzymes for metabolizing L-arabinose in *Saccharomyces cerevisiae*. *Enzyme Microb Technol* 2001;**28**:16-24.
Sedlak M, Ho NW. Production of ethanol from cellulosic biomass hydrolysates using genetically engineered *Saccharomyces* yeast capable of cofermenting glucose and xylose date last accessed).
Sedlak M, Ho NW. Characterization of the effectiveness of hexose transporters for transporting xylose during glucose and xylose co-fermentation by a recombinant *Saccharomyces* yeast. *Yeast* 2004;**21**:671-84.
Shi S, Liang Y, Zhang MM *et al.* A highly efficient single-step, markerless strategy for multi-copy chromosomal integration of large biochemical pathways in *Saccharomyces cerevisiae*. *Metab Eng* 2016;**33**:19-27.
Shin H, Nijland J, de Waal P *et al.* An engineered cryptic Hxt11 sugar transporter facilitates glucose-xylose co-consumption in *Saccharomyces cerevisiae*. *Biotechnol Biofuels* 2015;**8**:176.
Silveira MHL, Morais ARC, Da Costa Lopes AM *et al.* Current pretreatment technologies for the development of cellulosic ethanol and biorefineries. *ChemSusChem* 2015;**8**:3366-90.
Sims-Borre P. The economics of enzyme production. Ethanol Producer Magazine, 2010.
Sinha H, Nicholson BP, Steinmetz LM *et al.* Complex genetic interactions in a quantitative trait locus. *PLoS Genet* 2006;**2**:e13.
Sivers MV, Zacchi G, Olsson L *et al.* Cost analysis of ethanol production from willow using recombinant *Escherichia coli*. *Biotechnol Prog* 1994;**10**:555-60.
Smith J, van Rensburg E, Görgens JF. Simultaneously improving xylose fermentation and tolerance to lignocellulosic inhibitors through evolutionary engineering of recombinant *Saccharomyces cerevisiae* harbouring xylose isomerase. *BMC Biotechnol* 2014;**14**:41.
Snoek T, Nicolino MP, Van den Bremt S *et al.* Large-scale robot-assisted genome shuffling yields industrial *Saccharomyces cerevisiae* yeasts with increased ethanol tolerance. *Biotechnol Biofuels* 2015;**8**:32.
Sonderegger M, Sauer U. Evolutionary engineering of *Saccharomyces cerevisiae* for anaerobic growth on xylose. *Appl Environ Microbiol* 2003;**69**:1990-8.
Steiner G. Use of ethanol plant by-products for yeast propagation. US Patent 8183022 B2. 2008.
Stovicek V, Borodina I, Forster J. CRISPR–Cas system enables fast and simple genome editing of industrial *Saccharomyces cerevisiae* strains. *Metab Eng Commun* 2015;**2**:13-22.

Subtil T, Boles E. Improving L-arabinose utilization of pentose fermenting *Saccharomyces cerevisiae* cells by heterologous expression of L-arabinose transporting sugar transporters. *Biotechnol Biofuels* 2011;**4**:38.

Subtil T, Boles E. Competition between pentoses and glucose during uptake and catabolism in recombinant *Saccharomyces cerevisiae*. *Biotechnol Biofuels* 2012;**5**:1.

Swinnen S, Fernández-Niño M, González-Ramos D *et al.* The fraction of cells that resume growth after acetic acid addition is a strain-dependent parameter of acetic acid tolerance in *Saccharomyces cerevisiae*. *FEMS Yeast Res* 2014;**14**:642-53.

Swinnen S, Schaerlaekens K, Pais T *et al.* Identification of novel causative genes determining the complex trait of high ethanol tolerance in yeast using pooled-segregant whole-genome sequence analysis. *Genome Res* 2012;**22**:975-84.

Taherzadeh MJ, Eklund R, Gustafsson L *et al.* Characterization and fermentation of dilute-acid hydrolyzates from wood. *Ind Eng Chem Res* 1997;**36**:4659-65.

Taherzadeh MJ, Gustafsson L, Niklasson C *et al.* Conversion of furfural in aerobic and anaerobic batch fermentation of glucose by *Saccharomyces cerevisiae*. *J Biosci Bioeng* 1999;**87**:169-74.

Tantirungkij M, Nakashima N, Seki T *et al.* Construction of xylose-assimilating *Saccharomyces cerevisiae*. *J Ferment Bioeng* 1993;**75**:83-8.

Tenenbaum DJ. Food vs. fuel: Diversion of crops could cause more hunger. *Environ Health Perspect* 2008;**116**:A254-A7.

Thomas K, Ingledew W. Production of 21%(v/v) ethanol by fermentation of very high gravity (VHG) wheat mashes. *J Ind Microbiol Biotechnol* 1992;**10**:61-8.

Thompson OA, Hawkins GM, Gorsich SW *et al.* Phenotypic characterization and comparative transcriptomics of evolved *Saccharomyces cerevisiae* strains with improved tolerance to lignocellulosic derived inhibitors. *Biotechnol Biofuels* 2016;**9**.

Tsai C-S, Kong II, Lesmana A *et al.* Rapid and marker-free refactoring of xylose-fermenting yeast strains with Cas9/CRISPR. *Biotechnol Bioeng* 2015;**112**:2406-11.

Ulbricht R, Northup S, Thomas J. A review of 5-hydroxymethylfurfural (HMF) in parenteral solutions. *Fundam Appl Toxicol* 1984;**4**:843-53.

Ullah A, Chandrasekaran G, Brul S *et al.* Yeast adaptation to weak acids prevents futile energy expenditure. *Front Microbiol* 2013;**4**.

UNCTAD. Second generation biofuel markets: state of play, trade and developing country perspectives. *United Nations Conference on Trade and Development* 2016.

United States Environmental Protection Agency. Public data for the Renewable Fuel Standard. https://www.epa.gov/fuels-registration-reporting-and-compliance-help/public-data-renewable-fuel-standard (4 Jan 2017, date last accessed).

Van den Brink J, de Vries RP. Fungal enzyme sets for plant polysaccharide degradation. *Appl Microbiol Biotechnol* 2011;**91**:1477.

Van Dijken JP, Scheffers WA. Redox balances in the metabolism of sugars by yeasts. *FEMS Microbiol Lett* 1986;**32**:199-224.

Van Hazendonk JM, Reinerik EJM, de Waard P *et al.* Structural analysis of acetylated hemicellulose polysaccharides from fibre flax (*Linum usitatissimum L.*). *Carbohydr Res* 1996;**291**:141-54.

Van Maris AJ, Abbott DA, Bellissimi E *et al.* Alcoholic fermentation of carbon sources in biomass hydrolysates by *Saccharomyces cerevisiae*: current status. *Antonie Van Leeuwenhoek* 2006;**90**:391-418.

Van Maris AJ, Winkler AA, Kuyper M *et al.* Development of efficient xylose fermentation in *Saccharomyces cerevisiae*: xylose isomerase as a key component.*in: Biofuels*: Springer, 2007,179-204.

van Rossum HM, Kozak BU, Niemeijer MS *et al.* Requirements for carnitine shuttle-mediated translocation of mitochondrial acetyl moieties to the yeast cytosol. *mBio* 2016;**7**.

Van Vleet JH, Jeffries TW, Olsson L. Deleting the para-nitrophenyl phosphatase (pNPPase), *PHO13,* in recombinant *Saccharomyces cerevisiae* improves growth and ethanol production on D-xylose. *Metab Eng* 2008;**10**:360-9.

Verduyn C, Postma E, Scheffers WA *et al.* Effect of benzoic acid on metabolic fluxes in yeasts: A continuous-culture study on the regulation of respiration and alcoholic fermentation. *Yeast* 1992;**8**:501-17.

Verhoeven MD, Lee M, Kamoen L *et al.* Mutations in *PMR1* stimulate xylose isomerase activity and anaerobic growth on xylose of engineered *Saccharomyces cerevisiae* by influencing manganese homeostasis. *Sci Rep* 2017;**7**.

Verhoeven MD, de Vals SC, Daran J-MG *et al.* Fermentation of glucose-xylose-arabinose mixtures by a synthetic consortium of single-sugar-fermenting *Saccharomyces cerevisiae strains*. *FEMS Yeast Res* 2018.

Vohra M, Manwar J, Manmode R *et al.* Bioethanol production: Feedstock and current technologies. *Journal of Environmental Chemical Engineering* 2014;**2**:573-84.

Wahlbom CF, Hahn–Hägerdal B. Furfural, 5-hydroxymethyl furfural, and acetoin act as external electron acceptors during anaerobic fermentation of xylose in recombinant *Saccharomyces cerevisiae*. *Biotechnol Bioeng* 2002;**78**:172-8.

Wang BL, Ghaderi A, Zhou H *et al.* Microfluidic high-throughput culturing of single cells for selection based on extracellular metabolite production or consumption. *Nat Biotechnol* 2014;**32**:473-8.

Wang C, Bao X, Li Y *et al.* Cloning and characterization of heterologous transporters in *Saccharomyces cerevisiae* and identification of important amino acids for xylose utilization. *Metab Eng* 2015;**30**:79-88.

Wang C, Li Y, Qiu C *et al.* Identification of Important Amino Acids in Gal2p for Improving the L-arabinose Transport and Metabolism in *Saccharomyces cerevisiae*. *Front Microbiol* 2017;**8**.

Wang M, Yu C, Zhao H. Directed evolution of xylose specific transporters to facilitate glucose-xylose co-utilization. *Biotechnol Bioeng* 2016;**113**:484-91.

Watanabe S, Kodaki T, Makino K. Complete reversal of coenzyme specificity of xylitol dehydrogenase and increase of thermostability by the introduction of structural zinc. *J Biol Chem* 2005;**280**:10340-9.

Watanabe S, Saleh AA, Pack SP *et al.* Ethanol production from xylose by recombinant *Saccharomyces cerevisiae* expressing protein-engineered NADH-preferring xylose reductase from *Pichia stipitis*. *Microbiology* 2007;**153**:3044-54.

Wei N, Quarterman J, Kim SR *et al.* Enhanced biofuel production through coupled acetic acid and xylose consumption by engineered yeast. *Nat Commun* 2013;**4**.

Weierstall T, Hollenberg CP, Boles E. Cloning and characterization of three genes (*SUT1–3*) encoding glucose transporters of the yeast *Pichia stipitis*. *Mol Microbiol* 1999;**31**:871-83.

Wiedemann B, Boles E. Codon-optimized bacterial genes improve L-arabinose fermentation in recombinant *Saccharomyces cerevisiae*. *Appl Environ Microbiol* 2008;**74**:2043-50.

Wilkening S, Lin G, Fritsch ES *et al.* An evaluation of high-throughput approaches to QTL mapping in *Saccharomyces cerevisiae*. *Genetics* 2014;**196**:853-65.

Wimalasena TT, Greetham D, Marvin ME *et al.* Phenotypic characterisation of *Saccharomyces spp.* yeast for tolerance to stresses encountered during fermentation of lignocellulosic residues to produce bioethanol. *Microb Cell Fact* 2014;**13**:47.

Wisselink HW, Cipollina C, Oud B *et al.* Metabolome, transcriptome and metabolic flux analysis of arabinose fermentation by engineered *Saccharomyces cerevisiae*. *Metab Eng* 2010;**12**:537-51.

Wisselink HW, Toirkens MJ, del Rosario Franco Berriel M *et al.* Engineering of *Saccharomyces cerevisiae* for efficient anaerobic alcoholic fermentation of L-arabinose. *Appl Environ Microbiol* 2007;**73**:4881-91.

Wisselink HW, Toirkens MJ, Wu Q *et al.* Novel evolutionary engineering approach for accelerated utilization of glucose, xylose, and arabinose mixtures by engineered *Saccharomyces cerevisiae* strains. *Appl Environ Microbiol* 2009;**75**:907-14.

Wisselink HW, Van Maris AJA, Pronk JT *et al.* Polypeptides with permease activity. US Patent 9034608 B2. 2015.

Wright J, Bellissimi E, de Hulster E *et al.* Batch and continuous culture-based selection strategies for acetic acid tolerance in xylose-fermenting *Saccharomyces cerevisiae*. *FEMS Yeast Res* 2011;**11**:299-306.

Xia PF, Zhang GC, Liu JJ *et al.* GroE chaperonins assisted functional expression of bacterial enzymes in *Saccharomyces cerevisiae*. *Biotechnol Bioeng* 2016.

Xu H, Kim S, Sorek H *et al.* PHO13 deletion-induced transcriptional activation prevents sedoheptulose accumulation during xylose metabolism in engineered *Saccharomyces cerevisiae*. *Metab Eng* 2016;**34**:88-96.

Yan X, Inderwildi OR, King DA. Biofuels and synthetic fuels in the US and China: A review of Well-to-Wheel energy use and greenhouse gas emissions with the impact of land-use change. *Energy & Environmental Science* 2010;**3**:190-7.

Yang Y, Lang N, Yang G *et al.* Improving the performance of solventogenic *clostridia* by reinforcing the biotin synthetic pathway. *Metab Eng* 2016;**35**:121-8.

Young EM, Comer AD, Huang H *et al.* A molecular transporter engineering approach to improving xylose catabolism in *Saccharomyces cerevisiae*. *Metab Eng* 2012;**14**:401-11.

Young EM, Tong A, Bui H *et al.* Rewiring yeast sugar transporter preference through modifying a conserved protein motif. *Proc Natl Acad Sci USA* 2014;**111**:131-6.

Yu KO, Kim SW, Han SO. Engineering of glycerol utilization pathway for ethanol production by *Saccharomyces cerevisiae*. *Bioresour Technol* 2010;**101**:4157-61.

Zhang GC, Kong II, Wei N *et al.* Optimization of an acetate reduction pathway for producing cellulosic ethanol by engineered yeast. *Biotechnol Bioeng* 2016a;**113**:2587-96.

Zhang K, Zhang L-J, Fang Y-H *et al.* Genomic structural variation contributes to phenotypic change of industrial bioethanol yeast *Saccharomyces cerevisiae*. *FEMS Yeast Res* 2016b;**16**:fov118.

Zhou H, Cheng J-s, Wang BL *et al.* Xylose isomerase overexpression along with engineering of the pentose phosphate pathway and evolutionary engineering enable rapid xylose utilization and ethanol production by *Saccharomyces cerevisiae*. *Metab Eng* 2012;**14**:611-22.

2.

The *Penicillium chrysogenum* transporter *Pc*AraT enables high affinity, glucose insensitive L-arabinose transport in *Saccharomyces cerevisiae*

Jasmine M. Bracher*, Maarten D. Verhoeven*, H. Wouter Wisselink, Barbara Crimi, Jeroen G. Nijland, Arnold J.M. Driessen, Paul Klaassen, Antonius J.A. van Maris, Jean-Marc G. Daran and Jack T. Pronk

*These authors contributed equally to the presented work.

Essentially as published in **Biotechnology for Biofuels** 2018;11:63
(https://doi.org/10.1186/s13068-018-1047-6)

Abstract

L-Arabinose occurs at economically relevant levels in lignocellulosic hydrolysates. Its low-affinity uptake via the *Saccharomyces cerevisiae* Gal2 galactose transporter is inhibited by D-glucose. Especially at low concentrations of L-arabinose, uptake is an important, rate-controlling step in the complete conversion of these feedstocks by engineered, pentose-metabolizing *S. cerevisiae* strains. Chemostat-based transcriptome analysis yielded 16 putative sugar transporter genes in the filamentous fungus *Penicillium chrysogenum* whose transcript levels were at least three-fold higher in L-arabinose-limited cultures than in D-glucose-limited and ethanol-limited cultures. Of five genes, that encoded putative transport proteins and showed an over 30-fold higher transcript level in L-arabinose-grown cultures compared to D-glucose-grown cultures, only one (Pc20g01790) restored growth on L-arabinose upon expression in an engineered L-arabinose-fermenting *S. cerevisiae* strain in which the endogenous L-arabinose transporter, *GAL2*, had been deleted. Sugar-transport assays indicated that this fungal transporter, designated as *Pc*AraT, is a high-affinity (K_m = 0.13 mM), high-specificity L-arabinose-proton symporter that does not transport D-xylose or D-glucose. An L-arabinose-metabolizing *S. cerevisiae* strain in which *GAL2* was replaced by *PcaraT* showed 450-fold lower residual substrate concentrations in L-arabinose-limited chemostat cultures than a congenic strain in which L-arabinose import depended on Gal2 ($4.2 \cdot 10^{-3}$ g L^{-1} and 1.8 g L^{-1}, respectively). Inhibition of L-arabinose transport by the most abundant sugars in hydrolysates, D-glucose and D-xylose, was far less pronounced than observed with Gal2. Expression of *Pc*AraT in a hexose-phosphorylation-deficient, L-arabinose-metabolizing *S. cerevisiae* strain enabled growth in media supplemented with both 20 g L^{-1} L-arabinose and 20 g L^{-1} D-glucose, which completely inhibited growth of a congenic strain in the same condition that depended on L-arabinose transport via Gal2. Its high affinity and specificity for L-arabinose, combined with limited sensitivity to inhibition by D-glucose and D-xylose make *Pc*AraT a valuable transporter for application in metabolic engineering strategies aimed at engineering *S. cerevisiae* strains for efficient conversion of lignocellulosic hydrolysates.

Introduction

At an annual production of 100 Mton (Renewable Fuels Association 2017), bioethanol produced by the yeast *Saccharomyces cerevisiae* is by volume the largest fermentation product in industrial biotechnology. Cane sugar and corn starch, which are still the predominant feedstocks for bioethanol production, almost exclusively yield sucrose and D-glucose as fermentable sugars. Alternative lignocellulosic feedstocks, derived from agricultural residues or energy crops, contain cellulose, hemicellulose and, in some cases, pectin (Hahn-Hägerdal *et al.* 2006). The pentoses D-xylose and L-arabinose typically represent 10 to 25% and 2 to 3%, respectively, of the monomeric sugars in lignocellulosic hydrolysates (Lynd 1996). Some industrially relevant hydrolysates, however, contain higher L-arabinose concentrations. For instance, in hydrolysates of corn fibre and sugar beet pulp, L-arabinose represents 16 and 26% of the total sugar content, respectively (Micard *et al.* 1996, Grohmann and Bothast 1997).

While pentose sugars are not natural substrates of *S. cerevisiae*, their efficient conversion to ethanol and, ultimately, other bulk products, is essential to ensure economically viable processes (Lin and Tanaka 2006). Extensive metabolic and evolutionary engineering has been applied to enable efficient xylose fermentation, based on expression of either a heterologous xylose reductase and xylitol dehydrogenase, or a heterologous xylose isomerase (reviewed by (Moysés *et al.* 2016) and (Jansen *et al.* 2017)). Construction of yeast strains capable of L-arabinose fermentation involved functional expression of bacterial genes encoding L-arabinose isomerase (AraA), L-ribulokinase (AraB), and L-ribulose-5-phosphate-4-epimerase (AraD) (Becker and Boles 2003, Wisselink *et al.* 2007, Bettiga *et al.* 2008, Wiedemann and Boles 2008, Xia *et al.* 2016). Additional overexpression of *S. cerevisiae* genes encoding enzymes of the non-oxidative pentose-phosphate pathway (*RPE1, RKI1, TAL1,* and *TKL1*) strongly improved rates of D-xylose and L-arabinose fermentation (Kuyper *et al.* 2005, Wisselink *et al.* 2007). In *S. cerevisiae* strains whose metabolic pathways have been intensively optimized for pentose fermentation by metabolic and evolutionary engineering, uptake of L-arabinose and D-xylose is an important rate-controlling step (Richard *et al.* 2003, Leandro *et al.* 2009, Young *et al.* 2010).

Several *S. cerevisiae* plasma-membrane hexose-transporter proteins are able to transport D-xylose and/or L-arabinose but invariably exhibit a high K_m for these pentoses (Kou *et al.* 1970, Reifenberger *et al.* 1997, Hamacher *et al.* 2002, Lee *et al.* 2002, Saloheimo *et al.* 2007, Subtil and Boles 2011, Subtil and Boles 2012, Farwick *et al.* 2014). This low affinity causes sluggish pentose conversion ('tailing') towards the end of anaerobic batch cultures. Amongst the set of 18 *S. cerevisiae* hexose transporters (Hxt1-17 and Gal2), only the galactose transporter Gal2 and, with much lower activities, Hxt9 and Hxt10 support L-arabinose import (Kou *et al.* 1970, Subtil and Boles 2011). Gal2 has a high affinity for D-glucose and

galactose but its affinity for L-arabinose is low (K_m= 57 – 371 mM) (Subtil and Boles 2011, Knoshaug *et al.* 2015). Consequently, engineered strains in which L-arabinose transport depends on Gal2 fail to grow at low L-arabinose concentrations (Subtil and Boles 2011). Moreover, even when D-glucose-induced transcriptional repression of *GAL2* (Horak and Wolf 1997, Özcan and Johnston 1999, Horak *et al.* 2002) is prevented, kinetic competition prevents L-arabinose consumption by such strains in the presence of D-glucose.

So far, few heterologous L-arabinose transporters have been functionally expressed and characterized in *S. cerevisiae* (Subtil and Boles 2011, Knoshaug *et al.* 2015, Li *et al.* 2015). In these previous studies, *S. cerevisiae* strains harbouring a functional L-arabinose fermentation pathway but no native hexose transporters proved to be excellent platforms for characterization of heterologous L-arabinose transporters. In such experiments, transporters from the yeasts *Scheffersomyces stipitis* (*Ss*AraT), *Pichia guilliermondii* (*Pg*Axt1) and from the plant *Arabidopsis thaliana* (*At*Stp2) were shown to support L-arabinose transport in *S. cerevisiae*. These transporters exhibited K_m values of 0.13 – 4.5 mM but low transport capacities, while also exhibiting severe D-glucose inhibition (Subtil and Boles 2011, Knoshaug *et al.* 2015). Inhibition by D-xylose was only studied for *Pg*Axt1, where it completely blocked L-arabinose uptake (Knoshaug *et al.* 2015). Conversely, L-arabinose transporters from the fungi *Neurospora crassa* (Lat-1) and *Myceliophthora thermophilum* (*Mt*Lat-1) supported high-capacity, low-affinity (K_m = 58 and 29 mM, respectively) L-arabinose uptake and were also strongly affected by D-glucose inhibition (Li *et al.* 2015). The strong inhibition of these transporters by D-glucose and/or D-xylose precludes the simultaneous utilization of D-glucose and L-arabinose in *S. cerevisiae* strains depending on these transporters for L-arabinose uptake.

The filamentous fungus *P. chrysogenum* and its genome have been intensively studied in relation to its role in the production of β-lactam antibiotics (Van Den Berg *et al.* 2008, Van den Berg 2011). *P. chrysogenum* is able to hydrolyse arabinoxylan to L-arabinose by its Axs5 extracellular arabinofuranohydrolase, followed by uptake and metabolism of L-arabinose as a carbon and energy source (Chiang and Knight 1960, Chiang and Knight 1961, Sakamoto *et al.* 2011). This ability implies the presence of one or more membrane transporters capable of importing L-arabinose across the plasma membrane of this fungus.

The goal of this study was to explore the *P. chrysogenum* genome for L-arabinose transporters that can be functionally expressed in *S. cerevisiae* and support D-glucose- and D-xylose insensitive, high-affinity transport of L-arabinose. To this end, transcriptomes of L-arabinose-, ethanol- and D-glucose-limited chemostat cultures of *P. chrysogenum* were compared, and putative L-arabinose transporter genes were tested for their ability to support L-arabinose transport upon expression in an *S. cerevisiae* strain engineered for L-arabinose fermentation in which *GAL2* had been deleted. A *P. chrysogenum* transporter

identified in this screen, *Pc*AraT, was subjected to more detailed analysis, including kinetic sugar-uptake studies with radiolabelled substrates, *in vivo* studies on uptake inhibition, and physiological studies with engineered *S. cerevisiae* strains in L-arabinose-limited chemostat cultures.

Materials and Methods

Microbial strains, growth media and maintenance

All *S. cerevisiae* strains constructed and used in this study (**Table 1**) are derived from the CEN.PK lineage (Entian and Kötter 2007). Yeast strains were grown on synthetic medium (SM) (Verduyn *et al.* 1992) or on YP medium (10 g L^{-1} Bacto yeast extract, 20 g L^{-1} Bacto peptone). For shake-flask cultures on synthetic medium, ammonium sulfate was replaced with urea as nitrogen source to minimize acidification. The resulting SM-urea contained 38 mmol L^{-1} urea and 38 mmol L^{-1} K_2SO_4 instead of $(NH_4)_2SO_4$.

SM and YP media were autoclaved at 121°C for 20 min, or filter sterilized using 0.2-µm bottle-top filters (Thermo Scientific, Waltham MA). Subsequently, synthetic media were supplemented with 1 ml L^{-1} of a sterile-filtered vitamin solution (Verduyn *et al.* 1992).

SM, SM-urea and YP media were further supplemented with 20 g L^{-1} D-glucose or L-arabinose, by adding concentrated solutions autoclaved at 110°C for 20 min, yielding SMD or SMA, SMD-urea or SMA-urea and YPD or YPA, respectively. Yeast cultures were grown in 100 mL medium in 500-mL shake flasks at 30 °C and at 200 rpm in an Innova Incubator (New Brunswick Scientific, Edison NJ). Solid SMD, SMA, YPD and YPA contained 1.5% Bacto agar and when indicated, 200 mg L^{-1} G418 (Invivogen, San Diego CA). Solid medium with ethanol and glycerol as carbon source (YPEG, SMEG, YPEG-G418) contained 2% ethanol and 3% glycerol. Selection and counter selection of the *amdSYM* marker cassette were performed as described previously (Solis-Escalante *et al.* 2013).

Escherichia coli strains were grown in 5 ml Lysogeny Broth (10 g L^{-1} Bacto tryptone, 5 g L^{-1} Bacto yeast extract, 5 g L^{-1} NaCl) supplemented with 100 mg L^{-1} ampicillin in 25 ml shake flasks at 37 °C and 200 rpm in an Innova 4000 shaker (New Brunswick Scientific). Before storage at -80 °C, yeast and *E. coli* cultures were mixed with glycerol (30% v/v). *Penicillium chrysogenum* DS17690 was kindly provided by DSM Anti-infectives (Delft, The Netherlands) and grown in mineral medium (pH 5.5), containing 3.5 g $(NH_4)_2SO_4$, 0.8 g KH_2PO_4, 0.5 g $MgSO_4{\cdot}7H_2O$ and 10 ml of trace element solution (15 g L^{-1} $Na_2EDTA{\cdot}2H_2O$, 0.5 g L^{-1} $Cu_2SO_4{\cdot}5H_2O$, 2 g L^{-1} $ZnSO_4{\cdot}7H_2O$, 2 g L^{-1} $MnSO_4{\cdot}H_2O$, 4 g L^{-1} $FeSO_4{\cdot}7H_2O$, and 0.5 g L^{-1} $CaCl_2{\cdot}2H_2O$) per liter of demineralised water. The mineral medium was

Table 1. *Saccharomyces cerevisiae* strains used in this study.

Strain	Relevant genotype	Reference
CEN.PK 113-7D	*MAT*a *URA3 HIS3 LEU2 TRP1 MAL2-8c SUC2*	(Entian and Kötter 2007)
CEN.PK 113-5D	*MAT*a *ura3-52 HIS3 LEU2 TRP1 MAL2-8c SUC2*	(Entian and Kötter 2007)
CEN.PK102-12A	*MAT*a *ura3-52 his3-D1 leu2-3,112 TRP1 MAL2-8c SUC2*	(Entian and Kötter 2007)
IMX080	CEN.PK102-12A *glk1::SpHis5, hxk1::KlLEU2*	(Solis-Escalante *et al.* 2015)
IMX581	CEN.PK113-5D *can1::cas9-natNT2*	(Mans *et al.* 2015)
IMX486	IMX080 *gal1::cas9-AMDS*	This study
IMX604	IMX486 *ura3-52 gre3::pTDH3-RPE1, pPGK1-TKL1, pTEF1-TAL1, pPGI1-NQM1, pTPI1-RKI1, pPYK1-TKL2*	This study
IMX658	IMX604 *ura3-52 gal80::(pTPI-AraA-tCYC1c)*9, pPYK-AraB-tPGI1, pPGK-AraD-tTDH3*	This study
IMX660	IMX658 *hxk2::KlURA3*	This study
IMX728	IMX658 *hxk2::PcaraT*	This study
IMX844	IMX660 *gal2::KANMX*	This study
IMX869	IMX728 *gal2::KanMX*	This study
IMX918	IMX581 *gre3::pTDH3-RPE1, pPGK1-TKL1, pTEF1-TAL1, pPGI1-NQM1, pTPI1-RKI1, pPYK1-TKL2*	This study
IMX928	IMX918 *gal80::(pTPI-AraA-tCYC1)*9, pPYK-AraB-tPGI1, pPGK-AraD-tTDH3*	This study
IMX929	IMX918 *gal80::(pTPI-AraA-tCYC1)*9, pPYK-AraB-tPGI1, pPGK-AraD-tTDH3*, pUDE348	This study
IMX1504	IMX928, *gal2Δ,* pUDR245	This study
IMX1505	IMX928 *gal2::Pc13g04640* (from pPWT111), pUDR245	This study
IMX1506	IMX928 *gal2::Pc13g08230* (from pPWT113), pUDR245	This study
IMX1507	IMX928 *gal2::Pc16g05670* (from pPWT116), pUDR245	This study
IMX1508	IMX928 *gal2::pc20g0179 (PcaraT)* (from pPWT118), pUDR246	This study
IMX1509	IMX928 *gal2::Pc22g14520* (from pPWT123), pUDR245	This study
DS68616	MATa, ura3-52, leu2-112, gre3::loxP, loxP-pTPI-TAL1, loxP-pTPI-RKI1, loxP-pTPI-TKL1, loxP-pTPI-RPE1, leu2::pADH1-XKS1-tCYC1-LEU2, ura3::URA3-pTPI1-XylA-tCYC1	DSM, The Netherlands
DS68625	DS68616 his3::loxP, hxt2::loxP-kanMX-loxP, hxt367::loxP-hphMX-loxP, hxt145::loxP-natMX-loxP, gal2::loxP-zeoMX-loxP	(Nijland *et al.* 2014)
DS68625-*PcaraT*	DS68625, pRS313-*PcaraT*	This study
DS68625-*GAL2*	DS68625, pRS313-*GAL2*	This study
DS68625-mcs	DS68625, pRS313-mcs (empty)	This study

supplemented with 7.5 g L^{-1} D-glucose. Precultivation for chemostat cultures was carried out on mineral medium with 7.5 g L^{-1} D-glucose, 7.5 g L^{-1} L-arabinose, or 5.8 g L^{-1} ethanol as carbon source.

Molecular biology techniques

DNA fragments were amplified by PCR amplification with Phusion Hot Start II High Fidelity Polymerase (Thermo Scientific) and desalted or PAGE-purified oligonucleotide primers (Sigma-Aldrich, St. Louis, MO) performed according to the manufacturers' instructions. Diagnostic PCR reactions were run with DreamTaq polymerase (Thermo Scientific). Oligonucleotide primers used in this study can be found in **Additional Table 1** that can be found at https://doi.org/10.1186/s13068-018-1047-6. PCR products were separated by electrophoresis on 1% (w/v) agarose gels (Thermo Scientific) in TAE buffer (Thermo Scientific) and, if required, purified with a Zymoclean Gel DNA Recovery kit (Zymo Research, Irvine, CA) or a GenElute PCR Clean-Up Kit (Sigma-Aldrich). Yeast or *E. coli* plasmids were isolated with a Zymoprep Yeast Plasmid Miniprep II kit (Zymo Research), or a Sigma GenElute Plasmid kit (Sigma-Aldrich), respectively. A YeaStar Genomic DNA kit (Zymo Research) or an SDS/lithium acetate protocol (Lõoke *et al.* 2011) was used to isolate yeast genomic DNA. Yeast strains were transformed using the lithium acetate/ polyethylene glycol method (Gietz and Woods 2002). Single-colony isolates were obtained from three consecutive re-streaks on selective solid agar plates, followed by analytical PCR analysis of the relevant genotype. *E. coli* DH5α cultures were transformed by chemical transformation (Inoue *et al.* 1990). After isolation, plasmids were verified by restriction analysis and analytical PCR.

Plasmid construction

Plasmids used in this study are shown in **Table 2**. All synthesized gene expression cassettes were constructed by GeneArt (Regensburg, Germany). Genes encoding the five putative transporters Pc13g04640 [Genbank: CAP91533.1] Pc13g08230 [Genbank: CAP91892.1], Pc16g05670 [Genbank: CAP93237.1], Pc20g01790 (*PcaraT*)[Genbank: CAP85508.1], Pc22g14520 [Genbank: CAP98740.1] were codon-pair optimized (Roubos and Van Peij 2008) for expression in *S. cerevisiae* and cloned into the plasmid pPWT007 (Klaassen *et al.* 2013) resulting in pPWT111, 113, 116, 118 and 123 respectively, harbouring each an expression cassette consisting of the *ADH1* promoter, the codon-optimized open-reading frame of a putative transporter gene, and the *PMA1* terminator. Expression cassettes for the coding regions of *L. plantarum* L-arabinose isomerase *araA* [Genbank: ODO63149.1], L-ribulose kinase *araB* [Genbank: ODO63147.1] and L-ribulose-5P epimerase *araD* [Genbank: ODO63148.1] were codon optimized according to the glycolytic codon preference in *S. cerevisiae* (Wiedemann and Boles 2008) and provided by GeneArt in pMK-RQ based cloning vectors named, pUDE354, pUDE355 and pUDE356, respectively. The episomal plasmids used to express guide RNAs (gRNAs) were constructed from PCR amplified fragments that were ligated using the Gibson Assembly Cloning Kit (New England

Table 2. Plasmids used in this study.

Plasmid	Characteristic	Source
p414-TEF1p-Cas9-CYC1t	CEN6/ARS4 ampR pTEF1-cas9-tCYC1	(DiCarlo *et al.* 2013)
pUG-amdSYM	Template for *amdSYM* marker	(Solis-Escalante *et al.* 2013)
pUG-72	Template for *KlURA3* marker	(Gueldener *et al.* 2002)
pUG6	Template for *KanMX* marker	(Güldener *et al.* 1996)
pUDE327	2 µm, KlURA3, pSNR52-gRNA.HXK2.Y	(Kuijpers *et al.* 2016)
pUDE335	2 µm, KlURA3, pSNR52-gRNA.GRE3.Y	(Verhoeven *et al.* 2017)
pUDE348	2 µm, KlURA3, pSNR52-gRNA.GAL80.Y	This study
pUDR246	2 µm, KlURA3, pSNR52-gRNA.GAL2.Y pSNR52-gRNA.GAL2.Y	This study
pUDR245	2 µm, KlURA3, pSNR52-gRNA.GAL2.Y pSNR52-gRNA.GAL2.Y	This study
pMEL10	pSNR52-gRNA.CAN1.Y-tSUP4	(Mans *et al.* 2015)
pROS10	2 µm, KlURA3, pSNR52-gRNA.CAN1.Y pSNR52-gRNA.ADE2.Y	(Mans *et al.* 2015)
pUD344	pJET1.2Blunt TagA-*pPGI1-NQM1*-TagB	(Verhoeven *et al.* 2017)
pUD345	pJET1.2Blunt TagB-*pTPI1-RKI1*-TagC	(Verhoeven *et al.* 2017)
pUD346	pJET1.2Blunt TagC-*pPYK1-TKL2*-TagF	(Verhoeven *et al.* 2017)
pUD347	pJET1.2BluntTagG-*pTDH3-RPE1*-TagH	(Verhoeven *et al.* 2017)
pUD348	pJET1.2Blunt TagH-*pPGK1-TKL1*-TagI	(Verhoeven *et al.* 2017)
pUD349	pJET1.2Blunt TagI-*pTEF1-TAL1*-TagA	(Verhoeven *et al.* 2017)
pUD405	pJET1.2Blunt Gal2-flanked KanMX	This study
pPWT111	ampR KanMX, amdSYM, *pADH1*-Pc13g04640-*tPMA1*	This study
pPWT113	ampR KanMX, amdSYM, *pADH1*-Pc13g08230-*tPMA1*	This study
pPWT116	ampR KanMX, amdSYM, *pADH1*-Pc16g05670-*tPMA1*	This study
pPWT118	ampR KanMX, amdSYM, *pADH1*-Pc20g01790 *(PcaraT)*-*tPMA1*	This study
pPWT123	ampR KanMX, amdSYM, *pADH1*-Pc22g14520-*tPMA1*	This study
pUD354	pMK-RQ-pTPI1-araA-tADH3	This study
pUD355	pMK-RQ-pPYK1-araB-tPGI1	This study
pUD356	pMK-RQ-pPGK1-araD-tTDH3	This study
pRS313-mcs	CEN6, ARSH4, HIS3-pHXT7, tHXT7	(Nijland *et al.* 2014)
pRS313-*PcaraT*	CEN6, ARSH4, HIS3, ampR, pHXT7-PcaraT-tHXT7	This study
pRS313-*GAL2*	CEN6, ARSH4, HIS3, ampR, pHXT7-GAL2-tHXT7	This study

Biolabs, Ipswich, MA). gRNA plasmids pUDR246 and pUDR245 were constructed using pROS10 as a template (Mans *et al.* 2015), with oligonucleotide primers listed in **Additional Table 1** that can be found at: https://doi.org/10.1186/s13068-018-1047-6. pUDE348 was derived from pMEL10 by first PCR amplifying the plasmid backbone using primers 5792-5980.

The gRNA sequence was introduced in the gRNA expression cassette with primers 6631-5979 using pMEL10 (Mans *et al.* 2015) as a template. Subsequently, both fragments were

combined using the Gibson Assembly Cloning kit. pUD405 was obtained by integration of a Gal2-flanked KanMX cassette obtained from pUG6 with primers 944 & 945 into a pJET1.2 blunt vector according to the manufacturers' instructions. Construction of the low-copy-number centromeric plasmid pRS313-mcs was described previously (Nijland *et al.* 2014). *GAL2* was amplified from genomic DNA of *S. cerevisiae* DS68616 (Nijland *et al.* 2014) and *PcaraT* was amplified from plasmid pPWT118 using primers F *GAL2* Xbai & R *GAL2* Cfr9i and primers F *PcaraT* Xbai & R *PcaraT* Cfr9i, respectively and cloned into pRS313-mcs, resulting in plasmids pRS313-*PcaraT* & pRS313-*GAL2*.

Strain construction

Gene expression cassettes were PCR amplified with oligonucleotide primers listed in **Additional Table 1** (https://doi.org/10.1186/s13068-018-1047-6) and genomic DNA of CEN.PK113-7D or plasmids described in **Table 2**. Gene knock-outs and construct integrations were introduced with a chimeric CRISPR/Cas9 editing system (Mans *et al.* 2015). To enable CRISPR/Cas9 mediated editing in strain IMX080, the SpCas9 expression cassette was amplified from p414-pTEF1-cas9-tCYC1 (Addgene plasmd # 43802) and integrated into the GAL1 locus via in-vivo assembly, together with the *amdSYM* marker, yielding strain IMX486.

For overexpression of the non-oxidative pentose-phosphate pathway (PPP), IMX486 and IMX581 were co-transformed with gRNA plasmid pUDE335 and repair fragments flanked with either 60 bp homologous to *GRE3* or with synthetic tags (Kuijpers *et al.* 2013) assisting homologous recombination of the PPP expression cassettes ($gre3_{flank}$-*pTDH3-RPE1- TagH, TagH-pPGK1-TKL1-TagI, TagI-pTEF1-TAL1-TagA, TagA-pPGI1-NQM1-TagB, TagB-pTPI1-RKI1-TagC, TagC-pPYK1-TKL2-*$gre3_{flank}$). After counter selection of the *URA3*-based plasmid pUDE335, the resulting strains, IMX604 and IMX918, respectively, were co-transformed with pUDE348 and repair fragments flanked with either 60 bp homologous to *GAL80* or with synthetic tags (Kuijpers *et al.* 2013) ($GAL80_{flank}$-*pTPI1-araA-TagG, TagG-pTPI1-araA-TagA, TagA- pTPI1-araA-TagB, TagB-pTPI1-araA-TagC, TagC-pTPI1-araA-TagD, TagD-pTPI1-araA-TagM, TagM-pTPI1-araA-TagN, TagN-pTPI1-araA-TagO, TagO-pTPI1-araA-TagI, TagI-pPYK1-araB-TagK, TagK-pPGK1-araD-*$GAL80_{flank}$) resulting in nine copies of *araA* and a single copy of *araB* and *araD* integrated in the *GAL80* locus. After verification of the resulting strains IMX929 and IMX658, respectively, plasmid pUDE348 was counter selected in IMX929 to yield strain IMX928. Disruption of *HXK2* in IMX658 was done by PCR-amplification and transformation of the *KlURA3*-based deletion cassette from pUG-72 (Gueldener *et al.* 2002) to obtain strain IMX660 upon transformation and plating in solid SMA. *GAL2* was disrupted in IMX660 by transformation with a KanMX cassette amplified from pUD405 with primers 944 & 945 flanked with 60 bp homologous

to *GAL2*. Transformants were incubated for 2 h in YPE before plating on YPEG-G418, yielding strain IMX844.

Expression of *PcaraT* in IMX658 was achieved by transforming IMX658 with the gRNA plasmid pUDE327 together with an expression cassette of *PcaraT* (pADH1-*PcaraT*-tPMA1) with flanking regions homologous to the *HXK2* locus amplified with the primer pair 7660 & 7676. Counter selection of the pUDE327 and subsequent transformation of a DNA fragment derived from CEN.PK113-7D using primers 2641-1522 repaired uracil auxotrophy and resulted in strain IMX728. *GAL2* was disrupted in IMX728 by transformation with a KanMX cassette amplified from pUD405 with primers 944 & 945 flanked with 60 bp homologous to *GAL2*. Transformants were incubated for 2 h in YPE before plating on YPEG-G418, yielding strain IMX869. Strains IMX1505-1509 were constructed by co-transforming pUDR245 or pUDR246 and a *GAL2*-flanked expression cassette (pADH1-ORF-tPMA1) amplified from pPWT111, 113, 116, 118 or 123, respectively, amplified with the primer pair 10585 & 10584. IMX1504, harbouring a knockout of *GAL2,* was constructed by co-transforming pUDR245 and a repair fragment based on the annealed primers 9563 and 9564. Transformation of *GAL2* and *PcaraT* plasmids, and the pRS313-mcs plasmid (as an empty plasmid/control) into the hexose transporter deletion strain DS68625 yielded strains DS68625-*GAL2*, DS68625-*PcaraT*, and DS68625-mcs.

Growth experiments in shake flasks

Thawed 1 ml aliquots from frozen stock cultures were used to inoculate shake-flask precultures on SM-urea supplemented with either D-glucose (20 g L^{-1}), L-arabinose (20 g L^{-1}), or both sugars (both 20 g L^{-1}). These precultures were used to inoculate a second culture, which was subsequently used to inoculate a third culture, which was inoculated at an initial OD_{660} of 0.1 and used to monitor growth. Optical densities at 660 nm were measured with a Libra S11 spectrophotometer (Biochrom, Cambridge, United Kingdom). Maximum specific growth rates (μ_{max}) were derived from at least four consecutive data points derived from samples taken during the exponential growth phase of each culture.

Spot plates

L-arabinose metabolising *S. cerevisiae* strains expressing putative *P. chrysogenum* L-arabinose transporter genes (IMX1504-1509) were grown on SMD medium and a total number of approximately 10^4, 10^3, 10^2, and 10^1 cells were spotted on duplicate agar plates as described previously (Van Rossum *et al.* 2016, Bracher *et al.* 2017) containing either 20 g L^{-1} L-arabinose or D-glucose as carbon source (pH 6). Cell numbers were estimated from calibration curves of OD_{660} versus cell counts determined with an Accuri flow cytometer

(Becton Dickinson B.V., Breda, The Netherlands), derived from exponentially growing shake-flask cultures of *S. cerevisiae* CEN.PK113-7D on SMD medium. SMA and SMD plates were incubated at 30 °C for 97 and 41 h, respectively.

Chemostat cultivation

Aerobic carbon-limited chemostat cultures of *P. chrysogenum* were grown at 25 °C in 3-L turbine-stirred bioreactors (Applikon, Schiedam, The Netherlands) with a working volume of 1.8 L and a dilution rate of 0.03 h^{-1} as described previously (Harris *et al.* 2009), with the exception that, in addition to cultures grown on 7.5 g L^{-1} D-glucose, chemostat cultures were also grown on either 7.5 g L^{-1} L-arabinose or 5.8 g L^{-1} ethanol. Aerobic, L-arabinose-limited chemostat cultures of *S. cerevisiae* were grown at 30 °C in 2-L Applikon bioreactors with a working volume of 1 L and at a dilution rate of 0.05 h^{-1}. SMA (7.5 g L^{-1} L-arabinose) supplemented with 0.15 g L^{-1} antifoam Pluronic PE 6100 was used as culture medium for the initial batch phase and for chemostat cultivation, with the exception of the initial batch phase of strain IMX929 which was grown on 20 g L^{-1} L-arabinose. Cultures were stirred at 800 rpm, kept at pH 5.0 by automatic addition of 2 M KOH, and sparged with 0.5 L min^{-1} air. Upon completion of the batch phase, chemostat cultivation was initiated, ensuring a constant culture volume with an electric level sensor. When after at least five volume changes, biomass dry weight and CO_2 production varied by less than 2% over two consecutive volume changes, the culture was considered to be in steady state.

Analytical methods

P. chrysogenum biomass dry weight was determined in duplicate by filtration of 10 mL culture sample over pre-weighed glass fibre filters (Type A/E, Pall Life Sciences, Hoegaarden, Belgium). After filtration, filters were washed with demineralised water and dried for 10 min at 600 W in a microwave oven (Bosch, Stuttgart, Germany) prior to reweighing. Biomass dry weight in *S. cerevisiae* culture samples was determined with a similar procedure using nitrocellulose filters (0.45 µm pore size; Gelman Laboratory, Ann Arbor, MI) and drying for 20 min in a microwave oven at 360 W output. Optical density (OD) of the cultures was determined at 660 nm with a Libra S11 spectrophotometer (Biochrom, Cambridge, United Kingdom). Determination of CO_2 and O_2 concentrations in the bioreactor exhaust gas and HPLC analysis of metabolite concentrations in culture supernatant samples were performed as described previously (Verhoeven *et al.* 2017).

Sampling, RNA extraction, microarrays analysis, and data analysis

Samples (60 mL) from *P. chrysogenum* chemostat cultures were rapidly filtered over a glass fibre filter (Type A/E, Pall Life Sciences) and further processed for total RNA extraction by phenol-chloroform extraction (Harris *et al.* 2009). The cRNA sample preparation (cDNA synthesis, purification, *in vitro* transcription, labelling, purification, fragmentation and biotinylation) was performed according to Affymetrix recommendations (Van Den Berg *et al.* 2008). Eventually cRNA samples were hybridized onto custom-made *P. chrysogenum* GeneChip microarrays (array code DSM_PENa520255F). Data Acquisition, hybridization, quantification of processed array images, and data filtering were performed using the Affymetrix GeneChip Operating Software (GCOS version 1.2).

Global array normalization was performed by scaling the global fluorescence intensity of each microarray to 100. The scaling factors of the individual arrays were highly similar and ranged from 0.21 to 0.35. Subsequently, significant variations in expression were statistically estimated by comparing replicate array experiments using the Significance Analysis of Microarray software (SAM version 2.0) (Tusher *et al.* 2001) with the multiclass setting. A false discovery rate of 1% was applied to minimize the chance of false positive hits. Genes with an over 3-fold higher transcript level in arabinose-grown cultures than in D-glucose-grown cultures and a less than 3-fold difference in ethanol- and D-glucose-grown cultures were deemed to show arabinose-specific expression. Transcriptome data of strain DS17690 grown on D-glucose, ethanol or arabinose are accessible at NCBI Genome Omnibus database [https://www.ncbi.nlm.nih.gov/geo/] under accession number GSE12632, GSE24212, and GSE10449, respectively (Harris *et al.* 2009).

Analysis of sugar uptake kinetics

Uptake experiments with [^{14}C] L-arabinose, [^{14}C] D-xylose, or [^{14}C] D-glucose, labelled at the first carbon atom (50-60 mCi/mmol) (ARC St. Louis, MO), were performed with *S. cerevisiae* hexose transporter deletion strains (DS68625) harbouring a low copy plasmid with constitutively expressed *PcaraT* (pRS313-*PcaraT*) or *GAL2* (pRS313-*GAL2*). The experimental workflow was carried out as described previously (Nijland *et al.* 2014) with [^{14}C] L-arabinose concentrations of 0.5 to 2000 mmol L^{-1}, [^{14}C] D-xylose concentrations of 0.5 to 500 mM, or [^{14}C] D-glucose concentrations of 0.1 to 500 mmol L^{-1}. Transport competition experiments were carried out in the presence of 50 mmol L^{-1} [^{14}C] L-arabinose and 0 to 500 mmol L^{-1} D-glucose or D-xylose, and at [^{14}C] L-arabinose concentration of 2 mmol L^{-1} together with increasing D-glucose and xylose concentrations of 0 to 20 mM. Maximum biomass-specific transport rates ('V_{max}') calculated from transport assays were expressed as nmol sugar transported per mg biomass dry weight per minute [nmol (mg biomass)$^{-1}$ min^{-1}]. As this V_{max} is influenced by the expression level of the relevant transporter,

it is not solely dependent on intrinsic transporter kinetics. The impact of proton-gradient uncoupling on transport activity was determined in 200 µl synthetic medium at a [^{14}C]-L-arabinose concentration of 2 mmol L^{-1}, by comparing transport rates upon addition of either 10 µmol L^{-1} CCCP (0.5 µl of a stock solution dissolved in 100% DMSO), 0.5 µL DMSO (control), or 0.5 µL water.

Phylogenetic methods

Protein sequences used for generation of a phylogenetic tree were derived from NCBI (https://www.ncbi.nlm.nih.gov/) and the *Saccharomyces* Genome Database (https://www.yeastgenome.org/). Mafft was used to generate a CLUSTAL format alignment of all sequences, using the L-INS-i method default settings (https://mafft.cbrc.jp/alignment/server/) (Katoh *et al.* 2005, Katoh *et al.* 2017). Alignments were further processed using neighbour-joing and a 500-times bootstrap. The resulting Newick tree file was visualized and midpoint rooted in iTOL (https://itol.embl.de/) (Letunic and Bork 2016). Gene accession numbers were: *ScGAL2*: P13181, *PcaraT*: CAP85508, *SsaraT*: XP_001382755, *Atstp2*: OAP13698, *Kmaxt1*: GZ791039, *Pgaxt1*: GZ791040, *Amlat1*: AY923868, *Amlat2*: AY923869, *Nclat-1*: EAA30346, *Mtlat-1*: XP_003663698.

Results

Chemostat-based transcriptome analysis of *Penicillium chrysogenum* for identification of possible L-arabinose transporter genes

Filamentous fungi exhibit a much broader range of carbon source utilization than *S. cerevisiae* and, similar to many other ascomycetous fungi, *P. chrysogenum* can grow on L-arabinose as sole carbon source (Chiang and Knight 1961, Sakamoto *et al.* 2003). To identify candidate structural genes for L-arabinose transporters in *P. chrysogenum*, carbon-limited chemostat cultures of strain DS17690 were grown at a dilution rate of 0.03 h^{-1} on different carbon sources. To discriminate between alleviation of carbon repression and L-arabinose induction, duplicate D-glucose-, L-arabinose-, and ethanol- limited chemostat cultures were performed. RNA was extracted from steady-state cultures and gene expression levels were obtained using Affymetrix DNA-arrays (Harris *et al.* 2009). A total of 540 genes were differentially expressed over the three conditions. Of these differentially expressed genes, 137 exhibited an over 3-fold higher transcript level in L-arabinose-limited cultures than in D-glucose-limited cultures, as well as a less than three-fold difference in transcript level between ethanol- and D-glucose-limited cultures (**Additional Figure 5**). Genes whose transcript levels in L-arabinose- and ethanol-limited cultures were both at least 2-fold higher than in D-glucose-grown cultures were not

Table 3. Putative transporter genes that showed higher relative transcript levels in aerobic, L-arabinose-limited chemostat cultures of *Penicillium chrysogenum* than in corresponding D-glucose- and ethanol-limited cultures. *P. chrysogenum* DS1769 was grown in L-arabinose-, D-glucose-, or ethanol-limited chemostat cultures (dilution rate = 0.03 h^{-1}, pH 6.5, T = 25 °C). Genes indicated in bold were selected for further analysis, based fold higher transcript level in L-arabinose-limited cultures than in D-glucose-limited cultures. Data represent average ± mean deviation of globally scaled (target 100) Affymetrix microarrays for independent duplicate chemostat cultures.

		Relative transcript levels under different nutrient limitations				
Gene	**Strong similarity to**	**Glucose**	**L-Arabinose**	**Ethanol**	**Ethanol / glucose**	**L-Arabinose / glucose**
Pc13g08230	*S. cerevisiae* maltose transport protein Mal31	13 ± 1	664 ± 3	17 ± 1	1.4	53
Pc16g05670	*Neurospora crassa* glucose transporter rco-3	63 ± 28	3176 ± 40	69 ± 1	1.1	51
Pc20g01790 (PcaraT)	*Kluyveromyces lactis* high-affinity glucose transporter HGT1	32 ± 6	1415 ± 42	46 ± 3	1.4	44
Pc22g14520	*S. cerevisiae* allantoate permease Dal5	19 ± 2	770 ± 104	28 ± 1	1.5	41
Pc13g04640	*K. lactis* high-affinity glucose transporter HGT1	29 ± 5	971 ± 32	53 ± 7	1.8	34
Pc21g10190	*K. lactis* high-affinity glucose transporter HGT1	12 ± 1	167 ± 26	12 ± 1	1.0	14
Pc12g00190	*Candida albicans* ABC transporter CDR4	13 ± 2	164 ± 24	29 ± 2	2.2	12
Pc14g01680	*Escherichia coli* L-fucose permease fucP	106 ± 14	1269 ± 172	68 ± 1	0.64	12.0
Pc21g12210	*Aspergillus nidulans* quinate transport protein qutD	12 ± 0	118 ± 1	12 ± 1	1	9.8
Pc06g01480	*S. cerevisiae* maltose transport protein Mal31	459 ± 85	3551 ± 102	226 ± 3	0.5	7.7
Pc13g10030	*S. cerevisiae* high-affinity nicotinic acid permease Tna1	125 ± 25	827 ± 33	216 ± 3	1.7	6.6
Pc21g09830	*K. lactis* high-affinity glucose transporter HGT1	185 ± 9	842 ± 1	126 ± 3	0.68	4.6
Pc16g02680	*S. cerevisiae* allantoate permease Dal5	80 ± 29	360 ± 14	113 ± 6	1.4	4.5
Pc12g05440	*S. cerevisiae* maltose transport protein Mal31	596 ± 201	2633 ± 64	104 ± 8	0.17	4.4
Pc13g15590	*S. cerevisiae* glucose permease Rgt2	12 ± 1	48.0 ± 1.0	12 ± 1	1	4.0
Pc13g06440	*S. cerevisiae* high-affinity nicotinic acid permease Tna1	66 ± 23	225 ± 11	48 ± 5	0.73	3.4

considered for further analysis as their regulation could have reflected unspecific D-glucose (de)repression.

An annotation screen indicated that 16 of the identified 'arabinose-induced' genes encoded putative transporters, whose transcript levels were 3.4 to 52-fold higher in the L-arabinose-limited cultures than in the D-glucose-limited cultures (**Table 3**). Five of these genes, whose transcript levels were at least 30-fold higher in L-arabinose-limited cultures than in D-glucose-limited cultures, shared similarity with the *S. cerevisiae* maltose transporter Mal31, the *Neurospora crassa* D-glucose transporter Rco-3, the *Kluyveromyces lactis* high-affinity D-glucose transporter Hgt1 and the *S. cerevisiae* allantoate transporter Dal5. These five transporter genes (Pc13g08230, Pc16g05670, Pc20g01790, Pc22g14520, and Pc13g04640, respectively) were selected for further functional analysis.

*Pc*AraT: A *P. chrysogenum* L-arabinose transporter that can be functionally expressed in *S. cerevisiae*

S. cerevisiae strains in which *HXT* transporter genes have been deleted and which express heterologous pathways for pentose metabolism have proven to be powerful platforms for screening and characterization of heterologous pentose transporter genes (Subtil and Boles 2011, Knoshaug *et al.* 2015, Sloothaak *et al.* 2016). To enable screening for *P. chrysogenum* L-arabinose transporters, *S. cerevisiae* strains were first engineered for L-arabinose consumption. Using CRISPR/Cas9-mediated *in-vivo* assembly (Mans *et al.* 2015), the overexpression cassettes for all structural genes involved in the non-oxidative pentose-phosphate pathway (*TAL1, NQM1, TKL1, TKL2, RKI1, RPE1*) were stably integrated into the *GRE3* locus, thereby inactivating synthesis of the Gre3 aldose reductase. Subsequently, nine copies of an expression cassette for overexpression of codon-optimized *Lactobacillus plantarum* L-arabinose isomerase AraA and single copies of *L. plantarum* AraB (L-ribulokinase) and AraD (L-ribulose-5-phosphate-4-epimerase) expression cassettes were integrated into the *GAL80* locus, using a strain-construction strategy previously described for expression of a D-xylose pathway into *S. cerevisiae* (Verhoeven *et al.* 2017). This integration inactivated *GAL80* and thereby alleviated transcriptional repression by D-glucose of *GAL2*, which encodes the major L-arabinose transporter in *S. cerevisiae* (Torchia *et al.* 1984, Tani *et al.* 2016). The resulting strain IMX929 was able to grow in liquid media supplemented with L-arabinose as sole carbon source and was used as a platform strain to test if any of the five selected putative *P. chrysogenum* transporter genes, placed under the control of the constitutive ADH1 promoter, could support L-arabinose transport in *S. cerevisiae*.

To this end, single copies of codon-optimized expression cassettes were integrated into the *GAL2* locus of the L-arabinose metabolizing *S. cerevisiae* strain IMX928, a uracil auxotrophic

daughter strain of IMX929, thereby inactivating the *GAL2* gene. Consistent with previous studies (Subtil and Boles 2011, Knoshaug *et al.* 2015), inactivation of *GAL2* in strain IMX928 yielded a strain (IMX1504) that was unable to grow on SMA plates (**Figure 1**). All five strains in which *GAL2* had been replaced by putative *P. chrysogenum* transporter genes (IMX1505-1509) showed vigorous growth on SMD plates. However, only strain IMX1508, which expressed the *P. chrysogenum* gene Pc20g01790, showed growth on L-arabinose (**Figure 1**). Based on this observation, Pc20g01790 was designated *PcaraT* (*P. chrysogenum* Arabinose Transporter). A Blast-p search revealed strong homology of Pc20g01790 with the *K. lactis* gene *HGT1*, which encodes a high-affinity D-glucose and galactose transporter (Billard *et al.* 1996, Baruffini *et al.* 2006).

PcaraT encodes a high-affinity, high-specificity L-arabinose transporter

Sugar transport kinetics of *Pc*AraT were analysed using ^{14}C-labelled L-arabinose, D-xylose and D-glucose. To dissect transporter kinetics of *Pc*AraT and Gal2, their structural genes were separately expressed in *S. cerevisiae* DS68625 (Nijland *et al.* 2014). Each gene was introduced on a centromeric plasmid and expressed from the *HXT7* promoter. In strain DS68625, the major hexose transporter genes (*HXT1-7* and *GAL2*) are deleted, while its inability to metabolize L-arabinose enables the specific analysis of sugar uptake rather than the combination of radioactive sugar uptake and metabolism. The negative control strain DS68625-mcs (DS68625 transformed with the 'empty' centromeric plasmid pRS313-

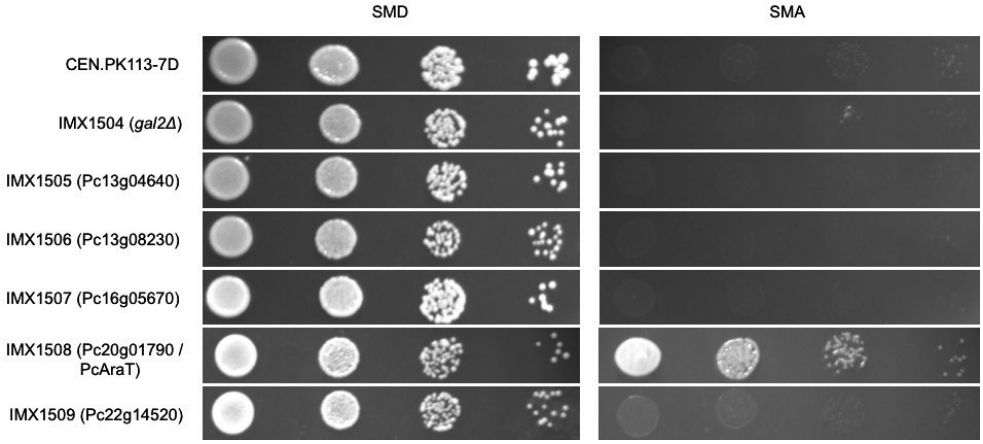


Figure 1. Impact of the expression of putative *P. chrysogenum* sugar transporter genes in an L-arabinose metabolizing *S. cerevisiae* strain in which *GAL2* was deleted. Strains were pregrown on liquid SMD and spotted on plates containing 20 g L^{-1} D-glucose (SMD, left) or L-arabinose (SMA, right) as carbon source. Codes on left hand side indicate *S. cerevisiae* strain names and, in brackets, the systematic name of the corresponding over-expressed *P. chrysogenum* gene. CEN.PK113-7D is a control strain that was not engineered for L-arabinose metabolism. SMD and SMA plates were incubated at 30 °C for 47 and 91 h, respectively. The experiment was performed in duplicate, data shown are from a single representative experiment.

mcs) did not show significant [^{14}C] L-arabinose uptake, while expression of either Gal2 or *Pc*AraT (strains DS68625-*GAL2* and DS68625-*PcaraT*, respectively) restored L-arabinose transport (**Table 4**).

In kinetic analyses, the K_m of *Pc*AraT for L-arabinose (0.13 mmol L^{-1}), was found to be three orders of magnitude lower than that of Gal2 (335 mmol L^{-1}), while its transport capacity (V_{max}) was 14-fold lower than that of Gal2 (5.3 and 75 nmol (mg biomass)$^{-1}$ min^{-1}, respectively) (**Table 4**). *Pc*AraT was found to be highly L-arabinose specific, as its expression in strain DS68625 did not support transport of either [^{14}C] D-glucose or [^{14}C] D-xylose. Consistent with earlier reports (Subtil and Boles 2011, Knoshaug *et al.* 2015), expression of Gal2 in strain DS68625 enabled transport of D-glucose (K_m = 1.9 mmol L^{-1}, V_{max} = 26 nmol (mg biomass)$^{-1}$ min^{-1}, while Gal2 has previously been shown to enable low-affinity D-xylose transport (K_m = 226 mmol L^{-1}; (Farwick *et al.* 2014)).

The impact of the presence of D-glucose and D-xylose on L-arabinose transport by Gal2 and *Pc*AraT was investigated in transport assays with 50 mmol L^{-1} [^{14}C] L-arabinose and increasing concentrations of non-radioactive D-glucose or D-xylose. In these assays, both transporters exhibited a reduced L-arabinose transport capacity in the presence of D-glucose or D-xylose (**Table 4, Additional Figure 1**). At a concentration of 100 mmol L^{-1} (i.e., twice the concentration of L-arabinose), D-xylose and D-glucose inhibited L-arabinose uptake rate via Gal2 by 29 and 85%, respectively. In contrast, L-arabinose transport via *Pc*AraT was less impaired at this concentration of D-xylose and, especially, D-glucose (22 and 63% inhibition, respectively). To study the transport mechanism of *Pc*AraT, the impact of the protonophore uncoupler CCCP on transport kinetics was tested. Transport

Table 4. Kinetic data for the *S. cerevisiae* transporter Gal2 and *P. chrysogenum Pc*AraT derived from uptake studies with ^{14}C-labelled L-arabinose, D-glucose, and D-xylose. Sugar transport kinetics were measured by uptake of ^{14}C-radiolabelled sugars by *S. cerevisiae* DS68625, an engineered strain lacking the Hxt1-7 and Gal2 transporters, expressing either *GAL2* or *PcaraT*. Transport inhibition was determined at 50 mmol L^{-1} [^{14}C] L-arabinose and 100 mmol L^{-1} of either D-glucose or D-xylose and expressed relative to the transport rate observed in the absence of D-xylose or D-glucose. Values are represented as average ± mean deviation of duplicate experiments. Graphs used to calculate kinetic parameters are shown in **Additional Figures 3-6**. ARA = L-arabinose; GLC = D-glucose; XYL = D-xylose.

	Gal2	*Pc*AraT
$K_{m, ARA}$ [mmol L^{-1}]	335 ± 21	0.13 ± 0.03
$V_{max, ARA}$ [nmol (mg biomass)$^{-1}$ min^{-1}]	75 ± 5.2	5.3 ± 0.2
$K_{m, GLC}$ [mmol L^{-1}]	1.9	-
$V_{max, GLC}$ [nmol (mg biomass)$^{-1}$ min^{-1}]	26	-
L-arabinose transport inhibition by glucose	85%	63%
$K_{m, XYL}$ [mmol L^{-1}]	226 (Farwick *et al.* 2014)	-
$V_{max, XYL}$ [nmol (mg biomass)$^{-1}$ min^{-1}]$^{+}$	91 (Farwick *et al.* 2014)	-
L-arabinose transport inhibition by D-xylose	29%	22%

of L-arabinose via Gal2, which mediates facilitated diffusion of sugars (Maier *et al.* 2002), was not affected by CCCP, while this uncoupler completely abolished transport via *Pc*AraT (**Additional Figure 5**). These results indicate that *Pc*AraT mediates proton-coupled import of L-arabinose.

Functional expression of *PcaraT* in an L-arabinose-fermenting *S. cerevisiae* strain enables L-arabinose consumption in the presence of D-glucose

The ability to transport L-arabinose in the presence of D-glucose is a highly relevant characteristic in the construction of platform *S. cerevisiae* strains for conversion of lignocellulosic hydrolysates (Jansen *et al.* 2017). To investigate whether expression of *PcaraT* can confer this ability, a set of three strains was constructed that (i) could not metabolize D-glucose due to the deletion of *HXK1, HXK2, GLK1* and *GAL1* [20, 62] (ii) (over)expressed non-oxidative PPP enzymes and the *L. plantarum* AraA, AraB and AraD genes to enable L-arabinose metabolism, and (iii) had different genotypes with respect to L-arabinose transport (*GAL2, PcaraT/gal2Δ* and *gal2Δ* in strains IMX660, IMX869 and IMX844, respectively). Since these 'arabinose specialist strains' cannot grow on D-glucose, the impact of the presence of D-glucose on L-arabinose metabolism can be directly measured via its effect on growth.

As anticipated, strain IMX844 (*gal2Δ*) was unable to grow on synthetic medium supplemented with either 20 $g \cdot L^{-1}$ L-arabinose or a mix of 20 $g \cdot L^{-1}$ of each, L-arabinose and D-glucose. In contrast, the L-arabinose specialist strains IMX660 (*GAL2*) and IMX869 (*PcaraT/gal2Δ*) grew on synthetic medium with L-arabinose as the sole carbon source at specific growth rates of 0.240 ± 0.001 h^{-1} and 0.099 ± 0.001 h^{-1}, respectively (**Figure 2A**). However, when 20 g L^{-1} D-glucose was added to the L-arabinose medium, strain IMX660 (*GAL2*) did not show growth during a 120 h batch cultivation experiment (**Figure 2B**), while strain IMX869 (*PcaraT/gal2Δ*) grew at 60% of the specific growth rate observed in the absence of D-glucose (μ = 0.057 ± 0.003 h^{-1} versus 0.099 ± 0.001 h^{-1}, **Figure 2B**). This result indicated that expression of *Pc*AraT in strain IMX869 enabled uptake of L-arabinose in the presence of D-glucose.

Low residual substrate concentrations in chemostat cultures confirm high-affinity L-arabinose transport kinetics of *Pc*AraT

To further evaluate the *in vivo* impact of L-arabinose transport via *Pc*AraT, biomass-specific L-arabinose consumption rates and residual substrate concentrations were analysed in L-arabinose-limited, aerobic chemostat cultures, grown at a dilution rate of 0.05 h^{-1}.

Under these conditions, the L-arabinose-metabolizing strain IMX1508 (*PcaraT/gal2Δ*) exhibited a residual L-arabinose concentration of only $4.2 \cdot 10^{-3}$ g L^{-1}, compared to 1.8 g L^{-1}

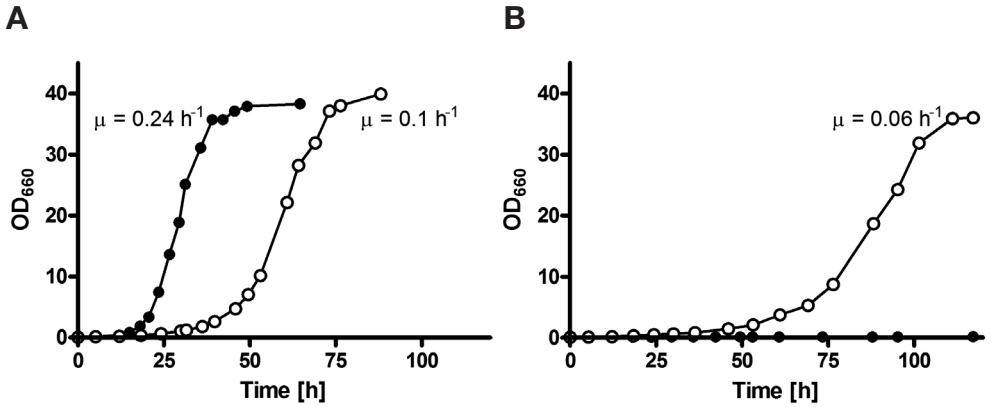


Figure 2. Growth curves of *S. cerevisiae* L-arabinose specialist strains, engineered for L-arabinose consumption and disabled for D-glucose consumption by deletion of the hexose kinase genes *HXK1*, *HXK2*, *GLK1* and *GAL1,* and expressing either *GAL2* (IMX660, filled circles) or the *P. chrysogenum* transporter *Pc*AraT (IMX869, open circles) as the sole L-arabinose transporter. To assess the ability of Gal2 and *Pc*AraT to support import of L-arabinose by growing cultures in absence (A) and presence (B) of D-glucose, specific growth rates were estimated from shake-flask cultures on synthetic media supplied with 20 g L^{-1} L-arabinose (A) and on synthetic media supplied with L-arabinose and d-glucose (20 g L^{-1} each, (B)).

in cultures of strain IMX929 (*GAL2*) (**Table 5**). In these growth experiments, different promoters were used for expression of *PcaraT* and *GAL2* (pADH1 and derepressed pGAL2, respectively). However, while this may moderately affect expression levels of the two transporters, this cannot explain the over 1000-fold difference in residual L-arabinose concentration. This difference was entirely consistent with the conclusion from the kinetic analyses of ^{14}C-L-arabinose uptake, in which both transporter genes were expressed from the same promoter (pHXT7) and which also indicated that *PcaraT* encodes an L-arabinose transporter with a much higher affinity for L-arabinose than Gal2. In shake-flask batch cultures grown on an initial L-arabinose concentration of 7.5 g L^{-1}, these strains exhibited initial specific growth rates of 0.085 h^{-1} and 0.13 h^{-1}, respectively. Based on this observation and on the high K_m of Gal2 for L-arabinose (Subtil and Boles 2011, Subtil and Boles 2012),

Table 5. Physiological data derived from steady-state chemostat cultures of engineered, L-arabinose-metabolizing *S. cerevisiae* strains. Strains expressing either *GAL2* (IMX929) or *PcaraT* (IMX1508) as sole functional L-arabinose transporter were grown in aerobic, L-arabinose-limited chemostat cultures (7.5 g L^{-1} L-arabinose, dilution rate = 0.05 h^{-1}, pH 5, T = 30 °C). Data are derived from independent triplicate experiments and presented as average ± mean deviation.

	IMX929 (*GAL2*)	IMX1508 (*PcaraT*)
Residual L-arabinose [g L^{-1}]	1.77 ± 0.19	0.004 ± 0.002
$Y_{X/S}$ [g biomass (g L-arabinose)$^{-1}$]	0.48 ± 0.06	0.40 ± 0.01
$q_{L\text{-arabinose}}$ [mmol g^{-1} h^{-1}]	0.70 ± 0.10	0.80 ± 0.08

this study), the *in vivo* activity of *Pc*AraT can be expected to exceed that of Gal2 when L-arabinose concentrations are below ca. 4 g L^{-1}.

In duplicate steady-state chemostat cultures, the biomass-specific L-arabinose consumption rate of strain IMX1508 (*PcaraT*) was approximately 14% higher than the one of strain IMX929 (*GAL2*; 0.8 ± 0.1 and 0.7 ± 0.1 mmol g^{-1} h^{-1}), reflecting the slightly lower biomass yield of the former strain. This difference in biomass yield is close to the difference of 8.1% that, based on published estimates of the P/O ratio and proton stoichiometry of the plasma-membrane ATPase in aerobic *S. cerevisiae* cultures (both close to 1.0, (Verduyn *et al.* 1990, Weusthuis *et al.* 1993)), would be expected if L-arabinose uptake via *Pc*AraT occurred via symport with a single proton.

Discussion

Chemostat-based transcriptome analysis of *Penicillium chrysogenum* proved to be an efficient method to identify candidate genes for L-arabinose transporters in this fungus. In comparison with similar studies in batch cultures, use of chemostat cultures offered several advantages. First, chemostat cultivation at a fixed dilution rate eliminated the impact of specific growth rate on transcriptional regulation (Daran-Lapujade *et al.* 2008). Furthermore, use of L-arabinose-limited chemostat cultures of *P. chrysogenum*, in which residual concentrations of this pentose were very low, enabled a focus on the identification of high-affinity transporters. Finally, the use of both D-glucose- and ethanol-limited cultures as references helped to eliminate transcriptional responses of *P. chrysogenum* that were specific to either of these two carbon sources, e.g. as a result of CreA-mediated D-glucose repression of relevant transporter genes (Cubero and Scazzocchio 1994, Cepeda-García *et al.* 2014). Although this study was focused on L-arabinose transport, the *P. chrysogenum* transcriptome dataset from D-glucose, ethanol and arabinose grown cultures generated in this study (available via GEO, [https://www.ncbi.nlm.nih.gov/geo/] under accession numbers GSE12632, GSE24212, and GSE104491, respectively) may contribute to studies on other aspects on metabolism and metabolic regulation in this industrially relevant fungus.

Of five putative transporter genes that showed an over 30-fold higher transcript level in L-arabinose-limited chemostat cultures of *P. chrysogenum* than in D-glucose-limited cultures, only *PcaraT* was shown to encode an L-arabinose transporter that is functional in *S. cerevisiae*. While the low K_m of this transporter observed upon its expression in *S. cerevisiae* is consistent with its upregulation in L-arabinose-limited cultures of *P. chrysogenum,* this observation does not necessarily imply that *Pc*AraT is the only or even the most important L-arabinose transporter active in these cultures. Problems in protein

folding, plasma-membrane (mis)targeting, post-translational modification and/or protein turnover (Hamacher *et al.* 2002, Jahn *et al.* 2002) may have affected expression of the other candidate genes. Indeed, in screening of cDNA libraries encoding putative heterologous transporters, typically only few of the candidate genes are found to enable transport of the substrate upon expression in *S. cerevisiae* (Du *et al.* 2010, Verho *et al.* 2011).

Several studies have used *gal2Δ* strains of *S. cerevisiae* to analyse transport kinetics of heterologous L-arabinose transporters (**Table 6, Figure 3**). Two studies that estimated K_m and V_{max} of Gal2 upon its reintroduction in such a strain found different results (**Table 6**) (Subtil and Boles 2011, Knoshaug *et al.* 2015). At L-arabinose concentrations of about 10 mmol L^{-1} these studies reported Gal2-mediated transport rates of 0.3 and 8.9 nmol (mg biomass)$^{-1}$ min^{-1}, respectively, as compared to a value of 2.5 nmol (mg biomass)$^{-1}$ min^{-1} observed in the present study. One of the previous studies (Knoshaug *et al.* 2015) used

2

Table 6. Comparison of key characteristics of Gal2, *Pc*AraT and heterologous L-arabinose transporters that were previously expressed in *S. cerevisiae*. n.d = not determined; GLC = D-glucose; XYL = D-xylose.

Protein	Origin	K_m [mM]	V_{max} [nmol (mg biomass)$^{-1}$ min^{-1}]	GLC trans-port	XYL trans-port	Mechanism	References
*Sc*Gal2	S. cerevisiae	335 ± 21.0[a] 57 ± 11[b] 371 ± 19[c]	75 ± 5 [a] 2.2 ± 0.3 [b] 18 ± 0.8[c]	✓	✓	facilitated diffusion	[a]This study [b](Subtil and Boles 2011) [c](Knoshaug *et al.* 2015)
*Pc*AraT	P. chrysogenum	0.13 ± 0.03	5.3 ± 0.2	×	×	H^+ symport	This study
*Ss*AraT	Scheffersomyces stipitis	3.8 ± 1.7	0.4 ± 0.1	✓	×	n.d	(Subtil and Boles 2011)
*At*Stp2	Arabidopsis thaliana	4.5 ± 2.2	0.6 ± 0.1	×	×	H^+ symport	(Subtil and Boles 2011)
*Km*Axt1	Kluyveromyces marxianus	263 ± 57	57 ± 6	×	✓	facilitated diffusion	(Knoshaug *et al.* 2015)
*Pg*Axt1	Pichia guilliermondii	0.13 ± 0.04	18 ± 0.8	×	✓	H^+ symport	(Knoshaug *et al.* 2015)
*Am*Lat1	Ambrosiozyma monospora	0.03*	0.2 ± 0.0	×	×	n.d	(Verho *et al.* 2011, Londesborough *et al.* 2014)
*Am*Lat2	A. monospora	n.d	4 ± 0	×	×	n.d	(Verho *et al.* 2011, Londesborough *et al.* 2014)
*Nc*Lat-1	Neurospora-crassa	58 ± 4	1945 ± 50	✓	n.d	H^+ symport	(Li *et al.* 2015)
*Mt*Lat-1	Myceliophthora. thermophila	29 ± 4	172 ± 6	×	n.d	H^+ symport	(Li *et al.* 2015)

*K_m of *Am*Lat1 was determined as a GFP-fusion protein(Londesborough *et al.* 2014)

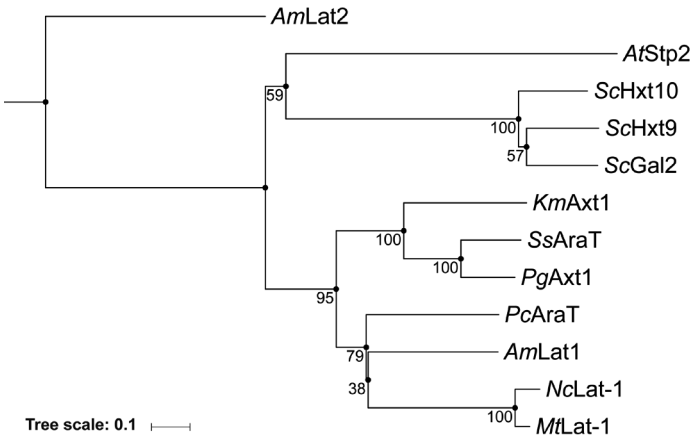


Figure 3. Phylogenetic tree of *S. cerevisiae* Gal2, *Pc*AraT, and other heterologous L-arabinose transporters that have previously been functionally expressed in *S. cerevisiae*. Species names are added in two-letter code in front of protein names. Numbers are derived from a 500-times bootstrap iteration. Characteristics and literature references for each transporter are provided in Table 6. Accession numbers: *Sc*Gal2: P13181, *Pc*AraT: CAP85508, *Ss*AraT: A3LQQ5-1, *At*Stp2: OAP13698, *Km*Axt1: GZ791039, *Pg*Axt1: GZ791040, *Am*Lat1: AY923868, *Am*Lat2: AY923869, *Nc*Lat-1: EAA30346, *Mt*Lat-1: G2QFT5-1.

a strain that also expressed a functional bacterial L-arabinose pathway, thereby raising the possibility that apparent uptake rates were enhanced by subsequent metabolism of L-arabinose. Moreover, in different studies, *GAL2* was expressed from different promoters (*pTDH3, pADH1, and pHXT7*) and either high-copy number (2μ) (Subtil and Boles 2011, Knoshaug *et al.* 2015) or low-copy number centromeric (this study) expression plasmids. D-Glucose transport kinetics via Gal2 determined in this study (K_m = 1.9 mmol L^{-1}, V_{max} = 26 nmol (mg biomass)$^{-1}$ min^{-1}) were similar to previously reported values (1.5 mmol L^{-1} and 27 nmol (mg biomass)$^{-1}$ min^{-1} (Farwick *et al.* 2014).

L-Arabinose transport rates in L-arabinose-limited chemostat cultures of both Gal2- and *Pc*AraT-dependent strains were higher than the V_{max} values calculated from transporter assays with radioactively labelled L-arabinose. A similar difference between transport assays and rates of L-arabinose uptake in growing cultures was reported by Knoshaug *et al.* (Knoshaug *et al.* 2015). These discrepancies suggest that either the transport assays did not accurately reflect zero-trans-influx kinetics (Teusink *et al.* 1998) or that differences in experimental conditions and/or cellular energy status between transport assays and chemostat cultures influenced L-arabinose uptake.

Assuming that *Pc*AraT mediates symport of L-arabinose with a single proton, the L-arabinose consumption rate in aerobic, L-arabinose-limited chemostat cultures of the strain IMX1508 (*Pc*AraT *gal2Δ*) (0.8 mmol (g biomass) h^{-1}; **Table 5**) would, under anaerobic

conditions, correspond to an ATP production rate of ca. 0.3 mmol ATP (g biomass)$^{-1}$ h^{-1}. This rate of ATP production is well below the reported ATP requirement of anaerobic *S. cerevisiae* cultures for cellular maintenance (ca. 1 mmol ATP (g biomass)$^{-1}$ h^{-1} (Boender *et al.* 2009). Consistent with this observation, no growth on L-arabinose as sole carbon source was observed in anaerobic shake flask cultures of the L-arabinose specialist strain IMX869 strain (*Pc*AraT *gal2Δ*) (data not shown).

Differences in experimental protocols for strain-construction and sugar-uptake studies, as well as the different kinetics observed in transport assays and growing cultures, complicate quantitative comparisons between different studies. Nevertheless, some important differences can be discerned between the heterologous L-arabinose transporters that have hitherto been expressed in *S. cerevisiae* (**Table 6, Figure 3**). Protein sequence alignment of *Pc*AraT and transporters that were previously shown to mediate L-arabinose import in *S. cerevisiae* showed that *Pc*AraT clusters with *Ambriosozyma monospora Am*Lat1 (**Figure 3**).

In terms of its low K_m, *Pc*AraT most closely resembled *Am*Lat1 and the *Pichia guillermondii Pg*Axt1 transporter. However, expression in *S. cerevisiae* of *Am*Lat1 (Verho *et al.* 2011, Londesborough *et al.* 2014) led to ~25-fold lower reported V_{max} of L-arabinose uptake than found in the present study for *Pc*AraT. In contrast to *Pc*AraT, *Pg*Axt was able to transport D-glucose, which might contribute to the strong inhibition of the latter transporter by D-glucose (Knoshaug *et al.* 2015). Although *Pc*AraT resembled *A. thaliana* Stp2 (Subtil and Boles 2011) in being partially inhibited by D-glucose despite an inability to transport this sugar, *Pc*AraT enabled consumption of L-arabinose in batch cultures containing 20 g L^{-1} D-glucose. In common with other high-affinity sugar transporters in yeasts (Knoshaug *et al.* 2015) and the observation that *Pc*AraT mediates L-arabinose-proton symport should be taken into account in future strain designs, since simultaneous activity of proton symport and facilitated diffusion, e.g. via Gal2, may result in energy-consuming futile cycles (Jansen *et al.* 2017).

In the lignocellulosic hydrolysates now used in the first industrial-scale plants for 'second generation' bioethanol production, L-arabinose generally represents between 2 and 3% of the total sugars (Jansen *et al.* 2017). At the resulting low concentrations of L-arabinose in the industrial processes, Gal2 operates far from substrate saturation and is, moreover, strongly inhibited by D-glucose. Based on its kinetic characteristics, as analysed in transport assays and growing cultures, *Pc*AraT represents an interesting candidate transporter for evaluation of L-arabinose co-consumption under industrial conditions. If the characteristics of *Pc*AraT determined in the present study can be reproduced in industrial strains and under industrial conditions, this transporter can contribute to a timely and efficient conversion of L-arabinose and, thereby to the overall process economics.

Conclusion

Transcriptome analyses of L-arabinose-limited *Penicillium chrysogenum* chemostat cultures proved valuable for identification of the high-affinity L-arabinose transporter *Pc*AraT. Functional expression and characterization in *Saccharomyces cerevisiae* revealed a high affinity and specificity of this transporter for L-arabinose (K_m = 0.13 mmol L^{-1}), combined with a limited sensitivity to inhibition by D-glucose and D-xylose, which are present at high concentrations in lignocellulosic hydrolysates. These characteristics differentiate *Pc*AraT from the endogenous *S. cerevisiae* transporter capable of L-arabinose transport (Gal2) and qualify it as a potentially valuable additional element in metabolic engineering strategies towards efficient and complete conversion of L-arabinose present in second generation feedstocks for yeast-based production of fuels and chemicals.

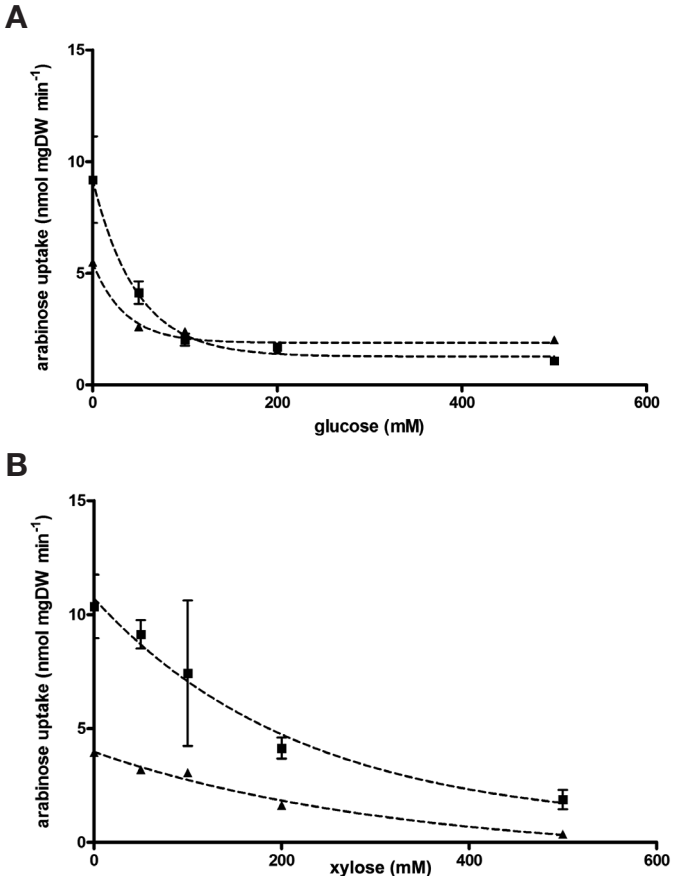
Acknowledgements

The authors thank Andreas K. Gombert and Isthar Snoek for help with the *P. chrysogenum* chemostat cultures, Aurin M. Vos for help with the generation of the phylogenetic tree (**Figure 3**) and Ioannis Papapetridis and Paul de Waal for valuable input in this project.

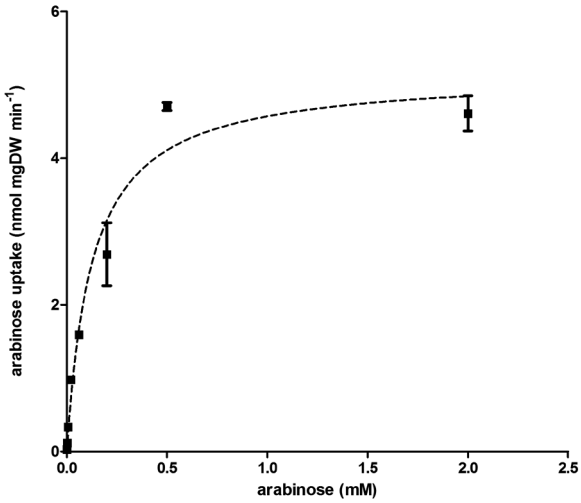
Additional material

Additional Table 1. Primers used in this study. File can be downloaded at: **https://doi.org/10.1186/s13068-018-1047-6**

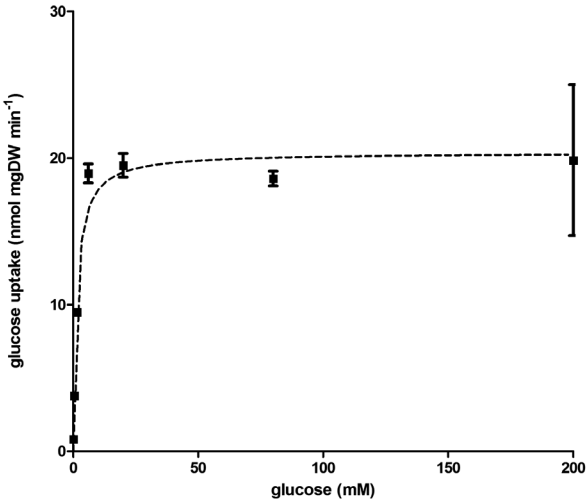
Additional Table 2. Differentially expressed genes in *P. chrysogenum* chemostat cultures over three different conditions (D-glucose-, L-arabinose, and ethanol-limited). The table includes 540 genes that were differentially expressed amongst D-glucose-, L-arabinose, and ethanol-limited chemostat cultures of *P. chrysogenum* (7.5 g L^{-1}, 7.5 g L^{-1}, 5.8 g L^{-1}, respectively, D = 0.03 h^{-1}, T = 30 °C). 137 transcripts exhibited a fold-change higher than three on L-arabinose and lower than three on ethanol relative to the D-glucose condition (the top 137 genes presented in the table). Shown are averages from triplicate (D-glucose) or duplicate chemostat cultures (L-arabinose, ethanol), the respective standard deviation and the ratio of gene expression levels in the presence of L-arabinose over D-glucose (Ara vs. Glc), as well as ethanol over D-glucose (EtOH vs. Glc). Excel file can be downloaded at: **https://doi.org/10.1186/s13068-018-1047-6**



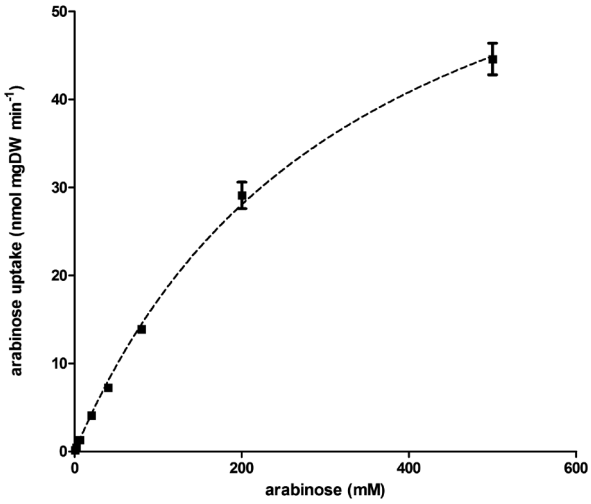
Additional Figure 1. Effect of D-glucose (A) and D-xylose (B) on the specific rate of L-arabinose uptake by *Pc*AraT (filled triangles) and Gal2 (filled squares). Uptake experiments were performed with 50 mmol L^{-1} [^{14}C-] l-arabinose in the presence of increasing concentrations of D-glucose (A) or D-xylose (B). Symbols indicate uptake rates observed with the Hxt1-7 and Gal2 deletion strain *S. cerevisiae* DS68625-PcaraT (closed triangles) and DS68625-GAL2, expressing either *Pc*AraT or Gal2, respectively. Data are derived from duplicate experiments and shown as the average ± mean deviation.



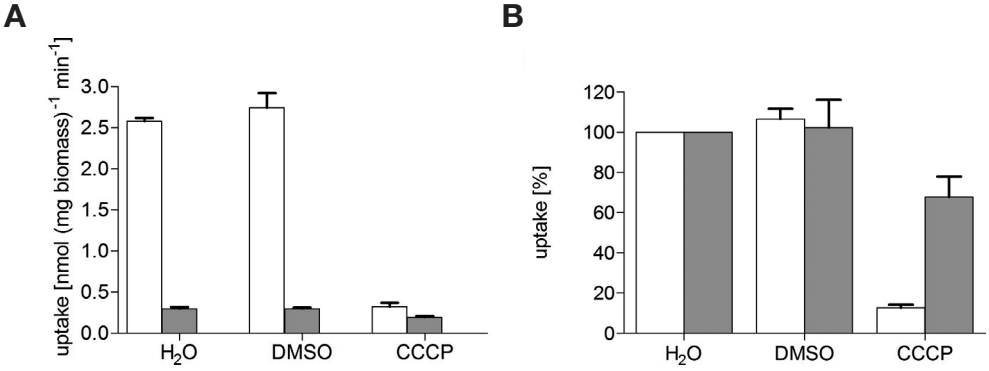
Additional Figure 2. Specific rate of L-arabinose uptake by *Pc*AraT. Uptake experiments were performed with increasing concentrations of [^{14}C-] L-arabinose with the Hxt1-7 and Gal2 deletion strain *S. cerevisiae* DS68625-*PcaraT* expressing *PcaraT* on a centromeric plasmid. No [^{14}C-] D-glucose uptake was observed for this strain. Data are derived from duplicate experiments and shown as the average ± mean deviation.



Additional Figure 3. Specific rate of D-glucose uptake by Gal2. Uptake experiments were performed with increasing concentrations of [^{14}C-] D-glucose with the Hxt1-7 and Gal2 deletion strain *S. cerevisiae* DS68625-GAL2 expressing GAL2 on a centromeric plasmid. Data are derived from duplicate experiments and shown as the average ± mean deviation.



Additional Figure 4. Specific rate of L-arabinose uptake by Gal2. Uptake experiments were performed with increasing concentrations of [^{14}C-] L-arabinose with the Hxt1-7 and Gal2 deletion strain *S. cerevisiae* DS68625-GAL2 expressing GAL2 on a centromeric plasmid. Data are derived from duplicate experiments and shown as the average ± mean deviation.



Additional Figure 5. Impact of proton-gradient uncoupling on transport activity. Transport rates of [^{14}C]-L-arabinose of the Hxt1-7 and Gal2 deletion strains DS68625-*PcaraT* and DS68625-*GAL2* expressing either *PcAraT* (DS68625-*PcaraT*, white bars) or Gal2 (DS68625-*GAL2*, grey bars) on a centromeric plasmid. Transport rates were determined in 200 µl synthetic medium at a [^{14}C]- L-arabinose concentration of 2 mmol L^{-1} upon addition of either 0.5 µL water, 0.5 µL DMSO, or 10 µM CCCP (0.5 µl of a stock solution dissolved in 100% DMSO) (**A**). Panel (**B**) shows the uptake capacity in % relative to the control (H_2O). Data are derived from duplicate experiments and shown as the average ± mean deviation.

References

Baruffini E, Goffrini P, Donnini C *et al.* Galactose transport in *Kluyveromyces lactis*: major role of the glucose permease Hgt1. *FEMS Yeast Res* 2006;**6**:1235-42.

Becker J, Boles E. A modified *Saccharomyces cerevisiae* strain that consumes L-arabinose and produces ethanol. *Appl Environ Microbiol* 2003;**69**:4144-50.

Bettiga M, Hahn-Hägerdal B, Gorwa-Grauslund MF. Comparing the xylose reductase/xylitol dehydrogenase and xylose isomerase pathways in arabinose and xylose fermenting *Saccharomyces cerevisiae* strains. *Biotechnol Biofuels* 2008;**1**:16.

Billard P, Ménart S, Blaisonneau J *et al.* Glucose uptake in K*luyveromyces lactis*: role of the *HGT1* gene in glucose transport. *J Bacteriol* 1996;**178**:5860-6.

Boender LG, de Hulster EA, van Maris AJ *et al.* Quantitative physiology of *Saccharomyces cerevisiae* at near-zero specific growth rates. *Appl Environ Microbiol* 2009;**75**:5607-14.

Bracher JM, de Hulster E, Koster CC *et al.* Laboratory evolution of a biotin-requiring *Saccharomyces cerevisiae* strain for full biotin prototrophy and identification of causal mutations. *Appl Environ Microbiol* 2017.

Cepeda-García C, Domínguez-Santos R, García-Rico RO *et al.* Direct involvement of the CreA transcription factor in penicillin biosynthesis and expression of the *pcbAB* gene in *Penicillium chrysogenum*. *Appl Microbiol Biotechnol* 2014;**98**:7113-24.

Chiang C, Knight S. A new pathway of pentose metabolism. *Biochem Biophys Res Commun* 1960;**3**:554-9.

Chiang C, Knight S. L-Arabinose metabolism by cell-free extracts of *Penicillium chrysogenum*. *Biochim Biophys Acta* 1961;**46**:271-8.

Cubero B, Scazzocchio C. Two different, adjacent and divergent zinc finger binding sites are necessary for CREA-mediated carbon catabolite repression in the proline gene cluster of *Aspergillus nidulans*. *EMBO J* 1994;**13**:407.

Daran-Lapujade P, Daran J-M, van Maris AJ *et al.* Chemostat-based micro-array analysis in baker's yeast. *Adv Microb Physiol* 2008;**54**:257-417.

DiCarlo JE, Norville JE, Mali P *et al.* Genome engineering in *Saccharomyces cerevisiae* using CRISPR-Cas systems. *Nucleic Acids Res* 2013:gkt135.

Du J, Li S, Zhao H. Discovery and characterization of novel D-xylose-specific transporters from *Neurospora crassa* and *Pichia stipitis*. *Mol Biosyst* 2010;**6**:2150-6.

Entian K-D, Kötter P. 25 Yeast genetic strain and plasmid collections. *Methods Microbiol* 2007;**36**:629-66.

Farwick A, Bruder S, Schadeweg V *et al.* Engineering of yeast hexose transporters to transport D-xylose without inhibition by D-glucose. *Proc Natl Acad Sci USA* 2014;**111**:5159-64.

Gietz RD, Woods RA. Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method. *Methods Enzymol* 2002;**350**:87-96.

Grohmann K, Bothast RJ. Saccharification of corn fibre by combined treatment with dilute sulphuric acid and enzymes. *Process Biochem* 1997;**32**:405-15.

Gueldener U, Heinisch J, Koehler G *et al.* A second set of loxP marker cassettes for Cre-mediated multiple gene knockouts in budding yeast. *Nucleic Acids Res* 2002;**30**:e23-e.

Güldener U, Heck S, Fiedler T *et al.* A new efficient gene disruption cassette for repeated use in budding yeast. *Nucleic Acids Res* 1996;**24**:2519-24.

Hahn-Hägerdal B, Galbe M, Gorwa-Grauslund M-F *et al.* Bio-ethanol–the fuel of tomorrow from the residues of today. *Trends Biotechnol* 2006;**24**:549-56.

Hamacher T, Becker J, Gárdonyi M *et al.* Characterization of the xylose-transporting properties of yeast hexose transporters and their influence on xylose utilization. *Microbiology* 2002;**148**:2783-8.

Harris DM, van der Krogt ZA, Klaassen P *et al.* Exploring and dissecting genome-wide gene expression responses of *Penicillium chrysogenum* to phenylacetic acid consumption and penicillinG production. *BMC Genomics* 2009;**10**:75.

Horak J, Regelmann J, Wolf DH. Two distinct proteolytic systems responsible for glucose-induced degradation of fructose-1,6-bisphosphatase and the Gal2p transporter in the yeast *Saccharomyces cerevisiae* share the same protein components of the glucose signaling pathway. *J Biol Chem* 2002;**277**:8248-54.

Horak J, Wolf DH. Catabolite inactivation of the galactose transporter in the yeast *Saccharomyces cerevisiae*: ubiquitination, endocytosis, and degradation in the vacuole. *J Bacteriol* 1997;**179**:1541-9.

Inoue H, Nojima H, Okayama H. High efficiency transformation of *Escherichia coli* with plasmids. *Gene* 1990;**96**:23-8.

Jahn TP, Schulz A, Taipalensuu J *et al.* Post-translational modification of plant plasma membrane H+-ATPase as a requirement for functional complementation of a yeast transport mutant. *J Biol Chem* 2002;**277**:6353-8.

Jansen ML, Bracher JM, Papapetridis I *et al. Saccharomyces cerevisiae* strains for second-generation ethanol production: from academic exploration to industrial implementation. *FEMS Yeast Res* 2017.

Katoh K, Kuma K-i, Toh H *et al.* MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Res* 2005;**33**:511-8.

Katoh K, Rozewicki J, Yamada KD. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Brief Bioinform* 2017:bbx108.

Klaassen P, Van Der Laan JM, Gielesen BEM *et al.* Pentose sugar fermenting cell. US Patent 8.399.215 B2. 2013.

Knoshaug EP, Vidgren V, Magalhães F *et al.* Novel transporters from *Kluyveromyces marxianus* and *Pichia guilliermondii* expressed in *Saccharomyces cerevisiae* enable growth on L-arabinose and D-xylose. *Yeast* 2015;**32**:615-28.

Kou S, Christensen MS, Cirillo VP. Galactose transport in *Saccharomyces cerevisiae* II. Characteristics of galactose uptake and exchange in galactokinaseless cells. *J Bacteriol* 1970;**103**:671-8.

Kuijpers N, Solis-Escalante D, Bosman L *et al.* A versatile, efficient strategy for assembly of multi-fragment expression vectors in *Saccharomyces cerevisiae* using 60 bp synthetic recombination sequences. *Microb Cell Fact* 2013;**12**:47.

Kuijpers NG, Solis-Escalante D, Luttik MA *et al.* Pathway swapping: Toward modular engineering of essential cellular processes. *Proceedings of the National Academy of Sciences* 2016:201606701.

Kuyper M, Hartog MM, Toirkens MJ *et al.* Metabolic engineering of a xylose-isomerase-expressing *Saccharomyces cerevisiae* strain for rapid anaerobic xylose fermentation. *FEMS Yeast Res* 2005;**5**:399-409.

Leandro MJ, Fonseca C, Gonçalves P. Hexose and pentose transport in ascomycetous yeasts: an overview. *FEMS Yeast Res* 2009;**9**:511-25.

Lee W, Kim M, Ryu Y *et al.* Kinetic studies on glucose and xylose transport in *Saccharomyces cerevisiae*. *Appl Microbiol Biotechnol* 2002;**60**:186-91.

Letunic I, Bork P. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res* 2016;**44**:W242-W5.

Li J, Xu J, Cai P *et al.* Functional analysis of two L-arabinose transporters from filamentous fungi reveals promising characteristics for improved pentose utilization in *Saccharomyces cerevisiae*. *Appl Environ Microbiol* 2015;**81**:4062-70.

Lin Y, Tanaka S. Ethanol fermentation from biomass resources: current state and prospects. *Appl Microbiol Biotechnol* 2006;**69**:627-42.

Londesborough J, Richard P, Valkonen M *et al.* Effect of C-terminal protein tags on pentitol and L-arabinose transport by *Ambrosiozyma monospora Lat1* and *Lat2* transporters in *Saccharomyces cerevisiae*. *Appl Environ Microbiol* 2014;**80**:2737-45.

Lõoke M, Kristjuhan K, Kristjuhan A. Extraction of genomic DNA from yeasts for PCR-based applications. *BioTechniques* 2011;**50**:325.

Lynd LR. Overview and evaluation of fuel ethanol from cellulosic biomass: technology, economics, the environment, and policy. *Annu Rev Energy Env* 1996;**21**:403-65.

Maier A, Völker B, Boles E *et al.* Characterisation of glucose transport in *Saccharomyces cerevisiae* with plasma membrane vesicles (countertransport) and intact cells (initial uptake) with single Hxt1, Hxt2, Hxt3, Hxt4, Hxt6, Hxt7 or Gal2 transporters. *FEMS Yeast Res* 2002;**2**:539-50.

Mans R, van Rossum HM, Wijsman M *et al.* CRISPR/Cas9: a molecular Swiss army knife for simultaneous introduction of multiple genetic modifications in *Saccharomyces cerevisiae*. *FEMS Yeast Res* 2015;**15**::fov004.

Micard V, Renard C, Thibault J-F. Enzymatic saccharification of sugar-beet pulp. *Enzyme Microb Technol* 1996;**19**:162-70.

Moysés DN, Reis VCB, de Almeida JRMM, Lidia Maria Pepe de *et al.* Xylose fermentation by *Saccharomyces cerevisiae*: challenges and prospects. *Int J Mol Sci* 2016;**17**:207.

Nijland JG, Shin HY, de Jong RM *et al.* Engineering of an endogenous hexose transporter into a specific D-xylose transporter facilitates glucose-xylose co-consumption in *Saccharomyces cerevisiae*. *Biotechnol Biofuels* 2014;**7**:168.

Özcan S, Johnston M. Function and regulation of yeast hexose transporters. *Microbiol Mol Biol Rev* 1999;**63**:554-69.

Reifenberger E, Boles E, Ciriacy M. Kinetic characterization of individual hexose transporters of *Saccharomyces cerevisiae* and their relation to the triggering mechanisms of glucose repression. *Eur J Biochem* 1997;**245**:324-33.

Renewable Fuels Association. World fuel ethanol production. http://ethanolrfa.org/resources/industry/statistics/ (24 March 2017, date last accessed).

Richard P, Verho R, Putkonen M *et al.* Production of ethanol from L-arabinose by *Saccharomyces cerevisiae* containing a fungal L-arabinose pathway. *FEMS Yeast Res* 2003;**3**:185-9.

Roubos JA, Van Peij NNME. A method for achieving improved polypeptide expression date last accessed).

Sakamoto T, Ihara H, Kozaki S *et al.* A cold-adapted endo-arabinanase from *Penicillium chrysogenum*. *Biochim Biophys Acta* 2003;**1624**:70-5.

Sakamoto T, Ogura A, Inui M *et al.* Identification of a GH62 α-l-arabinofuranosidase specific for arabinoxylan produced by *Penicillium chrysogenum*. *Appl Microbiol Biotechnol* 2011;**90**:137-46.

Saloheimo A, Rauta J, Stasyk V *et al.* Xylose transport studies with xylose-utilizing *Saccharomyces cerevisiae* strains expressing heterologous and homologous permeases. *Appl Microbiol Biotechnol* 2007;**74**:1041-52.

Sloothaak J, Tamayo-Ramos JA, Odoni DI *et al.* Identification and functional characterization of novel xylose transporters from the cell factories *Aspergillus niger* and *Trichoderma reesei*. *Biotechnol Biofuels* 2016;**9**:1-15.

Solis-Escalante D, Kuijpers NG, Barrajon-Simancas N *et al.* A minimal set of glycolytic genes reveals strong redundancies in *Saccharomyces cerevisiae* central metabolism. *Eukaryot Cell* 2015;**14**:804-16.

Solis-Escalante D, Kuijpers NG, Bongaerts N *et al. amdSYM*, a new dominant recyclable marker cassette for *Saccharomyces cerevisiae*. *FEMS Yeast Res* 2013;**13**:126-39.

Subtil T, Boles E. Improving L-arabinose utilization of pentose fermenting *Saccharomyces cerevisiae* cells by heterologous expression of L-arabinose transporting sugar transporters. *Biotechnol Biofuels* 2011;**4**:38.

Subtil T, Boles E. Competition between pentoses and glucose during uptake and catabolism in recombinant *Saccharomyces cerevisiae*. *Biotechnol Biofuels* 2012;**5**:1.

Tani T, Taguchi H, Fujimori KE *et al.* Isolation and characterization of xylitol-assimilating mutants of recombinant *Saccharomyces cerevisiae*. *J Biosci Bioeng* 2016;**122**:446-55.

Teusink B, Diderich JA, Westerhoff HV *et al.* Intracellular glucose concentration in derepressed yeast cells consuming glucose is high enough to reduce the glucose transport rate by 50%. *J Bacteriol* 1998;**180**:556-62.

Torchia T, Hamilton R, Cano C *et al.* Disruption of regulatory gene *GAL80* in *Saccharomyces cerevisiae*: effects on carbon-controlled regulation of the galactose/melibiose pathway genes. *Mol Cell Biol* 1984;**4**:1521-7.

Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci* 2001;**98**:5116-21.

Van den Berg MA. Impact of the *Penicillium chrysogenum* genome on industrial production of metabolites. *Appl Microbiol Biotechnol* 2011;**92**:45-53.

Van Den Berg MA, Albang R, Albermann K *et al.* Genome sequencing and analysis of the filamentous fungus *Penicillium chrysogenum*. *Nat Biotechnol* 2008;**26**:1161-8.
Van Rossum HM, Kozak BU, Niemeijer MS *et al.* Requirements for carnitine shuttle-mediated translocation of mitochondrial acetyl moieties to the yeast cytosol. *MBio* 2016;**7**:e00520-16.
Verduyn C, Postma E, Scheffers WA *et al.* Energetics of *Saccharomyces cerevisiae* in anaerobic glucose-limited chemostat cultures. *Microbiology* 1990;**136**:405-12.
Verduyn C, Postma E, Scheffers WA *et al.* Effect of benzoic acid on metabolic fluxes in yeasts: a continuous-culture study on the regulation of respiration and alcoholic fermentation. *Yeast, 81992 nr 7* 1992.
Verho R, Penttilä M, Richard P. Cloning of two genes (*LAT1, 2*) encoding specific L-arabinose transporters of the L-arabinose fermenting yeast *Ambrosiozyma monospora*. *Appl Biochem Biotechnol* 2011;**164**:604-11.
Verhoeven MD, Lee M, Kamoen L *et al.* Mutations in *PMR1* stimulate xylose isomerase activity and anaerobic growth on xylose of engineered *Saccharomyces cerevisiae* by influencing manganese homeostasis. *Sci Rep* 2017;**7**.
Weusthuis RA, Adams H, Scheffers WA *et al.* Energetics and kinetics of maltose transport in *Saccharomyces cerevisiae*: a continuous culture study. *Appl Environ Microbiol* 1993;**59**:3102-9.
Wiedemann B, Boles E. Codon-optimized bacterial genes improve L-arabinose fermentation in recombinant *Saccharomyces cerevisiae*. *Appl Environ Microbiol* 2008;**74**:2043-50.
Wisselink HW, Toirkens MJ, del Rosario Franco Berriel M *et al.* Engineering of *Saccharomyces cerevisiae* for efficient anaerobic alcoholic fermentation of L-arabinose. *Appl Environ Microbiol* 2007;**73**:4881-91.
Xia PF, Zhang GC, Liu JJ *et al.* GroE chaperonins assisted functional expression of bacterial enzymes in *Saccharomyces cerevisiae*. *Biotechnol Bioeng* 2016.
Young E, Lee S-M, Alper H. Optimizing pentose utilization in yeast: the need for novel tools and approaches. *Biotechnol Biofuels* 2010;**3**:24.

3.

Laboratory evolution of a glucose-phosphorylation-deficient, arabinose-fermenting *S. cerevisiae* strain reveals mutations in *GAL2* that enable glucose-insensitive L-arabinose uptake

Maarten D. Verhoeven, Jasmine M. Bracher, Jeroen G. Nijland, Jonna Bouwknegt, Jean-Marc G. Daran, Arnold J.M. Driessen, Antonius J.A. van Maris and Jack T. Pronk

Essentially as published in **FEMS Yeast Research** 2018, foy062
(https://doi.org/10.1093/femsyr/foy062)

Abstract

Cas9-assisted genome editing was used to construct an engineered glucose-phosphorylation-negative *S. cerevisiae* strain, expressing the *Lactobacillus plantarum* L-arabinose pathway and the *Penicillium chrysogenum* transporter *Pc*AraT. This strain, which showed a growth rate of 0.26 h^{-1} on L-arabinose in aerobic batch cultures, was subsequently evolved for anaerobic growth on L-arabinose in the presence of D-glucose and D-xylose. In four strains isolated from two independent evolution experiments the galactose-transporter gene *GAL2* had been duplicated, with all alleles encoding $Gal2^{N376T}$ or $Gal2^{N376I}$ substitutions. In one strain, a single *GAL2* allele additionally encoded a $Gal2^{T89I}$ substitution, which was subsequently also detected in the independently evolved strain IMS0010. In ^{14}C-sugar-transport assays, $Gal2^{N376S}$, $Gal2^{N376T}$ and $Gal2^{N376I}$ substitutions showed a much lower glucose sensitivity of L-arabinose transport and a much higher K_m for D-glucose transport than wild-type Gal2. Introduction of the $Gal2^{N376I}$ substitution in a non-evolved strain enabled growth on L-arabinose in the presence of D-glucose. $Gal2^{N376T, T89I}$ and $Gal2^{T89I}$ variants showed a lower K_m for L-arabinose and a higher K_m for D-glucose than wild-type Gal2, while reverting $Gal2^{N376T, T89I}$ to $Gal2^{N376}$ in an evolved strain negatively affected anaerobic growth on arabinose. This study indicates that optimal conversion of mixed-sugar feedstocks may require complex 'transporter landscapes', consisting of sugar transporters with complementary kinetic and regulatory properties.

Introduction

Saccharomyces cerevisiae is used on a massive scale for industrial production of ethanol from cane sugar and hydrolysed corn starch, in which fermentable sugars predominantly occur as hexoses or hexose dimers (Renewable Fuels Association 2017). In contrast, hydrolysis of lignocellulosic feedstocks such as the agricultural residues corn stover, corn cobs and wheat straw, yields mixtures of pentose and hexose sugars (Lynd 1996, Van Maris *et al.* 2006). Industrial ethanol production from lignocellulosic feedstocks therefore requires efficient conversion of pentose and hexose sugars into ethanol (Lynd 1996). In most lignocellulosic hydrolysates, D-xylose represents up to 25% of the total sugar monomers while another pentose, L-arabinose, typically accounts for up to 3% of the sugar content. In some industrially relevant hydrolysates, such as those derived from corn-fibre hydrolysates and sugar-beet pulp, L-arabinose can account for up to 26% of the sugar content (Grohmann and Bothast 1994, Grohmann and Bothast 1997, Van Maris *et al.* 2006).

Since wild-type *S. cerevisiae* strains cannot ferment pentoses, metabolic engineering is required to enable utilization of these substrates. Strategies to engineer *S. cerevisiae* for D-xylose fermentation are either based on heterologous expression of a fungal D-xylose reductase and xylitol dehydrogenase or on expression of a heterologous D-xylose isomerase (Jeffries 1983, Kuyper *et al.* 2003, Kuyper *et al.* 2005). Current L-arabinose-fermenting *S. cerevisiae* strains are based on functional expression of genes encoding L-arabinose isomerase (AraA), L-ribulokinase (AraB), and L-ribulose-5-phosphate-4-epimerase (AraD) from bacteria such as *Escherichia coli*, *Bacillus subtilis* or *Lactobacillus plantarum* (Sedlak and Ho 2001, Becker and Boles 2003, Wisselink *et al.* 2007). Unlike the fungal L-arabinose pathway (Richard *et al.* 2002), the bacterial pathway enables conversion of L-arabinose to D-xylulose-5-phosphate via redox-cofactor-independent reactions (Becker and Boles 2003, Wisselink *et al.* 2007). As previously shown for engineered D-xylose-consuming strains (Kuyper *et al.* 2005), overexpression of the *S. cerevisiae* genes encoding xylulokinase (Xks1, EC 2.7.1.17), ribulose 5-phosphate epimerase (Rpe1, EC 5.3.1.1), ribulose 5-phosphate isomerase (Rki1, EC 5.3.1.6), transketolase (Tkl1, EC 2.2.1.1) and transaldolase (Tal1, EC 2.2.1.2) enabled efficient coupling of the heterologously expressed L-arabinose pathway to glycolysis and alcoholic fermentation (Wisselink *et al.* 2007). Subsequent laboratory evolution yielded an efficient L-arabinose fermenting strain, which combined a high ethanol yield (0.43 g g^{-1}) with an ethanol production rate of 0.29 g h^{-1} [g biomass] $^{-1}$ and an arabinose consumption rate of 0.70 g h^{-1} [g biomass]$^{-1}$) (Wisselink *et al.* 2007). Transcriptome studies identified that increased expression of the genes encoding 'secondary' isoenzymes of transaldolase (*NQM1*) and transketolase (*TKL2*) contributed to improved L-arabinose fermentation of this laboratory-evolved strain (Wisselink *et al.* 2010). Improvements in L-arabinose fermentation

were also achieved by codon optimization of bacterial L-arabinose genes to match the codon preference of highly expressed glycolytic yeast genes (Wiedemann and Boles 2008).

Transport of pentose sugars in *S. cerevisiae* occurs by facilitated diffusion, mediated by native HXT hexose transporters (Hamacher *et al.* 2002, Becker and Boles 2003). Transport of L-arabinose predominantly occurs via the hexose transporter Gal2 which, however, exhibits a much lower affinity for L-arabinose than for D-glucose or galactose (Kou *et al.* 1970, Becker and Boles 2003, Subtil and Boles 2011, Subtil and Boles 2012). Of the other *S. cerevisiae* HXT transporters, only Hxt9p and Hxt10p show low rates of L-arabinose transport (Subtil and Boles 2011). In practice, expression of *GAL2* has been found to be essential for L-arabinose fermentation by strains that do not express heterologous transporters (Becker and Boles 2003, Wisselink *et al.* 2010, Subtil and Boles 2011, Wang *et al.* 2017). Recently, an *in silico* model of the three-dimensional structure of Gal2, based on the crystal structure of the *E. coli* xylose permease XylEp, was used to predict mutations in *GAL2* that improve L-arabinose docking in Gal2. Based on this theoretical analysis, single-amino-acid substitutions were introduced at position 85 of Gal2 and shown to significantly increase L-arabinose transport activity (Wang *et al.* 2017).

Since transcription of *GAL2* is both induced by D-galactose and repressed by D-glucose, design of L-arabinose-fermenting *S. cerevisiae* strains often includes expression of *GAL2* behind a strong constitutive promoter (Tani *et al.* 2016, Wang *et al.* 2017). Alternatively, laboratory evolution can result in upregulation of *GAL2* (Wisselink *et al.* 2009, Wisselink *et al.* 2010). However, L-arabinose transport by Gal2 is also subject to strong competitive inhibition by D-glucose (Subtil and Boles 2012, Knoshaug *et al.* 2015), which precludes simultaneous utilization of L-arabinose and D-glucose in batch cultures grown on sugar mixtures (Wisselink *et al.* 2009, Wisselink *et al.* 2010).

To improve the kinetics of L-arabinose transport, transporter genes from other yeasts, filamentous fungi and plants have been functionally expressed in L-arabinose-metabolizing *S. cerevisiae* strains (Subtil and Boles 2011, Verho *et al.* 2011, Knoshaug *et al.* 2015, Li *et al.* 2015, Bracher *et al.* 2018). K_m values of these heterologous transporters for L-arabinose ranged from 0.13 to 263 mM, while reported V_{max} values ranged from 0.4 to 171 [nmol L-arabinose (mg biomass)$^{-1}$ min^{-1}]. Several heterologous transporters, including the high-affinity L-arabinose proton symporter *Pc*AraT from *P. chrysogenum* (Bracher *et al.* 2018) allowed for L-arabinose uptake in the presence of D-glucose (Subtil and Boles 2011, Knoshaug *et al.* 2015, Li *et al.* 2015, Bracher *et al.* 2018).

Laboratory evolution of pentose-fermenting *S. cerevisiae* strains in which D-glucose phosphorylation was abolished has been successfully used to select strains in which L-arabinose or D-xylose uptake is less sensitive to D-glucose inhibition (Farwick *et al.*

2014, Nijland *et al.* 2014, Wisselink *et al.* 2015). Single-amino-acid substitutions, found at corresponding positions in the hexose transporters Hxt7 (N370) and Gal2 (N376), as well as in a chimera of Hxt3 and Hxt6 (N367), were shown to substantially decrease inhibition of D-xylose transport by D-glucose (Farwick *et al.* 2014, Nijland *et al.* 2014).

The aim of the present study was to identify mutations that enable anaerobic growth of engineered, L-arabinose-fermenting *S. cerevisiae* strains on L-arabinose in the presence of D-glucose and D-xylose, with a focus on mutations that affect uptake of L-arabinose. To this end, we constructed a glucose-phosphorylation-negative, L-arabinose-fermenting *S. cerevisiae* strain by Cas9-assisted genome editing. This strain was then subjected to prolonged laboratory evolution in sequential batch bioreactor cultures, grown on mixtures of L-arabinose, D-xylose and D-glucose. Subsequently, causal mutations for glucose-insensitive growth on L-arabinose were identified by whole-genome sequencing and functional analysis of mutant alleles in non-evolved strain backgrounds.

Materials and Methods

Strains and maintenance

The *S. cerevisiae* strains used and constructed in this study **(Table 1**) were derived from the CEN.PK lineage (Entian and Kötter 2007, Nijkamp *et al.* 2012a). All stock cultures were grown in shake flasks on synthetic medium (SM, see below) or, in case of auxotrophic strains, on yeast-extract/peptone (YP) medium (10 g L^{-1} Bacto yeast extract (Becton Dickinson, Franklin Lakes, NJ) and 20 g L^{-1} Bacto Peptone (Becton Dickinson)) (Verduyn *et al.* 1992). These media were either supplemented with 20 g L^{-1} D-glucose or, in case of the constructed glucose-phosphorylation-negative strains, 20 g L^{-1} L-arabinose. For storage and in pre-cultures, single-colony isolates obtained after laboratory evolution were grown in medium containing 20 g L^{-1} of each D-glucose, D-xylose and L-arabinose. After adding glycerol (30% vol/vol) to these cultures, 1 mL aliquots were stored at -80 °C.

Cultivation and media

All strain characterisation studies were performed in synthetic medium (SM) prepared as described previously (Verduyn *et al.* 1992). Carbon source and vitamin solutions were added after autoclaving the medium for 20 min at 121 °C. Concentrated solutions (50% w/w) of D-glucose, D-xylose and L-arabinose were autoclaved separately at 110 °C for 20 min. Prior to inoculation 20 g L^{-1} L-arabinose (SMA), 20 g L^{-1} D-glucose (SMG), 20 g L^{-1} L-arabinose and 20 g L^{-1} D-glucose (SMAG) or 20 g L^{-1} L-arabinose, 20 g L^{-1} D-glucose and 20 g L^{-1} D-xylose (SMAGX) were added to SM as carbon sources. Aerobic shake-

Table 1. *Saccharomyces cerevisiae* strains used in this study.

Strain	Relevant genotype	Reference
CEN.PK 113-7D	*MATa MAL2-8c SUC2*	(Entian and Kötter 2007)
IMX080	*MATa ura3-52 his3-1 leu2-3,112 MAL2-8c SUC2 glk1::Sphis5, hxk1::KILEU2*	(Solis-Escalante *et al.* 2015)
IMX486	*MATa ura3-52 his3-1 leu2-3,112 MAL2-8c SUC2 glk1::Sphis5, hxk1::KILEU2 gal1::cas9-amdS*	This study
IMX604	*MATa ura3-52 his3-1 leu2-3,112 MAL2-8c SUC2 glk1::Sphis5, hxk1::KILEU2 gal1::cas9- amdS gre3::pTDH3_RPE1 pPGK1_ TKL1, pTEF1_TAL1 pPGI1_NQM1 pTPI1_RKI1 pPYK1_TKL2*	This study
IMX658	*MATa ura3-52 his3-1 leu2-3,112 MAL2-8c SUC2 glk1::Sphis5, hxk1::KILEU2 gal1::cas9- amdS gre3::pTDH3_RPE1 pPGK1_ TKL1, pTEF1_TAL1 pPGI1_NQM1 pTPI1_RKI1 pPYK1_TKL2 gal80::(pTPI_araA_tCYC)*9 pPYK-araB-tPGI1 pPGK-araD-tTDH3*	This study
IMX660	*MATa ura3-52 his3-1 leu2-3,112 MAL2-8c SUC2 glk1::Sphis5, hxk1::KILEU2 gal1::cas9- amdS gre3::pTDH3_RPE1 pPGK1_ TKL1, pTEF1_TAL1 pPGI1_NQM1 pTPI1_RKI1 pPYK1_TKL2 gal80::(pTPI_araA_tCYC)*9 pPYK-araB-tPGI1 pPGK-araD-tTDH3 hxk2::KIURA3*	This study
IMX728	*MATa ura3-52 his3-1 leu2-3,112 MAL2-8c SUC2 glk1::Sphis5, hxk1::KILEU2 gal1::cas9- amdS gre3::pTDH3_RPE1 pPGK1_ TKL1, pTEF1_TAL1 pPGI1_NQM1 pTPI1_RKI1 pPYK1_TKL2 gal80::(pTPI_araA_tCYC)*9 pPYK-araB-tPGI1 pPGK-araD-tTDH3 hxk2:: PcaraT*	This study
IMX1106	*MATa ura3-52 his3-1 leu2-3,112 MAL2-8c SUC2 glk1::Sphis5, hxk1::KILEU2 gal1::cas9- amdS gre3::pTDH3_RPE1 pPGK1_ TKL1, pTEF1_TAL1 pPGI1_NQM1 pTPI1_RKI1 pPYK1_TKL2 gal80::(pTPI_araA_tCYC)*9 pPYK-araB-tPGI1 pPGK-araD-tTDH3 hxk2::KIURA3 gal2::gal2^{N376I}*	This study
IMX1386	*MATa ura3-52 his3-1 leu2-3,112 MAL2-8c SUC2 glk1::Sphis5, hxk1::KILEU2 gal1::cas9- amdS gre3::pTDH3_RPE1 pPGK1_ TKL1, pTEF1_TAL1 pPGI1_NQM1 pTPI1_RKI1 pPYK1_TKL2 gal80::(pTPI_araA_tCYC)*9 pPYK-araB-tPGI1 pPGK-araD-tTDH3 hxk2::KIURA3 CAN1::pGAL2_GAL2_tGAL2*	This study
IMS0514	IMX728 that has undergone laboratory evolution on a mixture of arabinose and glucose under aerobic conditions	This study
IMS0520	IMX728 that has undergone laboratory evolution on mixture of arabinose, glucose and xylose under anaerobic conditions	This study
IMS0521	IMX728 that has undergone laboratory evolution on mixture of arabinose, glucose and xylose under anaerobic conditions	This study
IMS0522	IMX728 that has undergone laboratory evolution on mixture of arabinose, glucose and xylose under anaerobic conditions	This study
IMS0523	IMX728 that has undergone laboratory evolution on mixture of arabinose, glucose and xylose under anaerobic conditions	This study
IMW088	*As IMS0522; PcaraTΔ*	This study
IMW091	*As IMS0522; gal2^{I89T}*	This study
DS68616	*Mat a, ura3-52, leu2-112, gre3::loxP, loxP-Ptpi:TAL1, loxP-Ptpi::RKI1, loxP-Ptpi-TKL1, loxP-Ptpi-RPE1, delta::Padh1XKS1Tcyc1-LEU2, delta::URA3-Ptpi-xylA-Tcyc1*	DSM, The Netherlands

Table 1 continues on next page.

Table 1 – *Continued*

Strain	Relevant genotype	Reference
DS68625	DS68616, *his3::loxP, hxt2::loxP-kanMX-loxP, hxt367::loxP-hphMX-loxP, hxt145::loxP-natMX-loxP, gal2::loxP-zeoMX-loxP*	(Nijland *et al.* 2014)
DS68625-Gal2	DS68616, *his3::loxP, hxt2::loxP-kanMX-loxP, hxt367::loxP-hphMX-loxP, hxt145::loxP-natMX-loxP, gal2::loxP-zeoMX-loxP*	This study
DS68625-Gal2^{N376I}	DS68616, *his3::loxP, hxt2::loxP-kanMX-loxP, hxt367::loxP-hphMX-loxP, hxt145::loxP-natMX-loxP, gal2::loxP-zeoMX-loxP*	This study
DS68625-Gal2^{N376S}	DS68616, *his3::loxP, hxt2::loxP-kanMX-loxP, hxt367::loxP-hphMX-loxP, hxt145::loxP-natMX-loxP, gal2::loxP-zeoMX-loxP*	This study
DS68625-Gal2^{N376T}	DS68616, *his3::loxP, hxt2::loxP-kanMX-loxP, hxt367::loxP-hphMX-loxP, hxt145::loxP-natMX-loxP, gal2::loxP-zeoMX-loxP*	This study
DS68625-Gal2$^{N376T / T89I}$	DS68616, *his3::loxP, hxt2::loxP-kanMX-loxP, hxt367::loxP-hphMX-loxP, hxt145::loxP-natMX-loxP, gal2::loxP-zeoMX-loxP*	This study
DS68625-Gal2^{T89I}	DS68616, *his3::loxP, hxt2::loxP-kanMX-loxP, hxt367::loxP-hphMX-loxP, hxt145::loxP-natMX-loxP, gal2::loxP-zeoMX-loxP*	This study
DS68625-mcs	DS68616, *his3::loxP, hxt2::loxP-kanMX-loxP, hxt367::loxP-hphMX-loxP, hxt145::loxP-natMX-loxP, gal2::loxP-zeoMX-loxP*	This study

flask cultures were grown in an orbital shaker at 200 rpm set at 30 °C, using 500-ml flasks containing 100 ml medium.

For plates 20 g L^{-1} agar (BD) was added prior to autoclaving at 121 °C for 20 min. Aerobic shake-flask cultures used as pre-cultures for anaerobic cultures were inoculated with frozen stocks and, in late-exponential phase, a 1 mL sample was used to start a second aerobic pre-culture. Anaerobic batch cultures used for characterization were inoculated from these cultures to obtain an initial OD_{660} of 0.5.

Anaerobic shake-flask cultures were incubated at 30 °C in an Innova anaerobic chamber (5% H_2, 6% CO_2, and 89% N_2, New Brunswick Scientific, Edison, NJ) in 50 mL shake flasks placed on an orbital shaker set at 200 rpm. Bioreactor batch cultures were performed at 30 °C in 2-L laboratory bioreactors (Applikon, Delft, The Netherlands) with a working volume of 1 L. Culture pH was controlled at 5.0 by automatic addition of 2 M KOH and cultures were stirred at 800 rpm. Anaerobic bioreactors were equipped with Viton O-rings and Norprene tubing (Cole Palmer Instrument Company, Vernon Hills, IL) to minimize oxygen diffusion and continuously sparged with nitrogen gas (<10 ppm oxygen) at 0.5 L min^{-1}. After autoclaving, synthetic media were supplemented with 0.2 g L^{-1} antifoam C (Sigma-Aldrich, St. Louis, MO). Furthermore, the anaerobic growth factors Tween 80 (420 mg L^{-1}) and ergosterol (10 mg L^{-1}) were added to the medium as described previously (Verduyn *et al.* 1990).

Laboratory evolution experiments were performed in sequential batch reactors (SBRs). On-line measurement of CO_2 concentrations in the off gas of SBRs was used as input for a control routine programmed in MFCS/win 3.0 (Sartorius AG, Göttingen, Germany). An

empty-refill cycle was automatically initiated when the CO_2 concentration in the off gas had first exceeded a threshold value of 0.2% and subsequently declined below 0.1% as fermentable sugars were depleted. After the emptying phase, when approximately 7% of the initial culture volume was left in the reactor, the reactor was automatically refilled with fresh medium from a 20 L reservoir, which was continuously sparged with nitrogen gas. After ca. 450 generations of selective growth, single colony isolates were obtained by plating culture samples on SMAGX agar, supplemented with Tween 80 (420 mg L^{-1}) and ergosterol (10 mg L^{-1}). Plates were incubated in an anaerobic chamber for 6 days and restreaked three times to obtain single-cell lines.

Plasmid and strain construction

The plasmids constructed and used in this study are shown in **Additional Table 1**. The open reading frames of *araA* [Genbank: ODO63149.1, encoding L-arabinose isomerase], *AraB* [Genbank: ODO63147.1, encoding L-ribulose kinase] and *AraD* [Genbank: ODO63147.1, encoding L-ribulose-5-phosphate epimerase] were codon optimized based on the codon usage in glycolytic genes in *S. cerevisiae* (Wiedemann and Boles 2008), synthesized behind the promoter of *TPI1*, *PGK1* and *PYK1* originating from CEN.PK113-7D and cloned in pMK-RQ based vectors by GenArt GmbH (Regensburg, Germany). These plasmids were named pUD354-pUD356 and transformed into *E. coli* DH5a cells. Plasmid DNA was isolated from *E. coli* cultures using a GenElute Plasmid kit (Sigma-Aldrich, St. Louis, MO). PCR amplification of expression cassettes and plasmid fragments was performed using Phusion High Fidelity DNA Polymerase (Thermo Scientific, Waltham, MA). *S. cerevisiae* strains were transformed as described by Gietz and Woods (Gietz *et al.* 1995). To facilitate CRISPR/Cas9-mediated genome editing in strain *S. cerevisiae* IMX080, a Cas9 expression cassette PCR-amplified from p414-TEF1p-cas9-CYCt using primers 4653-5981 as well as a second fragment containing the amdSYM marker cassette, PCR amplified from pUG-amdSYM with primers 3093-1678 (Solis-Escalante *et al.* 2013), were *in vivo* co-assembled and integrated (Kuijpers *et al.* 2013) into the *GAL1* locus through a double cross over mediated by homologous recombination (**Additional Figure 1A**).

After each transformation round the cells were restreaked thrice and correct integration of all the fragments was confirmed by diagnostic PCR. Transformants were selected on solid synthetic medium with acetamide (Solis-Escalante *et al.* 2013) as sole nitrogen source and checked by diagnostic PCR using DreamTaq polymerase (Thermo Scientific), according to the manufacturer's protocol. The resulting strain IMX486 was subsequently transformed with 200 pmol each of eight DNA fragments containing expression cassettes for the genes encoding the enzymes of the non-oxidative branch of the pentose phosphate pathway, which were PCR amplified from pUD344-346 and two 500bp fragments containing flanking

regions of *GRE3* amplified from CEN.PK113-7D genomic DNA, using oligonucleotide primers shown in **Additional Table 2** (available at https://doi.org/10.1093/femsyr/foy062). These fragments were co-transformed with 500 ng of plasmid pUDE335 to induce a double strand break in the *GRE3* locus using CRISPR-Cas9 (**Additional Figure 1B**). The cells were plated on SMD plates and correct assembly of all six fragments in the *GRE3* locus was verified by diagnostic colony PCR. Plasmid pUDE335 was counter selected on YP with 20 g L^{-1} D-glucose (YPD) agar with 5-fluoroorotic acid (5-FOA) as described previously (Mans *et al.* 2015). The resulting strain IMX604 was then transformed with nine DNA fragments that carried expression cassettes for *araA* and single fragments for *araB* and *araD* amplified from pUD354-356. The primers used for PCR amplification of these fragments added homologous regions required for *in-vivo* assembly and integration into the *GAL80* locus (**Additional Figure 1C**). The fragments were co-transformed with gRNA-plasmid pUDE348, which was constructed as described previously (Verhoeven *et al.* 2017). After verification of the correct assembly and integration of the fragments by diagnostic PCR, plasmid pUDE348 was counter selected with 5-FOA, yielding strain IMX658.

HXK2 was deleted by inducing a double strand break using the gRNA plasmid pUDE327, and using an expression cassette for the *PcaraT* gene, obtained by PCR amplification from plasmid Pwt118, as the repair fragment (**Additional Figure 1D**). After counter selection of pUDE327 with 5-FOA, a DNA fragment carrying the *URA3* gene from CEN.PK113-7D was amplified using primers 2641-1522 and transformed to restore a wild-type *URA3* gene, yielding strain IMX728. Strain IMX660 was constructed from IMX658 by transforming a KlURA3 based knock-out cassette targeting the *HXK2* locus.

Strain IMX1386 was obtained by co-transforming IMX728 with plasmid pROS13 and a fragment containing the promoter region, ORF and terminator of *GAL2*, PCR amplified from genomic DNA of strain CEN.PK113-7D with primer pair 7285-7286. Transformants were selected on YP with 20 g L^{-1} L-arabinose (YPA) supplemented with G418. After verifying correct integration of the *GAL2* expression cassette in the *CAN1* locus, the gRNA plasmid was counter selected by serial plating on YPA as described previously (Mans *et al.* 2015). The gRNA plasmids pUDR172 and pUDR187 were constructed using pROS13 as template by Gibson assembly according to the method described by Mans *et al.* (2015). The single nucleotide polymorphism (SNP) at position 1127 in *GAL2* was inserted by co-transforming strain IMX660 with pUDR172 and a repair fragment containing the single nucleotide change resulting in the amino acid substitution from N to I at position 376. Expression cassettes of the *GAL2*, *GAL2*^{N376I} and *GAL2*^{N376S} variants were amplified using genomic DNA from CEN.PK113-7D, IMS0520 and IMS0514 as template using the primers F_Gal2_XbaI and R_Gal2_Cfr9I (**Additional Table 2,** available at https://doi.org/10.1093/femsyr/foy062) and subsequently cloned into plasmid pRS313-P7T7.

To obtain the *GAL2*^{N376T}, *GAL2*$^{N376T\ /\ T89I}$ and *GAL2*^{T89I} mutants, fragments were PCR amplified from CEN.PK113-7D genomic DNA using the primers shown in **Additional Table 2** (available at https://doi.org/10.1093/femsyr/foy062). Subsequently, overlap PCR amplifications were done to obtain the full-length genes encoding for Gal2^{N376T}, Gal2$^{N376T\ /\ T89I}$ and Gal2^{T89I}. All genes were cloned into pRS313-P7T7 and subsequently confirmed by Sanger sequencing (Baseclear).

The Gal2 variants and the pRS313-P7T7-mcs plasmid (as an empty plasmid/control) were transformed to the hexose transporter deletion strain DS68625 and positive colonies were named DS68625-Gal2, DS68625-Gal2 N376I, DS68625-Gal2 N376S, DS68625-Gal2 N376T, DS68625-Gal2 N376T / T89I, DS68625-Gal2 T89I and DS68625-mcs. To revert the altered amino acid position 89 in Gal2 from isoleucine to threonine, pUDR187 and a repair fragment containing the original coding sequence were co-transformed. Inactivation of *PcaraT* in IMS0522 was achieved by in-vivo assembling of the pROS13 backbone and the gRNA expression cassette, both PCR amplified from pROS13 using the primers shown **in Additional Table 2** (available at https://doi.org/10.1093/femsyr/foy062) and providing a repair oligo for *HXK2*. After verifying correct transformants the plasmids were removed by counter selection on YPAGX plates and the resulting strains were named IMW091 and IMW088 respectively

Analytical methods

To monitor growth of batch cultures, a Libra S11 spectrometer (Biochrom, Cambridge, United Kingdom) was used for optical density measurements at 660 nm. To calculate specific growth rates, biomass dry weight measurements were performed on at least six samples taken during the exponential growth phase. Dry weight measurements were performed by filtering 10 mL culture samples over pre-weighed nitrocellulose filter (pore size, 0.45 µm; Gelman Laboratory, Ann Arbor, MI). The filter was washed with demineralized water and dried in a microwave oven (Bosch, Stuttgart, Germany) for 20 min at 360 W. The correlation between these CDW measurement and the corresponding OD data was used to estimate CDW in samples for which no direct CDW measurements were done.

Metabolite concentrations in culture supernatants, obtained by centrifugation, were determined by high-performance liquid chromatography (HPLC)on an Agilent 1260 HPLC (Agilent Technologies, Santa Clara, CA) equipped with a Bio-Rad HPX 87 H column (Bio-Rad, Hercules, CA). The column was eluted at 60 °C with 0.5 g L^{-1} H_2SO_4 at a flow rate of 0.6 ml min^{-1}. Detection was by means of an Agilent G1362A refractive-index detector and an Agilent G1314F VWD detector. CO_2 and O_2 concentrations were measured in bioreactor off gas using an NGA 2000 analyzer (Rosemount Analytical, Orrville, OH) after it was first cooled by a condenser (2 °C) and dried with a Permapure MD-110-48P-4

dryer (Permapure, Toms River, NJ). Correction for ethanol evaporation was done for all bioreactor experiments as described previously (Guadalupe Medina *et al.* 2010).

DNA sequence analysis

Genomic DNA for sequencing was isolated using the QIAGEN Blood & Cell Culture Kit With 100/G Genomic-tips (QIAGEN, Valencia, CA) according to the manufacturer's protocol. DNA libraries were prepared using the Nextera XT DNA library Preparation Kit (Illumina, San Diego, CA), yielding 300bp fragments. The Libraries were sequenced using the Illumina MiSeq platform (Illumina, San Diego, CA, USA). Data were aligned and mapped to the CEN.PK113-7D genome using the Burrows-Wheeler alignment tool (Li and Durbin 2009). Variant calling by Pilon (Walker *et al.* 2014) was checked with the Integrated Genomics Viewer (IGV) (Thorvaldsdóttir *et al.* 2013). The Poisson mixture model based algorithm Magnolya (Nijkamp *et al.* 2012b) was used to detect and quantify chromosomal copy number variations (CNV). Copy numbers of *araA, araB, araD* and *PcaraT* were estimated as described previously using the average read depth of the chromosomes that did not contain any duplication according to the CNV analysis done with Magnolya (Verhoeven *et al.* 2017). Genome sequence data of strains IMS0520, IMS0521, IMS0522, IMS0523, IMS514 and parental strains IMX728 and IMX660 have been deposited at the NCBI Sequence Read archive (www.ncbi.nlm.nih.gov/sra) with the corresponding BioProject ID PRJNA414371. *GAL2* ORFs from strains IMS0002 and IMS0010 and the nucleotides surrounding position T89 in IMS0003 and IMS0007 were PCR amplified from genomic DNA using primers as listed in **Additional Table 2** (available at https://doi.org/10.1093/femsyr/foy062) and sanger sequenced (Baseclear).

Sugar transport assays

Strains expressing *GAL2* alleles in the DS68625 strain background were pre-grown in aerobic shake flasks on synthetic medium with 20 g L^{-1} D-xylose, after which cells were collected by centrifugation (3000 g, 3 min), washed and re-suspended in SM without sugar. Uptake experiments were initiated by adding [^{14}C] L-arabinose or [^{14}C] D-glucose (ARC St. Louis, MO, USA) to the cell suspension at concentrations ranging from 0.2 to 500 mM. [^{14}C] L-arabinose and [^{14}C] D-glucose (50–60 mCi $mmol^{-1}$) were added at concentrations of 0.1 mCi ml^{-1}. At set time points, uptake was arrested by adding 5 mL of ice-cold 0.1 M LiCl, filtration over 0.45-µm HV membrane filters (Millipore, France) and washing with 5 mL ice-cold 0.1 M LiCl. Radioactivity on the filters was then counted using a Liquid Scintillation Counter (PerkinElmer, Waltham, MA) in Ultima Gold MV Scintillation cocktail (PerkinElmer). D-Glucose inhibition experiments were measured using 50 mM [^{14}C] L-arabinose with [^{14}C] D-glucose added at concentrations between 50 and 500 mM.

Availability of supporting data

The genome sequence data of strains IMS0520, IMS0521, IMS0522, IMS0523, IMS514 and parental strains IMX728 and IMX660 have been deposited in Genbank (http://www.ncbi.nlm.nih.gov/), BioProject ID (PRJNA414371).

Results

Construction of a *S. cerevisiae* L-arabinose specialist strain

To investigate uptake and growth on L-arabinose in the presence of D-glucose, an L-arabinose 'specialist strain' was constructed by Cas9-assisted genome editing. First, a Cas9 expression cassette (DiCarlo *et al.* 2013) and the counter-selectable marker cassette amdSYM (Solis-Escalante *et al.* 2013) were *in vivo* assembled and integrated at the *GAL1* locus in *S. cerevisiae* IMX080 (*hxk1Δ glk1Δ*) (**Additional Figure 1A**). In the resulting strain IMX486 (*hxk1Δ glk1Δ gal1Δ::*{*Spcas9*-amdSYM}, *HXK2* was the sole remaining functional hexose kinase.

Subsequently, cassettes for constitutive expression of *RPE1, RKI1, TKL1, TKL2, TAL1* and *NQM1*, which encode (iso)enzymes of the non-oxidative pentose-phosphate pathway (NPPP), were *in vivo* assembled and integrated at the *GRE3* locus of strain IMX486, yielding strain IMX658 (*hxk1Δ glk1Δ gal1Δ::*{*Spcas9*-amdSYM} *gre3Δ::*{NPPP}) (**Additional Figure 1B**). Inactivation of *GRE3* prevents xylitol formation by the Gre3 non-specific aldose reductase (Träff *et al.* 2001, Kuyper *et al.* 2005).

To introduce a functional bacterial L-arabinose pathway (ARAP) (Wisselink *et al.* 2007), the *Lactobacillus plantarum* genes encoding L-arabinose isomerase (AraA), ribulokinase (AraB) and ribulose-5P epimerase (AraD) were codon optimized and placed under the control of the strong constitutive promoters of *TPI1*, *PGI1* and *TDH3*, respectively. For high expression of *araA*, we used the multi-copy tandem integration strategy previously described for high-level expression of xylose isomerase which, in laboratory evolution experiments, facilitates rapid adaptation of copy number by homologous recombination (Verhoeven *et al.* 2017). Using this approach, nine copies of the *araA* cassette and single copies of the *AraB* and *AraD* cassettes were integrated at the *GAL80* locus by *in vivo* assembly and Cas9-mediated integration (**Additional Figure 1C**). Inactivation of *GAL80* eliminates the need for galactose induction of *GAL2* expression (Torchia *et al.* 1984, Ostergaard *et al.* 2001, Tani *et al.* 2016). Expression of the high-affinity L-arabinose transporter *Pc*AraT from *P. chrysogenum* enables L-arabinose uptake in the presence of glucose (Bracher *et al.* 2018). To simultaneously abolish the ability to metabolize D-glucose, an expression cassette for *Pc*AraT (Bracher *et al.* 2018) was integrated into the *HXK2* locus. Subsequent introduction of a functional *URA3* allele from *S. cerevisiae* CEN.PK113-

7D (**Additional Figure 1D**) yielded the prototrophic 'arabinose specialist' *S. cerevisiae* strain IMX728 (*hxk1Δ glk1Δ gal1Δ::*{*Spcas9*-amdSYM} *gre3Δ::*{NPPP} *hxk2Δ::PcaraT gal80Δ*::{ARAP}).

In aerobic shake flasks on SMA, strain IMX728 exhibited a specific growth rate of 0.26 h^{-1} (**Figure 1**). As anticipated after deletion of all four genes encoding glucose-phosphorylating enzymes (Farwick *et al.* 2014, Nijland *et al.* 2014, Wisselink *et al.* 2015), no growth was observed on SMD. Consistent with the incomplete inhibition of *Pc*AraT by glucose (Bracher *et al.* 2018) strain IMX728 grew at a specific growth rate of 0.04 h^{-1} in aerobic batch cultures on SMAGX (**Figure 1**).

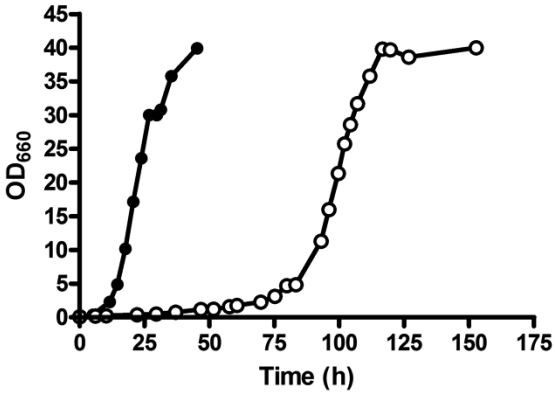


Figure 1. Growth curves of *S. cerevisiae* strain IMX728 (engineered, non-evolved arabinose-consuming, glucose-phosphorylation-negative, expressing *PcaraT*) in aerobic batch cultures on SM containing 20 g L^{-1} L-arabinose (SMA) as sole carbon and energy source (•) and in SM with 20 g L^{-1} L-arabinose, 20 g L^{-1} D-glucose and 20 g L^{-1} D-xylose added (○). Data shown in the figure represent from a single growth experiment, data from a replicate experiment are shown in **Additional Figure 2**. Derived biomass-specific conversion rates in the replicate experiments differed by less than 5%.

Anaerobic laboratory evolution of an engineered L-arabinose specialist in the presence of D-glucose

The ability of strain IMX728 to grow anaerobically on L-arabinose as sole carbon source was investigated in anaerobic bioreactors. In contrast to the fast and instantaneous growth in aerobic shake flasks (**Figure 1**), slow anaerobic growth on L-arabinose was only observed after approximately 150 h (**Additional Figure 5**). To select for spontaneous mutants that combined an increased anaerobic growth rate on L-arabinose with a decreased sensitivity to growth inhibition by D-glucose and D-xylose, strain IMX728 was grown in duplicate anaerobic sequential batch reactors (SBRs) on SMAGX. Continuous measurements of CO_2 concentrations in the off gas of the SBRs were used to automatically initiate empty-refill cycles and monitor the increase of specific growth rate over time (**Additional Figure 4**).

After an initial cultivation cycle under aerobic conditions, the first anaerobic cycle required a 17 d period before growth was observed. Only then, within 24 h, both replicate cultures showed an exponential increase in CO_2 production, corresponding to an estimated specific growth rate of 0.07 h^{-1}. Both SBRs were operated for 254 d (**Figure 2**), after which specific growth rates had increased to above 0.13 h^{-1}. Specific growth rates on SMAGX of 31 single-colony isolates obtained from the SBR cultures were then measured in anaerobic shake flasks (**Additional Figure 5**). Four selected isolates, named IMS0520 to IMS0523, showed specific growth rates of 0.13 to 0.17 h^{-1} in anaerobic bioreactor batch cultures on SMAGX, which corresponded closely to the growth rates estimated from on-line CO_2 analysis of the evolving SBR cultures (**Additional Figure 4**).

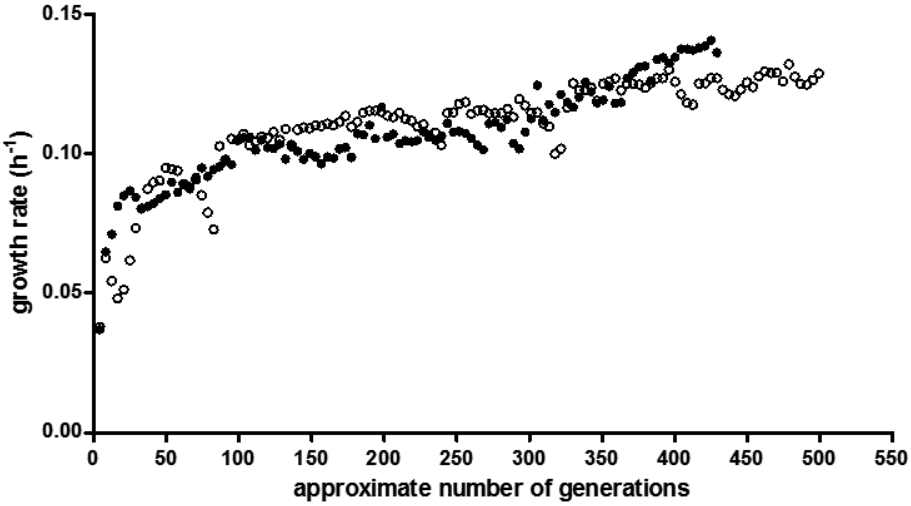


Figure 2. Specific growth rate (h^{-1}) estimated from CO_2 production profiles in two parallel laboratory evolution experiments in anaerobic SBRs: SBR 1 (○) and SBR 2 (•). Both cultures were inoculated with the L-arabinose consuming, glucose-phosphorylation negative *S. cerevisiae* strain IMX728 and grown on SM with 20 g L^{-1} L-arabinose, 20 g L^{-1} D-glucose and 20 g L^{-1} D-xylose. The first data point for each experiment corresponds to the initial aerobic batch cultivation, all subsequent data points represent estimated growth rates in anaerobic SBR cycles.

Evolved strains show mutations, duplication and allelic variation in *GAL2*

To identify the mutations that enabled growth of laboratory-evolved isolates IMS0520-IMS0523 on L-arabinose in the presence of D-glucose and D-xylose, their genomes were sequenced and compared to that of their common parental strain IMX728. Variant screening for non-synonymous single-nucleotide mutants and insertion/deletions revealed that, together, the four strains carried 9 single-nucleotide mutations in open reading frames (**Table 2**). All four isolates carried mutations in *GAL2* that changed the asparagine

Table 2. Single-nucleotide mutations in L-arabinose-metabolizing, glucose-phosphorylation-negative *S. cerevisiae* strains IMS0520-IMS0523, evolved for anaerobic growth on L-arabinose in the presence of D-xylose and D-glucose in biological duplicate (reactors 1 and 2). Mutants were identified alignment of the whole-genome sequence data to *S. cerevisiae* IMX728. Read coverage observed for each SNP was higher than 99% unless stated otherwise in the table. Descriptions of gene functions are derived from the *Saccharomyces* Genome Database (as of 19-11-2017).

Gene and strain	Nucleotide change	Amino acid change	Description
GAL2			
IMS0520 (reactor 1)	A1127T	N376I	Galactose permease; required for utilization of galactose and also able to transport glucose (Kou *et al*. 1970).
IMS0521 (reactor 1)	A1127T	N376I	
IMS0522 (reactor 2)	A1127C	N376T	
	C266T (57%)	T89I	
IMS0523 (reactor 2)	A1127C	N376T	
BMH1			
IMS0520 (reactor 1)	G383T	G128V	14-3-3 protein, major isoform; controls, involved in regulation of exocytosis, vesicle transport, Ras/MAPK and rapamycin-sensitive signalling, aggresome formation, spindle position checkpoint
IMS0522 (reactor 2)	G64T	E22*	
IMS0523 (reactor 2)	G64T	E22*	
DCK1			
IMS0522 (reactor 2)	T2890C (42%)	Y964H	Dock family protein (Dedicator Of CytoKinesis), homolog of human DOCK1; interacts with Ino4p; cytoplasmic protein that relocates to mitochondria under oxidative stress
IPT1			
IMS0522 (reactor 2)	C938T	S313F	Inositolphosphotransferase; involved in synthesis of mannose-(inositol-P)2-ceramide (M(IP)2C) sphingolipid; can mutate to resistance to the antifungals syringomycin E and DmAMP1 and to *K. lactis* zymocin
UPC2			
IMS0523(reactor 2)	A2648G	Y883C	Sterol regulatory element binding protein; induces sterol biosynthetic genes, upon sterol depletion; acts as a sterol sensor, binding ergosterol in sterol rich conditions
RPL6B			
IMS0520 (reactor 1)	T784A	E261G	Ribosomal 60S subunit protein L6B; binds 5.8S rRNA; homologous to mammalian ribosomal protein L6

in position 376 of Gal2, into an isoleucine ($Gal2^{N376I}$) or a threonine ($Gal2^{N376T}$) in isolates originating from SBR 1 and 2, respectively. Isolate IMS0522 contained an additional single-nucleotide mutation in *GAL2* resulting in a $Gal2^{T89I}$ substitution, which, however, was only observed in 57% of the reads already suggesting a copy number increase of *GAL2* locus.

In addition to mutations in *GAL2*, single-nucleotide changes were identified in *DCK1*, *IPT1* *UPC2* and *RPL6B*, which each occurred in only one of the four isolates. *IPT1* encodes

an inositol phosphotransferase involved in sphingolipid production (Dickson *et al.* 1997), while *UPC2* plays a key role in regulation of sterol metabolism and uptake (Crowley *et al.* 1998). Both mutations may have affected membrane composition and, thereby, the uptake of L-arabinose. Mutations in *BMH1* were observed in strains IMS0520, IMS0522 and IMS0523. In view of the pleiotropic phenotype of *bmh1* mutants (van Heusden *et al.* 1992, Roberts *et al.* 1997, Van Hemert *et al.* 2001, Wang *et al.* 2009), we decided to not investigate the impact of these mutations on L-arabinose fermentation. Since, *GAL2* was the only transporter gene whose coding region was found to be mutated in strain IMS0521, further analysis was focused on the *GAL2* mutations found in the four strains.

As frequently observed in laboratory evolution experiments with *S. cerevisiae* (Oud *et al.* 2012, Gorter de Vries *et al.* 2017), read depth analysis revealed changes in copy number of specific regions in the genomes of the evolved strains (**Additional Figure 7**).

Notably, regions of Chromosome XII, which harbors *GAL2*, were found to be duplicated in all four isolates, with strain IMS0522 showing a probable duplication of this entire chromosome (**Figure 3**). To test the physiological significance of the resulting duplication of *GAL2*, an additional copy of the wild-type *GAL2* allele was integrated in the genome of the parental 'arabinose specialist' IMX728, yielding strain IMX1386. Similar to *S. cerevisiae* IMX728, the strain with two copies of *GAL2* required 120 h before growth on L-arabinose was observed in anaerobic batch cultures (**Additional Figure 8**), indicating that duplication of *GAL2* was not, in itself, sufficient to enable instantaneous anaerobic growth on L-arabinose.

Read-depth comparisons of the genes encoding the enzymes of the bacterial L-arabinose pathway (*araA, araB, araD*) revealed some differences in copy number of *araA* and *araB* between evolved strains and the parental strain (**Table 3**).

Table 3. Estimated copy numbers of the heterologously expressed *Lactobacillus plantarum* genes *araA, araB* and *araD* and *Penicillium chrysogenum PcaraT* in the evolved, anaerobically growing and glucose-tolerant 'arabinose specialist' strains *S. cerevisiae* IMS0520-IMS0523, obtained from two independent evolution experiments in SBRs (1 and 2), and in the parental strain IMX728. Copy numbers were estimated by comparing gene read depths to the average of all chromosomes.

	Reference	*Reactor 1*		*Reactor 2*	
Gene	**IMX728**	**IMS0520**	**IMS0521**	**IMS0522**	**IMS0523**
araA	17.5	8.7	19.3	19.5	16.4
araB	0.8	0.9	3.1	2.2	1.2
araD	0.9	1.0	1.2	1.0	1.1
PcaraT	0.9	1.1	1.0	1.1	1.2
[1]Excluded chromosomes:	*3, 10, 12, 16*	*2, 8, 12*		*12*	*12, 16*

[1]Chromosomes excluded from the average read depth, based on duplication of chromosomal regions identified with the Magnolya algorithm (Nijkamp *et al.* 2012).

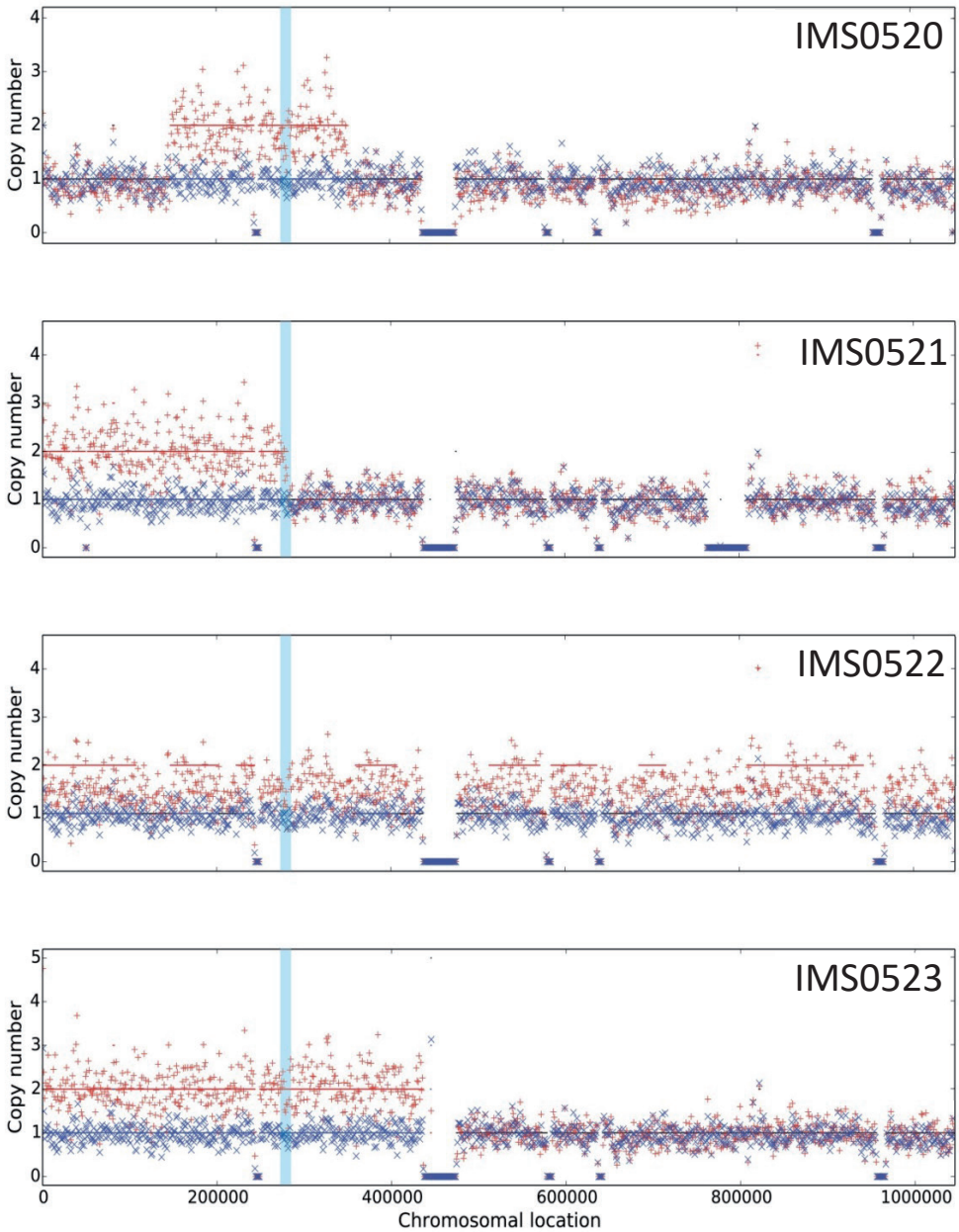


Figure 3. Chromosomal copy number estimations for chromosome 12 in the evolved anaerobically L-arabinose-growing, glucose-tolerant *S. cerevisiae* strains IMS0520 to IMS0523. Chromosomal copy numbers were estimated using the Magnolya algorithm (Nijkamp *et al.* 2012). Reads of the four strains (red crosses) show duplications relative to the corresponding chromosome 12 sequences in the common parental strain IMX728 (blue crosses). The light blue region corresponds to the location of *GAL2*.

However, these differences were not consistent across the four evolved strains. Read-depth analysis of the *PcaraT* transporter gene revealed no copy number changes in any of the four evolved isolates. To examine the contribution of *PcaraT* to the physiology of the evolved strains, it was deleted in evolved strain IMS0522, resulting in strain IMW088. In duplicate anaerobic bioreactor cultures of strains IMS0522 and IMW088 on SMAGX, both completely fermented the L-arabinose in the medium. The strain lacking *PcaraT* (IMW088) did, however, exhibit a lower specific growth rate than strain IMS0522 (0.10 h^{-1} and 0.12h^{-1}, respectively; **Table 4**) and, consequently, required substantially more time for complete conversion of L-arabinose in the anaerobic batch cultures (**Figure 4**).

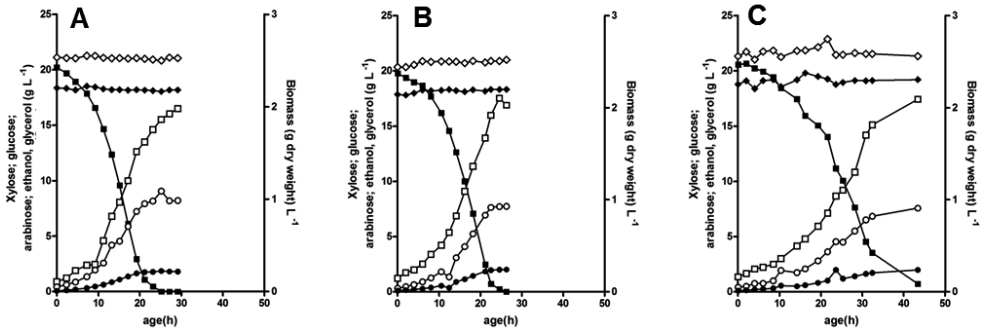


Figure 4. Growth and extracellular metabolite concentrations in anaerobic cultures of the evolved L-arabinose consuming, glucose-phosphorylation negative *S. cerevisiae* strain IMS0522 and two derived strains with specific genetic modifications. Cultures were grown in bioreactors containing synthetic medium with 20 g L^{-1} L-arabinose, 20 g L^{-1} D-glucose and 20 g L^{-1} D-xylose. **A**. IMS0522, **B**. IMW088 (*PcaraT*Δ), **C**. IMW091 (Gal2^{89I} restored to Gal2^{89T}). ▪ L-arabinose; □ biomass dry weight; ● glycerol; ○ ethanol; ◆D-glucose; ◇D-xylose. Data shown in the figure represent from a single growth experiment for each strain, data from independent replicate experiments are shown in **Additional Figure 9**. Derived biomass-specific conversion rates in the replicate experiments differed by less than 5%.

Amino acid substitutions in Gal2 at position N376 enable L-arabinose uptake in the presence of D-glucose and D-xylose

A previous laboratory evolution study with D-xylose-fermenting, glucose-phosphorylation-deficient *S. cerevisiae* strains also identified mutations that led to a single-amino-acid change in Gal2 at position 376, which enabled growth on D-xylose in the presence of D-glucose (Farwick *et al.* 2014). In the present study, the parental strain IMX728 could not only import L-arabinose via Gal2, but also via the fungal L-arabinose transporter *Pc*AraT, which is much less sensitive to glucose inhibition than Gal2 (Bracher *et al.* 2018).

To investigate whether the observed mutations in *GAL2* were also required for growth on L-arabinose in the presence of D-glucose in a strain background without *PcaraT*, a short laboratory evolution experiment was conducted with the 'arabinose specialist' strain

IMX660, which differed from strain IMX728 by the absence of *Pc*araT. After approximately 100 h incubation in an aerobic shake flask on 20 g L^{-1} L-arabinose and D-glucose, growth was observed and strain IMS0514 was obtained as a single colony isolate and confirmed to have a stable phenotype for growth on SMAG (data not shown). As in strains IMS0520-IMS0523, whole-genome sequencing of strain IMS0514 revealed a single mutation in *GAL2* at position 1128, in this case resulting in a $Gal2^{N376S}$ substitution. Additionally, introduction of the $Gal2^{N376I}$ substitution in both *GAL2* alleles of strain IMS0520 enabled immediate growth on SMAG of the arabinose specialist strain lacking *Pc*AraT (IMX1106, **Figure 5**).

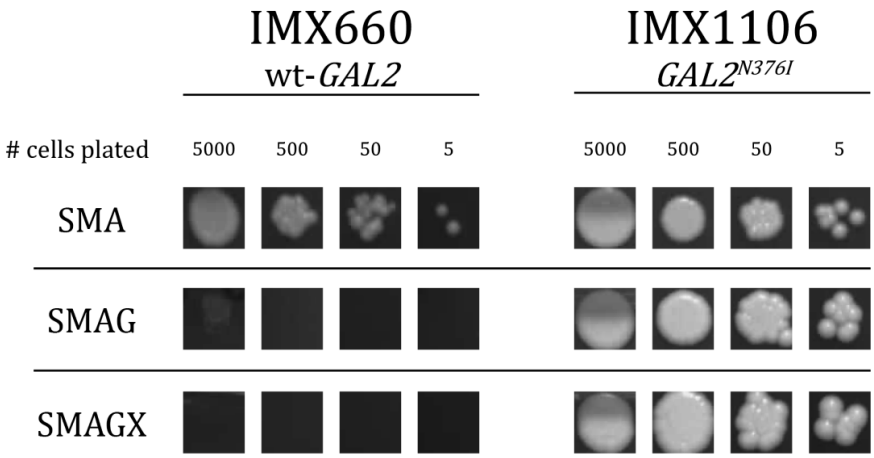


Figure 5. Growth of engineered glucose-phosphorylation-negative *S. cerevisiae* strains on L-arabinose in the presence and absence of D-glucose and D-xylose. An approximate set numbers of cells of strain *S. cerevisiae* IMX1106, expressing Gal2 with amino acid substitution N376 and parental strain IMX660 were plated on solid synthetic medium with L-arabinose (SMA), L-arabinose and D-glucose (SMAG) or L-arabinose, D-glucose and D-xylose (SMAGX). Plates were incubated aerobically at 30°C for 4 days.

Sugar transport kinetics of Gal2 variants obtained by laboratory evolution

To examine inhibition by D-glucose of L-arabinose transport by different Gal2 variants found in evolved strains, the corresponding evolved *GAL2* alleles were expressed in *S. cerevisiae* DS68625. In this strain, which does not express the pathway required for arabinose metabolism, the hexose-transporter genes *HXT1* to *HXT7* and *GAL2* have been disrupted (Nijland *et al.* 2014). To examine the extent to which D-glucose inhibited L-arabinose uptake by the different Gal2 variants, [^{14}C]-L-arabinose-uptake experiments were performed at different concentrations of D-glucose (**Figure 6**). Consistent with earlier reports (Subtil and Boles 2011, Farwick *et al.* 2014), uptake of L-arabinose (50 mM) by wild-type Gal2 was inhibited by ca. 80% in the presence of 100 mM D-glucose (**Figure 6**).

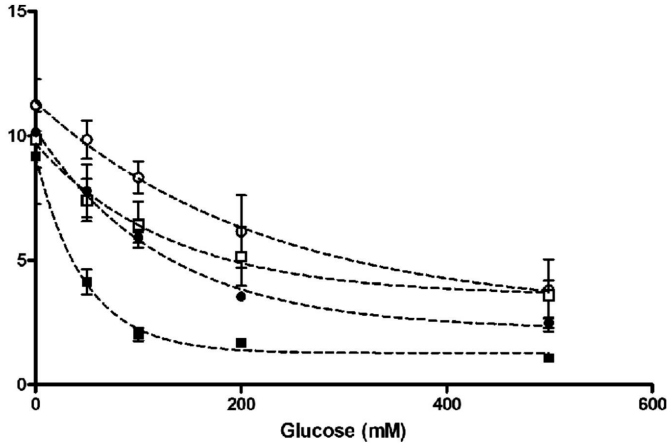


Figure 6. Effect of D-glucose on specific rates of L-arabinose uptake by different Gal2 variants. Uptake experiments were performed with 50 mM [^{14}C-] L-arabinose in the presence of different D-glucose concentrations. Symbols indicate uptake rates observed with strain *S. cerevisiae* DS68625 expressing Gal2^{N376T} (○), Gal2^{N376S} (●), Gal2^{N376I} (□) and wild-type Gal2 (■). DW = Cell dry weight.

D-glucose inhibition was not significantly affected by the Gal2^{T89I} substitution but, under the same conditions, L-arabinose uptake by Gal2^{N376I}, Gal2^{N376S} and Gal2^{N376T} was inhibited by only 20–30% (**Figure 6 and Additional Figure 10**).

The Gal2^{N376I}, Gal2^{N376S} and Gal2^{N376T} substitutions yielded 25-60% lower K_m values for L-arabinose than wild-type Gal2, while their K_m values for D-glucose were one to two orders of magnitude higher than those of wild-type Gal2 (**Table 5**). As a result, the ratios of the K_m values for L-arabinose vs. D-glucose was 2 orders of magnitude lower for the strains expression Gal2 variants with substitutions at position 376. For Gal2 variants that only carried an amino acid substitution at this position, transport capacities (V_{max}) for L-arabinose and D-glucose differed by less than two-fold from those of wild-type Gal2 (**Table 5**). These changes in transport kinetics are consistent with a strongly reduced competitive inhibition of L-arabinose transport by D-glucose.

A single Gal2^{T89I} substitution increased the K_m for D-glucose transport from 1.9 to 7 mM, while it decreased the K_m for L-arabinose from 335 to 99 mM. For both sugars, the Gal2^{T89I} substitution caused a 2–3 fold reduction of V_{max} (**Table 5**). These kinetic properties suggest that Gal2^{T89I} may, by itself, confer a selective advantage at low extracellular L-arabinose concentrations. Remarkably, a Gal2 variant that harboured both the Gal2^{N376T} and Gal2^{T89I} substitutions no longer transported D-glucose, while K_m and V_{max} values for L-arabinose were similar to those of Gal2^{T89I} (**Table 5**).

Table 5. K_m and V_{max} values for L-arabinose and D-glucose for Gal2 variants with amino acid substitutions at positions N376 and T89. Transport kinetics were measured by uptake studies with radioactive sugars after expression of *GAL2* alleles in *S. cerevisiae* DS68625. Values are average and mean deviation of two independent sets of uptake experiments. The detection limit for D-glucose uptake (V_{max}) was 1.8 nmol^{-1} (mg biomass)$^{-1}$ min^{-1}. Data used to calculate K_m and V_{max} values are shown in **Additional Figure 11.**

	K_m (mM)		K_m ratio	V_{max} (nmol (mg biomass)$^{-1}$ min^{-1})	
Gal2 variant	**L-arabinose**	**D-glucose**	**Ara/Glc**	**L-arabinose**	**D-glucose**
Gal2[1]	335 ± 21	1.9 ± 0.2	176	75 ± 5	26 ± 1
Gal2^{N376I}	117 ± 16	101 ± 47	1	39 ± 3	32 ± 18
Gal2^{N376S}	186 ± 33	38 ± 1	5	64 ± 2	28 ± 1
Gal2^{N376T}	171 ± 17	57 ± 1	3	65 ± 2	17 ± 4
Gal2$^{T89I, N376T}$	103 ± 40	-	-	30 ± 2	< 1.8
Gal2^{T89I}	99 ± 18	7 ± 0.2	15	22 ± 3	13 ± 0.1

[1]Kinetic data for wild-type Gal2 were derived from (Bracher *et al.* 2018).

The Gal2^{T89I} substitution also occurs in a previously evolved pentose-fermenting strain

Strain *S. cerevisiae* IMS0010 is an engineered pentose-fermenting strain that was previously evolved for improved anaerobic fermentation kinetics on mixtures of D-glucose, D-xylose and L-arabinose (Wisselink *et al.* 2009). Sequence analysis of *GAL2* in IMS0010 showed that this strain contains the same Gal2^{T89I} substitution found in one of the two *GAL2* alleles in the independently evolved glucose-phosphorylation-negative strain IMS0522. Analysis of *GAL2* sequences in two intermediate strains in the construction and laboratory evolution of strain IMS0010 (strains IMS0002 and IMS0007; (Wisselink *et al.* 2009)) showed that this mutation was acquired during the evolution for improved kinetics of mixed-substrate utilisation.

Reverting the Gal2^{T89I} substitution negatively affects L-arabinose consumption

CO_2 production profiles of anaerobic bioreactor batch cultures (**Additional Figure 6**) indicated that strain IMS0522, in which one of the two *GAL2* alleles encodes both the Gal2^{N376T} and Gal2^{T89I} substitutions, grew faster on L-arabinose in the presence of D-glucose and D-xylose than strain IMS0523, in which both copies of *GAL2* only encode the Gal2^{N376T} substitution. To investigate whether the additional Gal2^{T89I} substitution contributed to this phenotype, the allele encoding this substitution was restored to 'Gal2^{N376T} only'. Anaerobic batch cultures of the resulting strain IMW091 on SMAGX showed a substantially reduced specific growth rate relative to its parental strain IMS0522 (0.069 h^{-1} and 0.12 h^{-1}, respectively, **Table 4**). As a consequence, the time needed to completely ferment L-arabinose in the presence of D-glucose and D-xylose was much shorter for the parental strain IMS0522 (Gal2^{N376T}/Gal2$^{N376T\ T89I}$) than for the strain IMW091 (Gal2^{N376T}/Gal2$^{N376T\ T}$); 25 and 45 h, respectively) (**Figure 4**).

Discussion

The *S. cerevisiae* strain used for this evolutionary engineering study combined genetic modifications that were previously shown to enable or stimulate growth on L-arabinose. Combined expression of a bacterial L-arabinose pathway (Sedlak and Ho 2001, Becker and Boles 2003, Wisselink *et al.* 2007), overexpression of native non-oxidative pentose-phosphate-pathway enzymes (Kuyper *et al.* 2005, Wisselink *et al.* 2007), deletion of the non-specific aldose reductase gene *GRE3* (Träff *et al.* 2001), deregulated expression of the Gal2 galactose/L-arabinose transporter (Tani *et al.* 2016, Wang *et al.* 2017), here accomplished by deletion of *GAL80* (Torchia *et al.* 1984, Ostergaard *et al.* 2001, Tani *et al.* 2016), and expression of the fungal L-arabinose transporter *Pc*AraT (Bracher *et al.* 2018) yielded a strain that, in aerobic shake-flask cultures, exhibited a specific growth rate on L-arabinose of 0.26 h^{-1}. Moreover, deletion of all four genes encoding enzymes with glucose-phosphorylating activity rendered this 'arabinose-specialist' strain *S. cerevisiae* IMX728 unable to grow on glucose. Construction of this strain was completed within 3 months, which illustrates the power of Cas9-assisted genome editing combined with *in vivo* and *in vitro* DNA assembly methods (DiCarlo *et al.* 2013, Kuijpers *et al.* 2013, Mans *et al.* 2015, Tsai *et al.* 2015).

The specific growth rate of strain IMX728 on L-arabinose in aerobic batch cultures, which was achieved without evolutionary engineering or random mutagenesis, corresponded to 65% of the specific growth rate on glucose of the parental strain CEN.PK113-7D (Solis-Escalante *et al.* 2015) and is the highest growth rate on L-arabinose hitherto reported for an engineered *S. cerevisiae* strain (Wisselink *et al.* 2007, Wiedemann and Boles 2008, Wang *et al.* 2013). However, the genetic modifications in this strain did not enable anaerobic growth on L-arabinose. Anaerobic growth on L-arabinose requires that the biomass-specific L-arabinose consumption rate ($q_{arabinose}$) is sufficiently high to meet the ATP requirement for cellular maintenance (ca. 1 mmol ATP (g biomass)$^{-1}$ h^{-1}, (Boender *et al.* 2009)). If L-arabinose uptake occurs via Gal2-mediated facilitated diffusion, anaerobic, fermentative metabolism of L-arabinose yields 1.67 ATP, while transport of L-arabinose exclusively via *Pc*AraT-mediated proton symport would result in a 60% lower ATP yield (Weusthuis *et al.* 1994).

Even if L-arabinose metabolism in the aerobic batch cultures of strain IMX728 would occur exclusively via the energetically much more favorable process of respiratory dissimilation (Bakker *et al.* 2001, Verhoeven *et al.* 2017), its specific growth rate of 0.26 h^{-1} would still correspond to a $q_{arabinose}$ of at least 3 mmol g^{-1} h^{-1}. This rate of L-arabinose metabolism should be sufficient to meet anaerobic maintenance requirements via fermentative metabolism. The inability of strain IMX728 to instantaneously grow on L-arabinose in anaerobic cultures therefore suggests that it cannot achieve the same $q_{arabinose}$ under those

conditions, which might reflect suboptimal expression, folding, membrane environment and/or stability of wild-type Gal2 in anaerobic cultures.

The observation that the two *GAL2* copies in individual strains evolved for anaerobic growth in L-arabinose in the presence of D-glucose and D-xylose encoded the same Gal2^{N376} substitution, indicated that these substitutions most probably preceded gene duplication. The importance of these substitutions is further underlined by the observation that introduction of a single copy of the *GAL2*^{N376I} allele enabled growth of an 'arabinose specialist' strain on L-arabinose in the presence of glucose and/or L-arabinose under aerobic conditions (**Figure 5**). *GAL* genes are transcriptionally regulated via pathways involving Mig1 and Gal80 (Ostergaard *et al.* 2001). In the present study, Gal80-mediated repression was eliminated by deletion of its encoding gene. The observed growth of glucose-phosphorylation-negative strains expressing an evolved *GAL2*^{N376} allele on L-arabinose in glucose-containing media indicates that any Mig1-mediated glucose repression of *GAL2* was incomplete. In the glucose-phosphorylation-negative strains, absence of Hxk2, which is involved in Mig1-dependent glucose repression (Ahuatzi *et al.* 2007) is likely to have ~~to~~ glucose derepression of *GAL* genes. Indeed, deletion of *HXK2* has been reported to stimulate co-consumption of glucose and galactose (Raamsdonk *et al.* 2001). Additionally, in the same *S. cerevisiae* genetic background, combined deletion of *MIG1*, *GAL6* and *GAL80* was shown not to significantly affect *GAL2* transcript levels (Bro *et al.* 2005).

Two previous studies used D-glucose-phosphorylation-negative D-xylose-fermenting strains to evolve for D-glucose tolerance of D-xylose utilization (Farwick *et al.* 2014, Nijland *et al.* 2014). These studies identified mutations in hexose transporters that reduced D-glucose sensitivity of D-xylose transport. In Gal2, which also transports D-xylose (Farwick *et al.* 2014) and in a chimera of Hxt3 and Hxt6 (Hxt36) (Nijland *et al.* 2014), these mutations caused amino acid substitutions at the corresponding positions N376 and N367, respectively. Using *in silico* models of the three-dimensional structure of Gal2 and Hxt36, based on the crystal structure of the *E. coli* XylE D-xylose/H^+ symporter, changes of the amino acid at these positions were predicted to confine the space in the substrate binding pocket for the hydroxymethyl group of D-glucose (Sun *et al.* 2012, Farwick *et al.* 2014, Nijland *et al.* 2014). Both studies showed that size, charge, polarity and presence of hydrophobic side chains for the substituted amino acid strongly influenced relative transport activities with D-glucose and D-xylose. The present study shows that modifications of Gal2 at this position also enable transport of L-arabinose in the presence of D-glucose. Of the Gal2^{N376} substitutions identified and kinetically analyzed in our study, Gal2^{N376T} showed the highest K_m and lowest V_{max} for D-glucose. Moreover, this substitution decreased the K_m for L-arabinose without affecting V_{max} for uptake of this pentose.

Of the four D-glucose-tolerant evolved strains isolated from the evolution experiments, strain IMS0522 showed significantly better L-arabinose fermentation performance than the other strains. Its L-arabinose consumption rate of 1.6 ± 0.08 g (g biomass)$^{-1}$ h^{-1} and ethanol production rate of 0.61 ± 0.01 g (g biomass)$^{-1}$ h^{-1} in anaerobic cultures are the highest reported to date for an engineered *S. cerevisiae* strain (Wisselink *et al.* 2007, Wiedemann and Boles 2008, Wang *et al.* 2013). In addition to a Gal2^{N376T} substitution, one of the *GAL2* copies in this strain encoded an additional Gal2^{T89I} substitution, whose restoration to wild type resulted in a fermentation performance similar to that of the other strains. The physiological relevance of the mutation was further supported by its identification in a strain that was previously evolved for mixed-sugar fermentation (Wisselink *et al.* 2009). By itself, the Gal2^{T89I} showed a reduced K_m for L-arabinose and an increased K_m for D-glucose, while the V_{max} for both sugars was decreased relative to that of wild-type Gal2. The Gal2^{T89I} variant was also investigated in a recent study that systematically investigated amino acid substitutions at positions within or close to the predicted L-arabinose binding pocket of Gal2 (Wang *et al.* 2017). Subsequent expression of the Gal2^{T89I} substitution did not show a significant effect on the growth rate on L-arabinose (Wang *et al.* 2017). However, our results suggest that, by itself, this substitution is advantageous for L-arabinose uptake in the presence of glucose and/or at low L-arabinose concentrations. Additionally, in the evolved strain IMS0010, in which Gal2 contained only this substitution, the mutations may have contributed to anaerobic co-consumption of L-arabinose and D-xylose (Wisselink *et al.* 2009).

Transport assays with the Gal2^{N376T,T89I} variant showed complete loss of D-glucose transport capacity, combined with a reduced K_m and V_{max} for L-arabinose. In the context of the SBR protocol in which this variant evolved, these kinetic properties are particularly relevant during the growth phase in which the L-arabinose concentration declined, while D-glucose in still present at 20 g L^{-1}. The complete loss of glucose transport activity of the Gal2^{N376T,T89I}, which was encoded by only one of two copies of *GAL2* in the laboratory-evolved strain originating from a duplication of Chromosome XII, underline the importance of chromosomal copy number variations in laboratory evolution and neofunctionalization of duplicated genes (Gorter de Vries *et al.* 2017). These results from a 500-generation laboratory evolution experiment illustrate how, during evolution in dynamic natural environments, similar duplication and neofunctionalization events, may have contributed to the diverse kinetic characteristics and substrate specificities of the Hxt transporter family in *S. cerevisiae* (Reifenberger *et al.* 1997, Biswas *et al.* 2013, Jordan *et al.* 2016).

In addition to Gal2, the parental strain used in the evolution experiments expressed the fungal L-arabinose transporter *Pc*AraT, which upon expression in *S. cerevisiae* mediates high-affinity, L-arabinose-proton symport (K_m = 0.13 mM) and is much less sensitive

to glucose inhibition than wild-type Gal2 (Bracher *et al.* 2018). During the evolution experiments, no loss of function mutations were observed in *Pc*AraT. This observation indicates that its presence did not confer a strong selective disadvantage, for example as a result of the activity of a futile cycle caused by simultaneous facilitated diffusion and proton symport of L-arabinose. Moreover, the observations that deletion of *Pc*AraT resulted in decreased specific growth rate on L-arabinose and that presence of *Pc*AraT coincided with a faster consumption of L-arabinose towards the end of anaerobic batch cultures (**Figure 4b**), are in line with its reported low V_{max} and low K_m for L-arabinose upon expression in *S. cerevisiae* (Bracher *et al.* 2018).

These results demonstrate the potential of high-affinity L-arabinose transporters to efficiently convert low concentrations of L-arabinose towards the end of fermentation processes, thereby preventing prolonged 'tailing' of industrial fermentation processes.

Conclusions

Laboratory evolution of engineered, L-arabinose-metabolizing and glucose-phosphorylation-negative *S. cerevisiae* yielded evolved strains that anaerobically fermented L-arabinose in the presence of D-glucose and D-xylose. Amino acid substitutions in Gal2 that affected the kinetics of L-arabinose and D-glucose transport played a key role in this evolution. The best performing evolved strain contained two different mutated alleles of *GAL2*, encoding Gal2 variants with distinct kinetic properties. This result demonstrates the importance of engineering 'transporter landscapes' for uptake of individual sugars, consisting of transporters with complementary kinetic and/or regulatory properties, for efficient sugar conversion in dynamic, mixed-sugar fermentation processes.

Acknowledgements

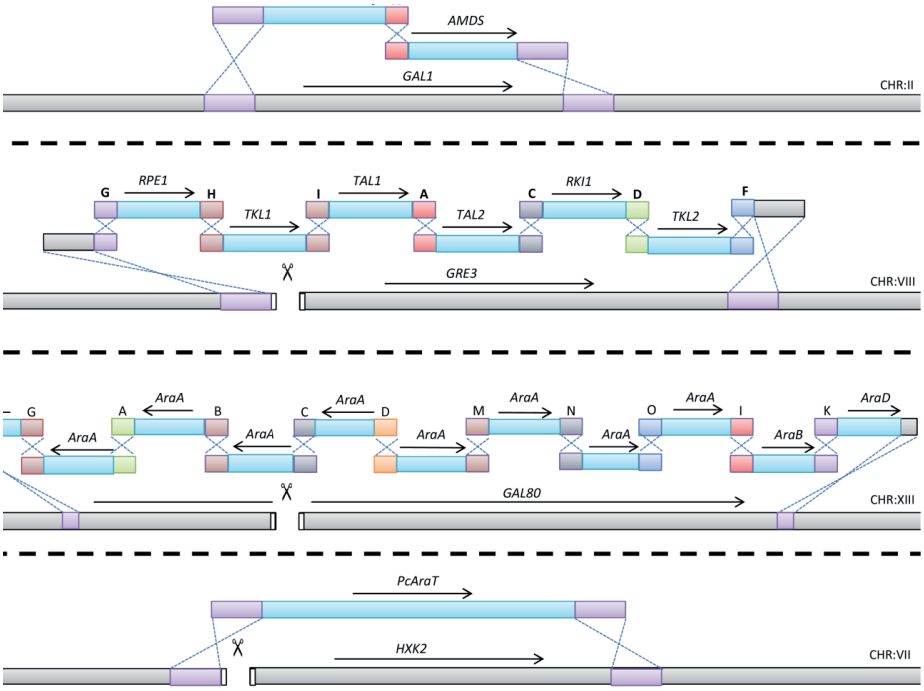
This work was performed within the BE-Basic R&D Program (http://www.be-basic.org/), which is financially supported by an EOS Long Term grant from the Dutch Ministry of Economic Affairs, Agriculture and Innovation (EL&I). The authors thank Marcel van den Broek (TUD), Pilar de la Torre (TUD), Ioannis Papapetridis (TUD), Laura Valk (TUD), Renzo Rozenbroek (TUD), Paul de Waal (DSM), Hans de Bruijn (DSM) and Paul Klaassen (DSM) for their input in this project.

Additional material

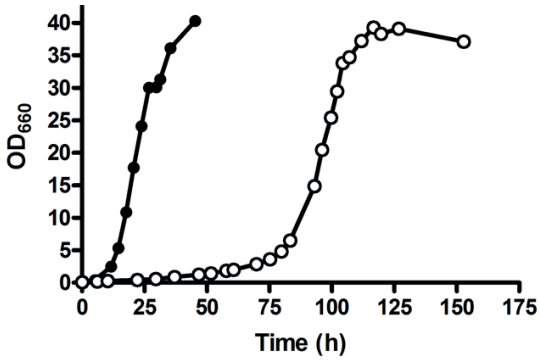
Additional Table 1. Plasmids used in this study.

Plasmid	Characteristic	Source
p414-TEF1p-cas9-CYC1t	*CEN6/ARS4* ampR *pTEF1-cas9-tCYC1*	(DiCarlo *et al.* 2013)
pPWT118	KanMX, amdSYM, pADH1_*PcaraT*_tPMA1	(Bracher *et al.* 2018)
pUG-amdSYM	Template for amdS marker	(Solis-Escalante *et al.* 2013)
pUG-72	Template for *KlURA3* marker	(Gueldener *et al.* 2002)
pUDE327	2 µm ori, *KlURA3*, pSNR52-gRNA.HXK2.Y	(Kuijpers *et al.* 2016)
pUDE335	2 µm ori, *KlURA3*, pSNR52-gRNA.*GRE3*.Y	(Verhoeven *et al.* 2017)
pMEL10	pSNR52-gRNA.*CAN1*.Y-tSUP4	(Mans *et al.* 2015)
pROS10	2 µm ori, *KlURA3*, pSNR52-gRNA.*CAN1*.Y pSNR52-gRNA.*ADE2*.Y	(Mans *et al.* 2015)
pROS10	2 µm ori, *kanMX*, pSNR52-gRNA.*CAN1*.Y pSNR52-gRNA.*ADE2*.Y	(Mans *et al.* 2015)
pUD344	pJET1.2Blunt + TagA_pPGI1_*NQM1*_TagB	(Verhoeven *et al.* 2017)
pUD345	pJET1.2Blunt + TagB_pTPI1_*RKI1*_Tagc	(Verhoeven *et al.* 2017)
pUD346	pJET1.2Blunt + TagC_pPYK1_*TKL2*_TagF	(Verhoeven *et al.* 2017)
pUD347	pJET1.2Blunt + TagG_pTDH3_*RPE1*_TagH	(Verhoeven *et al.* 2017)
pUD348	pJET1.2Blunt + TagH_pPGK1_*TKL1*_TagI	(Verhoeven *et al.* 2017)
pUD349	pJET1.2Blunt + TagI_pTEF1_*TAL1*_TagA	(Verhoeven *et al.* 2017)
pUD354	pMK-RQ_pTPI1_*araA*_tADH3	This study
pUD355	pMK-RQ_pPYK1_*araB*_tPGI1	This study
pUD356	pMK-RQ_pPGK1_*araD*_tTDH3	This study
pUDE348	2 µm ori, *KlURA3*, pSNR52-gRNA.*GAL80*.Y	This study
pUDR172	2 µm ori, *KlURA3*, pSNR52-gRNA.*GAL2*^{N376}.Y pSNR52-gRNA.*GAL2*^{N376}.Y	This study
pUDR187	2 µm ori, *kanMX*, pSNR52-gRNA.*GAL2*^{I89}.Y pSNR52-gRNA.*GAL2*^{I89}.Y	This study
pRS313-Gal2	CEN6/ARS4 ampR pHXT7-*GAL2*-tHXT7	This study
pRS313-Gal2-T	CEN6/ARS4 ampR pHXT7- *GAL2*^{N376T}-tHXT7	This study
pRS313-Gal2-S	CEN6/ARS4 ampR pHXT7- *GAL2*^{N376S}-tHXT7	This study
pRS313-Gal2-I	CEN6/ARS4 ampR pHXT7- *GAL2*^{N376I}-tHXT7	This study
pRS313-Gal2-TI	CEN6/ARS4 ampR pHXT7- *GAL2*^{N376T,T89I}-tHXT7	This study
pRS313-*PcaraT*	CEN6/ARS4 ampR pHXT7-*PcaraT*-tHXT7	(Bracher *et al.* 2018)
pRS313-mcs	CEN6/ARS4 ampR pHXT7-mcs-tHXT7	(Nijland *et al.* 2014)

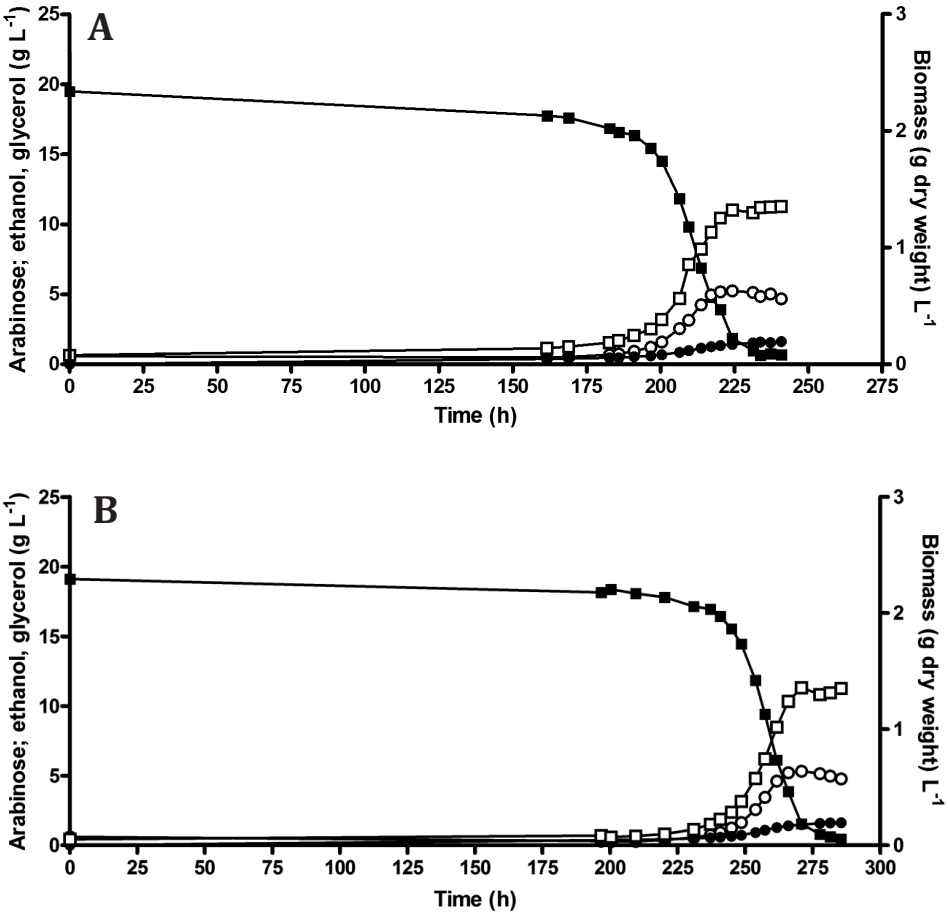
Additional Table 2. Primers used in this study are available at: https://doi.org/10.1093/femsyr/foy062



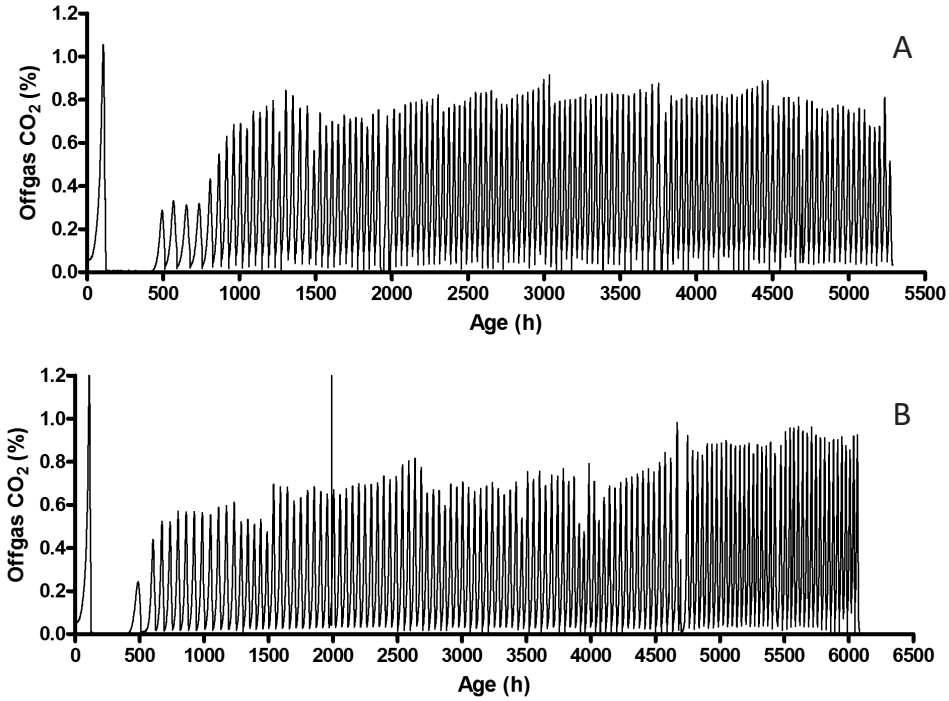
Additional Figure 1. Schematic overview of the integrated constructs that result in strain IMX728 (*hxk1Δ glk1Δ gal1Δ*::{Cas9_amdSYM} *gre3Δ*::{NPPP} *hxk2Δ::PcaraT gal80Δ*::{ARAP}). A. fragments with expression cassettes for Cas9 and amdSYM were cotrasnformed in IMX080. **B.** the fragments with six expression cassettes for genes in the non-oxidative PPP were transformed with pUDE335 allowing for a Cas9-induced double-strand break in *GRE3*. **C**. 11 fragments with 9 times an expression cassette for *araA* and single fragments for *AraB* and *AraD* were transformed with plasmid pUDE348. **D.** *PcaraT* was integrated in the *HXK2* locus using plasmid pUDE327. Correct integration of all the fragments was verified by diagnostic PCR.



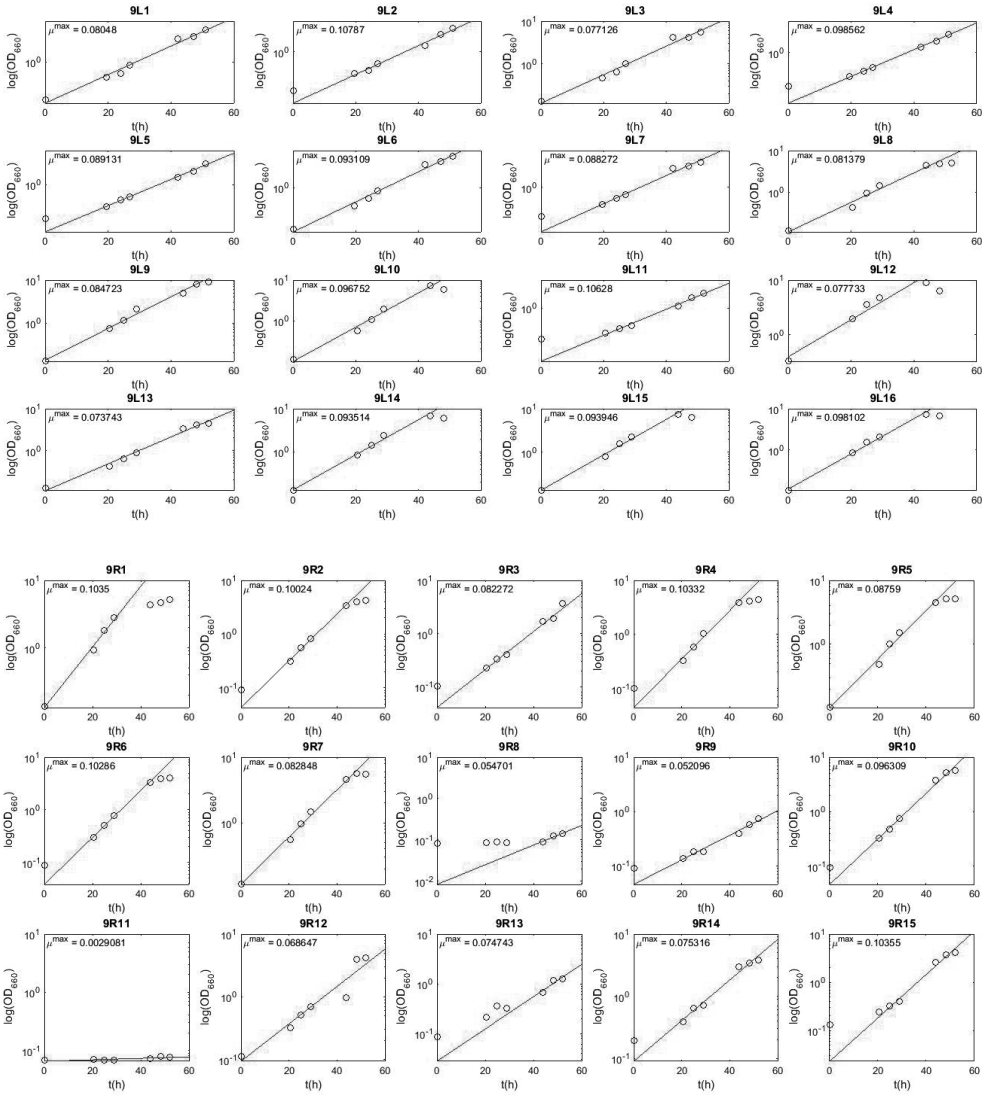
Additional Figure 2. Growth curves for replicate experiments of *S. cerevisiae* strain IMX728 (engineered, non-evolved arabinose-consuming, glucose-phosphorylation-negative, expressing *PcaraT*) in aerobic batch cultures on synthetic medium containing 20 g L^{-1} L-arabinose (SMA) as sole carbon and energy source (●) and in synthetic medium with 20 g L^{-1} L-arabinose, 20 g L^{-1} D-glucose and 20 g L^{-1} D-xylose added (○).



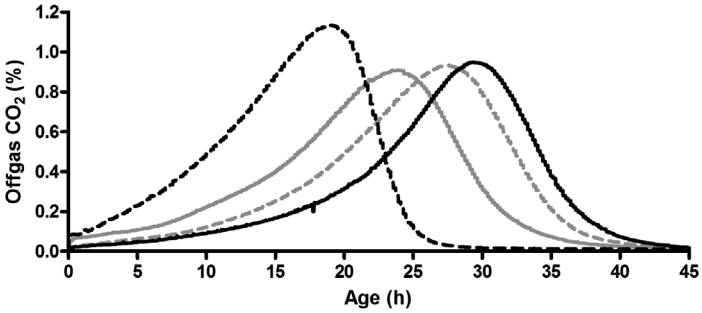
Additional Figure 3. Growth and extracellular metabolite concentrations for duplicate anaerobic cultures (A. reactor 1 and B. reactor 2) of engineered L-arabinose consuming glucose-phosphorylation negative strain IMX728. The culture was grown in bioreactors containing synthetic medium containing 20 g L^{-1} L-arabinose. ▪ L-arabinose; □ biomass dry weight; ● glycerol; ○ ethanol.



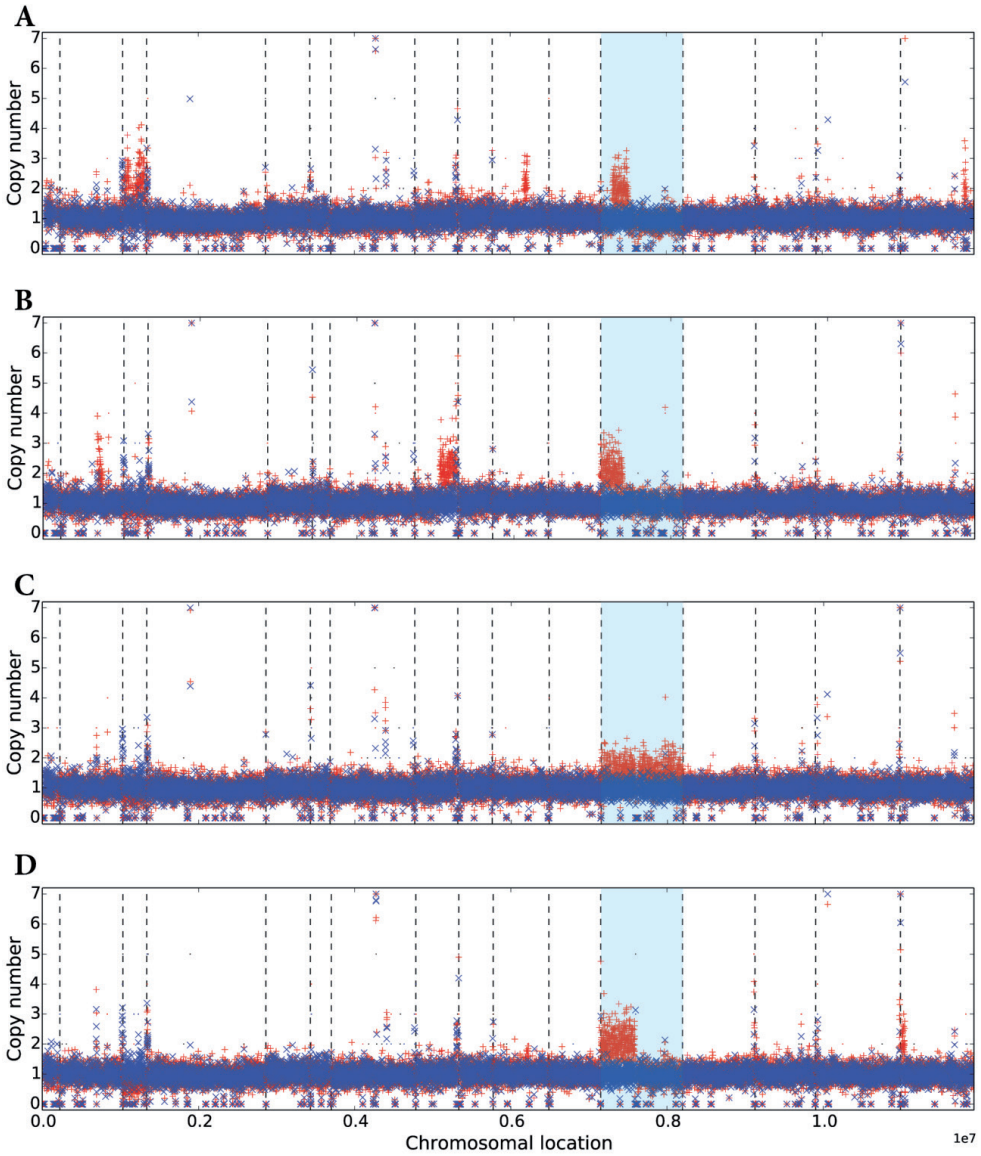
Additional Figure 4. CO_2 offgas measurements (%) for both anaerobic sequential batch A. reactor 1 and B. reactor 2 during laboratory evolution of strain IMX728 (L-arabinose consuming, glucose-phosphorylation-negative) on synthetic medium with 20 g L^{-1} L-arabinose, 20 g L^{-1} D-glucose and 20 g L^{-1} D-xylose. The first CO_2 curve corresponds to the initial aerobic batch prior to switching sparging from air to nitrogen.



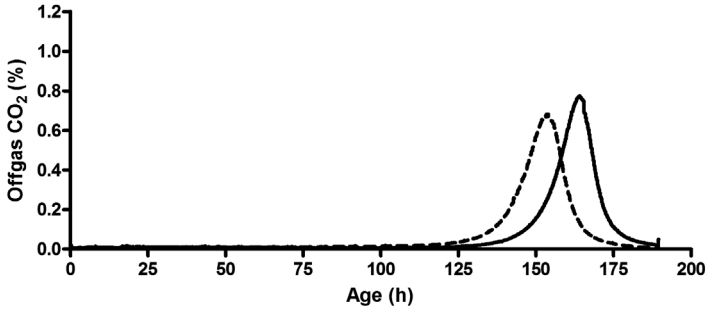
Additional Figure 5. Logarithmic plots of OD_{660} measurements for anaerobic shake flask cultures using single colony isolates obtained after laboratory evolution of strain IMX728. Shake flasks were incubated at 30°C in an anaerobic chamber. 9L1-9L16 and 9R1-9R15 correspond to isolates originating from laboratory evolution reactor 1 or 2 respectively. Values shown are averaged from two independent duplicates.



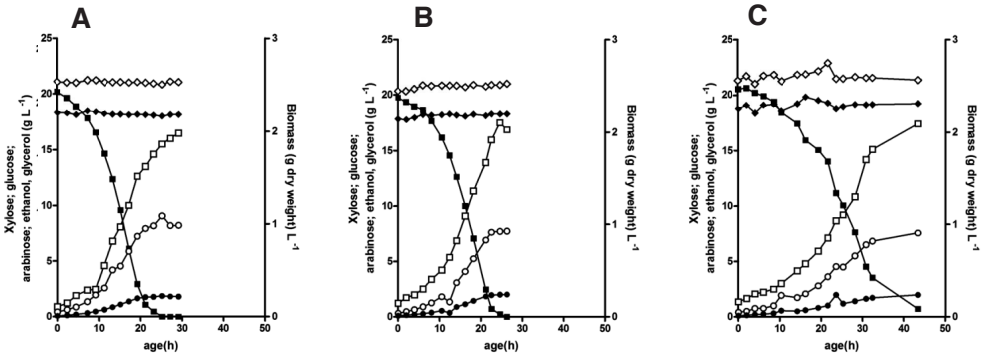
Additional Figure 6. CO_2 offgas measurements (%) for four isolates selected after laboratory evolution of IMX728 (L-arabinose consuming, glucose-phosphorylation-negative) on synthetic medium with 20 g L^{-1} L-arabinose, 20 g L^{-1} D-glucose and 20 g L^{-1} D-xylose (SMAGX). Single anaerobic bioreactor cultures of IMS0520 (solid grey), IMS0521 (dashed grey), IMS0522 (dashed black) and IMS523 (solid black) were conducted and monitored on SMAGX at 30°C and pH5.0.



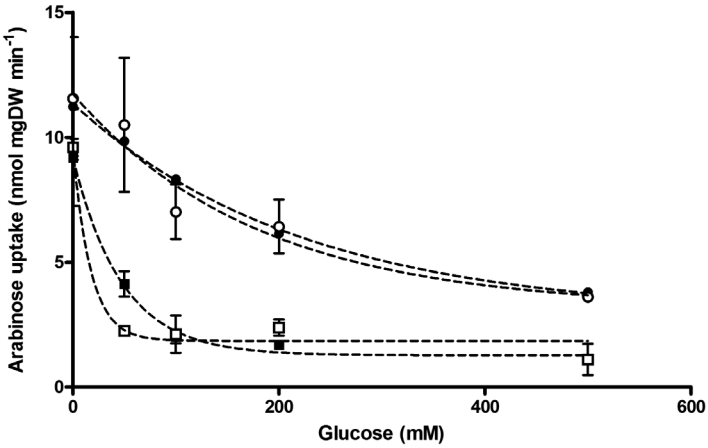
Additional Figure 7. Chromosomal copy number prediction of the full *S. cerevisiae* genome estimated using the Magnolya algorithm. Graphs A to D correspond to single colony isolated IMS0520 to IMS0523 respectively. Reads from each isolate (red crosses) show duplications in chromosome 12 compared to parental strain IMX728 (blue crosses). The light blue region corresponds to chromosome 12.



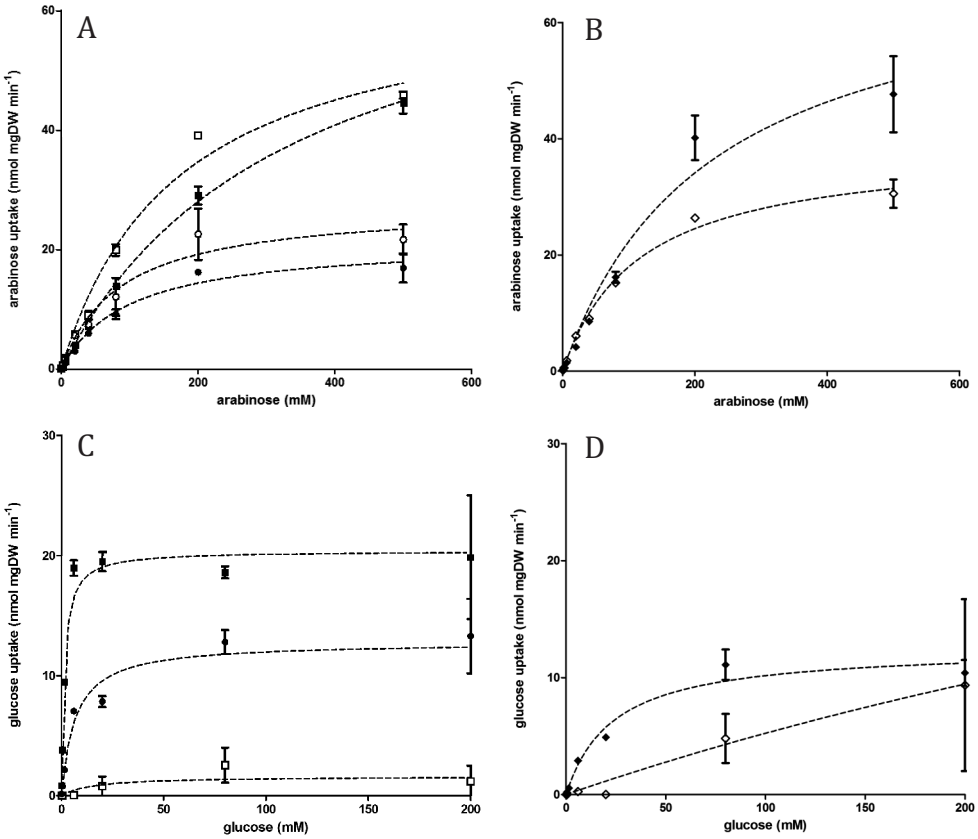
Additional Figure 8. CO_2 offgas measurements (%) for strain IMX1286 (L-arabinose consuming, glucose-phosphorylation-negative, with an additional copy of Gal2 integrated in the SGA1 locus) on synthetic medium with 20 g L^{-1} L-arabinose. The two curves represent the measurement data of two independent replicate experiments.



Additional Figure 9. Growth and extracellular metabolite concentrations of replicate anaerobic cultures of the evolved L-arabinose consuming, glucose-phosphorylation negative *S. cerevisiae* strain IMS0522 and two derived strains with specific genetic modifications. Cultures were grown in bioreactors containing synthetic medium with 20 g L^{-1} L-arabinose, 20 g L^{-1} D-glucose and 20 g L^{-1} D-xylose. **A.** IMS0522, **B.** IMW088 (*PcaraT*Δ), **C.** IMW091 (Gal2^{89I} reverted to Gal2^{89T}). ▪ L-arabinose; □ biomass dry weight; ● glycerol; ○ ethanol; ◆ D-glucose; ◇ D-xylose.



Additional Figure 10. Uptake experiments using 50 mM [^{14}C-] L-arabinose in the presence of a range of D-glucose concentrations using strain DS68625 expressing Gal2 variants: N376T+T89I (○), N376T (●), T89I (□) and Gal2 wild-type (▪).



Additional Figure 11. Uptake experiments using various concentrations of [^{14}C-] L-arabinose (graph A. and B.) or [^{14}C-] D-glucose (graph C. and D.) to determine K_m and V_{max} of strain DS68625 expressing Gal2 variants: T89I (○), N376T+T89I (●), N376T (□), Gal2 wild-type (▪), N376S (◆) and N376I (◇).

References

Ahuatzi D, Riera A, Peláez R *et al.* Hxk2 regulates the phosphorylation state of Mig1 and therefore its nucleocytoplasmic distribution. *J Biol Chem* 2007;**282**:4485-93.

Association RF. World fuel ethanol production (24 March 2017, date last accessed).

Bakker BM, Overkamp KM, van Maris AJ *et al.* Stoichiometry and compartmentation of NADH metabolism in *Saccharomyces cerevisiae*. *FEMS Microbiol Rev* 2001;**25**:15-37.

Becker J, Boles E. A modified *Saccharomyces cerevisiae* strain that consumes L-arabinose and produces ethanol. *Appl Environ Microbiol* 2003;**69**:4144-50.

Biswas C, Djordjevic JT, Zuo X *et al.* Functional characterization of the hexose transporter Hxt13p: an efflux pump that mediates resistance to miltefosine in yeast. *Fungal Genet Biol* 2013;**61**:23-32.

Boender LG, de Hulster EA, van Maris AJ *et al.* Quantitative physiology of Saccharomyces cerevisiae at near-zero specific growth rates. *Appl Environ Microbiol* 2009;**75**:5607-14.

Bracher JM, Verhoeven MD, Wisselink HW *et al.* The *Penicillium chrysogenum* transporter *Pc*AraT enables high-affinity, glucose-insensitive L-arabinose transport in *Saccharomyces cerevisiae*. *Biotechnol Biofuels* 2018a;**11**:63.

Bro C, Knudsen S, Regenberg B *et al.* Improvement of galactose uptake in *Saccharomyces cerevisiae* through overexpression of phosphoglucomutase: example of transcript analysis as a tool in inverse metabolic engineering. *Appl Environ Microbiol* 2005;**71**:6465-72.

Crowley JH, Leak FW, Shianna KV *et al.* A mutation in a purported regulatory gene affects control of sterol uptake in *Saccharomyces cerevisiae*. *J Bacteriol* 1998;**180**:4177-83.

Gorter de Vries AR, Pronk JT, Daran J-MG. Industrial relevance of chromosomal copy number variation in *Saccharomyces* yeasts. *Appl Environ Microbiol* 2017;**83**:e03206-16.

DiCarlo JE, Norville JE, Mali P *et al.* Genome engineering in *Saccharomyces cerevisiae* using CRISPR-Cas systems. *Nucleic Acids Res* 2013;**41**:4336-43.

Dickson RC, Nagiec EE, Wells GB *et al.* Synthesis of mannose-(inositol-P) 2-ceramide, the major sphingolipid in *Saccharomyces cerevisiae*, requires the *IPT1* (YDR072c) gene. *J Biol Chem* 1997;**272**:29620-5.

Entian K-D, Kötter P. 25 Yeast genetic strain and plasmid collections. *Method Microbiol* 2007;**36**:629-66.

Farwick A, Bruder S, Schadeweg V *et al.* Engineering of yeast hexose transporters to transport D-xylose without inhibition by D-glucose. *Proc Natl Acad Sci* 2014;**111**:5159-64.

Gietz RD, Schiestl RH, Willems AR *et al.* Studies on the transformation of intact yeast cells by the LiAc/SS-DNA/PEG procedure. *Yeast* 1995;**11**.

Grohmann K, Bothast R. Pectin-rich residues generated by processing of citrus fruits, apples, and sugar beets. ACS Publications, 1994.

Grohmann K, Bothast RJ. Saccharification of corn fibre by combined treatment with dilute sulphuric acid and enzymes. *Process Biochem* 1997;**32**:405-15.

Guadalupe Medina V, Almering MJ, van Maris AJ *et al.* Elimination of glycerol production in anaerobic cultures of a *Saccharomyces cerevisiae* strain engineered to use acetic acid as an electron acceptor. *Appl Environ Microbiol* 2010;**76**:190-5.

Gueldener U, Heinisch J, Koehler G *et al.* A second set of *loxP* marker cassettes for Cre-mediated multiple gene knockouts in budding yeast. *Nucleic Acids Res* 2002;**30**:e23-e.

Hamacher T, Becker J, Gárdonyi M *et al.* Characterization of the xylose-transporting properties of yeast hexose transporters and their influence on xylose utilization. *Microbiology* 2002;**148**:2783-8.

Jeffries TW. Utilization of xylose by bacteria, yeasts, and fungi.*Pentoses and Lignin*: Springer, 1983,1-32.

Jordan P, Choe J-Y, Boles E *et al.* Hxt13, Hxt15, Hxt16 and Hxt17 from *Saccharomyces cerevisiae* represent a novel type of polyol transporters. *Sci Rep* 2016;**6**.

Knoshaug EP, Vidgren V, Magalhães F *et al.* Novel transporters from *Kluyveromyces marxianus* and *Pichia guilliermondii* expressed in *Saccharomyces cerevisiae* enable growth on L-arabinose and D-xylose. *Yeast* 2015;**32**:615-28.

Kou S-C, Christensen MS, Cirillo VP. Galactose transport in *Saccharomyces cerevisiae* II. Characteristics of galactose uptake and exchange in galactokinaseless cells. *J Bacteriol* 1970;**103**:671-8.

Kuijpers NG, Chroumpi S, Vos T *et al.* One-step assembly and targeted integration of multigene constructs assisted by the I-SceI meganuclease in *Saccharomyces cerevisiae*. *FEMS Yeast Res* 2013;**13**:769-81.

Kuijpers NG, Solis-Escalante D, Luttik MA *et al.* Pathway swapping: Toward modular engineering of essential cellular processes. *Proc Natl Acad Sci* 2016:201606701.

Kuyper M, Harhangi HR, Stave AK *et al.* High-level functional expression of a fungal xylose isomerase: the key to efficient ethanolic fermentation of xylose by *Saccharomyces cerevisiae FEMS Yeast Res* 2003;**4**.

Kuyper M, Hartog MM, Toirkens MJ *et al.* Metabolic engineering of a xylose-isomerase-expressing *Saccharomyces cerevisiae* strain for rapid anaerobic xylose fermentation. *FEMS Yeast Res* 2005;**5**:399-409.

Li H, Durbin R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* 2009;**25**:1754-60.

Li J, Xu J, Cai P *et al.* Functional analysis of two L-arabinose transporters from filamentous fungi reveals promising characteristics for improved pentose utilization in *Saccharomyces cerevisiae*. *Appl Environ Microbiol* 2015;**81**:4062-70.

Lynd LR. Overview and evaluation of fuel ethanol from cellulosic biomass: technology, economics, the environment, and policy. *Annu Rev Energy Env* 1996;**21**:403-65.

Mans R, van Rossum HM, Wijsman M *et al.* CRISPR/Cas9: a molecular Swiss army knife for simultaneous introduction of multiple genetic modifications in *Saccharomyces cerevisiae*. *FEMS Yeast Res* 2015a;**15**:fov004.

Nijkamp JF, van den Broek M, Datema E *et al.* De novo sequencing, assembly and analysis of the genome of the laboratory strain *Saccharomyces cerevisiae* CEN. PK113-7D, a model for modern industrial biotechnology. *Microb Cell Fact* 2012a;**11**:36.

Nijkamp JF, van den Broek MA, Geertman J-MA *et al. De novo* detection of copy number variation by co-assembly. *Bioinformatics* 2012b;**28**:3195-202.

Nijland JG, Shin HY, de Jong RM *et al.* Engineering of an endogenous hexose transporter into a specific D-xylose transporter facilitates glucose-xylose co-consumption in *Saccharomyces cerevisiae*. *Biotechnol Biofuels* 2014;**7**:168.

Ostergaard S, Walløe KO, Gomes CS *et al.* The impact of GAL6, GAL80, and MIG1 on glucose control of the GAL system in *Saccharomyces cerevisiae*. *FEMS Yeast Res* 2001;**1**:47-55.

Oud B, van Maris AJ, Daran J-M *et al.* Genome-wide analytical approaches for reverse metabolic engineering of industrially relevant phenotypes in yeast. *FEMS Yeast Res* 2012;**12**:183-96.

Raamsdonk LM, Diderich JA, Kuiper A *et al.* Co-consumption of sugars or ethanol and glucose in a *Saccharomyces cerevisiae* strain deleted in the *HXK2* gene. *Yeast* 2001;**18**:1023-33.

Reifenberger E, Boles E, Ciriacy M. Kinetic characterization of individual hexose transporters of *Saccharomyces cerevisiae* and their relation to the triggering mechanisms of glucose repression. *FEBS J* 1997;**245**:324-33.

Richard P, Putkonen M, Väänänen R *et al.* The missing link in the fungal L-arabinose catabolic pathway, identification of the L-xylulose reductase gene. *Biochem* 2002;**41**:6432-7.

Roberts RL, Mösch H-U, Fink GR. 14-3-3 proteins are essential for RAS/MAPK cascade signaling during pseudohyphal development in *S. cerevisiae*. *Cell* 1997;**89**:1055-65.

Sedlak M, Ho NW. Expression of *E. coli* araBAD operon encoding enzymes for metabolizing L-arabinose in *Saccharomyces cerevisiae*. *Enzyme Microb Technol* 2001;**28**:16-24.

Solis-Escalante D, Kuijpers NG, Barrajon-Simancas N *et al.* A minimal set of glycolytic genes reveals strong redundancies in *Saccharomyces cerevisiae* central metabolism. *Eukaryot Cell* 2015; **14**:804-16.

Solis-Escalante D, Kuijpers NG, Bongaerts N *et al. amdSYM*, a new dominant recyclable marker cassette for *Saccharomyces cerevisiae*. *FEMS Yeast Res* 2013;**13**:126-39.

Subtil T, Boles E. Improving L-arabinose utilization of pentose fermenting *Saccharomyces cerevisiae* cells by heterologous expression of L-arabinose transporting sugar transporters. *Biotechnol Biofuels* 2011;**4**:38.

Subtil T, Boles E. Competition between pentoses and glucose during uptake and catabolism in recombinant *Saccharomyces cerevisiae*. *Biotechnol Biofuels* 2012;**5**.

Sun L, Zeng X, Yan C *et al.* Crystal structure of a bacterial homologue of glucose transporters GLUT1-4. *Nature* 2012;**490**:361.

Tani T, Taguchi H, Fujimori KE *et al.* Isolation and characterization of xylitol-assimilating mutants of recombinant *Saccharomyces cerevisiae*. *J Biosci Bioeng* 2016;**122**:446-55.

Thorvaldsdóttir H, Robinson JT, Mesirov JP. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief Bioinform* 2013;**14**:178-92.

Torchia T, Hamilton R, Cano C *et al.* Disruption of regulatory gene *GAL80* in *Saccharomyces cerevisiae*: effects on carbon-controlled regulation of the galactose/melibiose pathway genes. *Mol Cell Biol* 1984;**4**:1521-7.

Träff K, Cordero RO, Van Zyl W *et al.* Deletion of the *GRE3* aldose reductase gene and its influence on xylose metabolism in recombinant strains of *Saccharomyces cerevisiae* expressing the *xylA* and *XKS1* genes. *Appl Environ Microbiol* 2001;**67**:5668-74.

Tsai C-S, Kong II, Lesmana A *et al.* Rapid and marker-free refactoring of xylose-fermenting yeast strains with Cas9/CRISPR. *Biotechnol Bioeng* 2015;**112**:2406-11.

Van Hemert M, van Heusden G, Steensma H. Yeast 14-3-3 proteins. *Yeast* 2001;**18**:889-95.

van Heusden GPH, Wenzel TJ, Lagendijk EL *et al.* Characterization of the yeast *BMH1* gene encoding a putative protein homologous to mammalian protein kinase II activators and protein kinase C inhibitors. *FEBS Lett* 1992;**302**:145-50.

Van Maris AJ, Abbott DA, Bellissimi E *et al.* Alcoholic fermentation of carbon sources in biomass hydrolysates by *Saccharomyces cerevisiae*: current status. *Antonie Van Leeuwenhoek* 2006;**90**: 391-418.

Verduyn C, Postma E, Scheffers WA *et al.* Physiology of *Saccharomyces cerevisiae* in anaerobic glucose-limited chemostat cultures. *Microbiology* 1990;**136**:395-403.

Verduyn C, Postma E, Scheffers WA *et al.* Effect of benzoic acid on metabolic fluxes in yeasts: A continuous-culture study on the regulation of respiration and alcoholic fermentation. *Yeast* 1992;**8**:501-17.

Verho R, Penttilä M, Richard P. Cloning of two genes (*LAT1, 2*) encoding specific L-arabinose transporters of the L-arabinose fermenting yeast *Ambrosiozyma monospora*. *Appl Biochem Biotechnol* 2011;**164**:604-11.

Verhoeven MD, Lee M, Kamoen L *et al.* Mutations in *PMR1* stimulate xylose isomerase activity and anaerobic growth on xylose of engineered *Saccharomyces cerevisiae* by influencing manganese homeostasis. *Sci Rep* 2017;**7**.

Walker BJ, Abeel T, Shea T *et al.* Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One* 2014;**9**:e112963.

Wang C, Li Y, Qiu C *et al.* Identification of important amino acids in Gal2p for improving the L -arabinose transport and metabolism in *Saccharomyces cerevisiae*. *Front Microbiol* 2017;**8**.

Wang C, Shen Y, Zhang Y *et al.* Improvement of L-arabinose fermentation by modifying the metabolic pathway and transport in *Saccharomyces cerevisiae*. *Biomed Res Int* 2013;**2013**.

Wang C, Skinner C, Easlon E *et al.* Deleting the 14-3-3 protein Bmh1 extends life span in *Saccharomyces cerevisiae* by increasing stress response. *Genetics* 2009;**183**:1373-84.

Weusthuis RA, Pronk JT, Van Den Broek P *et al.* Chemostat cultivation as a tool for studies on sugar transport in yeasts. *Microbiol Rev* 1994;**58**:616-30.

Wiedemann B, Boles E. Codon-optimized bacterial genes improve L-arabinose fermentation in recombinant *Saccharomyces cerevisiae*. *Appl Environ Microbiol* 2008;**74**:2043-50.

Wisselink HW, Cipollina C, Oud B *et al.* Metabolome, transcriptome and metabolic flux analysis of arabinose fermentation by engineered *Saccharomyces cerevisiae*. *Metab Eng* 2010;**12**:537-51.

Wisselink HW, Toirkens MJ, Del Rosario Franco Berriel M *et al.* Engineering of *Saccharomyces cerevisiae* for efficient anaerobic alcoholic fermentation of L-arabinose. *Appl Environ Microbiol* 2007;**73**.

Wisselink HW, Toirkens MJ, Wu Q *et al.* Novel evolutionary engineering approach for accelerated utilization of glucose, xylose, and arabinose mixtures by engineered *Saccharomyces cerevisiae* strains. *Appl Environ Microbiol* 2009;**75**.

Wisselink HW, Van Maris AJA, Pronk JT *et al.* Polypeptides with permease activity. US Patent 9034608 B2. 2015.

4.

Reassessment of requirements for anaerobic xylose fermentation by engineered, non-evolved *Saccharomyces cerevisiae* strains

Jasmine M. Bracher, Oscar A. Martinez-Rodriguez, Wijb J.C. Dekker, Maarten D. Verhoeven, Antonius J.A. van Maris and Jack T. Pronk

Essentially as published as in **FEMS Yeast Research** 2018; foy104
(https://doi.org/10.1093/femsyr/foy104)

Abstract

Expression of a heterologous xylose isomerase, deletion of the *GRE3* aldose-reductase gene and overexpression of genes encoding xylulokinase (*XKS1*) and non-oxidative pentose-phosphate-pathway enzymes (*RKI1*, *RPE1*, *TAL1*, *TKL1*) enables aerobic growth of *Saccharomyces cerevisiae* on D-xylose. However, literature reports differ on whether anaerobic growth on D-xylose requires additional mutations. Here, CRISPR-Cas9-assisted reconstruction and physiological analysis confirmed an early report that this basic set of genetic modifications suffices to enable anaerobic growth on D-xylose in the CEN.PK genetic background. Strains that additionally carried overexpression cassettes for the transaldolase and transketolase paralogs *NQM1* and *TKL2* only exhibited anaerobic growth on D-xylose after a 7–10 day lag phase. This extended lag phase was eliminated by increasing inoculum concentrations from 0.02 to 0.2 g biomass L^{-1}. Alternatively, a long lag phase could be prevented by sparging low-inoculum-density bioreactor cultures with a CO_2/N_2-mixture, thus mimicking initial CO_2 concentrations in high-inoculum-density, nitrogen-sparged cultures, or by using L-aspartate instead of ammonium as nitrogen source. This study resolves apparent contradictions in the literature on the genetic interventions required for anaerobic growth of CEN.PK-derived strains on D-xylose. Additionally, it indicates the potential relevance of CO_2 availability and anaplerotic carboxylation reactions for anaerobic growth of engineered *S. cerevisiae* strains on D-xylose.

Introduction

Over the past decades, major industrial and academic research efforts have been devoted to engineering *Saccharomyces cerevisiae*, a key microbial cell factory (Liu *et al.* 2013, Nielsen *et al.* 2013), for efficient conversion of lignocellulosic feedstocks into fuel ethanol, the largest-volume product of industrial biotechnology (reviewed by Alper and Stephanopoulos 2009, Young *et al.* 2010, Moysés *et al.* 2016, Jansen *et al.* 2017). A large part of this effort was geared towards enabling *S. cerevisiae* strains to ferment D-xylose and L-arabinose, two pentoses that comprise a substantial fraction of the potentially fermentable sugars in lignocellulosic feedstocks (Lynd 1996, Olsson and Hahn-Hägerdal 1996, Moysés *et al.* 2016, Jansen *et al.* 2017). In lignocellulosic feedstocks, D-xylose is typically the most abundant sugar after glucose, often accounting for 10-25% of the carbohydrate content (Lynd 1996). Wild-type *S. cerevisiae* strains can grow, albeit slowly, on the D-xylose isomer D-xylulose, whose metabolism is linked to glycolysis via the combined action of xylulokinase (Xks1) and the enzymes of the non-oxidative pentose-phosphate pathway (PPP) (Wang and Schneider 1980, Hsiao *et al.* 1982).

In naturally D-xylose-metabolizing, non-*Saccharomyces* yeasts such as *Scheffersomyces stipitis* (Toivola *et al.* 1984), *Candida shehatae* (Toivola *et al.* 1984), *Pachysolen tannophilus* (Smiley and Bolen 1982, Toivola *et al.* 1984), *Hansenula polymorpha* (Ryabova *et al.* 2003) and *Kluyveromyces lactis* (Margaritis and Bajpai 1982), D-xylose is first converted into D-xylulose by the combined action of pyridine-nucleotide-dependent xylose reductases (XR) and xylitol dehydrogenases (XDH). Metabolic engineering strategies for enabling anaerobic fermentation of D-xylose by *S. cerevisiae* based on expression of heterologous XR/XDH genes, which continue to be intensively investigated, are complicated by the different redox cofactor preferences of these two oxido-reductases (Kötter and Ciriacy 1993, Hahn-Hägerdal *et al.* 2001, Jeffries 2006).

An alternative metabolic engineering strategy is based on redox-cofactor-independent isomerization of D-xylose to D-xylulose by a heterologously expressed xylose isomerase (XI). Implementation of this strategy was long hindered by insufficient functional expression of heterologous XI genes, under physiologically relevant conditions, in *S. cerevisiae* (Walfridsson *et al.* 1996). This situation changed when Kuyper *et al.* (2003) demonstrated functional expression of the *xylA* gene of the anaerobic fungus *Piromyces* sp. E2 in *S. cerevisiae* (Kuyper *et al.* 2003). Multi-copy overexpression of *xylA*, combined with overexpression of the native xylulokinase gene (*XKS1*) and the major paralogs encoding enzymes of the non-oxidative pentose-phosphate pathway (*RKI1*, *RPE1*, *TAL1*, *TKL1*), enabled fast aerobic growth of *S. cerevisiae* on D-xylose (Kuyper *et al.* 2005). This result has since been reproduced, both with *xylA* and with other heterologous XI genes, and in different *S. cerevisiae* genetic backgrounds (**Table 1**). In many studies, this metabolic engineering strategy was combined

with deletion of *GRE3*, which encodes a non-specific aldose reductase whose activity could lead to loss of carbon to xylitol (Träff *et al.* 2001, Kuyper *et al.* 2003, Kuyper *et al.* 2005, Lee *et al.* 2012). Moreover, accumulation of xylitol, a known inhibitor of XI enzymes (Yamanaka 1969), might inhibit heterologously expressed XI activity.

The literature is entirely consistent where it concerns aerobic growth on D-xylose of XI-based *S. cerevisiae* strains that carry the abovementioned genetic modifications. Conversely, reports on the ability of such strains to grow anaerobically on D-xylose appear to be conflicting. In their original paper, Kuyper *et al.* (2005) reported that *S. cerevisiae* RWB217, which was constructed in the CEN.PK genetic background (Entian and Kötter 2007) and harboured the complete set of genetic modifications described above as well as a *gre3Δ* mutation, grew anaerobically on D-xylose in synthetic media without a need for additional mutagenesis or laboratory evolution (Kuyper *et al.* 2005). However, when other groups constructed similar *S. cerevisiae* strains, anaerobic growth on D-xylose was reported to require additional laboratory evolution, mutagenesis and/or genetic engineering (**Table 1**). Until recently, the different reported anaerobic growth performances of engineered *S. cerevisiae* strains might have been attributed to differences in strain background and/or XI genes (**Table 1**).

However, the recent single-step, CRISPR-Cas9-mediated construction of a D-xylose-metabolizing strain in the CEN.PK background, based on *Piromyces xylA*, yielded a strain that showed instantaneous, fast aerobic growth on D-xylose but reproducibly required a 12day adaptation period before growth on D-xylose in anaerobic bioreactor cultures set in (Verhoeven *et al.* 2017). Whole-genome resequencing showed that this adaptation reflected selection for mutants carrying mutations in *PMR1*, which encodes a Golgi Mn^{2+}/Ca^{2+} ATPase. These mutations increased cellular contents of Mn^{2+}, the preferred metal cofactor of XylA (Lee *et al.* 2017, Verhoeven *et al.* 2017). Since Verhoeven *et al.* (2017) used the same *S. cerevisiae* genetic background and XI gene as Kuyper *et al.* (2005), and, moreover, also overexpressed *XKS1* and PPP genes while deleting *GRE3* (**Table 2**), their study raised questions on the genetic requirements for anaerobic D-xylose metabolism by CEN.PK-derived strains (Kuyper *et al.* 2005, Verhoeven *et al.* 2017).

The goal of the present study was to critically re-examine and reproduce the genetic modifications required for anaerobic growth of *S. cerevisiae* on D-xylose reported by Kuyper *et al.* (2005) and Verhoeven *et al.* (2017). To this end, we constructed new engineered, D-xylose-fermenting strains to investigate the impact of subtle differences in strain construction strategies applied in the two studies. Subsequently, we compared growth of the resulting strains on D-xylose in anaerobic bioreactors, with special attention to the impact of inoculum density, initial carbon dioxide concentration and nitrogen source on anaerobic growth.

Table 1. Literature data on anaerobic growth of metabolically engineered, xylose-isomerase-based *S. cerevisiae* strains. The table summarizes sets of targeted genetic modifications, aerobic specific growth rates on D-xylose and any additional optimization by laboratory evolution or mutagenesis required for anaerobic growth on D-xylose. NIA = no information available; * specific growth rate estimated from exponential increase of CO_2 concentration in bioreactor off gas.

Strain back-ground	Strain	XI gene	Native genes overexpressed	Other targeted modifications	Aerobic growth rate (h^{-1})	Anaerobic growth	Reference
CEN.PK	RWB202	*Piromyces XylA* (2 micron plasmid)	none	No	0.005	After extensive aerobic, oxygen-limited and anaerobic selection ($\mu = 0.03\ h^{-1}$)	(Kuyper *et al.* 2003)
CEN.PK	RWB217	*Piromyces XylA* (2 micron plasmid)	*RKI1, RPE1, TAL1, TKL1 (all pTPI1), XKS1 (pADH1)*	*gre3Δ*	0.22	Anaerobic growth after ca. 35 h when inoculated at low cell densities (0.02 g biomass L^{-1}, $\mu = 0.09\ h^{-1}$). Immediate anaerobic growth when inoculated at high biomass concentration (0.2 g biomass L^{-1}, $\mu = 0.09\ h^{-1}$)	(Kuyper *et al.* 2005), This study
CEN.PK	YEp-opt. XI-Clos-K	*Clostridium phytofermentans, codon optimized*	*XKS1, RKI1, RPE1, TAL1, TKL1*	*GAL2* over-expression	0.057	No	(Brat *et al.* 2009)
CEN.PK	YEp-opt. XI-Piro	*Piromyces XylA, codon-optimized*	*XKS1, RKI1, RPE1, TAL1, TKL1*	*GAL2* over-expression	0.056	No	(Brat *et al.* 2009)
BarraGrande	BWY10Xyl	*Clostridium phytofermentans, codon optimized*	NIA	NIA	0.04	Serial aerobic shake flask cultures (6) on D-xylose until anaerobic xylose consumption observed upon aerobic biomass production	(Brat *et al.* 2009)
CEN.PK	IMX696	*Piromyces XylA, codon-optimized*	*XKS1, RKI1, RPE1, TAL1, TKL1, NQM1, TKL2*	*gre3Δ*	0.21	After 12-day anaerobic adaptation phase (mutations in *PMR1*)	(Verhoeven *et al.* 2017)
Ethanol Red (diploid)	HDY.GUF5	*Clostridium phytofermentans*, codon-optimized	*XKS1, RKI1, RPE1, TAL1, TKL1, TKL2, NQM1*	*HXT7*, *S. stipitis* AraA, *B. licheniformis* AraA, *E. coli* AraD, AraB	NIA	After extensive mutagenesis-, genome-shuffling, and oxygen-limited selection experiments (no specific growth rates reported)	(Demeke *et al.* 2013)

Table 1 continues on next page.

Table 1 – *Continued*

Strain back-ground	Strain	XI gene	Native genes overexpressed	Other targeted modifications	Aerobic growth rate (h^{-1})	Anaerobic growth	Reference
Natural isolate YB-210/ GLBRCY0	Y22-3 (haploid spore)	*Clostridium phytofermentans XylA (ScTDH3p)*	-	*S. cerevisiae TAL1, S. stipitis XYL3*	NIA	No initial anaerobic growth observed. Aerobic selection in glucose-xylose media (34 transfers), anaerobic selection on same media (14 transfers). Evolved strain showed anaerobic growth in YPX medium (mutation in *GRE3* obtained).	(Parreiras *et al.* 2014)
BF264-15Dau (Sun *et al.* 1989)	H131-A3	*Piromyces XylA, codon-optimized* (2 micron plasmid)	*RPE1, RKI1, TKL1*	*P. stipitis XKS1& TAL1*	0.031 ± 0.022	Aerobic cultivation in anaerobic sequential batch reactors (SBRs) on SMX (2% xylose) (ca. 70 transfers), transfer to microaerobic SBRs (YNBX, 60 transfers), transfer to anaerobic SBRs (YNBX, 60 transfers), transfer to anaerobic chemostat with increasing dilution rate over time for ca. 60 generation (YNBX, d=0.02 h^{-1} to 0.12 h^{-1}), 20 more generations with YNBX with 10% xylose until dilution rate of 0.148 h^{-1}.	(Zhou *et al.* 2012)
CEN.PK	TMB3361	*Piromyces XylA* (2 micron plasmid)	*TAL1, TKL1, RPE1, RKI1, XKS1 (all pPGK1)*	*E. coli* XK (*xylB*), *gre3Δ*	0.089 ± 0.002	Anaerobic fermentations inoculated with very high cell densities (5 g L^{-1} CDW) resulting in partial conversion of the supplied xylose to ethanol (without an adaptation time) but without measurable growth (due to high initial cell densities).	(Parachin *et al.* 2011)
CEN.PK	YRH631 (naive), YRH1114 (evolved)	*Prevotella ruminicola TC2-24 XI (codon-optimized)*	*XKS1*	No	0.06 (naive) 0.23 (evolved)	Six transfers in microaerobic conditions (μ unknown).	(Hector *et al.* 2013)

Strain back-ground	Strain	XI gene	Native genes overexpressed	Other targeted modifications	Aerobic growth rate (h^{-1})	Anaerobic growth	Reference
INVSc1 (Invitrogen, USA)	INVSc1/ pRS406XKS/ pILSUT1/ pWOXYLA (XKS, Sut1, XylA)	*Orpinomyces xylA (2µm plasmid)*	*XKS1*	*P. stipitis SUT1 over-expression*	NIA	CO_2-flushed bottles inoculated with 5 g biomass L^{-1} showed consumption of 15.5 g L^{-1} xylose from a total of 50 g L^{-1} within 140 h.	(Madhavan *et al.* 2009)
CEN.PK	TMB3066	*Piromyces XylA (2µm plasmid)*	*TAL1, TKL1, RPE1, RKI1, XKS1 (all pPGK1*	*gre3Δ*	0.02	Anaerobic cultures resulting in partial conversion of the supplied xylose to ethanol (16.8 g of 50 g L^{-1} within 100 h, without an adaptation time) at high biomass density, no anaerobic growth reported.	(Karhumaa *et al.* 2007)
CEN.PK	IMU078	*Piromyces XylA (2µm plasmid)*	*TAL1, TKL1, RPE1, RKI1, NQM1, TKL2, XKS1*	*gre3Δ*	NIA	Anaerobic growth after ca. 7-8 d when inoculated at low biomass concentration (0.02 g biomass L^{-1}), µ = ca. 0.09 h^{-1}*. Immediate anaerobic growth when (i) inoculated at high biomass density (0.2 g biomass L^{-1}, µ = 0.05 h^{-1}*), (ii) upon supplementation with 0.1% CO_2 in N_2 used from sparging of bioreactors (µ = 0.05 h^{-1}), or (iii) when L-aspartate is supplied as nitrogen source (µ = 0.05 h^{-1}).	This study
CEN.PK	IMU079	*Piromyces XylA (2µm plasmid)*	*TAL1, TKL1, RPE1, RKI1, XKS1*	*gre3Δ*	NIA	Anaerobic growth after ca. 40 h when inoculated at low cell densities (0.02 g biomass L^{-1}, µ = 0.08 h^{-1}). Immediate anaerobic growth when inoculated at high biomass concentration (0.2 g biomass L^{-1}, µ = 0.07 h^{-1}*).	This study

Table 1 continues on next page.

Table 1 – *Continued*

Strain back-ground	Strain	XI gene	Native genes overexpressed	Other targeted modifications	Aerobic growth rate (h^{-1})	Anaerobic growth	Reference
PE-2	LVY27 LVY34.4 (evolved) LVY41.5 (evolved)	*Orpinomyces sp. xylA* (codon-optimized) (flanked with δ LTR sequences for high copy integration)	XKS1*2, *TAL1, RKI1, TKL1, RPE1*	*gre3Δ*	Evolved: μ = 0.23 and 0.129 h^{-1}	Very slow aerobic growth with 1 copy of *XylA*. No anaerobic growth upon integration of 1 copy of *xylA*. Selection in semi-anaerobic conditions with 5 g L^{-1} glucose and 40 g L^{-1} xylose (12 transfers). Faster growth upon selection for increased *xylA* copy numbers, resulting in 36 and 26 copies.	(dos Santos *et al.* 2016)
BY4741	BY4741-S2A3K	Mutated *Piromyces xylA3*(2μm plasmid)*	*XKS1*	*gre3Δ* , *S. stipitis TAL1*over-expression	0.061 h^{-1}	Xylose fermentation possible in high-cell density, micro-aerobic conditions (no growth rates available).	(Lee *et al.* 2012)
BY4741	SXA-R2P	Mutated *Piromyces xylA3**(2 copies) (Lee *et al.* 2012)	*XKS1*	*gre3Δ* , *pho13 Δ, S. stipitis TAL1*over-expression (2 copies)	0.105 h^{-1} and 0.128 h^{-1} (evolved)	Naïve strain slowly consumed xylose in microaerobic conditions. Adaptive evolution in closed falcon tubes with media containing 20 g L^{-1} xylose (12 transfers). Evolved strain was capable of fast xylose consumption when inoculated at high biomass concentration in non-purged anaerobic bioreactors where initial oxygen was consumed within 12h (no growth rates available).	(Lee *et al.* 2014)
CEN.PK	BSPC095	*Piromyces xylA (2μm plasmid)*	*TAL1, TKL1, RPE1, RKI1, XKS1*	*gre3Δ* , *cox4Δ* ,	No initial aerobic growth	Weak aerobic growth observed in liquid xylose medium upon 10 days of aerobic incubation. Serial transfers of aerobic cultures with xylose during 1000 h resulted in aerobic growth rate of , μ = 0.11 h^{-1}.	(Shen *et al.* 2012)

Materials and Methods

Yeast strains and maintenance

All *S. cerevisiae* strains used and constructed in this study belong to the CEN.PK lineage (Entian and Kötter 2007, Nijkamp *et al.* 2012, Salazar *et al.* 2017). Yeast cultures were grown on synthetic medium (SM, prepared and sterilized as described previously (Verduyn *et al.* 1992)) or YP medium (10 g L^{-1} Bacto yeast extract (Becton Dickinson, Sparks, MD), 20 g L^{-1} Bacto peptone (Becton Dickinson); autoclaved at 121 °C for 20 min. SM was supplemented with 1 mL L^{-1} filter-sterilized vitamin solution (Verduyn *et al.* 1992). Concentrated sterile D-glucose or D-xylose solutions (autoclaved at 110 °C for 30 min) were added to SM and YP media to a concentration of 20 g L^{-1}, resulting in SMD or SMX and YPD or YPX, respectively. Shake-flask cultures were grown in 500-mL round-bottom flasks containing 100 mL medium and incubated in an Innova incubator (Brunswick Scientific, Edison, NJ) at 30 °C and 200 rpm. Solid media contained 2% (w/v) Bacto agar (Becton Dickinson). *Escherichia coli* strains were grown in LB-ampicillin medium (10 g L^{-1} Bacto tryptone, 5 g L^{-1} Bacto yeast extract, 5 g L^{-1} NaCl, 100 mg L^{-1} ampicillin). For storage, sterile glycerol was added to stationary-phase cultures of yeast and *E. coli* to a final concentration of 30% (v/v), after which aliquots were stored at -80 °C.

Molecular biology

Analytical PCR was performed with Dreamtaq polymerase (Thermo Scientific, Waltham, MA) according to the manufacturer's instructions. DNA fragments for cloning were amplified with Phusion Hot Start II High Fidelity Polymerase (Thermo Scientific, Waltham, MA) and with the desalted or PAGE-purified oligonucleotide primers (Sigma-Aldrich, St. Louis, MO) listed in **Additional Table 1** which can be found at https://doi.org/10.1093/femsyr/foy104. DNA fragments were separated by gel electrophoresis (90 V, 35 min) on 1% (w/v) agarose gels (Thermo Scientific) buffered with 1x TAE (Thermo Scientific). When required, DNA fragments were excised from gels and purified with a Zymoclean Gel DNA Recovery Kit (Zymo Research, Irvine, CA) or, when no unspecific products were detected, directly purified from the PCR mix with a GenElute PCR Clean-Up Kit (Sigma-Aldrich). Yeast genomic DNA was extracted with a YeaStar Genomic DNA kit (Zymo Research) and DNA for diagnostic colony PCR was extracted by boiling cells picked from colonies in 10 µl 0.2 N NaOH for 5 min. After removal of debris by centrifugation (1 min at 2000 x *g*), 2 µl of the supernatant was used as template in a 20 µl PCR reaction (Dreamtaq).

Yeast and *E. coli* plasmids were extracted with a Zymoprep Yeast Plasmid Miniprep II kit (Zymo Research, Irvine, CA) and with a Sigma GenElute Plasmid kit (Sigma-Aldrich), respectively. Yeast transformation was carried out with the lithium-acetate method

(Gietz and Woods 2002). When *natNT2* was used as a marker gene, the transformation procedure included an overnight incubation step in SMG prior to selection on solid SMG plates containing 1 g L^{-1} glutamic acid as the sole nitrogen source, which were supplemented with 100 mg L^{-1} nourseothricin (pH 6). Single-colony isolates were obtained by three consecutive re-streaks on solid selective medium and genotype confirmation by analytical PCR. *E. coli* DH5α was used for chemical transformation (Inoue *et al.* 1990) or electroporation in 2 mm cuvettes (165 – 2086, Bio-Rad, Hercules, CA) using a Gene Pulser Xcell Electroporation system (Biorad). Isolated plasmids were routinely checked by analytical PCR and restriction analysis.

Strain and plasmid construction

Detailed genotypes of all strains and plasmids used or generated in this study are listed in **Table 2** and **3**, respectively. Genomic DNA of *S. cerevisiae* CEN.PK113-7D was used as a template for amplification of *S. cerevisiae* gene-, promoter-, and terminator fragments. The CEN.PK strain lineage and construction of the derived *S. cerevisiae* strains RWB217, IMX581, IMX994, and IMU079 are described elsewhere (Kuyper *et al.* 2005, Entian and Kötter 2007, Mans *et al.* 2015, Papapetridis *et al.* 2018).

Integration of overexpression cassettes for *NQM1* and *TKL2* (pPGI1-*NQM1*-TagB-pPYK1-*TKL2*) was achieved by targeted integration into IMX994 using 60 bp flanks homologous to *SGA1*. Fragments were obtained by amplification of the SGA1flank-pPGI1-*NQM1*-TagB-cassette from pUD344 using primers 11357 & 3276 and amplification of the TagB-pPYK1-*TKL2*-SGA1flank-cassette from pUD346 using primer pair 11356 and 7607. Both fragments were co-transformed with pUDR103, a plasmid expressing a gRNA targeting *SGA1,* into IMX994. pUDR103 was subsequently counter selected by non-selective growth in YPD medium followed by plating on solid SMG plates supplemented with uracil and 5-fluoroorotic acid (0.15 and 1 g L^{-1} final concentration, respectively), resulting in strain IMX1456 (IMX994 *sga1::NQM1, TKL2*). This strain was subsequently transformed with the *xylA* plasmid pAKX002, resulting in strain IMU081.

Construction of strain IMX800 (IMX581 *gre3::RPE1, TKL1, NQM1, RKI1, TKL2, XKS1)* was similar to the construction of strain IMX994 (IMX581 *gre3::RPE1, TKL1, TAL1, RKI1, XKS1)* (Papapetridis *et al.* 2018) with the difference that, in the former strain, the chromosomally integrated cluster of overexpression cassettes for pentose-phosphate-pathway genes included *NQM1* and *TKL2.* The expression cassette for *TAL1* was amplified from pUD349 with primer pair 3274 & 3275, yielding the tag-flanked expression cassette TagI-pTEF1-*TAL1*-TagA; *NQM1* was amplified from pUD344 with primers 3847 & 3276 to yield the expression cassette TagA-pPGI1-*NQM1*-TagB; the TagB-pTPI1-*RKI1*-TagC cassette was amplified from pUD345 with primers 4672 & 3277; the TagC-pPYK1-*TKL2*-TagL

Table 2. *Saccharomyces cerevisiae* strains used in this study.

Strain	Relevant genotype	Description	Reference
CEN.PK113-7D	*MATa*	Reference strain	(Entian and Kötter 2007)
CEN.PK113-5D	*MATa ura3-52*	Uracil auxotrophic reference strain	(Entian and Kötter 2007)
CEN.PK102-3A	*MATA ura3-52 leu2-112*	Uracil and leucine auxotrophic strain	(van Dijken *et al.* 2000)
RWB217	CEN.PK102-3A *loxP-pTPI::(–266, –1)TAL1 gre3Δ::hphMX pUGpTPI-TKL1 pUGpTPI–RPE1 KanloxP–pTPI::(–40, –1)RKI1 {pAKX002, p415ADHXKS}*	Metabolically engineered, non-evolved xylose consuming strain	(Kuyper *et al.* 2005)
IMX975	RWB217 *can1::CAS9-tagA-loxP-natNT2-loxP*	RWB217 expressing Cas9	This study
IMX1366	IMX975, *sga1:: pPGI1-NQM-TagB-pPYK1-TKL2*	IMX975 over-expressing *NQM1* and *TKL2*	This study
IMX581	CEN.PK113-5D *can1::CAS9-tagA-loxP-natNT2-loxP*	CEN.PK113-5D expressing *S. pyogenes* Cas9	(Mans *et al.* 2015)
IMX696	IMX581 *gre3:: pTDH3-RPE1-pPGK1-TKL1 –pTEF1-TAL1-pPGI1-NQM1 – pTPI1-RKI1 –pPYK1-TKL2-(pTPI1-xylA-tCYC1)*36)-pTEF-XKS1,* pUDE335	IMX581 over-expressing PPP genes incl. *NQM1* and *TKL2,* expression of *Piromyces xylA* based on 36 integrated copies	(Verhoeven *et al.* 2017)
IMX994	IMX581 *gre3:: pTDH3-RPE1-tagH-pPGK1-TKL1-TagI-pTEF1-TAL1-TagA- pTPI1-RKI1-TagL-pTEF-XKS1*	IMX581 over-expressing genes from the non-oxidative pentose phosphate pathway	(Papapetridis *et al.* 2018)
IMU079	IMX994, pAKX002 (2µm *xylA*)	IMX994 over-expressing *xylA* from a 2µm plasmid	(Papapetridis *et al.* 2018)
IMX1456	IMX994*, sga1:: pPGI1-NQM-TagB-pPYK1-TKL2*	IMX994 over-expressing *NQM1* and *TKL2*	This study
IMU081	IMX1456, pAKX002 (2µm *xylA*)	IMX1456 over-expressing *xylA* from a 2µm plasmid	This study
IMX800	IMX581, *gre3:: pTDH3-RPE1-TagH-pPGK1-TKL1-TagI-pTEF1-TAL1-TagA-pPGI1-NQM1-TagB-pTPI1-RKI1-TagC-pPYK1-TKL2-TagL- pTEF-XKS1*	IMX581 over-expressing PPP genes incl. *NQM1* and *TKL2*	This study
IMU078	IMX800, pAKX002 (2µm *xylA*)	IMX800 over-expressing *xylA* from a 2µm plasmid	This study
IMX1736	IMU078, *pck1 Δ*	IMU078 with a deletion in *PCK1*	This study

cassette was amplified from pUD346 with primers 3283 & 8285, and the TagL-pTEF-*XKS1*-GRE3flank cassette was amplified from pUD353 with the primer pair 7222 & 7135. In overexpression cassettes, *S. cerevisiae* genes retained their endogenous terminators. Transformation of strain IMX800 with the high-copy-number *xylA* plasmid pAKX002, which was isolated from *S. cerevisiae* RWB217 (Kuyper *et al.* 2005), yielded strain IMU078. *PCK1*

Table 3. Plasmids used in this study.

Name	Relevant characteristics	Reference
pROS15	2µm ampR *natNT2* gRNA-*CAN1*.Y gRNA-*ADE2*.Y	(Mans *et al.* 2015)
pROS11	2µm ampR *amdSYM* gRNA-*CAN1*.Y gRNA-*ADE2*.Y	(Mans *et al.* 2015)
pROS13	2µm ampR *kanMX* gRNA-*CAN1*.Y gRNA-*ADE2*.Y	(Mans *et al.* 2015)
pAKX002	2µm, *URA3, pTPI1-XylA* (*Piromyces spp.* E2)	(Kuyper *et al.* 2003)
p414 -pTEF1-Cas9-tCYC1	*CEN6/ARS4* ampR *pTEF1-cas9-tCYC1*	(DiCarlo *et al.* 2013)
pMEL10	2µm, *KlURA3, pSNR52*-gRNA.*CAN1*.Y-*tSUP4*	(Mans *et al.* 2015)
pJET1.2Blunt	Multi-purpose cloning vector for storage of assembled cassettes	ThermoFisher
pUD344	pJET1.2Blunt TagA-*pPGI1-NQM1-tNQM1*-TagB	(Verhoeven *et al.* 2017)
pUD345	pJET1.2Blunt TagB-*pTPI1-RKI1-tRKI1*-TagC	(Verhoeven *et al.* 2017)
pUD346	pJET1.2Blunt TagC-*pPYK1-TKL2-tTKL2*-TagF	(Verhoeven *et al.* 2017)
pUD347	pJET1.2Blunt TagG-*pTDH3-RPE1-tRPE1*-TagH	(Verhoeven *et al.* 2017)
pUD348	pJET1.2Blunt TagH-*pPGK1-TKL1-tTKL1*-TagI	(Verhoeven *et al.* 2017)
pUD349	pJET1.2Blunt TagI-*pTEF1-TAL1-tTAL1*-TagA	(Verhoeven *et al.* 2017)
pUD350	pJET1.2Blunt *pTPI1-XylA-tCYC1*	(Verhoeven *et al.* 2017)
pUD353	pJET1.2Blunt *pTEF1-XKS1-tXKS1*	(Verhoeven *et al.* 2017)
pUDE335	2µm ori, KlURA3, pSNR52-gRNA.GRE3.Y-tSUP4	(Verhoeven *et al.* 2017)
pUDR119	2µm, *amdSYM, pSNR52*-gRNA.*SGA1*.Y-*tSUP4*	(van Rossum *et al.* 2016)
pUDR103	2µm, *KlURA3, pSNR52*-gRNA.*SGA1*.Y-*tSUP4*	(Papapetridis *et al.* 2017)

was deleted in strain IMU078 according to the protocol described by Mans *et al.* (2015), resulting in strain IMX1736 (Mans *et al.* 2015). Two plasmid fragments were transformed directly into IMU078 for *in vivo* assembly: a 2µm fragment amplified from pROS13 with a double-binding primer adding a gRNA flank guiding a cut in *PCK1* (primer 14234) and a backbone harbouring a *kanMX* marker amplified from pROS13 using the double-binding primer 6005. Plasmid fragments were transformed together with an annealed double-strand repair fragment (oligonucleotide 14235 & 14236) homologous to 60bp up-and downstream of *PCK1* to delete the ORF. Mutants were selected on solid YPD plates supplemented with 200 mg L^{-1} G418. Primer pair 14237 & 14238 was used to screen for positive colonies. A positive colony was grown in non-selective liquid media (SMG), plated on solid SMG, and subsequently replica-plated on solid SMG and YPD-G418 plates to check for successful plasmid loss. Cells from a colony that lost the G418-marker harbouring plasmid were double checked for a *pck1Δ* genotype and subsequently stocked as strain IMX1736.

To enable CRISPR-Cas9-based editing of RWB217, a *cas9*-cassette with a *can1*- and a TagA- overhang amplified from p414 (DiCarlo *et al.* 2013) with primers 2873 & 3093 and a *natNT2* cassette with the same tags, amplified from pROS15 (Mans *et al.* 2015) with primers 4653 & 5542 was integrated into the *CAN1* locus by homologous recombination, resulting in strain IMX975. Integration of *sga1::NQM1*, *TKL2* into IMX975, resulting in strain IMX1366 (IMX975 *sag1::NQM1, TKL2*) was done as described for strain IMU081 with the difference that the gRNA plasmid pUDR119 with an *amdSYM* marker was used and no pAKX002 plasmid had to be transformed due to its presence in IMX975. Selection and counter selection of the *amdSYM* marker cassette were performed as described previously (Solis-Escalante *et al.* 2013). *PMR1* genes, including promoters and terminators, were amplified from duplicate 200 h anaerobic bioreactor cultures of strain IMU078, with primer pair 8790 & 8791 and subsequently Sanger sequenced (BaseClear, Leiden, The Netherlands) with primers 8790, 8791, and 13459-13475.

Bioreactor cultivation

S. cerevisiae strains were physiologically characterized in anaerobic 2 L laboratory bioreactors (Applikon, Delft, The Netherlands) with a 1 L working volume. Bioreactors filled with synthetic medium (SM) were autoclaved at 121 °C for 20 min. Synthetic medium with L-aspartate as nitrogen source was prepared as described previously (Zelle *et al.* 2010) and filter sterilized (Nalgene Rapid-Flow, 0.2 µm, Thermo Scientific) prior to addition to autoclaved bioreactors. Separately prepared solutions were subsequently added to final concentrations of 20 g L^{-1} D-xylose, 1 mL L^{-1} vitamin solution, 0.2 g L^{-1} antifoam C, as well as 10 and 420 mg L^{-1} of the anaerobic growth factors ergosterol and Tween-80, respectively. Cultures were sparged with nitrogen (0.5 L min^{-1}) to maintain anaerobic conditions.

For physiological characterization of strain IMU078 with CO_2 supplementation, nitrogen as a sparging gas was replaced by an analytically certified (2% tolerance) mixture of 99.9% N_2 and 0.1% CO_2 (Linde Group, Munich, Germany). Oxygen diffusion into bioreactors was minimized by equipping them with Norprene tubing (Saint-Gobain, Courbevoie, France) and Viton O-rings (Eriks, Alkmaar, The Netherlands). Evaporation was minimized by cooling the outlet gas to 4 °C. Cultures were stirred at 800 rpm, maintained at pH 5 by automatic addition of 2 M KOH, and kept at 30 °C. Separate inocula were prepared for each bioreactor. These inocula were obtained by growth of the respective strain, starting with a frozen stock culture, in two consecutive aerobic 100 mL shake flask cultures on SMX, at 30 °C and 200 rpm. When the OD_{660} of the pre-cultures was between 3.5 and 5.5, they were used to inoculate the anaerobic bioreactors. The initial OD_{660} of bioreactor cultures was either 0.1 or 1 (0.02 or 0.2 g biomass L^{-1}, respectively) as indicated.

Analytical techniques

Biomass dry weight measurements and optical density measurements at a wavelength of 660 nm were performed as described previously (Verhoeven *et al.* 2017). A correlation between biomass dry weight concentration and optical density, based on at least 10 measurement points during the exponential growth phase, was used to estimate the biomass concentrations for the first few bioreactor culture samples, when the cell density was too low to allow for accurate biomass dry weight determinations ($OD_{660} < 0.9$). Specific growth rates were calculated from at least six biomass dry-weight measurements during the exponential growth phase. Concentrations of biomass, ethanol, glycerol, and D-xylose, at the same sampling times, were used to calculate yields (gram product per gram xylose) from the slopes of the linear D-xylose concentration decrease versus the concentrations of the respective metabolites or biomass dry weight as described previously (Papapetridis *et al.* 2016). CO_2 and O_2 concentrations in the exhaust gas of bioreactors and metabolite concentrations were measured as described previously (Verhoeven *et al.* 2017).

Results

Genetic requirements for anaerobic growth on xylose

S. cerevisiae RWB217 was the first XI-based engineered yeast strain that was reported to grow anaerobically on D-xylose without prior laboratory evolution or mutagenesis (Kuyper *et al.* 2005) (**Figure 1A**). Recently, anaerobic growth on D-xylose of a strain that was newly reconstructed in the same genetic background and contained a similar set of genetic modifications (strain IMX696 (Verhoeven *et al.* 2017), **Table 2**) was reported to require mutations in *PMR1*. In view of this apparent discrepancy and other literature reports on additional genetic requirements for anaerobic growth of D-xylose-metabolising strains in other genetic backgrounds (**Table 1**), we re-investigated the requirement for anaerobic growth of CEN.PK-based *S. cerevisiae* strains.

While, in strain RWB217 as well as in strain IMX696, *xylA* was expressed from the *TPI1* promoter, the *xylA* coding region was codon optimized in strain IMX696 (Verhoeven *et al.* 2017) but not in strain RWB217 (Kuyper *et al.* 2005). Moreover, in strain IMX696 the *xylA* expression cassettes were chromosomally integrated, resulting in ca. 36 copies per haploid genome (Verhoeven *et al.* 2017) whereas the *xylA* expression cassette in strain RWB217 was expressed from a multi-copy episomal vector (Kuyper *et al.* 2005). To investigate whether these differences affected anaerobic growth on D-xylose of these two strains, a new strain was constructed that carried the same set of genetic modifications as strain IMX696 but expressed *xylA* from the same episomal vector (pAKX002) as strain RWB217. Growth on D-xylose of the resulting strain, IMU078 (*gre3Δ, RPE1↑, RKI1↑,*

TAL1↑, *TKL1*↑, *NQM1*↑, *TKL2*↑, *XKS1*↑; pAKX002 (2μm *xylA*)), was investigated in anaerobic, nitrogen-sparged bioreactor cultures, exactly as described previously (Verhoeven *et al.* 2017). In these cultures, strain IMU078 showed a 7–8 day adaptation phase before anaerobic growth on D-xylose set in (**Figure 1B**).

Under the same conditions, strain RWB217 showed anaerobic growth on D-xylose within 32 ± 6 h after inoculation (**Figure 1A**), consistent with the original report by Kuyper *et al.* (2005) (Kuyper *et al.* 2005). These results indicate that codon usage and/or mode of *xylA* overexpression were not decisive factors in causing the presence and absence of a prolonged anaerobic adaptation phase in strains IMX696 (Verhoeven *et al.* 2017) and RWB217 (Kuyper *et al.* 2005), respectively. To check if the integrated Cas9 expression cassette that was present in strains IMX696 and IMU078 but not in RWB217 affected anaerobic growth on D-xylose, the Cas9 cassette was integrated at the *CAN1* locus of strain RWB217, yielding strain IMX975. This strain exhibited the same anaerobic growth profile as strain RWB217 (**Figure 2A**).

When, after twelve days, replicate anaerobic bioreactor cultures of strain IMU078 on D-xylose had reached stationary phase, their *PMR1* loci were PCR amplified and Sanger sequenced. In contrast to observations in similar experiments with strain IMX696 (Verhoeven *et al.* 2017), no mutations were found in promoter, terminator or coding region of *PMR1.* This result suggested that other mutations were responsible for the eventual anaerobic growth on D-xylose of strain IMU078.

Inoculum density and resulting initial CO_2 concentration affect anaerobic growth on D-xylose

In the original characterization of strain RWB217 in anaerobic batch cultures (Kuyper *et al.* 2005), an inoculum concentration of 0.2 g biomass L^{-1} was used. The experiments discussed above and the previous characterization of strain IMX696 (Verhoeven *et al.* 2017), were inoculated with a 10-fold lower inoculum density. Use of an 0.2 g L^{-1} inoculum density completely abolished the 7–8 day adaptation phase of strain IMU078 (*gre3Δ, RPE1*↑, *RKI1*↑, *TAL1*↑, *TKL1*↑, *NQM1*↑, *TKL2*↑, *XKS1*↑; pAKX002 (2μm *xylA*)) in anaerobic, D-xylose-grown batch cultures (**Figure 1C**).

The anaerobic bioreactor cultures in this study were sparged with N_2 (0.5 L min^{-1}). As was to be expected, inoculum density positively correlated with initial CO_2 concentrations in the off gas of the cultures and therefore, by inference, with concentrations of CO_2 in the culture broth (**Figure 1A-F**). In cultures of strain IMU078 inoculated at 0.02 g biomass L^{-1}, the initial CO_2 content in the off gas was 0.004 ± 0.001%, as compared to 0.04 ± 0.005% in cultures started at a ten-fold higher biomass concentration. To investigate whether the

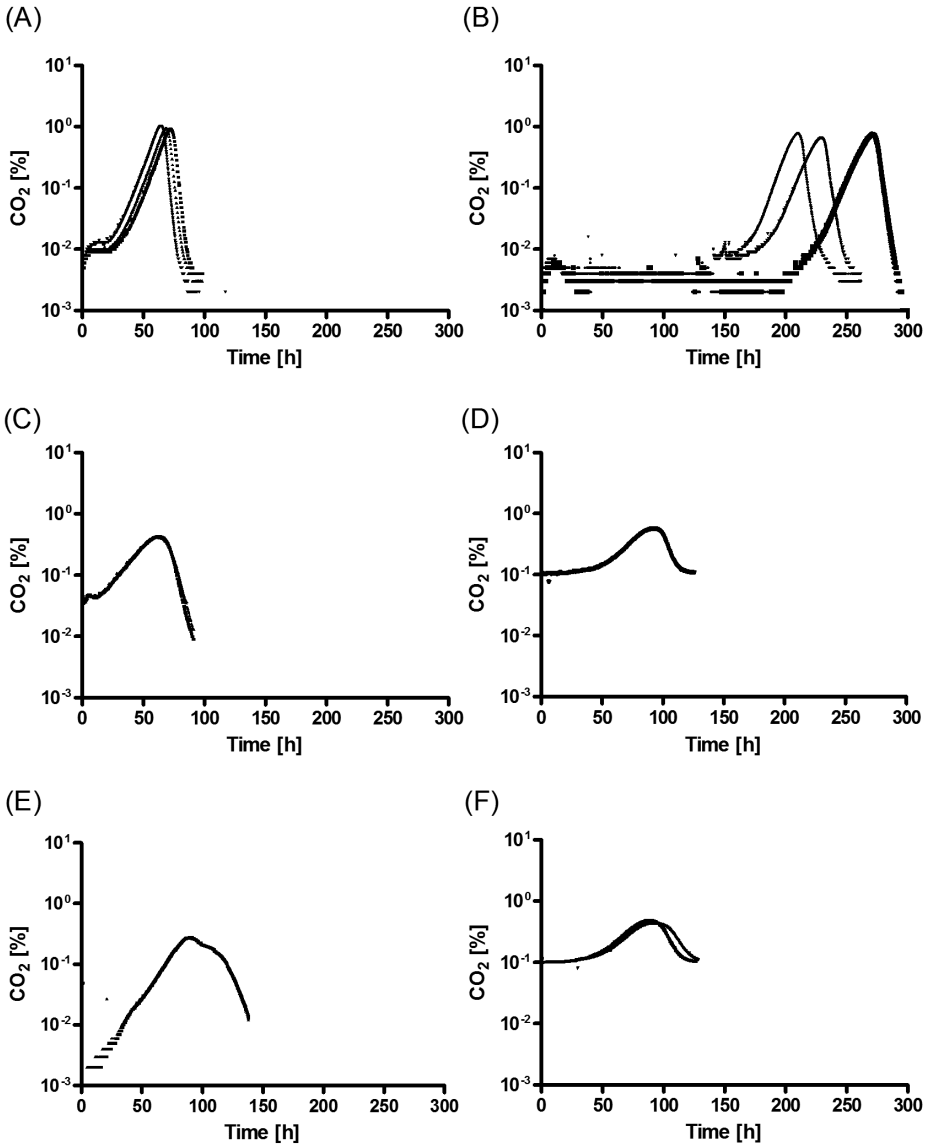


Figure 1. Fermentation profiles, indicated as percentage of CO_2 in off gas over time, of metabolically engineered, non-evolved D-xylose-metabolizing *S. cerevisiae* strains grown in anaerobic bioreactor batch cultures on synthetic medium supplemented with 20 g L^{-1} D-xylose. Unless indicated otherwise, cultures were inoculated at a biomass concentration of 0.02 g L^{-1} and sparged with 0.5 vvm N_2. (A) Strain RWB217 (*gre3Δ, RPE1*↑*, RKI1*↑*, TAL1*↑*, TKL1*↑, *XKS1*↑, *xylA* (pAKX002)) (Kuyper et al. 2005). (B) Strain IMU078 (*gre3Δ, RPE1*↑*, RKI1*↑*, TAL1*↑*, TKL1*↑*, NQM1*↑*, TKL2*↑, *XKS1*↑, *xylA* (pAKX002)). (C) Strain IMU078 inoculated at 0.2 g biomass L^{-1}. (D) Strain IMU078 inoculated at 0.02 g biomass L^{-1} and sparged with a mixture of 99.9% N_2 and 0.1% CO_2 at 0.5 vvm. (E) Strain IMU078 inoculated at 0.02 g biomass L^{-1} in media containing L--aspartate as nitrogen source instead of ammonium sulfate. Sampling of this culture for metabolite analyses (**Figure 2C**) affected the CO_2 profile. (F) Strain IMX1736 (IMU078 *pck1*Δ) inoculated at 0.02 g biomass L^{-1} and sparged with a mixture of 99.9% N_2 and 0.1% CO_2 at 0.5 vvm. The panels show results from 4 (panel B, D), 3 (panel A), and 2 (panel C, E, F) independent experiments, respectively. Detailed information on strain genotypes is provided in **Table 2**.

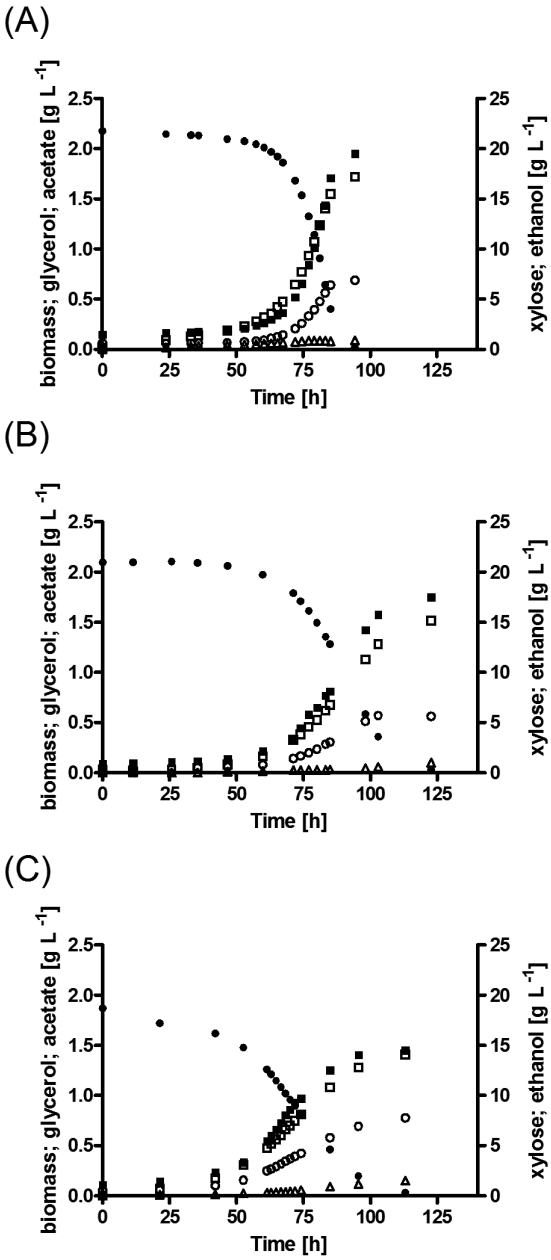


Figure 2. Growth and product formation in anaerobic bioreactor cultures of metabolically engineered, non-evolved, D-xylose-metabolizing *S. cerevisiae* strains. Cultures were inoculated at a biomass concentration of 0.02 g L^{-1} and, unless otherwise stated, were sparged with 0.5 vvm N_2. (A) Strain IMX975 (RWB217 (*gre3Δ, RPE1*↑, *RKI1*↑, *TAL1*↑, *TKL1*↑, *XKS1*↑, *xylA* (pAKX002), *can1::Cas9*)). (B) Strain IMU078 (*gre3Δ, RPE1*↑, *RKI1*↑, *TAL1*↑, *TKL1*↑, *NQM1*↑, *TKL2*↑, *XKS1*↑, *xylA* (pAKX002)) sparged with 0.5 vvm of a mixture of 0.1% CO_2 and 99.9% N_2. (C) Strain IMU078 grown in media containing L-aspartate as nitrogen source instead of ammonium sulfate. • = D-Xylose; ▪ = Biomass; □ = Glycerol; ○ = Ethanol; Δ = Acetate. The panels show data from single representative cultures from a set of two independent duplicate cultures for each strain. Data from replicate cultures are shown in **Additional Figure 1**.

initial CO_2 concentration influenced the onset of anaerobic growth on D-xylose, anaerobic bioreactor cultures of strain IMU078 (*gre3Δ, RPE1*↑*, RKI1*↑*, TAL1*↑*, TKL1*↑*, NQM1*↑*, TKL2*↑, *XKS1*↑*;* pAKX002 (2µm *xylA*)), inoculated at 0.02 g biomass L^{-1}, were sparged with a mixture of 0.1% CO_2 and 99.9% N_2 instead of with pure N_2.

This change, which led to an initial CO_2 concentration in the outlet gas of 0.1%, abolished the 7-8 d lag phase observed for this strain in low-inoculum-density cultures sparged with pure N_2 (**Figures 1D** and **2B**). These results indicate that anaerobic growth of engineered, D-xylose-metabolizing *S. cerevisiae* can strongly depend on the concentration of CO_2.

Use of L-aspartate as nitrogen source can replace CO_2 supplementation in low-inoculum-density cultures of strain IMU078

As recently proposed for xylose-fermenting *E. coli* (Gonzalez *et al.* 2017), the observed CO_2 requirement of engineered *S. cerevisiae* strains for anaerobic growth on D-xylose might reflect a pivotal role of the anaplerotic carboxylation of pyruvate or phosphoenolpyruvate (PEP) to oxaloacetate. To test this hypothesis, strain IMU078 (*gre3Δ, RPE1*↑*, RKI1*↑*, TAL1*↑*, TKL1*↑*, NQM1*↑*, TKL2*↑, *XKS1*↑*;* pAKX002 (2µm *xylA*)) was grown on D-xylose in low-inoculum density (0.02 g biomass L^{-1}) bioreactor cultures sparged with pure N_2, in which L-aspartate instead of ammonium sulfate was used as the sole nitrogen source. In cultures grown on L- aspartate, which can be directly transaminated to oxaloacetate in a CO_2-independent manner, strain IMU078 showed immediate anaerobic growth (**Figures 1E** and **2C**).

In *S. cerevisiae*, oxaloacetate can be formed from pyruvate by pyruvate carboxylase (Pyc1, Pyc2) or via PEP-carboxykinase (Pck1). *PCK1* expression is repressed by glucose (Gancedo and Schwerzmann 1976, Daran-Lapujade *et al.* 2004) and elevated concentrations of CO_2 are required to enable Pck1 to act as sole anaplerotic enzyme in *S. cerevisiae* (Zelle *et al.* 2010). A previous transcriptome study on an evolved XI-based, xylose-fermenting strain showed an 8-fold higher transcript level of *PCK1* than its non-evolved parental strain (Zhou *et al.* 2012). However, deletion of *PCK1* in strain IMU078 (*gre3Δ, RPE1*↑*, RKI1*↑*, TAL1*↑*, TKL1*↑*, NQM1*↑*, TKL2*↑, *XKS1*↑*,* pAKX002 (2µm *xylA*)), resulting in strain IMX1736 (IMU078, *pck1Δ*), did not abolish the positive effect of external CO_2 supplementation on anaerobic growth on D-xylose (**Figure 1F**).

Omitting the over-expression of pentose-phosphate pathway paralogs enables lag-phase free anaerobic xylose fermentation at low inoculum density | In addition to codon-optimization and *xylA* expression vector, a third genetic difference exists between RWB217 (*gre3Δ, RPE1*↑*, RKI1*↑*, TAL1*↑*, TKL1*↑, *XKS1*↑*;* pAKX002 (2µm *xylA*)) (Kuyper *et al.* 2005), **Table 2**) and the two strains requiring a multi-day anaerobic adaptation phase

prior to xylose fermentation (IMX696 (*gre3Δ, RPE1*↑*, RKI1*↑*, TAL1*↑*, TKL1*↑*, NQM1*↑*, TKL2*↑, *XKS1*↑, *xylA*36* (Verhoeven *et al.* 2017) and IMU078 (*gre3Δ, RPE1*↑*, RKI1*↑*, TAL1*↑*, TKL1*↑*, NQM1*↑*, TKL2*↑, *XKS1*↑, pAKX002 (2µm *xylA*), **Table 2**). This difference concerns the over-expression of minor paralogs of the PPP genes *TAL1* and *TKL1* (*NQM1* and *TKL2*, respectively) in the latter two strains.

To evaluate the potential impact of the presence of expression cassettes for *NQM1* and *TKL2* in strains IMX696 and IMU078, a strain was constructed that was congenic to IMU078 except for the omission of the expression cassettes corresponding to *NQM1* and *TKL2* in the synthetic gene cluster harbouring the genes encoding non-oxidative PPP enzymes and *XKS1* (**Table 2**). Although the topology of the expression cassettes was different, the relevant genotype of the resulting strain IMU079 (*gre3Δ, RPE1*↑*, RKI1*↑*, TAL1*↑*, TKL1*↑*, XKS1*↑*;* pAKX002 (2µm *xylA*), **Table 2**) mimicked that of strain RWB217. Just like strain RWB217, strain IMU079 initiated anaerobic growth on D-xylose within 35 ± 5 h in low-inoculum-density, nitrogen-sparged cultures grown with ammonium sulfate as nitrogen source (**Figure 3A** and **B**). Furthermore, specific growth rate and other physiological parameters of strain IMU079 were similar to those of the Cas9-expressing strain derived from RWB217 (IMX975, **Table 4**). When inoculated at a high cell density (0.2 g biomass L^{-1}), strains IMU079 and IMU078 showed closely corresponding CO_2 production profiles (**Additional Figure 3**).

To further investigate the effects of over-expression of *NQM1* and/or *TKL2* on anaerobic growth on D-xylose, over-expression cassettes for *NQM1* and *TKL2* were introduced in the *SGA1* loci of strains IMX975 (RWB217, *can1::Cas9*) and IMU079, neither of which

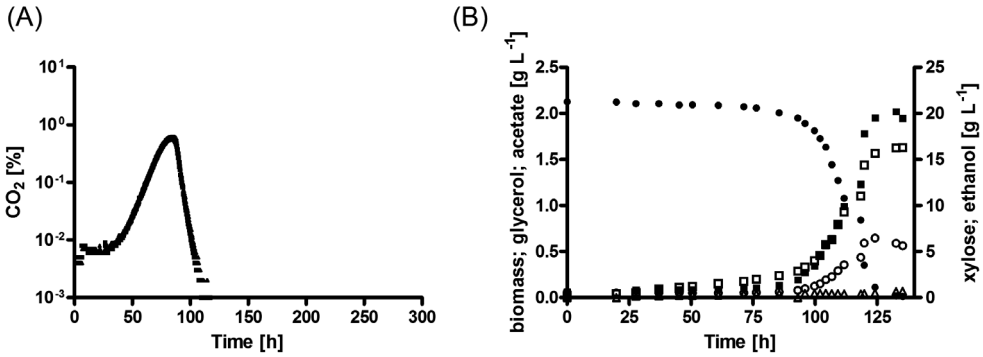


Figure 3. Fermentation profiles of the metabolically engineered, non-evolved D-xylose metabolizing *S. cerevisiae* strain IMU079 (*gre3Δ, RPE1*↑*, RKI1*↑*, TAL1*↑*, TKL1*↑*, XKS1*↑*, xylA* (pAKX002)) grown in anaerobic bioreactor batch cultures on synthetic medium supplemented with 20 g L^{-1} D-xylose. Cultures were inoculated at a biomass concentration of 0.02 g L^{-1} and sparged with 0.5 vvm N_2. (A) CO_2 concentration in off gas. (B) Growth and product formation. • = D-Xylose; ▪ = Biomass; □ = Glycerol; ○ = Ethanol; Δ = Acetate. Panel (A) shows data from duplicate cultures, panel (B) shows data from a single representative culture from a set of two independent duplicate cultures. Data from replicate cultures are shown in **Additional Figure 2**.

Table 4. Specific growth rates and yields of biomass, ethanol and glycerol of the metabolically engineered, xylose-fermenting *S. cerevisiae* strains IMX975 (RWB217-Cas9 (*gre3Δ, RPE1*↑, *RKI1*↑, *TAL1*↑, *TKL1*↑, *XKS1*↑, *can1::CAS9, 2µm xylA* (pAKX002)), IMU079 (*gre3Δ, RPE1*↑, *RKI1*↑, *TAL1*↑, *TKL1*↑, *XKS1*↑, *2µm xylA* (pAKX002)), and IMU078 (*gre3Δ, RPE1*↑, *RKI1*↑, *TAL1*↑, *TKL1*↑, *NQM1*↑, *TKL2*↑, *XKS1*↑, *2µm xylA* (pAKX002)) in anaerobic bioreactor batch cultures, inoculated with an initial biomass density of 0.02 g biomass L^{-1} and grown on 20 g L^{-1} xylose. Bioreactor cultures of strains IMX975, IMU079, and IMU078 (L-Asp) were sparged with pure N_2 (0.5 vvm). Cultures of strain IMU078 indicated as "IMU078 (+ 0.1% CO_2)" were sparged with a mixture of 0.1% CO_2 and 99.9% N_2. "L-Asp" indicates that the medium contained L-aspartate instead of ammonium sulfate as sole nitrogen source. Data represent average ± SE of two independent cultures for each strain. Detailed information on strain genotypes is provided in **Table 2**.

	IMX975	IMU079	IMU078 (+ 0.1% CO_2)	IMU078 (L-Asp)
Specific growth rate (h^{-1})	0.09 ± 0.01	0.08 ± 0.00	0.05 ± 0.01	0.054 ± 0.001
Biomass yield (g g^{-1})	0.088 ± 0.001	0.096 ± 0.004	0.093 ± 0.002	0.104 ± 0.002
Ethanol yield (g g^{-1})	0.395 ± 0.013	0.395 ± 0.002	0.382 ± 0.003	0.406± 0.002
Glycerol yield (g g^{-1})	0.085 ± 0.001	0.081 ± 0.001	0.074 ± 0.003	0.076 ± 0.000

previously harboured these cassettes. However, characterization of the resulting strains (IMU081 (IMU079 *sga1*::*NQM1,TKL2*), IMX1366 (IMX975 *sga1*::*NQM1,TKL2*)) in anaerobic bioreactor cultivations with xylose led to inconsistent lag-phase phenotypes (**Additional Figure 4**). In particular, introduction of the two cassettes in strains IMX975 (RWB217, *can1::Cas9*) and IMU079 did not yield prolonged anaerobic adaptation phases. These observations indicate that the effect of these cassettes on anaerobic growth may be related to or even caused by their immediate genetic context, e.g. by affecting expression of neighbouring genes in the pentose-pathway gene cluster in strains IMX696 and IMU078.

Discussion

Repeatability of results is a key requirement for progress in all scientific research (Nature 536, 373; (25 August 2016). This reassessment of requirements for anaerobic growth on D-xylose of engineered, XI-based *S. cerevisiae* strains was prompted by results from multiple laboratories, including our own (**Table 1**), which appeared to contradict an early report by our group (Kuyper *et al.* 2005). The results confirm the conclusions of Kuyper *et al.* (2005) by showing that the defined set of targeted genetic modifications reported in their study suffices to enable anaerobic growth of CEN.PK-based strains on D-xylose. Additionally, this study shows how anaerobic growth on D-xylose can critically depend on seemingly small differences in strain design and cultivation conditions.

Verhoeven *et al.* (2017) included overexpression of *NQM1* and *TKL2*, the 'minor' paralogs of the PPP genes *TAL1* and *TKL1*, respectively, in their strategy for single-step construction of a xylose-metabolising *S. cerevisiae* strain. Increased expression of these paralogs in

evolved strains had previously been shown to contribute to improved anaerobic growth L-arabinose (Wisselink *et al.* 2010), whose metabolism also proceeds via the non-oxidative PPP. This study shows that the mere omission of these two overexpression cassettes from the strain design of Verhoeven *et al.* (2017) suffices to eliminate a prolonged (12 d) adaptation phase prior to initiation of anaerobic growth on D-xylose in N_2-sparged cultures grown at a low inoculum density.

A recent study showed reduced protein levels of Tkl2 and Nqm1 in *xylA*-based strains evolved for improved D-xylose fermentation (Sato *et al.* 2016). However, no consistent effect on anaerobic growth performance was observed when we integrated overexpression cassettes for *NQM1* and *TKL2* outside rather than inside the synthetic cluster of pentose-metabolism genes that was central to the strain design used by Verhoeven *et al.* (2017) and in the present study. We therefore cannot exclude that the negative impact of these expression cassettes in strains IMU078 and IMX696 was related to their physical location, e.g. by transcriptional interference with other neighbouring genes in a tightly packed cluster of highly expressed genes (Kuijpers *et al.* 2016). Transcriptome analysis would provide a logical first step in attempts to elucidate the mechanism by which integration of the *NQM1* and/or *TKL2* cassettes prevents anaerobic growth on xylose in strains IMU078 and IMX696. However, the requirement for controlled bioreactor batch cultivation of multiple strains, some of which exhibit long lag phases, will make this a major experimental effort. Until this issue has been resolved, the work of Sato *et al.* (2016) and the present study indicate that it is prudent to omit *NQM1* and *TKL2* overexpression from initial designs for D-xylose-fermenting strains (Sato *et al.* 2016).

Even within the CEN.PK genetic background, literature reports differ on the ability of XI-based engineered strains to grow anaerobically on D-xylose without additional laboratory evolution (Kuyper *et al.* 2004, Karhumaa *et al.* 2007, Brat *et al.* 2009, Parachin *et al.* 2011, Shen *et al.* 2012, Hector *et al.* 2013, Verhoeven *et al.* 2017) (**Table 1**). In this study, increasing the initial concentration of biomass or CO_2 in N_2-sparged, anaerobic bioreactor cultures eliminated the extended anaerobic adaptation phase of *NQM1*/*TKL2* overexpressing strains. This observation showed that the inability of these strains to grow anaerobically on D-xylose, without first acquiring additional mutations, was conditional rather than absolute.

CO_2 and bicarbonate are essential for carboxylation reactions in biosynthesis. In several bacteria, including *E. coli*, anaerobic growth without a long lag phase requires external supply of CO_2 (Valley and Rettger 1927, Repaske *et al.* 1974, Repaske and Clayton 1978). The critical role of biosynthetic carboxylation reactions in *S. cerevisiae* is illustrated by the essentiality of carbonic anhydrase (Nce103), which interconverts CO_2 and HCO_3^-, for growth on glucose in aerated cultures at ambient atmospheric pressure (Aguilera *et al.*

2005). Despite the key roles of inorganic carbon in microbial metabolism, sparging with CO_2 -free N_2 is commonly applied in anaerobic laboratory bioreactor cultivation of *S. cerevisiae*, as sparging with CO_2 complicates quantification of its production in yeast metabolism. In glucose-grown cultures of wild-type *S. cerevisiae* strains, endogenous CO_2 production by vigorous alcoholic fermentation likely provides sufficiently high 'sparking' levels of CO_2 to enable growth initiation even at low initial biomass concentrations. Conversely, under the same conditions, the lower fermentation rates in engineered, non-evolved xylose-fermenting strains might not be able to provide the required CO_2 levels. The striking impact of endogenous CO_2 generation on anaerobic performance of engineered pentose-fermenting strains indicates that, especially for strains with low endogenous CO_2 production rates, supplementing anaerobic, nitrogen-sparged bioreactor batch cultures with CO_2 is a complicating but necessary measure. While relevant in laboratory settings, it should be borne in mind that industrial scale bioreactors are not sparged and have increased hydrostatic pressure, which in turn increases the CO_2 partial pressure. Additionally, CO_2 limitation is unlikely to occur during anaerobic fermentation of lignocellulosic hydrolysates, in which pentose fermentation is typically preceded by a vigorous glucose fermentation phase (Jansen *et al.* 2017).

A recent ^{13}C-flux analysis study showed that anaerobic fermentation of D-xylose, but not of glucose, by *E. coli* required lipid turn-over by β-oxidation to provide CO_2 for anaplerotic (pyruvate or PEP to oxaloacetate) carboxylation reactions (Gonzalez *et al.* 2017). Use of L-aspartate, whose transamination yields oxaloacetate, as the nitrogen source, completely eliminated the long anaerobic adaptation phase of *NQM1/TKL2*-expressing strains in N_2-sparged, low-inoculum density cultures. This strong effect of bypassing a single carboxylation reaction might reflect a CO_2 sparging effect that simply reduces the overall CO_2 requirement for growth. Alternatively, the link between anaplerotic synthesis of oxaloacetate and D-xylose metabolism may be more specific. At physiological pH values, Mn^{2+}, the preferred metal cofactor of *Piromyces* XylA (Lee *et al.* 2017), is also a much better metal cofactor for yeast pyruvate carboxylase (Pyc1 and Pyc2) than Mg^{2+} (Cazzulo and Stoppani 1969). High-level expression of *Piromyces xylA* might therefore result in a competition between the two enzymes for Mn^{2+}. Alternatively, lower intracellular concentrations of acetyl-CoA, a key activator of pyruvate carboxylase, in xylose-grown cultures than in glucose-grown cultures might compromise *in vivo* activity of Pyc1 and Pyc2 (Bergdahl *et al.* 2012).

In comparison with other strain backgrounds, the CEN.PK lineage may have at least one specific advantage for enabling anaerobic growth on D-xylose. Using a different *S. cerevisiae* genetic background, Sato et al. (2016) evolved a *Clostridium phytofermentans xylA*-expressing strain for anaerobic xylose fermentation and analysed causal mutations

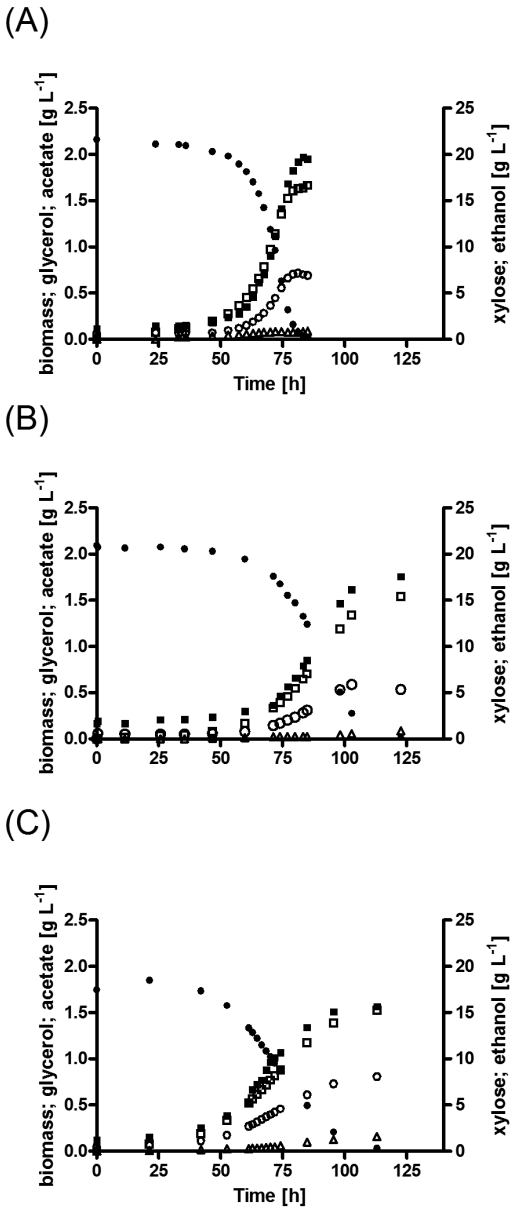
(Sato *et al.* 2016). In addition to mutations that affected *GRE3*, whose deletion was already in the strain design of Kuyper *et al*. (2005), these included mutations in *IRA2* (Sato *et al.* 2016). Loss of Ira2*,* an inhibitor of cAMP-PKA signalling, results in increased protein kinase A (PKA) activity (Tanaka *et al.* 1990). The PKA/cAMP pathway is activated by glucose and regulates multiple cellular processes (Rolland *et al.* 2002, Santangelo 2006). In contrast to glucose, xylose does not fully activate the cAMP/PKA pathway (Osiro *et al.* 2018). In comparison with other *S. cerevisiae* genetic backgrounds, CEN.PK strains carry many sequence differences in genes involved in this signal-transduction pathway (Vanhalewyn *et al.* 1999, Nijkamp *et al.* 2012). Moreover, they exhibit a higher basal PKA activity than other laboratory *S. cerevisiae* strains (Kümmel *et al.* 2010). The PKA/cAMP pathway regulates the glucose-dependent expression of the genes encoding several xylose-transporting members of the *S. cerevisiae* HXT (hexose transporter) family (i.e *HXT1, HXT2*, *HXT4, HXT5* and *HXT7* (Hamacher *et al.* 2002, Lee *et al.* 2002, Kim and Johnston 2006, Saloheimo *et al.* 2007)). A higher basal PKA activity might result in Hxt transporter landscapes that are conducive for fast xylose uptake and, thereby, enable the high rates of xylose fermentation that are required to sustain anaerobic growth.

We hope that this study will help colleagues with the design of pentose-fermenting strains and the experimental design of anaerobic yeast cultivation experiments, as well as that it will inspire further studies on the molecular basis for fast pentose fermentation in *S. cerevisiae*.

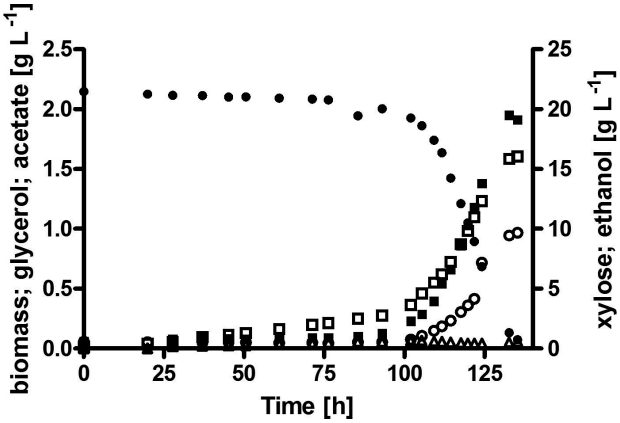
Acknowledgements

The authors thank Jolanda ter Horst and Erik de Hulster for help with fermentation set-up and sampling as well as Ioannis Papapetridis for valuable input in this project.

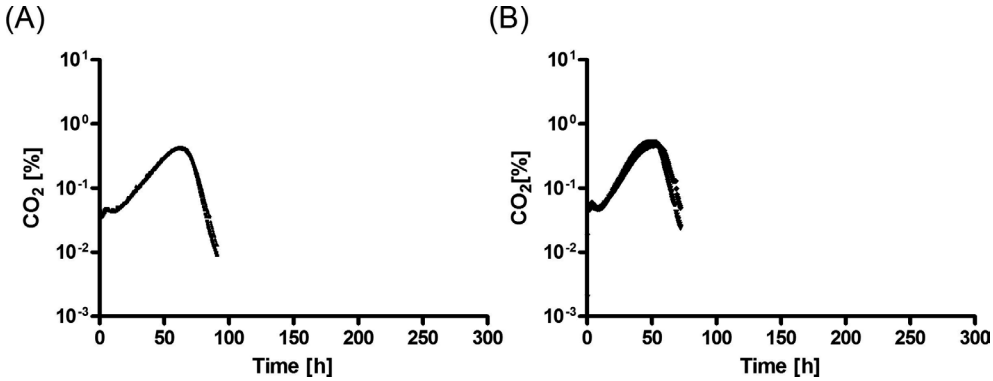
Additional material



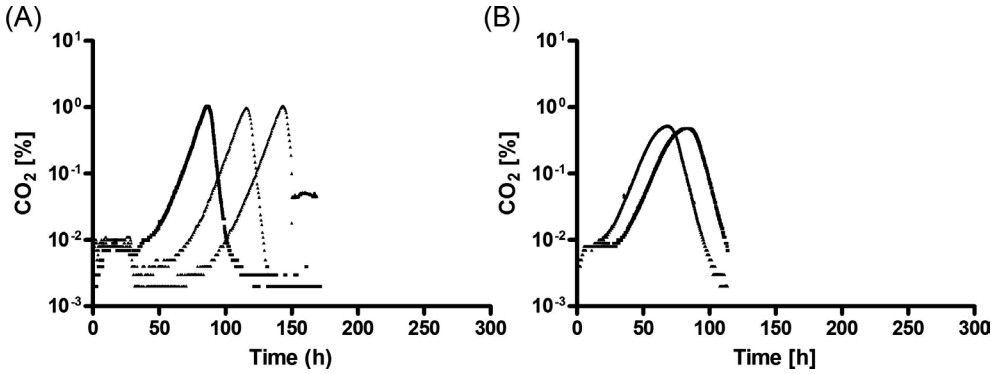
Additional Figure 1. Biomass and product formation in anaerobic bioreactor cultures of metabolically engineered, non-evolved D-xylose-metabolizing *S. cerevisiae* strains. Cultures were inoculated at a biomass concentration of 0.02 g L^{-1} and, unless stated otherwise, were sparged with 0.5 vvm N_2. (A) Strain IMX975 (RWB217 (*gre3Δ, RPE1*↑*, RKI1*↑*, TAL1*↑*, TKL1*↑, *XKS1*↑, *xylA* (pAKX002), *can1::Cas9*). (B) Strain IMU078 (*gre3Δ, RPE1*↑*, RKI1*↑*, TAL1*↑*, TKL1*↑*, NQM1*↑*, TKL2*↑, *XKS1*↑, *xylA* (pAKX002)) sparged with 0.5 vvm of a mixture of 0.1% CO_2 and 99.9% N_2. (C) Strain IMU078 grown in media containing L-aspartate as nitrogen source instead of ammonium sulfate. • = D-Xylose; ▪ = Biomass; □ = Glycerol; ○ = Ethanol; Δ = Acetate. Panels show data from single representative cultures from a set of two independent duplicate cultures for each strain. Data from replicate cultures are shown in **Figure 2.**



Additional Figure 2. Growth and product formation in anaerobic bioreactor cultures of the metabolically engineered, non-evolved D-xylose-metabolizing *S. cerevisiae* strain IMU079 (*gre3Δ, RPE1*↑, *RKI1*↑, *TAL1*↑, *TKL1*↑, *XKS1*↑, *xylA* (pAKX002)). Cultures were inoculated at a biomass concentration of 0.02 g L^{-1} and were sparged with 0.5 vvm N_2. • = D-Xylose; ▪ = Biomass; □ = Glycerol; ○ = Ethanol; Δ = Acetate. Data from a single representative culture from a set of two independent duplicate cultures; data from the replicate culture are shown in **Figure 3**.



Additional Figure 3. Fermentation profiles (CO_2 [%] in bioreactor off gas over time [h]) of metabolically engineered, non-evolved D-xylose metabolizing *S. cerevisiae* strains grown in anaerobic bioreactor batch cultures on synthetic medium supplemented with 20 g L^{-1} D-xylose. Cultures were inoculated at a biomass concentration of 0.2 g L^{-1} and sparged with 0.5 vvm N_2. (A) Strain IMU078 (*gre3Δ, RPE1*↑, *RKI1*↑, *TAL1*↑, *TKL1*↑, *NQM1*↑, *TKL2*↑, *XKS1*↑ *xylA* (pAKX002)) (B) Strain IMU079 (*gre3Δ, RPE1*↑, *RKI1*↑, *TAL1*↑, *TKL1*↑, *XKS1*↑, *xylA* (pAKX002)). The panels show results from 2 independent experiments. For detailed information on strain genotypes see **Table 2**.



Additional Figure 4. Fermentation profiles (CO_2 [%] in bioreactor off gas over time [h]) of metabolically engineered, non-evolved D-xylose -metabolizing *S. cerevisiae* strains grown in anaerobic bioreactor batch cultures on synthetic medium supplemented with 20 g L^{-1} D-xylose. Cultures were inoculated at a biomass concentration of 0.02 g L^{-1} and sparged with 0.5 vvm N_2. (A) Strain IMX1366 (RWB217, *can1::Cas9, sga1::NQM1, TKL2)* (B) Strain IMU081 (IMU079 *sga1::NQM1, TKL2)*. The panels show results from 3 (A) and 2 (B) independent experiments. For detailed information on strain genotypes see **Table 2**.

Additional Table 1. Oligonucleotides used in this study be found at: https://doi.org/10.1093/femsyr/foy104

References

Aguilera J, Van Dijken JP, De Winde JH *et al.* Carbonic anhydrase (Nce103p): an essential biosynthetic enzyme for growth of *Saccharomyces cerevisiae* at atmospheric carbon dioxide pressure. *Biochem J* 2005;**391**:311-6.

Alper H, Stephanopoulos G. Engineering for biofuels: exploiting innate microbial capacity or importing biosynthetic potential? *Nature Reviews Microbiology* 2009;**7**:715-23.

Bergdahl B, Heer D, Sauer U *et al.* Dynamic metabolomics differentiates between carbon and energy starvation in recombinant *Saccharomyces cerevisiae* fermenting xylose. *Biotechnol Biofuels* 2012;**5**:34.

Brat D, Boles E, Wiedemann B. Functional expression of a bacterial xylose isomerase in *Saccharomyces cerevisiae*. *Appl Environ Microb* 2009;**75**:2304-11.

Cazzulo J, Stoppani A. Effects of magnesium, manganese and adenosine triphosphate ions on pyruvate carboxylase from baker's yeast. *Biochem J* 1969;**112**:747-54.

Daran-Lapujade P, Jansen ML, Daran J-M *et al.* Role of transcriptional regulation in controlling fluxes in central carbon metabolism of *Saccharomyces cerevisiae:* A chemostat culture study. *J Biol Chem* 2004;**279**:9125-38.

Demeke MM, Dietz H, Li Y *et al.* Development of a D-xylose fermenting and inhibitor tolerant industrial *Saccharomyces cerevisiae* strain with high performance in lignocellulose hydrolysates using metabolic and evolutionary engineering. *Biotechnol Biofuels* 2013;**6**:89.

DiCarlo JE, Norville JE, Mali P *et al.* Genome engineering in *Saccharomyces cerevisiae* using CRISPR-Cas systems. *Nucleic Acids Res* 2013:gkt135.

dos Santos LV, Carazzolle MF, Nagamatsu ST *et al.* Unraveling the genetic basis of xylose consumption in engineered *Saccharomyces cerevisiae* strains. *Sci Rep* 2016;**6**:38676.

Entian K-D, Kötter P. 25 Yeast genetic strain and plasmid collections. *Methods Microbiol* 2007;**36**:629-66.

Gancedo C, Schwerzmann K. Inactivation by glucose of phosphoenolpyruvate carboxykinase from *Saccharomyces cerevisiae*. *Arch Microbiol* 1976;**109**:221-5.

Gietz RD, Woods RA. Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method. *Methods Enzymol* 2002;**350**:87-96.

Gonzalez JE, Long CP, Antoniewicz MR. Comprehensive analysis of glucose and xylose metabolism in *Escherichia coli* under aerobic and anaerobic conditions by ^{13}C metabolic flux analysis. *Metab Eng* 2017;**39**:9-18.

Hahn-Hägerdal B, Wahlbom CF, Gárdonyi M *et al.* Metabolic engineering of *Saccharomyces cerevisiae* for xylose utilization.*Metab Eng*,73: Springer, 2001,53-84.

Hamacher T, Becker J, Gárdonyi M *et al.* Characterization of the xylose-transporting properties of yeast hexose transporters and their influence on xylose utilization. *Microbiology* 2002;**148**:2783-8.

Hector RE, Dien BS, Cotta MA *et al.* Growth and fermentation of D-xylose by *Saccharomyces cerevisiae* expressing a novel D-xylose isomerase originating from the bacterium *Prevotella ruminicola* TC2-24. *Biotechnol Biofuels* 2013;**6**:84.

Hsiao H-Y, Chiang L-C, Chen L-F *et al.* Effects of borate on isomerization and yeast fermentation of high xylulose solution and acid hydrolysate of hemicellulose. *Enzyme Microb Technol* 1982;**4**:25-31.

Inoue H, Nojima H, Okayama H. High efficiency transformation of *Escherichia coli* with plasmids. *Gene* 1990;**96**:23-8.

Jansen ML, Bracher JM, Papapetridis I *et al. Saccharomyces cerevisiae* strains for second-generation ethanol production: from academic exploration to industrial implementation. *FEMS Yeast Res* 2017.

Jeffries TW. Engineering yeasts for xylose metabolism. *Curr Opin Biotechnol* 2006;**17**:320-6.

Karhumaa K, Sanchez RG, Hahn-Hägerdal B *et al.* Comparison of the xylose reductase-xylitol dehydrogenase and the xylose isomerase pathways for xylose fermentation by recombinant *Saccharomyces cerevisiae*. *Microb Cell Fact* 2007;**6**:5.

Kim J-H, Johnston M. Two glucose-sensing pathways converge on Rgt1 to regulate expression of glucose transporter genes in *Saccharomyces cerevisiae*. *J Biol Chem* 2006;**281**:26144-9.

Kötter P, Ciriacy M. Xylose fermentation by *Saccharomyces cerevisiae*. *Appl Microbiol Biotechnol* 1993;**38**:776-83.

Kuijpers NG, Solis-Escalante D, Luttik MA *et al.* Pathway swapping: Toward modular engineering of essential cellular processes. *PNAS* 2016:201606701.

Kümmel A, Ewald JC, Fendt S-M *et al.* Differential glucose repression in common yeast strains in response to *HXK2* deletion. *FEMS Yeast Res* 2010;**10**:322-32.

Kuyper M, Harhangi HR, Stave AK *et al.* High-level functional expression of a fungal xylose isomerase: the key to efficient ethanolic fermentation of xylose by *Saccharomyces cerevisiae*? *FEMS Yeast Res* 2003;**4**:69-78.

Kuyper M, Hartog MM, Toirkens MJ *et al.* Metabolic engineering of a xylose-isomerase-expressing *Saccharomyces cerevisiae* strain for rapid anaerobic xylose fermentation. *FEMS Yeast Res* 2005;**5**:399-409.

Kuyper M, Winkler AA, Dijken JP *et al.* Minimal metabolic engineering of *Saccharomyces cerevisiae* for efficient anaerobic xylose fermentation: a proof of principle. *FEMS Yeast Res* 2004;**4**:655-64.

Lee M, Rozeboom HJ, de Waal PP *et al.* Metal dependence of the xylose isomerase from *Piromyces sp. E2* axplored by activity profiling and protein crystallography. *Biochemistry* 2017;**56**:5991-6005.

Lee SM, Jellison T, Alper HS. Directed evolution of xylose isomerase for improved xylose catabolism and fermentation in the yeast *Saccharomyces cerevisiae*. *Appl Environ Microbiol* 2012;**78**:5708-16.

Lee SM, Jellison T, Alper HS. Systematic and evolutionary engineering of a xylose isomerase-based pathway in *Saccharomyces cerevisiae* for efficient conversion yields. *Biotechnol Biofuels* 2014;**7**:122.

Lee W, Kim M, Ryu Y *et al.* Kinetic studies on glucose and xylose transport in *Saccharomyces cerevisiae*. *Appl Microbiol Biotechnol* 2002;**60**:186-91.

Liu L, Redden H, Alper HS. Frontiers of yeast metabolic engineering: diversifying beyond ethanol and *Saccharomyces*. *Curr Opin Biotechnol* 2013;**24**:1023-30.

Lynd LR. Overview and evaluation of fuel ethanol from cellulosic biomass: technology, economics, the environment, and policy. *Annu Rev Energy Env* 1996;**21**:403-65.

Madhavan A, Tamalampudi S, Ushida K *et al.* Xylose isomerase from polycentric fungus *Orpinomyces*: gene sequencing, cloning, and expression in *Saccharomyces cerevisiae* for bioconversion of xylose to ethanol. *Appl Microbiol Biotechnol* 2009;**82**:1067-78.

Mans R, van Rossum HM, Wijsman M *et al.* CRISPR/Cas9: a molecular Swiss army knife for simultaneous introduction of multiple genetic modifications in *Saccharomyces cerevisiae*. *FEMS Yeast Res* 2015;**15**::fov004.

Margaritis A, Bajpai P. Direct fermentation of D-xylose to ethanol by *Kluyveromyces marxianus* strains. *Appl Environ Microbiol* 1982;**44**:1039-41.

Moysés DN, Reis VCB, de Almeida JRMM, Lidia Maria Pepe de *et al.* Xylose fermentation by *Saccharomyces cerevisiae*: challenges and prospects. *Int J Mol Sci* 2016;**17**:207.

Nature **536**, 373 (25 August 2016), doi:10.1038/536373a.

Nielsen J, Larsson C, van Maris A *et al.* Metabolic engineering of yeast for production of fuels and chemicals. *Curr Opin Biotechnol* 2013;**24**:398-404.

Nijkamp JF, van den Broek M, Datema E *et al. De novo* sequencing, assembly and analysis of the genome of the laboratory strain *Saccharomyces cerevisiae* CEN. PK113-7D, a model for modern industrial biotechnology. *Microb Cell Fact* 2012;**11**:36.

Olsson L, Hahn-Hägerdal B. Fermentation of lignocellulosic hydrolysates for ethanol production. *Enzyme Microb Technol* 1996;**18**:312-31.

Osiro KO, Brink DP, Borgström C *et al.* Assessing the effect of D-xylose on the sugar signaling pathways of *Saccharomyces cerevisiae* in strains engineered for xylose transport and assimilation. *FEMS Yeast Res* 2018.

Papapetridis I, Dijk M, Dobbe AP *et al.* Improving ethanol yield in acetate-reducing *Saccharomyces cerevisiae* by cofactor engineering of 6-phosphogluconate dehydrogenase and deletion of *ALD6*. *Microb Cell Fact* 2016;**15**:1.

Papapetridis I, Dijk M, Maris AJ *et al.* Metabolic engineering strategies for optimizing acetate reduction, ethanol yield and osmotolerance in *Saccharomyces cerevisiae*. *Biotechnol Biofuels* 2017;**10**:107.

Papapetridis I, Verhoeven MD, Wiersma SJ *et al.* Laboratory evolution for forced glucose-xylose co-consumption enables identification of mutations that improve mixed-sugar fermentation by xylose-fermenting *Saccharomyces cerevisiae*. *FEMS Yeast Res* 2018.

Parachin NS, Bergdahl B, van Niel EW *et al.* Kinetic modelling reveals current limitations in the production of ethanol from xylose by recombinant *Saccharomyces cerevisiae*. *Metab Eng* 2011;**13**:508-17.

Parreiras LS, Breuer RJ, Narasimhan RA *et al.* Engineering and two-stage evolution of a lignocellulosic hydrolysate-tolerant *Saccharomyces cerevisiae* strain for anaerobic fermentation of xylose from AFEX pretreated corn stover. *PloS one* 2014;**9**:e107499.

Repaske R, Clayton M. Control of *Escherichia coli* growth by CO_2. *J Bacteriol* 1978;**135**:1162.

Repaske R, Repaske AC, Mayer RD. Carbon dioxide control of lag period and growth of *Streptococcus sanguis*. *J Bacteriol* 1974;**117**:652-9.

Rolland F, Winderickx J, Thevelein JM. Glucose-sensing and-signalling mechanisms in yeast. *FEMS Yeast Res* 2002;**2**:183-201.

Ryabova OB, Chmil OM, Sibirny AA. Xylose and cellobiose fermentation to ethanol by the thermotolerant methylotrophic yeast *Hansenula polymorpha*. *FEMS Yeast Res* 2003;**4**:157-64.

Salazar AN, Gorter de Vries AR, van den Broek M *et al.* Nanopore sequencing enables near-complete de novo assembly of *Saccharomyces cerevisiae* reference strain CEN. PK113-7D. *FEMS Yeast Res* 2017;**17**.

Saloheimo A, Rauta J, Stasyk V *et al.* Xylose transport studies with xylose-utilizing *Saccharomyces cerevisiae* strains expressing heterologous and homologous permeases. *Appl Microbiol Biotechnol* 2007;**74**:1041-52.

Santangelo GM. Glucose signaling in *Saccharomyces cerevisiae*. *Microbiol Mol Biol Rev* 2006;**70**:253-82.

Sato TK, Tremaine M, Parreiras LS *et al.* Directed evolution reveals unexpected epistatic interactions that alter metabolic regulation and enable anaerobic xylose use by *Saccharomyces cerevisiae*. *PLoS Genet* 2016;**12**:e1006372.

Shen Y, Chen X, Peng B *et al.* An efficient xylose-fermenting recombinant *Saccharomyces cerevisiae* strain obtained through adaptive evolution and its global transcription profile. *Appl Microbiol Biotechnol* 2012;**96**:1079-91.

Smiley KL, Bolen PL. Demonstration of D-xylose reductase and D-xylitol dehydrogenase in *Pachysolen tannophilus*. *Biotechnol Lett* 1982;**4**:607-10.

Solis-Escalante D, Kuijpers NG, Bongaerts N *et al. amdSYM*, a new dominant recyclable marker cassette for *Saccharomyces cerevisiae*. *FEMS Yeast Res* 2013;**13**:126-39.

Sun H, Treco D, Schultes NP *et al.* Double-strand breaks at an initiation site for meiotic gene conversion. *Nature* 1989;**338**:87.

Tanaka K, Nakafuku M, Satoh T *et al. S. cerevisiae* genes *IRA1* and *IRA2* encode proteins that may be functionally equivalent to mammalian ras GTPase activating protein. *Cell* 1990;**60**:803-7.

Toivola A, Yarrow D, Van Den Bosch E *et al.* Alcoholic fermentation of D-xylose by yeasts. *Appl Environ Microbiol* 1984;**47**:1221-3.

Träff K, Cordero RO, Van Zyl W *et al.* Deletion of the *GRE3* aldose reductase gene and its influence on xylose metabolism in recombinant strains of *Saccharomyces cerevisiae* expressing the *xylA* and *XKS1* genes. *Appl Environ Microbiol* 2001;**67**:5668-74.

Valley G, Rettger LF. The influence of carbon dioxide on bacteria. *J Bacteriol* 1927;**14**:101.

van Dijken J, Bauer J, Brambilla L *et al.* An interlaboratory comparison of physiological and genetic properties of four *Saccharomyces cerevisiae* strains. *Enzyme Microb Technol* 2000;**26**:706-14.

van Rossum HM, Kozak BU, Niemeijer MS *et al.* Alternative reactions at the interface of glycolysis and citric acid cycle in *Saccharomyces cerevisiae*. *FEMS Yeast Res* 2016;**16**:fow017.

Vanhalewyn M, Dumortier F, Debast G *et al.* A mutation in *Saccharomyces cerevisiae* adenylate cyclase, *Cyr1K1876M*, specifically affects glucose-and acidification-induced cAMP signalling and not the basal cAMP level. *Mol Microbiol* 1999;**33**:363-76.

Verduyn C, Postma E, Scheffers WA *et al.* Effect of benzoic acid on metabolic fluxes in yeasts: a continuous-culture study on the regulation of respiration and alcoholic fermentation. *Yeast, 81992 nr 7* 1992.

Verhoeven MD, Lee M, Kamoen L *et al.* Mutations in *PMR1* stimulate xylose isomerase activity and anaerobic growth on xylose of engineered *Saccharomyces cerevisiae* by influencing manganese homeostasis. *Sci Rep* 2017;**7**.

Walfridsson M, Bao X, Anderlund M *et al.* Ethanolic fermentation of xylose with *Saccharomyces cerevisiae* harboring the *Thermus thermophilus xylA* gene, which expresses an active xylose (glucose) isomerase. *Appl Environ Microbiol* 1996;**62**:4648-51.

Wang PY, Schneider H. Growth of yeasts on D-xylulose. *Can J Microbiol* 1980;**26**:1165-8.

Wisselink HW, Cipollina C, Oud B *et al.* Metabolome, transcriptome and metabolic flux analysis of arabinose fermentation by engineered *Saccharomyces cerevisiae*. *Metab Eng* 2010;**12**:537-51.

Yamanaka K. Inhibition of D-xylose isomerase by pentitols and D-lyxose. *Arch Biochem Biophys* 1969;**131**:502-6.

Young E, Lee S-M, Alper H. Optimizing pentose utilization in yeast: the need for novel tools and approaches. *Biotechnol Biofuels* 2010;**3**:24.

Zelle RM, Trueheart J, Harrison JC *et al.* Phosphoenolpyruvate carboxykinase as the sole anaplerotic enzyme in *Saccharomyces cerevisiae*. *Appl Environ Microbiol* 2010;**76**:5383-9.

Zhou H, Cheng J-s, Wang BL *et al.* Xylose isomerase overexpression along with engineering of the pentose phosphate pathway and evolutionary engineering enable rapid xylose utilization and ethanol production by *Saccharomyces cerevisiae*. *Metab Eng* 2012;**14**:611-22.

5.

Laboratory evolution of a biotin-requiring *Saccharomyces cerevisiae* strain for full biotin prototrophy and identification of causal mutations

Jasmine M. Bracher, Erik de Hulster, Charlotte C. Koster, Marcel van den Broek, Jean-Marc G. Daran, Antonius J.A. van Maris and Jack T. Pronk

Essentially as published as in **Applied and Environmental Microbiology** 2017;
doi:10.1128/AEM.00892-17

Abstract

Biotin prototrophy is a rare, incompletely understood and industrially relevant characteristic of *Saccharomyces cerevisiae* strains. The genome of the haploid laboratory strain CEN.PK113-7D contains a full complement of biotin biosynthesis genes, but its growth in biotin-free synthetic media is extremely slow (specific growth rate $\mu \approx 0.01\ h^{-1}$). Four independent evolution experiments in repeated batch cultures and accelerostats yielded strains whose growth rates in biotin-free and biotin-supplemented medium were similar (μ up to $0.36\ h^{-1}$). Whole genome resequencing of these evolved strains revealed an up to 40-fold amplification of *BIO1*, which encodes pimeloyl-CoA synthetase. The additional copies of *BIO1* were found on different chromosomes and its amplification coincided with substantial chromosomal rearrangements. A key role of this gene amplification was confirmed by overexpression of *BIO1* in strain CEN.PK113-7D, which enabled growth in biotin-free medium ($\mu = 0.15\ h^{-1}$). Mutations in the membrane-transporter genes *TPO1* and/or *PDR12* were found in several of the evolved strains. Deletion of *TPO1* and *PDR12* in a *BIO1*-overexpressing strain increased its specific growth rate to $0.25\ h^{-1}$. The effects of null mutations in these genes, which have not been previously associated with biotin metabolism, were non-additive. This study demonstrates that *S. cerevisiae* strains that carry the basic genetic information for biotin synthesis can be evolved for full biotin prototrophy and identifies new targets for engineering biotin prototrophy into laboratory and industrial strains of this yeast.

Introduction

Biotin (vitamin H or B7) functions as an essential prosthetic group of enzymes in all three domains of life. The yeast *Saccharomyces cerevisiae* harbours four groups of biotin-dependent enzymes: (i) pyruvate carboxylases (Pyc1, Pyc2) catalyze anaplerotic formation of oxaloacetate (Gailiusis *et al.* 1964, Losada *et al.* 1964, Albertsen *et al.* 2003, Akao *et al.* 2011), (ii) acetyl-CoA carboxylases (Acc1, Hfa1) generate malonyl-CoA, the key precursor for lipid synthesis (Wakil *et al.* 1958), (iii) urea amidolyase (Dur1,2) releases ammonia from urea (Sumrada and Cooper 1982), and (iv) the aminoacyl-tRNA synthetase cofactor Arc1, which is the only biotin-dependent protein in *S. cerevisiae* that is not a carboxylase (Kim *et al.* 2004). Biotin is covalently linked to these enzymes by the biotin-protein ligase Bpl1 (Hoja *et al.* 1998).

While biotin prototrophy is widespread amongst prokaryotes and plants, animals and most fungi cannot synthesize this vitamin. The final four conserved steps in prokaryotic biotin biosynthesis are initiated by conversion of pimeloyl-CoA to 7-keto-8-aminopelargonic acid (KAPA) by KAPA synthase (BioF), after which DAPA aminotransferase (BioA) transaminates KAPA to 7,8-diaminopelargonic acid (DAPA). Subsequently, dethiobiotin synthetase (BioD) converts DAPA to dethiobiotin. Finally, sulfur insertion by biotin synthase (BioB) yields biotin (Streit and Entcheva 2003) (**Figure 1**). Two pathways for prokaryotic pimeloyl-CoA synthesis have been described. Some bacteria, including *Escherichia coli,* convert three molecules of malonyl-CoA to pimeloyl-CoA (Lin *et al.* 2010). Others, including *Bacillus subtilis,* generate this precursor by oxidative cleavage of fatty-acyl molecules (Kashiwagi and Igarashi 2011) (reviewed by (Lin and Cronan 2011)).

Biotin prototrophy is rare among yeasts and has a convoluted evolutionary history, with yeast strains of the same species sometimes exhibiting different biotin requirements. Most yeast strains isolated from nature, as well as laboratory strains, such as *S. cerevisiae* S288C, are biotin auxotrophs (Burkholder *et al.* 1944). Starting from pimelic acid, biotin biosynthesis in prototrophic yeast strains follows the prokaryotic pathway (**Figure 1**), but the yeast pathway for synthesis of this key precursor has not yet been fully elucidated (Lin and Cronan 2011, Magliano *et al.* 2011, Tanabe *et al.* 2011). Biotin biosynthesis genes in yeast are assumed to have been at least partially acquired from anaerobic bacteria by horizontal gene transfer, followed by gene duplication and neofunctionalization events (Hall *et al.* 2005, Hall and Dietrich 2007).

Most biotin auxotrophic *S. cerevisiae* strains do contain the genes encoding the last three enzymes of the biotin biosynthesis pathway (*BIO3, BIO4,* and *BIO2*, which are orthologs of *E. coli bioA*, *bioD*, and *bioB*). Some other yeasts, such as *Schizosaccharomyces pombe* and *Pichia pastoris,* only contain the biotin synthase (*BIO2*) gene, an ortholog

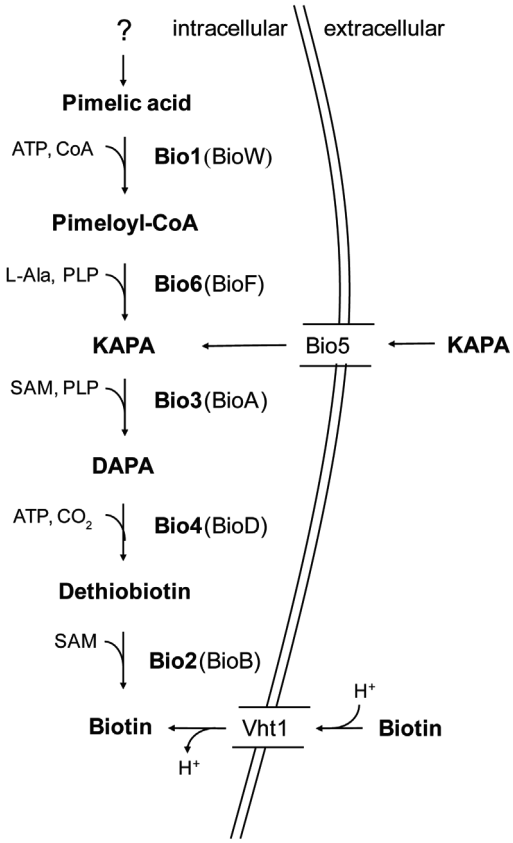


Figure 1. Biotin biosynthesis in *S. cerevisiae*. *BIO* genes encode the following enzymes: Bio1, pimeloyl-CoA synthetase; Bio6, 7-keto 8-amino-pelargonic acid (KAPA) synthetase; Bio3, 7,8-diaminopelargonic acid (DAPA) aminotransferase; Bio4, dethiobiotin synthase; Bio2, biotin synthase; SAM, S-adenosylmethionine; PLP, pyridoxal phosphate. Protein names in brackets indicate the corresponding bacterial enzymes. In *S. cerevisiae*, biotin and its precursor KAPA can be imported via the proton symporter Vht1 and the KAPA permease Bio5, respectively.

of prokaryotic *bioB* genes. In 2005, Wu *et al.* discovered *BIO6*, an ortholog of bacterial *bioF* genes, in biotin-prototrophic sake strains of *S. cerevisiae* (Wu *et al.* 2005). Hall and Dietrich completed the *S. cerevisiae* gene set for conversion of pimelic acid to biotin by discovering *BIO1*, which encodes pimeloyl-CoA synthetase and, in strains that carry either gene, is located adjacent to *BIO6*. Sequence similarity suggests that *BIO6* evolved by duplication and neofunctionalization of *BIO3,* thereby converting a DAPA synthase into a KAPA-synthetase. A similar evolutionary relationship was proposed for *BIO1* and YJR154W (Hall and Dietrich 2007).

Lack or loss of pathways for *de novo* vitamin biosynthesis in microbes has been proposed to reflect an evolutionary trade-off between fitness in biotin-scarce natural environments

and energy costs involved in biotin synthesis (Helliwell *et al.* 2013). Consistent with this notion, biotin prototrophic bacteria, archaea, and plants all harbour transporters for biotin uptake (Azhar *et al.* 2015). *S. cerevisiae* imports biotin via the high-affinity proton symporter Vht1 (Stolz *et al.* 1999). Some strains additionally harbour the Bio5 transporter, which imports the precursor 7-keto-8-aminopelargonic acid (KAPA), thus enabling biotin synthesis from exogenous KAPA (Phalip *et al.* 1999, Hall *et al.* 2005, Gasser *et al.* 2010). In *S. cerevisiae* strains that carry a *BIO5* gene, it is tightly linked to *BIO3* and *BIO4* in a gene cluster on Chromosome XIV (Hall and Dietrich 2007).

Most synthetic media for *S. cerevisiae* are routinely supplied with biotin. Vitamin addition increases production costs and decreases the shelf life of media. Biotin supplementation adds a delicate step in media preparation since the high pH required for dissolving biotin negatively affects its stability, thereby increasing the risk of batch-to-batch variations. Moreover, biotin is an expensive vitamin. It has been estimated that, at a reactor volume of 150 m^3, costs for large-scale industrial fermentation processes may be in the order of 1000 US$ per fermentation (Nijkamp *et al.* 2012a). Clearly, availability of fast-growing, biotin prototrophic strains could benefit process economics.

Several studies have focused on engineering microorganisms for biotin prototrophy. Expression of heterologous *BIO* genes in the biotin auxotrophic yeast *P. pastoris* reduced medium costs in fed-batch based production processes, even though the engineered biotin prototrophic strain grew slower than the reference strain (Gasser *et al.* 2010). Enhancing biotin synthesis in solventogenic clostridia improved production titres of acetone, butanol, and ethanol, predominantly by increasing cellular viability and performance (Yang *et al.* 2016).

The goal of this study was to identify key genetic determinants of biotin prototrophy in *S. cerevisiae*. To this end, the haploid strain *S. cerevisiae* CEN.PK113-7D, a popular model for systems biology and metabolic engineering research (Piper *et al.* 2002, Daran-Lapujade *et al.* 2007, Canelas *et al.* 2010, Nijkamp *et al.* 2012a), was subjected to laboratory evolution in biotin-free media. Evolved biotin-prototrophic cell lines were further characterised by whole-genome resequencing and by reverse engineering of identified mutations in the parental strain.

Materials and Methods

Strains, media, and maintenance

The *S. cerevisiae* strains used and constructed in this study (**Table 1**) belong to the CEN.PK lineage (Entian and Kötter 2007), with the exception of *Saccharomyces cerevisiae*

Table 1. *Saccharomyces cerevisiae* strains used in this study.

Strain	Relevant genotype	Description/use	Reference
CEN. PK113-7D	*MAT*a	Reference strain	(Entian and Kötter 2007)
CEN. PK113-5D	*MAT*a *ura3-52*	Uracil auxotrophic reference strain	(Entian and Kötter 2007)
S288C	*MAT*a	Marker strain for CHEF analysis	Yeast Genetic Stock Center, Berkeley, CA
IMX585	*MAT*a *can1::CAS9-tagA-loxP-natNT2-loxP*	CEN.PK113-7D expressing Cas9	(Mans *et al.* 2015)
IMS0478	*MAT*a *,* evolved	CEN.PK113-7D evolved for biotin prototrophy (accelerostat A)	This study
IMS0480	*MAT*a *,* evolved	CEN.PK113-7D evolved for biotin prototrophy (accelerostat B)	This study
IMS0481	*MAT*a *,* evolved	CEN.PK113-7D evolved for biotin prototrophy (accelerostat C)	This study
IMS0496	*MAT*a *,* evolved	CEN.PK113-7D evolved for biotin prototrophy (sequential batch reactor)	This study
IME327	*MAT*a *ura3-52,* pUDE446	CEN.PK113-5D, pUDE446 (*BIO1-BIO6*) (*URA3*)	This study
IME329	*MAT*a *ura3-52,* pUDE448	CEN.PK113-5D, pUDE448 (*BIO6*) (*URA3*)	This study
IME331	*MAT*a *ura3-52,* pUDE450	CEN.PK113-5D, pUDE450 (*BIO1*) (*URA3*)	This study
IME334	*MAT*a *ura3-52,* p426GPD	CEN.PK113-5D, p426GPD (empty) (*URA3*)	This study
IMK129	*MAT*a *ura3-52,tpo1::loxP-kanMX-loxP*	CEN.PK113-5D, *tpo1::loxP-kanMX-loxP*	This study
IMZ694	*MAT*a *ura3-52 tpo1::loxP-kanMX-loxP,* pUDE450	IMK129 ((*tpo1*Δ), pUDE450 (*BIO1*) (*URA3*)	This study
IMZ695	*MAT*a *ura3-52 tpo1::loxP-kanMX-loxP,* p426GPD	IMK129 (*tpo1*Δ), p426GPD (empty) (*URA3*)	This study
IMK163	*MAT*a *ura3-52 pdr12::loxP-kanMX-loxP*	CEN.PK113-5D, *pdr12::loxP-kanMX-loxP*	This study
IMZ704	*MAT*a *ura3-52 pdr12::loxP-kanMX-loxP,* pUDE450	IMK163 (*pdr12*Δ), pUDE450 (*BIO1*)	This study
IMZ705	*MAT*a *ura3-52 pdr12::loxP-kanMX-loxP,* p426GPD	IMK163 (*pdr12*Δ), p426GPD (empty)	This study
IMK773	*MAT*a *ura3-52 tpo1::loxP-kanMX-loxP, pdr12::hphNT1*	CEN.PK113-5D, *tpo1::loxP-kanMX-loxP, pdr12::hphNT1*	This study
IMZ701	*MAT*a *ura3-52 tpo1::loxP-kanMX-loxP, pdr12::hphNT1,* pUDE450	IMK773 (*tpo1*Δ*, pdr12*Δ), pUDE450 (*BIO1*)	This study
IMZ702	*MAT*a *ura3-52 tpo1::loxP-kanMX-loxP, pdr12::hphNT1,* p426GPD	IMK773 (*tpo1*Δ*, pdr12*Δ), p426GPD (empty)	This study

S288C (Mortimer and Johnston 1986, Entian and Kötter 2007). Yeast strains were grown on synthetic medium (SM) or YP medium (10 g·L^{-1} Bacto yeast extract, 20 g·L^{-1} Bacto peptone). Synthetic medium with urea as the nitrogen source (SM-urea) contained 38 mM urea and 38 mM K_2SO_4 instead of $(NH_4)_2SO_4$. After autoclaving SM at 121°C for 20 min,

or sterile-filtration of SM-urea using 0.2-µm bottle-top filters (Thermo Scientific, Waltham, Massachusetts), synthetic media were supplemented with 1 ml·L^{-1} of a filter-sterilized vitamin solution (0.05 g·L^{-1} D-(+)-biotin, 1.0 g·L^{-1} D-calcium pantothenate, 1.0 g·L^{-1} nicotinic acid, 25 g·L^{-1} myo-inositol, 1.0 g·L^{-1} thiamine hydrochloride, 1.0 g·L^{-1} pyridoxol hydrochloride, 0.20 g·L^{-1} 4-aminobenzoic acid). Biotin-free SM was prepared by omitting biotin from this solution. Unless specifically indicated, "SM" specifies synthetic medium with $(NH_4)_2SO_4$ as nitrogen source, whilst synthetic medium with urea as nitrogen source is abbreviated as "SM-urea". After autoclaving concentrated glucose solutions at 110°C for 20 min, glucose was added to SM, SM-urea, and YP media to a final concentration of 20 g·L^{-1}, yielding SMD, SMD-urea, and YPD, respectively. 500 ml shake flasks containing 100 ml medium, as well as 50 ml CELLreactor filter top tubes (Greiner Bio-One B.V., Alphen a/d Rijn, The Netherlands) containing 25 ml medium, were incubated at 30°C and at 200 rpm in an Innova Incubator (Brunswick Scientific, Edison, New Jersey). Solid media contained 1.5% Bacto agar and, when indicated, 200 mg·L^{-1} G418 or 200 mg·L^{-1} hygromycin. Selection and counter selection of the amdSYM cassette were performed as described earlier (Solis-Escalante *et al.* 2013). *Escherichia coli* strains were grown in LB (10 g·L^{-1} Bacto tryptone, 5 g·L^{-1} Bacto yeast extract, 5 g·L^{-1} NaCl) supplemented with 100 mg·L^{-1} ampicillin. Yeast and *E. coli* cultures were stored at -80°C after addition of (30% v/v) glycerol to stationary-phase shake-flask cultures.

Molecular biology techniques

PCR amplification of DNA fragments with Phusion Hot Start II High Fidelity Polymerase (Thermo Scientific) and desalted or PAGE-purified oligonucleotide primers (Sigma-Aldrich, St. Louis, Missouri) was performed according to the manufacturers' instructions. DreamTaq polymerase (Thermo Scientific) was used for diagnostic PCR. A table with the oligonucleotide primers used in this study can be found in **Additional Table 3,** at https://doi.org/10.1128/AEM.00892-17. Amplified DNA fragments were separated by electrophoresis on 1% (w/v) agarose gels (Thermo Scientific) in TAE buffer (Thermo Scientific) at 90 V for 35 min and purified with a GenElute PCR Clean-Up Kit (Sigma-Aldrich). Plasmids were isolated from yeast cultures with a Zymoprep Yeast Plasmid Miniprep II kit (Zymo Research, Irvine, California) and from *E. coli* with a Sigma GenElute Plasmid kit (Sigma-Aldrich). Yeast genomic DNA was isolated using a YeaStar Genomic DNA kit (Zymo Research) or with an SDS/LiAc protocol (Lõoke *et al.* 2011). Yeast strains were transformed by the lithium acetate method (Gietz and Woods 2002) and eight of the resulting colonies were re-streaked three consecutive times on biotin-supplemented selective media to obtain single colony isolates, followed by analytical PCR to verify their genotype. *E. coli* DH5α was used for chemical transformation (Inoue *et al.* 1990)

or for electroporation. Electroporation was done in a 2 mm cuvette (#1652086, Bio-Rad, Hercules, California) using a Gene Pulser Xcell Electroporation system (Bio-Rad). After isolation, plasmids were verified by restriction analysis and analytical PCR.

Plasmid construction

Promoter fragments of highly expressed yeast genes and promoter-less *BIO* genes were PCR amplified from genomic DNA of *S. cerevisiae* CEN.PK113-7D. Amplified *BIO* gene sequences included 0.5 kb terminator sequences. The 3' and 5' primers for amplification of promoter and terminator fragments, respectively, contained 60 bp synthetic extensions designed for efficient *in vivo* assembly of DNA fragments (Kuijpers *et al.* 2013). Promoters and coding regions were fused by fusion PCR (Shevchuk *et al.* 2004) and subsequently assembled into pJET1.2/blunt vectors with a CloneJet PCR Cloning kit (Thermo Scientific), resulting in the vector constructs pUD416 and pUD418 (**Table 2**). Yeast expression plasmids were assembled *in vivo* by amplifying promoter-gene-cassettes from pUD416, and pUD418, and vector fragments amplified from p426GPD (Mumberg *et al.* 1995). When necessary, oligonucleotide tags were changed to enable *in vivo* assembly with vector fragments. This assembly yielded the p426GPD-based plasmids pUDE446, pUDE448, and pUDE450 (**Table 2**).

Table 2. Plasmids used in this study. Indicated are the plasmid names and the corresponding relevant characteristics.

Name	Relevant characteristics	Reference
pUD416	ampR pJET1.2Blunt *TagB-pPYK-BIO1-tBIO1-TagA*	This study
pUD418	ampR pJET1.2Blunt *TagA-pPGK1-BIO6-tBIO6-5'Flank*$_{SGA1}$	This study
pUDE446	2µm ampR p426-GPD *pPYK-BIO1, pPGK1-BIO6*	This study
pUDE448	2µm ampR p426-GPD *pPGK1-BIO6*	This study
pUDE450	2µm ampR p426-GPD *pPYK-BIO1*	This study

Strain construction

S. cerevisiae strains with increased copy numbers of endogenous *BIO* genes (IME327, IME329, IME331, and IME334, **Table 1**) were constructed by transforming CEN.PK113-5D (*ura3*) with multi-copy plasmids (pUDE446, pUDE448, pUDE450, and p426GPD (*URA3*)). Geneticin (G418) resistance cassettes, PCR amplified from pUG6 (Güldener *et al.* 1996), were used to delete *TPO1* and *PDR12* in CEN.PK113-5D, resulting in strains IMK129 and IMK163, respectively. IMK129 was used to construct a *TPO1* and *PDR12* double deletion strain by transformation with a hygromycin resistance cassette (hphNT1) amplified from pMEL12 (Mans *et al.* 2015), resulting in IMK773. Deletion strains were subsequently

transformed with pUDE450 (*BIO1 URA3*), or with the empty reference vector p426GPD (*URA3*), resulting in IMZ694, IMZ695, IMZ701, IMZ702, IMZ704, and IMZ705 (**Table 1**).

Shake flask and plate growth experiments

Thawed aliquots of frozen stock cultures (1 ml) were inoculated in SMD shake-flask cultures and incubated for 12 h. This starter culture was used to inoculate a second shake-flask culture on SMD which, after another 12 h, was used to inoculate a third culture on SMD, at an initial OD_{660} of 0.05, 0.1, or 0.2. For biotin-free growth studies, all three cultures were grown on SMD without biotin. OD_{660} of the third culture was monitored with a Libra S11 spectrophotometer (Biochrom, Cambridge, United Kingdom). Specific growth rates were calculated from at least four time points in the exponential growth phase of each culture. Strain CEN.PK113-7D, which consistently failed to grow on biotin-free SMD in the third culture, was used as a negative control in all growth experiments. Cell numbers were estimated from calibration curves of OD_{660} versus cell counts in an Accuri flow cytometer (Becton Dickinson B.V., Breda, The Netherlands), generated with exponentially growing shake-flask cultures of strain CEN.PK113-7D on SMD medium. For plate assays, precultures were grown on SMD medium. Spot assays on SMD agar (pH 4.5), containing either 50 µM or 2 mM $MgSO_4$, were performed as described previously (Van Rossum *et al.* 2016). SMD agar plates were supplemented with either 0.05, 0.15, 0.5, 1, 1.5, 2, or 3 mM pimelic acid, 6 or 15 mM spermidine, or 0.5 mM potassium sorbate (**Additional Figure 1**). Liquid cultures of CEN.PK113-7D containing different concentrations of pimelic acid were prepared similar to shake-flask cultivations but were carried out in 50 ml CELLreactor filter top tubes with 25 ml of SMD-urea (pH 4.5) with an inoculation OD_{660} of 0.1. Filter-sterilized pimelic acid solution (1.25 M, pH 4.5) was added to duplicate cultures to a final concentration of 10, 12.5, 25, 50, 70, 80, 90, or 100 mM.

Laboratory evolution

Laboratory evolution of *S. cerevisiae* CEN.PK113-7D for biotin prototrophy was performed in accelerostats and in sequential bioreactor batch cultures. Accelerostat evolution was preceded by serial transfers in 3 parallel shake-flask experiments on biotin-free SMD, with an initial OD_{660} of 0.05 after each transfer. After 15 transfers (in 54 d), a specific growth rate of ca 0.1 h^{-1} was reached. 15 ml of each evolution culture was then used to inoculate separate 450 ml Multifors 2 parallel bioreactors (Infors Benelux, Doetinchem, The Netherlands) with a working volume of 100 ml. These bioreactors were subsequently operated as accelerostats, which are continuous cultures in which the dilution rate is increased over time (Paalme *et al.* 1995). During the initial batch phase, cells were grown on SM without biotin, supplemented with 20 $g \cdot L^{-1}$ glucose and 0.3 $g \cdot L^{-1}$ antifoam Pluronic

PE 6100 (BASF, Ludwigshafen, Germany). When a decrease in CO_2 production indicated glucose depletion (Bluesens CO_2 sensor, Herten, Germany), continuous cultivation was initiated at a dilution rate of 0.10 h^{-1} on SM without biotin, 7.5 $g \cdot L^{-1}$ glucose and 0.15 $g \cdot L^{-1}$ antifoam Pluronic PE 6100. Cultures were grown at 30°C, while a culture pH of 5 was maintained by automated addition of 2 M KOH. Cultures were sparged with air at a flow rate of 50 $ml \cdot min^{-1}$ and stirred at 1200 rpm. Dissolved oxygen concentrations remained above 40% of air saturation throughout the experiments. The dilution rate of the accelerostats was controlled by Iris Parallel Bioprocess Control software (Infors Benelux) based on a manually set threshold for the ratio between CO_2 concentration in the outlet gas and the feed rate of the medium supply pump. If this ratio was above the setpoint for at least 4 h, the medium feed rate was increased by 5%. When it was below the setpoint for the same period, the feed rate was decreased by 4.5%. This set-up ensured a steady increase in dilution rate as the culture evolved to reach higher specific growth rates, whilst preventing culture wash-out.

With increasing growth rates, the amount of CO_2 produced per glucose consumed increases due to an increased CO_2 production derived from the concomitantly increased CO_2 production during biomass formation (Verduyn *et al.* 1990). As during the evolution growth rates increased, the threshold ratio of CO_2 output and feed rate was manually adjusted accordingly. Accelerostat cultures were terminated when, within 3 months of accelerostat selection, dilution rates of 0.24 to 0.28 h^{-1} were reached.

Laboratory evolution in sequential batch reactors was preceded by shake-flask cultivation of strain CEN.PK113-7D in biotin-free SMD, using a frozen stock culture as inoculum. After two days, 0.1 ml of this culture was transferred to a fresh culture to deplete biotin stores and to generate a preculture from which, after ca. 12 h of incubation, a bioreactor culture was inoculated at an initial OD_{660} of 0.05.

Sequential batch cultivation was performed in 450 ml Multifors 2 parallel bioreactors (Infors Benelux) with a working volume of 100 ml. Aeration, pH, temperature and dissolved oxygen concentration thresholds were the same as in the accelerostat cultures. Growth was monitored based on CO_2 concentration in the off gas. When, after first having reached the CO_2 production peak, the CO_2 percentage in the off gas decreased below 0.02% a computer-controlled peristaltic pump automatically removed approximately 95% of the culture volume, leaving approximately 5% as inoculum for the next batch. The experiment was stopped when, after 11 batch cultivation cycles (47 d), no further increase in growth rate was observed over the following six consecutive batches and, moreover, cells with a multicellular phenotype were observed in the culture (Oud *et al.* 2013). Cells with a clumping phenotype, which facilitates sedimentation, are likely to have remained in the bioreactor during the empty-refill selection process due to their ability to evade wash-out

by clumping rather than by adapting for fast growth in absence of biotin (Oud *et al.* 2013). Hence, in order to ensure the selection of a fast-growing single cell line evolved for biotin prototrophy derived from a non-clumping cell, isolates obtained from the sequential batch reactor culture were selected for further analysis based on the absence of a clumping phenotype, determined by visual inspection with a Zeiss Imager D1 microscope. Single colony isolates were obtained from accelerostats and sequential batch reactor were obtained by plating on biotin-free SMD.

qPCR experiments

qPCR experiments were performed with duplicate cultures pregrown on SMD with or without biotin. RNA extraction was performed following the method of Schmitt *et al.* 1990 (Schmitt *et al.* 1990) whilst the sampling procedure was done as described elsewhere (Huisjes *et al.* 2012). cDNA was synthesized using a QuantiTect Reverse Transcription Kit (Qiagen, Düsseldorf, Germany) and concentrations were determined using a Qubit fluorometer (Life Technologies). qPCRs with cDNA from duplicate cultures were performed in technical triplicates on three dilutions of each sample using a QuantiTect SYBR Green PCR Kit (Qiagen) with a primer concentration of 0.5 µM in a total volume of 20 µl in the Rotor-Gene Q (Qiagen). All qPCR primers can be found at in **Additional Table 3,** at https://doi.org/10.1128/AEM.00892-17. *ACT1 e*xpression levels determined from the same duplicate culture were always included as an internal standard. The expression levels of the *BIO* genes relative to *ACT1* were determined using the ΔΔCT method (Schmittgen and Livak 2008).

For the determination of the normalized expression levels of the gene of interest (GOI), the average of the C(t)-values obtained from the triplicate measurements of the reference gene *ACT1* were subtracted from the averaged C(t)-values of the GOI ($\Delta C(t)_{Sample} = average\ C(t)_{GOI\ sample} - average\ C(t)_{ACT1}$). The normalized expression level of a GOI was then determined according to the ΔΔC(t) method ($Normalized\ target\ gene\ expression = 2^{-\Delta C(t)}$). The average between this value and the $\Delta C(t)_{Sample}$ value obtained from the duplicate culture and the standard deviation between those two values were then used to determine the normalized expression levels and the respective error (**Figure 3**, **Additional Table 2**).

Contour Clamped Homogenous Electric Field (CHEF) electrophoresis and Southern Blotting

Agarose plugs for all strains were prepared using a CHEF genomic plug kit following manufacturers' recommendations (Bio-Rad). One-third of each agarose plug was used per well of a 1% megabase agarose gel (Bio-Rad) buffered with 0.5 x TBE (5.4 g trizma base, 2.75 g boric acid, 2 ml 0.5 M EDTA, pH 8 in 1 L demineralized water). Chromosomes were separated in a CHEF-DR II system (Bio-Rad) for 28 h with a switch time of 60 s, an angle of 120 degrees and 5 $V \cdot cm^{-2}$, followed by 16 h with a switch time of 90 s. DNA was stained with ethidium bromide (3 $\mu g \cdot ml^{-1}$ in 0.5 x TBE buffer) and destained in 0.5 x TBE buffer for 20 min with gentle agitation. Imaging was done using an InGenius LHR Gel Imaging System (Syngene, Bangalore, India). Southern Blot probes (ca. 1 kb) were amplified from genomic DNA of CEN.PK113-7D and prepared with AlkPhos Direct Labelling and Detection system according to the manufacturers' protocol (Amersham Biosciences, Piscataway, NJ). Separated chromosomes were transferred to Hybond-N^+ nylon membranes (GE Healthcare, Diegem, Belgium) using vacuum transfer with a Vacuum Blotter (Bio-Rad) and fixed to the nylon membrane by UV exposure for 2 min. Signal generation and detection were performed using Amersham Gene Images AlkPhos Direct Labelling and Detection System together with CDP-Star chemiluminescent detection (Amersham Biosciences) and exposure to an Amersham Hyperfilm ECL (Amersham Biosciences) following manufacturers' recommendations.

DNA sequencing

Genomic DNA of strains IMS0478, IMS0480, IMS0481, and IMS0496 was isolated with a QIAGEN Blood & Cell Culture DNA Kit with 100/G Genomics-tips (QIAGEN, Valencia, California) according to the manufacturers' protocol. Paired-end sequencing (126 bp reads) was performed on 350-bp insert libraries with an Illumina HiSeq 2500 sequencer (Baseclear BV, Leiden, The Netherlands) with a minimum sample size of 550 Mbases accounting for a coverage of approximately 45x. Data mapping to the CEN.PK113-7D genome (Nijkamp *et al.* 2012a), data processing, and chromosome copy number variation determinations were done as described elsewhere (Verhoeven *et al.* 2017).

Copy numbers of *BIO1, BIO2, BIO3, BIO4,* and *BIO6* were estimated by comparing their read depth to the average read depth of the single-copy reference genes (YAL001C *(TFC3),* YBL015W *(ACH1),* YCL040W *(GLK1),* YDL029W *(ARP2),* YEL012W *(UBC8),* YER049W *(TPA1),* YBR196C *(PGI1),* YER178W *(PDA1),* YFL039C *(ACT1),* and YJL121C *(RPE1)*), processed with Pilon (Walker *et al.* 2014).

Accession numbers

The sequencing data are available under the Bioproject accession number PRJNA383023.

Results

Laboratory evolution of *S. cerevisiae* CEN.PK113-7D for full biotin prototrophy

After inoculation of biotin-free SMD with a biotin-depleted preculture of strain CEN. PK113-7D, it took ca. 20 d before slow growth, at a specific growth rate of ca. 0.01 h^{-1}, was observed. At a final cell number in the shake flasks of $4 \cdot 10^{11}$ cells·L^{-1} and a specific growth rate of 0.01 h^{-1} (i.e., a doubling time of 69 h), this implies an initial concentration of growing cells of at least $2 \cdot 10^{9}$ cells·L^{-1}. These results confirm an earlier report that strain CEN.PK113-7D is not completely auxotrophic for biotin, but can grow at very low rates in the absence of this vitamin (Nijkamp *et al.* 2012a).

To explore evolvability of full biotin prototrophy, i.e. a phenotype with identical specific growth rates in the presence and absence of biotin, strain CEN.PK113-7D was grown in accelerostats. These are continuous cultures in which the dilution rate is increased over time (Paalme *et al.* 1995). To select for fast-growing biotin-prototrophic mutants in triplicate glucose-limited accelerostat cultures on biotin-free SMD, the dilution rate was feedback-controlled based on the CO_2 concentration in the off gas (**Additional Figure 2).**

A fourth laboratory evolution experiment was performed in a sequential batch reactor (SBR), in which automated empty-refill cycles were based on the CO_2 concentration in the off gas, leaving ca. 5% of the culture as inoculum for each subsequent batch cycle (Kuyper *et al.* 2005b) (**Figure 2**). After 48 – 77 d of accelerostat cultivation, dilution rates of up to 0.27 h^{-1} were reached. Single-cell lines were isolated at the end of each of the accelerostat experiments and named IMS0478, IMS0480, and IMS0481 (**Table 1**). The SBR experiment was terminated when, after 47 d and 11 cycles, multicellular aggregates appeared in this culture due to the SBR setup-inherent empty-refill selection procedure that is prone for the unintended selection of cells developing a clumping phenotype enabling their persistence in the bioreactor without further improvement of their growth rate under selective conditions (Oud *et al.* 2013). At this point, the specific growth rate, as estimated from CO_2-production profiles, had reached 0.22 h^{-1} (**Table 3)**.

A non-aggregating single-colony isolate from the SBR culture was named IMS0496. All four single-cell isolates showed high specific growth rates in biotin-free SMD, ranging from 0.25 h^{-1} (strain IMS0478) to 0.36 h^{-1} (strain IMS0481; **Table 3**). This value is close to the specific growth rate of CEN.PK113-7D in biotin-supplemented SMD of 0.39 – 0.40 h^{-1} (Pronk 2002, Kuyper *et al.* 2005a, Hazelwood *et al.* 2006, Mans *et al.* 2015). Biotin-

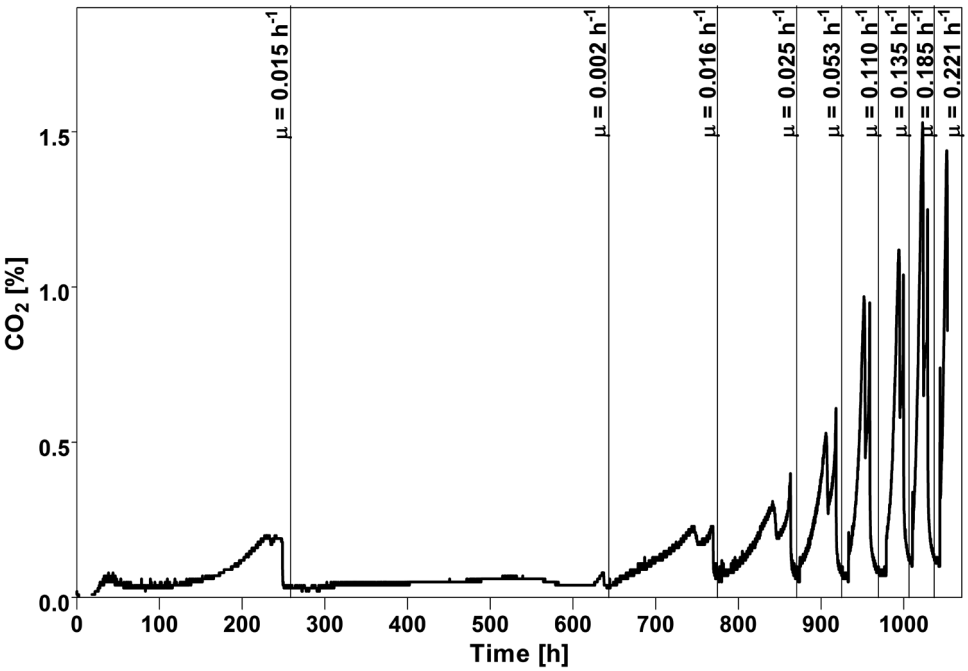


Figure 2. Laboratory evolution of *S. cerevisiae* CEN.PK113-7D cells in a sequential batch reactor (SBR) for improved growth in biotin-free synthetic media. Off-gas CO2 [%] profiles during an SBR experiment, in which automated empty-refill cycles were based on the CO_2 concentration in the off gas, leaving ca. 5% of the culture as inoculum for each subsequent batch cycle (Kuyper *et al.* 2005b). CO_2 production in the initial cycle reflects depletion of biotin pools in the inoculum. Specific growth rates were calculated from the exponential increase of the off-gas CO_2 concentration in each cycle. Graphs representing the increase of dilution rates in accelerostats over time are depicted in **Figure S2.**

supplemented cultures of the evolved strains showed an average specific growth rate of 0.35 h^{-1}, indicating that the evolved strain IMS0481 had acquired full biotin prototrophy **(Table 3**).

Massive amplification of *BIO1* and *BIO6* and increased expression of BIO genes in evolved biotin-prototrophic strains

To investigate the molecular basis of the acquired biotin prototrophy, expression levels and copy numbers of *BIO* genes were measured by qPCR analysis and whole-genome sequencing, respectively. In shake-flask cultures on biotin-supplemented SMD, transcript levels of *BIO* genes were lower than those of the reference gene *ACT1* in all strains tested (**Figure 3A, Additional Table 2**). Transcript levels of *BIO* genes in the evolved strains resembled those in the parental strain, except for *BIO1,* whose transcript levels were four- to eight-fold higher in the evolved strains. Growth in SMD without biotin resulted in transcriptional upregulation of *BIO* genes (**Figure 3B, Additional Table 2**).

Table 3. Specific growth rates of laboratory-evolved biotin prototrophic *S. cerevisiae* strains and of strains carrying defined mutations, grown in shake flask cultures containing synthetic media with (SMG + biotin) or without (SMG – biotin) biotin supplementation. Cells were grown in aerobic 100 ml shake flask cultures. Data represent averages and average deviation of the mean of measurements on independent duplicate cultures. *(Mans *et al.* 2015)

Strain	Relevant genotype	Growth rate[h^{-1}] in SMD	
		Without biotin	With biotin
CEN.PK113-7D	Haploid laboratory strain	< 0.01 h^{-1}	0.39 ± 0.01*
IMS0478	CEN.PK113-7D evolved for biotin prototrophy (accelerostat A)	0.25 ± 0.00	0.34 ± 0.02
IMS0480	CEN.PK113-7D evolved for biotin prototrophy (accelerostat B)	0.31 ± 0.00	0.34 ± 0.00
IMS0481	CEN.PK113-7D evolved for biotin prototrophy (accelerostat C)	0.36 ± 0.00	0.41 ± 0.00
IMS0496	CEN.PK113-7D evolved for biotin prototrophy (SBR)	0.32 ± 0.01	0.31 ± 0.02
IME327	CEN.PK113-5D + pUDE446 (2μ *BIO1-BIO6*)	0.16 ± 0.01	0.32 ± 0.01
IME329	CEN.PK113-5D + pUDE448 (2μ *BIO6*)	No growth	0.30 ± 0.00
IME331	CEN.PK113-5D + pUDE450 (2μ *BIO1*)	0.15 ± 0.02	0.32 ± 0.01
IME334	CEN.PK113-5D + p426GPD (empty plasmid)	No growth	0.34 ± 0.01
IMZ694	*tpo1*Δ, pUDE450 (2μ *BIO1*)	0.23 ± 0.00	0.32 ± 0.00
IMZ695	*tpo1*Δ, p426GPD (empty plasmid)	No growth	0.33 ± 0.01
IMZ704	*pdr12*Δ, pUDE450 (2μ *BIO1*)	0.25 ± 0.00	0.37 ± 0.00
IMZ705	*pdr12*Δ, p426GPD (empty plasmid)	No growth	0.40 ± 0.00
IMZ701	*pdr12*Δ*tpo1*Δ, pUDE450 (2μ *BIO1*)	0.25 ± 0.01	0.36 ± 0.01
IMZ702	*pdr12*Δ*tpo1*Δ, p426GPD (empty plasmid)	No growth	0.37 ± 0.00

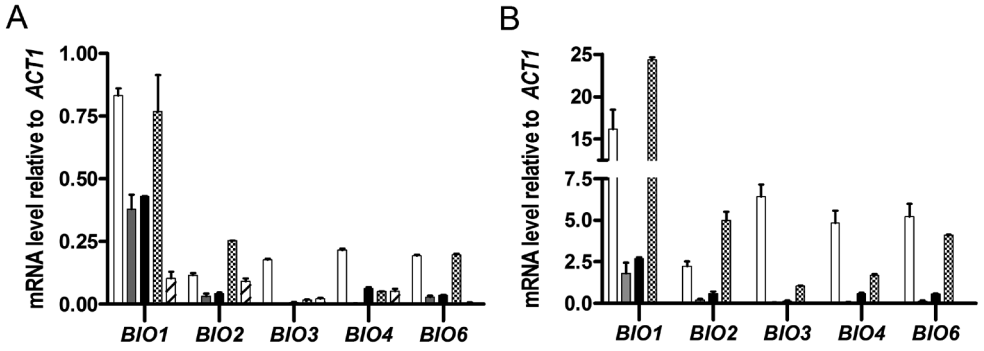


Figure 3. mRNA levels of BIO genes in strains evolved for biotin prototrophy. (A) Transcript levels in cultures grown on SMD with biotin | Transcript levels of *BIO1, BIO2, BIO3, BIO4*, and *BIO6* in the parent strain CEN. PK113-7D (diagonally hatched white bar) and in the evolved strains IMS0478 (white bar), IMS0480 (grey bar), IMS0481 (black bar), and IMS0496 (squared pattern), relative to *ACT1* expression **(B)** Transcript levels in cultures grown on SMD without biotin. Transcript levels of *BIO1, BIO2, BIO3, BIO4*, and *BIO6* in the evolved strains IMS0478 (white bar), IMS0480 (grey bar), IMS0481 (black bar), and IMS0496 (squared pattern), relative to *ACT1* expression levels. All qPCR experiments were carried out on duplicate cultures, with analytical triplicates for each culture. Relative expression levels were determined according to the ΔΔCT method (Schmittgen and Livak 2008). Error bars represent the standard deviation of the mean of duplicate analyses.

As observed in biotin-supplemented cultures, *BIO1* was expressed at higher levels than the other *BIO* genes, reaching transcript levels that were 1.5- to 24-fold higher than those of *ACT1* (**Figure 3B**). Analysis of sequencing read depths revealed massive amplification of *BIO1* and *BIO6* in the evolved strains, with copy numbers ranging from 8 to 42 (**Table 4**). The similar amplification of *BIO1* and *BIO6* observed in this analysis was consistent with their physical linkage (Nijkamp *et al.* 2012a). The evolved strains harboured one or two copies of *BIO2, BIO3,* or *BIO4,* except for the slowest growing strain IMS0478, in which *BIO3* and *BIO4* were amplified by eight and six fold, respectively.

Table 4. Estimated copy numbers of *BIO1, BIO2, BIO3, BIO4,* and *BIO6* in the evolved biotin-prototrophic strains and the parent strain CEN.PK113-7D. Copy numbers were calculated from sequence coverage depth of four strains independently evolved for biotin prototrophy in accelerostats (IMS0478, IMS0480, MS0481) and in a sequential batch reactor setup (IMS0496). All five genes indicated occur as single-copy genes in the parental strain CEN.PK113-7D **(Nijkamp *et al.* 2012a)**.

	Gene copy number per strain				
	BIO1	***BIO2***	***BIO3***	***BIO4***	***BIO6***
CEN.PK113-7D	1	1	1	1	1
IMS0478	8	2	8	6	8
IMS0480	42	2	1	1	43
IMS0481	19	2	1	1	19
IMS0496	19	2	2	1	20

Amplification of *BIO1* and *BIO6* involves chromosomal duplications, rearrangements and formation of neochromosomes

In three of the four evolved biotin-prototrophic strains, read depth analysis revealed a duplication of ChrI (**Figure 4**) which, in the parental strain, carries the *BIO1* and *BIO6* genes (Nijkamp *et al.* 2012a). Strain IMS0478, which also showed increased copy numbers of other *BIO* genes, carried additional duplications of ChrIX, the left arm of ChrVIII and an amplified region close to the telomere of the right arm of ChrXIV, where the *BIO3, 4, 5* cluster is located (**Figure 4A**). Electrophoretic karyotyping revealed strong differences in chromosome sizes between the evolved strains as well as with their common unevolved parental strain, while strains IMS0478, IMS0480, and IMS0481 contained additional, supernumerary isochromosomes (referred to as 'neochromosomes') (**Figure 5A**). Southern hybridization analysis showed that copies of the *BIO1-BIO6* gene cluster occurred on multiple chromosomes and neochromosomes in the evolved strains, indicating that, based on identical copy numbers and clustering of *BIO1* and *BIO6,* the *BIO1-BIO6* cluster multiplied its copy number via interchromosomal rearrangements (**Figure 5B**). This experiment confirmed that strain *S. cerevisiae* S288C, whose genomic DNA was used as

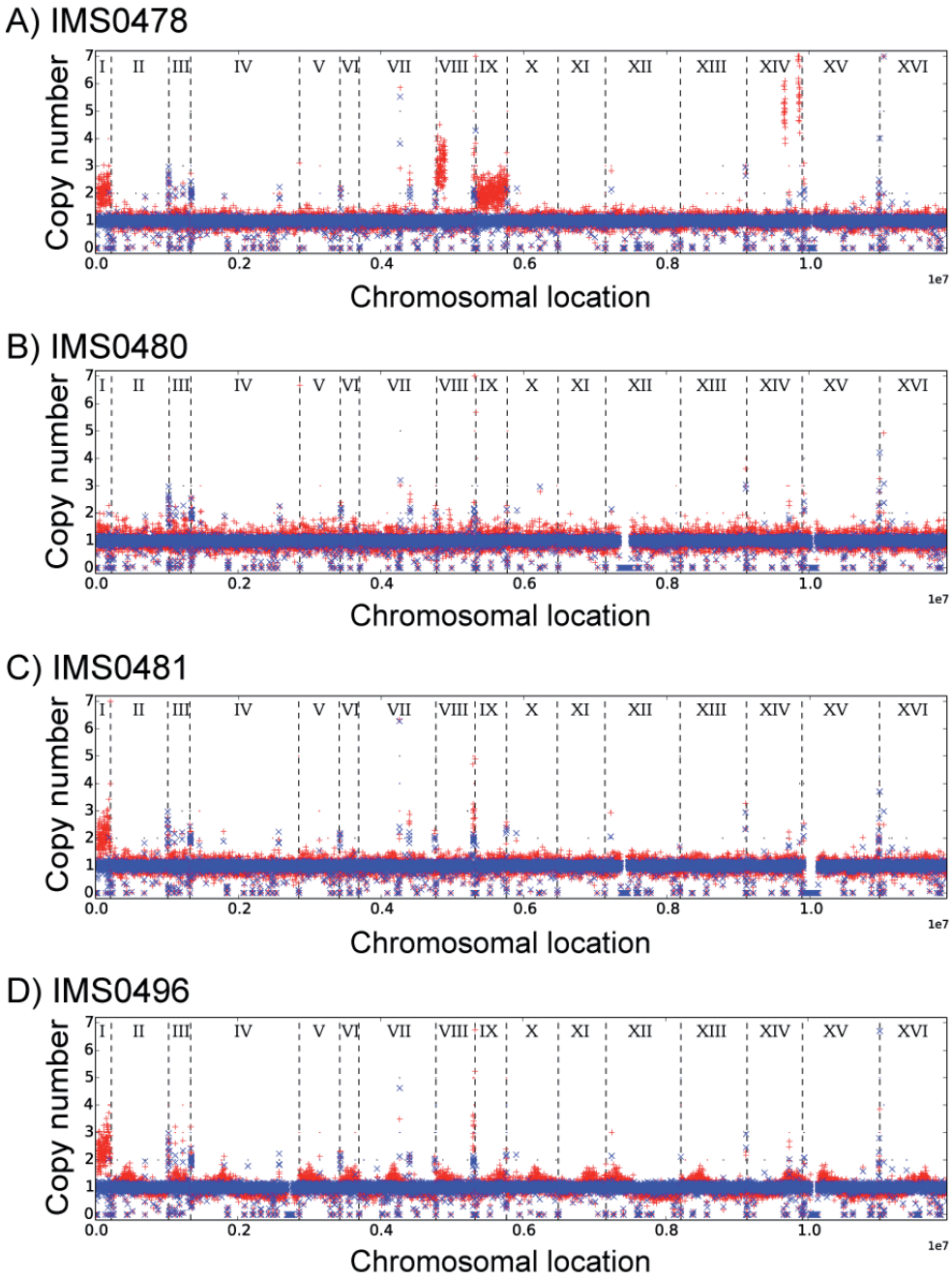


Figure 4. Chromosomal copy number variations in yeast strains evolved for full biotin prototrophy. Strains IMS0478, IMS0480, MS0481 were evolved in accelerostats, while strain IMS0496 was evolved in a sequential batch reactor **(A, B, C, D,** respectively). Copy numbers of chromosomes and chromosomal regions were calculated from sequence data with the Magnolya algorithm (Nijkamp *et al.* 2012b). Results for the parental strain *S. cerevisiae* CEN.PK113-7D and for the evolved strains are indicated in blue and red, respectively. Individual chromosomes, indicated by Roman numbers, are separated by dashed lines.

a size marker for chromosome identification, lacks the *BIO1* and consequently implies a lack of the whole *BIO1-BIO6* cluster (**Figure 5B**) (Wu *et al.* 2005, Nijkamp *et al.* 2012a).

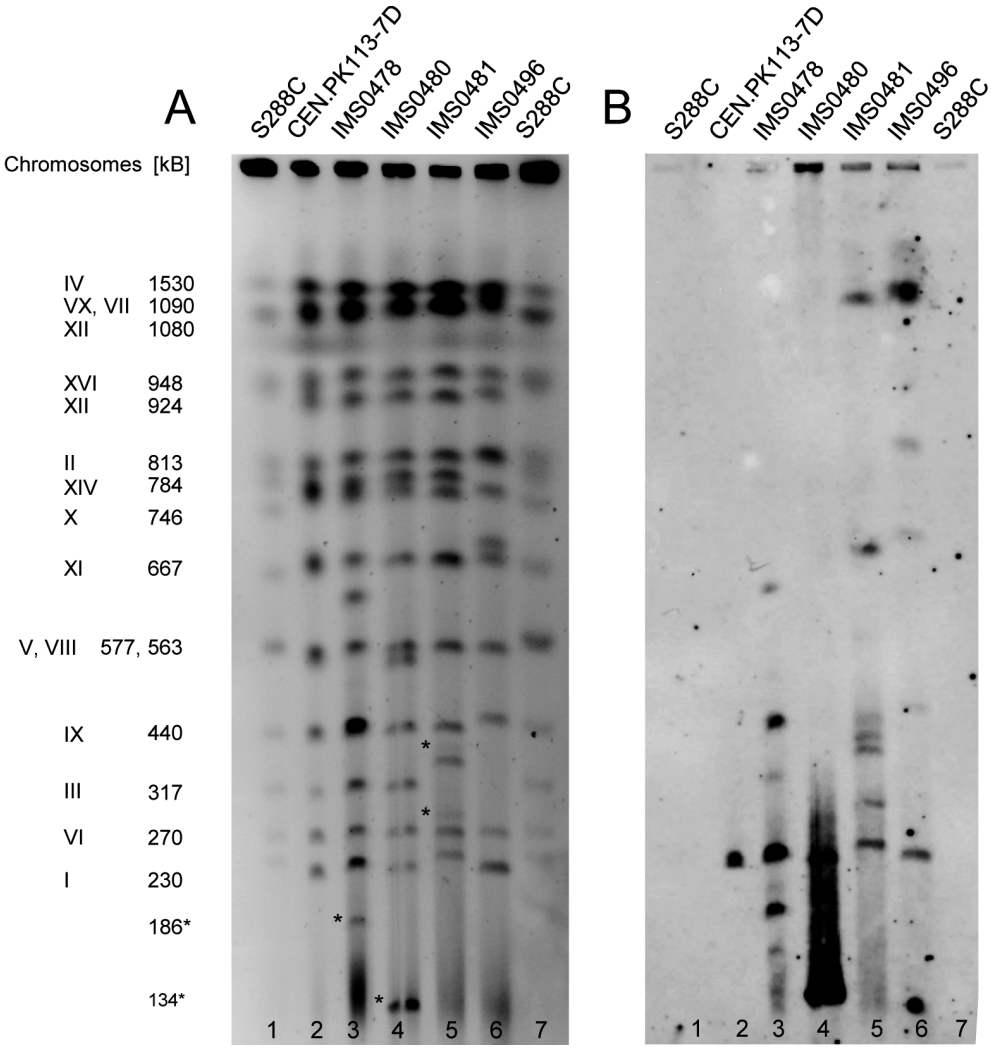


Figure 5. Karyotyping and chromosomal localization of *BIO1* in evolved biotin-prototrophic yeast strains | (A) Pulsed-field gel electrophoresis (PFGE) of the chromosomes of evolved biotin prototrophic strains MS0478, IMS0480, IMS0481, and IMS0496 and the parent strain *S. cerevisiae* CEN.PK113-7D**.** Indicated chromosome numbers and sizes (kbp) were obtained using *S. cerevisiae* S288C as ladder strain. Neochromosomes are indicated with asterisks (*). Chromosomes that changed in size and consequently in travel distance compared to the respective chromosome of the unevolved parent strain are not considered as neochromosomes. **(B)** Southern Blot of PFGE gel. Hybridization with *BIO1* probe reveals copies of *BIO1* on multiple chromosomes in evolved strains. Lanes are numbered at the bottom.

Increased copy numbers of *BIO1* play a pivotal role in acquired biotin prototrophy

To investigate if the high copy numbers of *BIO1* and *BIO6* contributed to the acquired biotin prototrophy of the evolved strains, these genes were expressed from multi-copy plasmids and under the control of strong, constitutive promoters, in a non-evolved, biotin-auxotrophic strain. Overexpression of *BIO1* together with *BIO6*, a situation that mimicked the amplification of both genes in the evolved strains, enabled growth in SMD without biotin, at a specific growth rate of 0.15 h^{-1} (**Table 3**). However, the same growth rate was obtained when only *BIO1* was overexpressed. Further, growth in biotin-free SMD was not observed in a strain that only overexpressed *BIO6*. These results indicated that increased copy numbers of *BIO1*, but not those of *BIO6*, contributed to the acquired biotin prototrophy of evolved strains.

Mutations in the membrane transporter genes *TPO1* and *PDR12* contribute to fast growth in biotin-free media

Although overexpression of *BIO1* in a non-evolved strain background enabled growth in biotin-free media, specific growth rates of the resulting strains were only about half of those observed in the fast-growing evolved strains. To further analyse the genetic basis for fast growth in biotin-free media, genome sequences of the evolved biotin-prototrophic strains were analysed in more detail. The four independently evolved strains harboured 6 to 11 single-nucleotide changes within coding regions relative to the parental strain (**Additional Table 1**). Three out of the four evolved strains harboured a nonsynonymous or nonsense mutation in *TPO1*, which encodes a plasma-membrane polyamine transporter (Albertsen *et al.* 2003). *TPO1* was the only gene that harboured SNPs in more than one strain. However, a nonsense mutation in the ABC-transporter gene *PDR12* in strain IMS0481 coincided with a nonsense mutation in *WAR1*, in strain IMS0480, which encodes a transcriptional activator of *PDR12* (Kren *et al.* 2003) (**Additional Table 1**). Inactivation of either *PDR12* or *WAR1* causes absence of Pdr12 from the plasma membrane (Kren *et al.* 2003). Pdr12 has been shown to export a wide range of monocarboxylic acids (C3 to C7) (Piper *et al.* 1998, Hazelwood *et al.* 2006). To investigate a possible role of the mutations in *PDR12* and/or *TPO1* in biotin prototrophy, *pdr12Δ, tpo1Δ*, and *pdr12Δ tpo1Δ* mutations were introduced in strains overexpressing *BIO1*. Deletion of *PDR12* or *TPO1* did not lead to biotin prototrophy in a strain with a single wild-type *BIO1* gene (**Table 3**). In contrast, deletion of either of these transporter genes greatly increased the specific growth rate of strains that overexpressed *BIO1* in biotin-free SMD, from 0.15 h^{-1} in a strain without transporter deletion, to up to 0.25 h^{-1} in a *pdr12Δ* strain (**Table 3**). Combination of both deletions did not lead to a further increase of the specific growth rate (**Table 3**).

A possible explanation for the observed impact of mutations in *TPO1* and *PDR12* is that they prevent export of an essential intermediate in the biotin biosynthesis pathway by the encoded transporters. Pimelic acid, a C7-dicarboxylic acid, is the substrate for Bio1 which, based on the amplification of *BIO1* observed in evolved strains, may catalyse a rate-controlling reaction in biotin synthesis. For many organic acids, a role of Pdr12 in their export from *S. cerevisiae* cells has been inferred from a strongly increased sensitivity of *pdr12Δ* to the acid at pH values below its pK_a due to weak organic acid uncoupling (Piper *et al.* 1998, Holyoak *et al.* 1999, Hazelwood *et al.* 2006). Pimelic acid supplementation of cultures grown at pH 4.5, which is below the pK_A values of pimelic acid ($pK_{A,1}$ = 4.71, $pK_{A,2}$ = 5.58), did not reveal an increased sensitivity of single or double deletion mutants in *PDR12* and/or *TPO1* (**Additional Figure 1**). However, growth of the reference strain CEN. PK113-7D at pH 4.5 was not inhibited by pimelic acid concentrations of up to 100 mM (data not shown), indicating that pimelic acid uptake rates were too low to cause weak organic acid uncoupling. The absence of an increased sensitivity in the deletion mutants therefore neither support nor exclude a possible role of Tpo1 and/or Pdr12 in pimelic acid export.

Discussion

Although the genome of *S. cerevisiae* CEN.PK113-7D contains a full set of *BIO* genes (*BIO1, BIO2, BIO3, BIO4, BIO5,* and *BIO6*)(Nijkamp *et al.* 2012a), the specific growth rate of this strain in biotin-free synthetic medium with glucose was at least 30-fold lower than in cultures supplemented with this vitamin. Although data obtained with this typical laboratory strain (Entian and Kötter 2007) cannot be directly extrapolated to natural isolates, such very low growth rates may well be relevant for survival in biotin-scarce natural environments. Independent laboratory evolution experiments with strain CEN. PK113-7D yielded four evolved strains, three of which grew equally fast in biotin-free medium as in biotin-supplemented cultures. Such a complete biotin prototrophy is rare amongst natural biotin prototrophic strains, whose growth on biotin-free media has been described to vary from weak to vigorous but typically is slower than in biotin-containing media (Wu *et al.* 2005, Hall and Dietrich 2007, Lin and Cronan 2011).

The mutations that were shown to contribute to full biotin prototrophy of the laboratory-evolved strains provide new insights into the genetic basis of this phenotype. An up to 40-fold amplification of the clustered *BIO1* and *BIO6* genes in the evolved strains was shown to reflect a key role of the copy number of *BIO1*, but not *BIO6*, in biotin prototrophy. Introduction of a multi-copy vector carrying *BIO1* in a non-evolved strain enabled it to grow on biotin-free medium, at a growth rate of ca. 40% of that observed in the fastest-

growing evolved strains. These results indicate that the pimeloyl-CoA synthetase Bio1 exerts a high degree of metabolic control over biotin biosynthesis in the non-evolved strain. Comparison of the predicted protein sequence of *BIO1* from strain CEN.PK113-7D did not identify amino acids that were not also found in *BIO1* genes of other *S. cerevisiae* strains, indicating that the requirement for *BIO1* amplification was unlikely to be due to a CEN.PK-specific, inferior *BIO1* allele.

Three of the four evolved strains contained a duplication of ChrI, which carries the native copies of *BIO1* and *BIO6*, while additional copies of these genes were found on neochromosomes resulting from translocation events. The plasticity of the yeast genome under selective growth conditions is further illustrated in one of the evolved strains (IMS0478), by the additional presence of two copies of ChrIX as well as of the left arm of ChrVIII. These results confirm that chromosomal rearrangements and copy number variation are key mechanisms for genetic adaptation in short-term laboratory evolution experiments (Wolfe and Shields 1997, de Kok *et al.* 2012). The biotin-prototrophic sake strain of *S. cerevisiae* in which *BIO6* was first discovered also contains copies of *BIO6* on multiple chromosomes (Wu *et al.* 2005). This genotypic similarity indicates that the genetic adaptations that enable sake yeasts to grow at the very low biotin concentrations in sake mash (Torigata 1968), at least partially overlap with those seen in the present laboratory-evolution study.

Surprisingly, loss-of-function mutations in the membrane-transporter genes *TPO1* and *PDR12*, as well as in *WAR1*, which encodes a positive regulator of *PDR12* (Kren *et al.* 2003, Schüller *et al.* 2004), had a strong positive effect on biotin prototrophy. Neither the polyamine transporter Tpo1 (Tomitori *et al.* 1999, Albertsen *et al.* 2003) nor the monocarboxylate exporter Pdr12 (Piper *et al.* 1998) have previously been associated with biotin synthesis. Although, in the case of Tpo1, direct mediation of monocarboxylic acid transport has not been demonstrated, both transporters have been implicated in this process, including, in the case of Pdr12, the transport of the C7-monocarboxylate heptatonic acid (Holyoak *et al.* 1999). If Pdr12 and/or Tpo1 export the C7-dicarboxylate pimelic acid, the precursor of biotin biosynthesis, the resulting decrease in intracellular pimelate concentration could reduce the flux through the rate-controlling pimeloyl-CoA synthetase (Bio1) reaction. Alternatively, Pdr12 and/or Tpo1 might catalyse the export of other key intermediates of biotin synthesis, such as the aminated compounds KAPA and DAPA. A mutation, in one of the evolved strains, in the KAPA transporter gene *BIO5* could also be consistent with this hypothesis.

In the case of Tpo1, an alternative explanation for its impact on biotin prototrophy is related to its role in spermidine homeostasis. Supplementation of this compound to a spermidine-deficient *S. cerevisiae* strain has been shown to cause an up to 14-fold upregulation of *BIO* genes (Chattopadhyay *et al.* 2009). At a pH of 5, which was used in the presented

laboratory evolution, Tpo1 has been reported to primarily catalyse polyamine export (Uemura *et al.* 2005). Moreover, inactivation of *TPO1* has been shown to lead to increased intracellular spermidine concentrations (Krüger *et al.* 2013, Kim *et al.* 2015). By enhancing the previously demonstrated induction of *BIO* genes upon biotin depletion (Wodicka *et al.* 1997, Wu *et al.* 2005, Madsen *et al.* 2015), accumulation of polyamines could contribute to increased growth rates in biotin-free media.

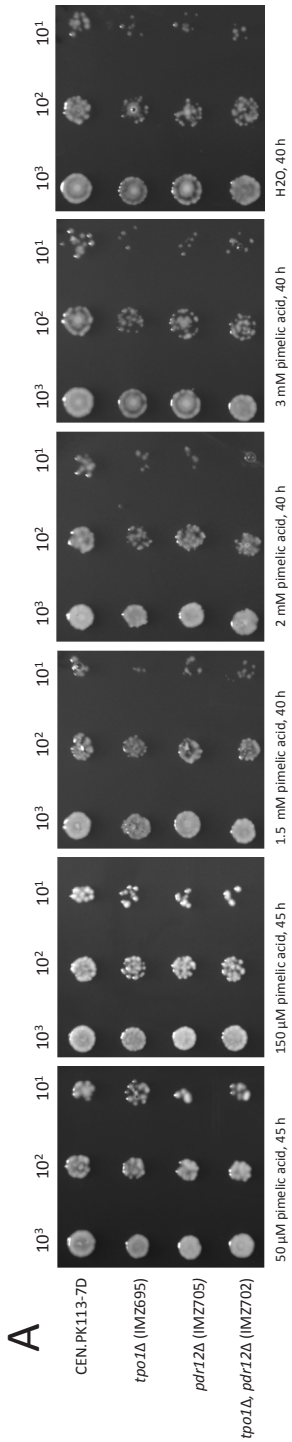
Elimination of vitamin requirements of *S. cerevisiae* could simplify design and scaling-up of fermentation processes and improve process economics. The demonstration that an *S. cerevisiae* strain that contains a basic complement of *BIO* genes can be evolved for complete biotin prototrophy opens up perspectives for the development of industrial *S. cerevisiae* strains that are completely prototrophic for this and, potentially, other vitamins. However, the extensive genomic rearrangements in the evolved strains complicates their direct use as metabolic engineering platforms. Overexpression of *BIO1*, combined with deletion of *TPO1* or *PDR12*, was sufficient to reach specific growth rates in biotin-free media that were only 40% lower than those observed in biotin-supplemented cultures. Null mutations in *TPO1* and *PDR12* had a similar effect on specific growth rates of *BIO1*-overexpressing strains in biotin-free media, but their effects were not additive (**Table 3**). While inactivation of *PDR12* strongly increases the sensitivity of *S. cerevisiae* to several apolar carboxylic acids (Piper *et al.* 1998), null mutations in *TPO1* have been reported to confer an increased tolerance to industrially relevant inhibitors, which in turn has been shown to result in higher productivities of industrial products (Kim *et al.* 2015, Kim *et al.* 2017). Inactivation of *TPO1* therefore appears to be the preferred intervention in strategies for eliminating biotin requirements in industrial yeast strains.

A further systematic analysis of the other mutations in evolved biotin prototrophic strains, combined with overexpression and/or codon optimization of *BIO1* and other *BIO* genes might allow additional improvements of rationally engineered biotin prototrophic *S. cerevisiae* strains. In addition, the fast-growing biotin prototrophic strains described in this study provide interesting experimental platforms for unravelling the elusive biochemistry of pimelate biosynthesis in *S. cerevisiae* and other yeasts (Lin and Cronan 2011, Magliano *et al.* 2011, Tanabe *et al.* 2011).

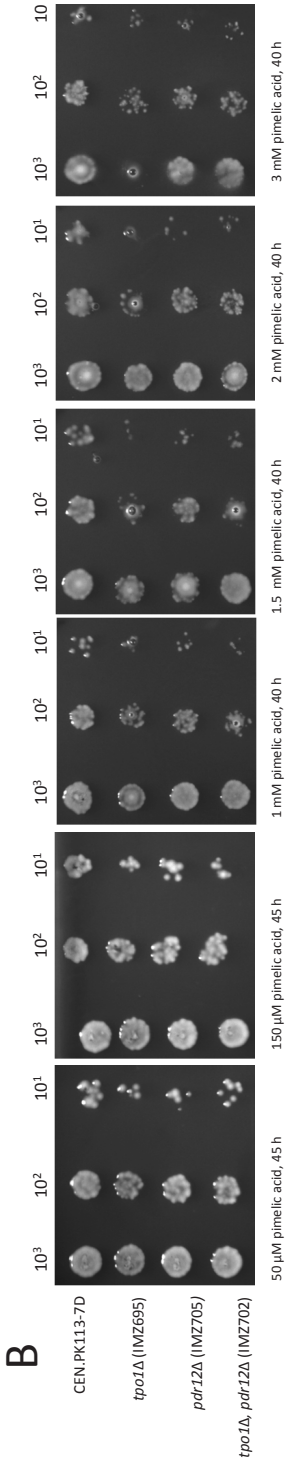
Acknowledgements

We thank Robert Mans for programming of the accelerostats, Erwin Suir for construction of strains IMK129 and IMK163, Sophie van der Horst for assisting with evolutionary engineering, Wijb Dekker and Stéphanie O'Herne for help with CHEF electrophoresis and Southern Blotting, and Maarten Verhoeven and Ioannis Papapetridis for fruitful discussions.

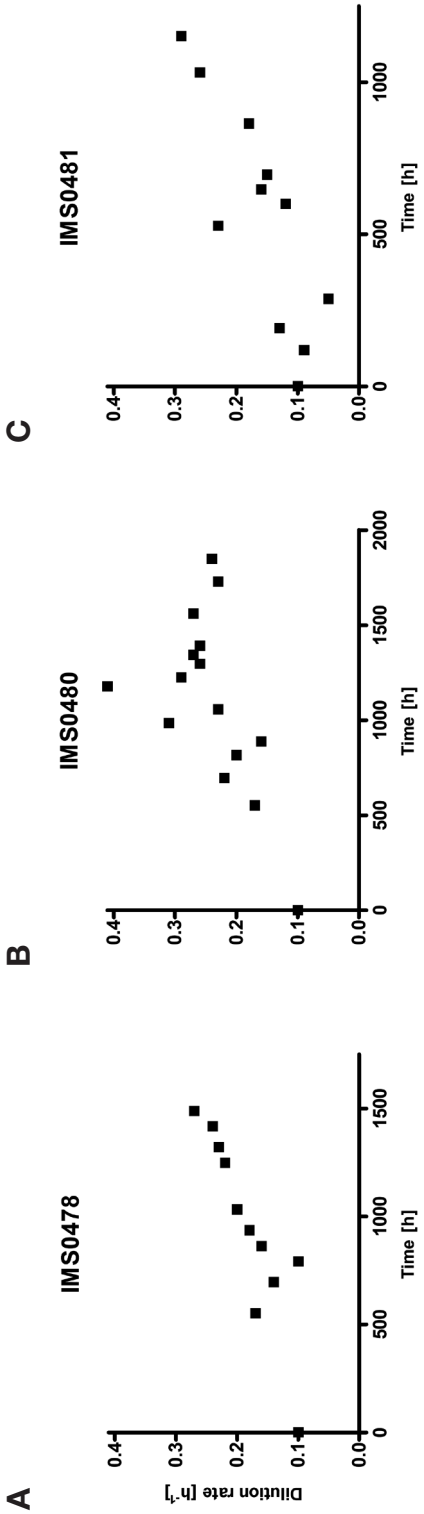
Additional material



Additional Figure 1. **Impact of tpo1Δ and pdr12Δ mutations on tolerance of S. *cerevisiae* to pimelic acid, sorbate, and spermidine.** S. *cerevisiae* strains were pregrown on liquid SMD and spotted on plates containing SMD (pH 4.5) **(A)** or in modified SMD (pH 4.5) in which the Mg^{2+} concentration was reduced to 50 µM to facilitate polyamine uptake **(B)**. Plates were supplemented with 50 µM, 150 µM, 1 mM, 1.5 mM, 2 mM or 3 mM pimelic acid, 0.5 mM potassium sorbate, 6 mM spermidine or water (control) as indicated. Strains and relevant genotypes are indicated on the left-hand side, estimated numbers of plated cells are indicated on top and incubation times at 30°C are indicated below the photographs. *Figure continues on next page*.



Additional Figure 1 – *Continued*



Additional Figure 2. Laboratory evolution of *S. cerevisiae* CEN.PK113-7D cells in accelerostat cultures for improved growth in biotin-free synthetic media. (A) IMS0478 **(B)** IMS0480 **(C)** IMS0481. Dilution rates were automatically feed-back controlled based on continuous analysis of the CO_2 output and the medium feed rate (see Methods). Dilution rates [h^{-1}] are shown for three independent experiments.

Additional Table 1. Mutations in evolved biotin prototrophic *S. cerevisiae* strains. Single nucleotide variants were called by comparison of whole-genome sequence data from the four strains evolved for biotin prototrophy with their parent strain CEN.PK113-7D. Listed are all nonsynonymous and nonsense mutations (*). Descriptions of gene functions were obtained from the *Saccharomyces* Genome Database website (www.yeastgenome.org). NT = nucleotide; AA = amino acid.

Gene	AA change	NT change	Description
IMS0478			
ULS1	Asp-130-Val	A389T	Swi2/Snf2-related translocase required for maintenance of NHEJ inhibition at telomeres. Plays role in antagonizing silencing during mating-type switching.
GPI11	Val-11-Gly	T32G	ER membrane protein involved in a late step of GPI anchor assembly. Involved in the addition of phosphoethanolamine to the multiply mannosylated glycosylphosphatidy-linositol (GPI) intermediate.
VIK1	Leu-577-Ile	T1729A	Protein that forms a kinesin-14 heterodimeric motor with Kar3p. Required for sister chromatid cohesion.
NMD2	Ser-567-*	C1700A	Protein involved in the nonsense-mediated mRNA decay (NMD) pathway. Involved in telomere maintenance.
THO2	Ser-53-*	C158A	Subunit of the THO complex. THO is required for efficient transcription elongation and involved in transcriptional elongation-associated recombination.
HSE1	Val-222-Ile	G664A	Subunit of the endosomal Vps27p-Hse1p complex. Required for sorting of ubiquitinated membrane proteins, as well as for recycling of Golgi proteins and formation of lumenal membranes.
IMS0480			
MAK10	Pro-42-Leu	C125T	Non-catalytic subunit of N-terminal acetyltransferase of the NatC type. Required for replication of dsRNA virus; expression is glucose-repressible.
TPO1	Asp-262-Gly	A785G	Polyamine transporter of the major facilitator superfamily. Member of the 12-spanner drug-H(+) antiporter DHA1 family. Recognizes spermine, putrescine, and spermidine. Catalyzes uptake of polyamines at alkaline pH and excretion at acidic pH.
ALD4	Gly-400-Ser	G1198A	Mitochondrial aldehyde dehydrogenase. Required for growth on ethanol and conversion of acetaldehyde to acetate.
BNI4	Ser-835-Pro	T2503C	Targeting subunit for Glc7p protein phosphatase. Localized to the bud neck, required for localization of chitin synthase III to the bud neck via interaction with the chitin synthase III regulatory subunit Skt5p.
WAR1	Glu-456-*	G1366T	Homodimeric Zn2Cys6 zinc finger transcription factor. Binds to a weak acid response element to induce transcription of *PDR12* and *FUN34*, encoding an acid transporter and a putative ammonia transporter, respectively.
CYR1	Glu-1879-Asp	A5673C	Adenylate cyclase. Required for cAMP production and cAMP-dependent protein kinase signalling.
BIO5	Ile-145-Thr	T434C	Putative transmembrane protein involved in the biotin biosynthesis. Responsible for uptake of 7-keto 8-aminopelargonic acid.

Additional Table 1 continues on next page.

Additional Table 1 – *Continued*

Gene	AA change	NT change	Description
IMS0481			
ASA1	Ile-436-Val	A1306G	Subunit of the ASTRA complex, involved in chromatin remodelling. Telomere length regulator involved in the stability or biogenesis of PIKKs such as TORC1.
TPO1	Asp-133-Asn	G397A	Polyamine transporter of the major facilitator superfamily. Member of the 12-spanner drug-H(+) antiporter DHA1 family. Recognizes spermine, putrescine, and spermidine. Catalyses uptake of polyamines at alkaline pH and excretion at acidic pH.
VPS15	Phe-452-Cys	T1355G	Serine/threonine protein kinase involved in vacuolar protein sorting. Functions as a membrane-associated complex with Vps34p.
SPT21	Glu-581-Lys	G1741A	Protein with a role in transcriptional silencing. Required for normal transcription at several loci including HTA2-HTB2 and HHF2-HHT2, but not required at the other histone loci.
VHT1	Gly-288-Ala Asn-174-Lys	G863C T522G	High-affinity plasma membrane H+-biotin (vitamin H) symporter. Mutation results in fatty acid auxotrophy.
BCY1	Val-262-Phe	G784T	Regulatory subunit of the cyclic AMP-dependent protein kinase (PKA). PKA is a component of a signalling pathway that controls a variety of cellular processes, including metabolism, cell cycle, stress response, stationary phase, and sporulation.
ALP1	Phe-422-Ser	T1265C	Arginine transporter. Expression is normally very low and it is unclear what conditions would induce significant expression.
PDE1	Gly-41-Cys	G121T	Low-affinity cyclic AMP phosphodiesterase. Controls glucose and intracellular acidification-induced cAMP signalling, target of the cAMP-protein kinase A (PKA) pathway.
PDR12	Trp-765-*	G2295A	Plasma membrane ATP-binding cassette (ABC) transporter. Weak-acid-inducible multidrug transporter required for weak organic acid resistance. Induced by sorbate and benzoate and regulated by War1p.
CIT2	Ala-289-Val	C866T	Citrate synthase, peroxisomal isozyme involved in glyoxylate cycle. Catalyzes condensation of acetyl coenzyme A and oxaloacetate to form citrate.
REC104	Ser-76-Tyr	C227A	Protein involved in early stages of meiotic recombination. Required for meiotic crossing over.
IMS0496			
CSS3	Asp-117-Tyr	G349T	Protein of unknown function, secreted when constitutively expressed. Deletion mutants are viable and have elevated levels of Ty1 retrotransposition and Ty1 cDNA.
TPO1	Trp-302-*	G905A	Polyamine transporter of the major facilitator superfamily. Member of the 12-spanner drug-H(+) antiporter DHA1 family. Recognizes spermine, putrescine, and spermidine. Catalyzes uptake of polyamines at alkaline pH and excretion at acidic pH.
TDA6	Pro-38-Arg	C113G	Putative protein of unknown function.
NCA3	Ala-289-Gly	C866G	Protein involved in mitochondrion organization.

Additional Table 1 continues on next page.

Additional Table 1 – *Continued*

Gene	AA change	NT change	Description
RPN1	Lys-943-*	A2827T	Non-ATPase base subunit of the 19S RP of the 26S proteasome.
FCJ1	Glu-234-Asp	A702C	Component of the MICOS complex. MICOS is a mitochondrial inner membrane complex that extends into the intermembrane space and has a role in the maintenance of crista junctions, inner membrane architecture, and formation of contact sites to the outer membrane.

Additional Table 2. mRNA levels of BIO genes in strains evolved for biotin prototrophy grown in the presence (+) or absence (-) of biotin prior to mRNA extraction | (+) Relative expression levels of *BIO1, BIO2, BIO3, BIO4*, and *BIO6* measured in the parent strain CEN.PK113-7D and the evolved strains IMS0478, IMS0480, IMS0481, and IMS0496 upon growth in the presence of biotin supplementation. **(-)** Relative expression levels of *BIO1, BIO2, BIO3, BIO4*, and *BIO6* measured in the evolved strains IMS0478, IMS0480, IMS0481, and IMS0496 upon growth in the absence of biotin supplementation. The parent strain was unable to grow in absence of biotin. All qPCR experiments were carried out on duplicate cultures, with analytical triplicates for each culture. Relative expression levels were determined according to the ΔΔCT method (Schmittgen and Livak 2008) and indicated with the standard error of the mean (SEM) of duplicate analyses. A graphical representation of the data can be found in **Figure 3.**

	Biotin	IMS0478	IMS0480	IMS0481	IMS0496	CEN.PK113-7D
BIO1	+	0.8 ± 0.0	0.4 ± 0.1	0.4 ± 0.0	0.8 ± 0.1	0.1 ± 0.0
	-	16.2 ± 2.3	1.8 ± 0.7	2.7 ± 0.1	24.4 ± 0.3	-
BIO2	+	0.1 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.3 ± 0.0	0.1 ± 0.0
	-	2.2 ± 0.3	0.2 ± 0.1	0.6 ± 0.1	5.0 ± 0.5	-
BIO3	+	0.2 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
	-	6.4 ± 0.7	0.0 ± 0.0	0.1 ± 0.0	1.1 ± 0.0	-
BIO4	+	0.2 ± 0.0	0.0 ± 0.0	0.1 ± 0.0	0.1 ± 0.0	0.1 ± 0.0
	-	4.8 ± 0.7	0.0 ± 0.0	0.6 ± 0.1	1.7 ± 0.1	
BIO6	+	0.2 ± 0.0	0.0 ± 0.0	0.0 ±0.0	0.2 ± 0.0	0.0 ± 0.0
	-	5.2 ± 0.8	0.1 ± 0.0	0.6 ± 0.0	4.1 ± 0.0	-

Additional Table 3. Primers used in this study can be found at: https://doi.org/10.1128/AEM.00892-17

References

Akao T, Yashiro I, Hosoyama A *et al.* Whole-genome sequencing of sake yeast *Saccharomyces cerevisiae* Kyokai no. 7. *DNA Res* 2011;**18**:423-34.

Albertsen M, Bellahn I, Krämer R *et al.* Localization and function of the yeast multidrug transporter Tpo1p. *J Biol Chem* 2003;**278**:12820-5.

Azhar A, Booker GW, Polyak SW. Mechanisms of Biotin Transport. *Biochem Anal Biochem* 2015;**4**:1.

Burkholder PR, McVeigh I, Moyer D. Studies on some growth factors of yeasts. *J Bacteriol* 1944;**48**:385.

Canelas AB, Harrison N, Fazio A *et al.* Integrated multilaboratory systems biology reveals differences in protein metabolism between two reference yeast strains. *Nat Commun* 2010;**1**:145.

Chattopadhyay MK, Chen W, Poy G *et al.* Microarray studies on the genes responsive to the addition of spermidine or spermine to a *Saccharomyces cerevisiae* spermidine synthase mutant. *Yeast* 2009;**26**:531-44.

Daran-Lapujade P, Rossell S, van Gulik WM *et al.* The fluxes through glycolytic enzymes in *Saccharomyces cerevisiae* are predominantly regulated at posttranscriptional levels. *Proc Natl Acad Sci USA* 2007;**104**:15753-8.

de Kok S, Nijkamp JF, Oud B *et al.* Laboratory evolution of new lactate transporter genes in a *jen1Δ* mutant of *Saccharomyces cerevisiae* and their identification as *ADY2* alleles by whole-genome resequencing and transcriptome analysis. *FEMS Yeast Res* 2012;**12**:359-74.

Entian K-D, Kötter P. 25 Yeast genetic strain and plasmid collections. *Methods Microbiol* 2007;**36**:629-66.

Gailiusis J, Rinne RW, Benedict C. Pyruvate—Oxaloacetate exchange reaction in baker's yeast. *Biochim Biophys Acta* 1964;**92**:595-601.

Gasser B, Dragosits M, Mattanovich D. Engineering of biotin-prototrophy in *Pichia pastoris* for robust production processes. *Metab Eng* 2010;**12**:573-80.

Gietz RD, Woods RA. Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method. *Methods Enzymol* 2002;**350**:87-96.

Güldener U, Heck S, Fiedler T *et al.* A new efficient gene disruption cassette for repeated use in budding yeast. *Nucleic Acids Res* 1996;**24**:2519-24.

Hall C, Brachat S, Dietrich FS. Contribution of horizontal gene transfer to the evolution of *Saccharomyces cerevisiae*. *Eukaryot Cell* 2005;**4**:1102-15.

Hall C, Dietrich FS. The reacquisition of biotin prototrophy in *Saccharomyces cerevisiae* involved horizontal gene transfer, gene duplication and gene clustering. *Genetics* 2007;**177**:2293-307.

Hazelwood LA, Tai SL, Boer VM *et al.* A new physiological role for Pdr12p in *Saccharomyces cerevisiae*: export of aromatic and branched-chain organic acids produced in amino acid catabolism. *FEMS Yeast Res* 2006;**6**:937-45.

Helliwell KE, Wheeler GL, Smith AG. Widespread decay of vitamin-related pathways: coincidence or consequence? *Trends Genet* 2013;**29**:469-78.

Hoja U, Wellein C, Greiner E *et al.* Pleiotropic phenotype of acetyl-CoA-carboxylase-defective yeast cells. *Eur J Biochem* 1998;**254**:520-6.

Holyoak CD, Bracey D, Piper PW *et al.* The *Saccharomyces cerevisiae* weak-acid-inducible ABC transporter Pdr12 transports fluorescein and preservative anions from the cytosol by an energy-dependent mechanism. *J Bacteriol* 1999;**181**:4644-52.

Huisjes EH, Luttik MA, Almering MJ *et al.* Toward pectin fermentation by *Saccharomyces cerevisiae*: expression of the first two steps of a bacterial pathway for D-galacturonate metabolism. *J Biotechnol* 2012;**162**:303-10.

Inoue H, Nojima H, Okayama H. High efficiency transformation of *Escherichia coli* with plasmids. *Gene* 1990;**96**:23-8.

Kashiwagi K, Igarashi K. Identification and assays of polyamine transport systems in *Escherichia coli* and *Saccharomyces cerevisiae*. *Methods Mol Biol* 2011;**720**:295-308.

Kim HS, Hoja U, Stolz J *et al.* Identification of the tRNA-binding protein Arc1p as a novel target of *in vivo* biotinylation in *Saccharomyces cerevisiae*. *J Biol Chem* 2004;**279**:42445-52.

Kim SK, Jin YS, Choi IG *et al.* Enhanced tolerance of *Saccharomyces cerevisiae* to multiple lignocellulose-derived inhibitors through modulation of spermidine contents. *Metab Eng* 2015;**29**:46-55.

Kim SK, Jo JH, Jin YS *et al.* Enhanced ethanol fermentation by engineered *Saccharomyces cerevisiae* strains with high spermidine contents. *Bioprocess Biosyst Eng* 2017:1-9.

Kren A, Mamnun YM, Bauer BE *et al.* War1p, a novel transcription factor controlling weak acid stress response in yeast. *Mol Cell Biol* 2003;**23**:1775-85.

Krüger A, Vowinckel J, Mülleder M *et al.* Tpo1-mediated spermine and spermidine export controls cell cycle delay and times antioxidant protein expression during the oxidative stress response. *EMBO rep* 2013;**14**:1113-9.

Kuijpers N, Solis-Escalante D, Bosman L *et al.* A versatile, efficient strategy for assembly of multi-fragment expression vectors in *Saccharomyces cerevisiae* using 60 bp synthetic recombination sequences. *Microb Cell Fact* 2013;**12**:47.

Kuyper M, Hartog MM, Toirkens MJ *et al.* Metabolic engineering of a xylose-isomerase-expressing *Saccharomyces cerevisiae* strain for rapid anaerobic xylose fermentation. *FEMS Yeast Res* 2005a;**5**:399-409.

Kuyper M, Toirkens MJ, Diderich JA *et al.* Evolutionary engineering of mixed-sugar utilization by a xylose-fermenting *Saccharomyces cerevisiae* strain. *FEMS Yeast Res* 2005b;**5**:925-34.

Lin S, Cronan JE. Closing in on complete pathways of biotin biosynthesis. *Mol Biosyst* 2011;**7**:1811-21.

Lin S, Hanson RE, Cronan JE. Biotin synthesis begins by hijacking the fatty acid synthetic pathway. *Nat Chem Biol* 2010;**6**:682-8.

Lõoke M, Kristjuhan K, Kristjuhan A. Extraction of genomic DNA from yeasts for PCR-based applications. *BioTechniques* 2011;**50**:325.

Losada M, Canovas J, Ruiz A. Oxaloacetate, citramalate and glutamate formation from pyruvate in baker's yeast. *Biochem Z* 1964;**340**:60-74.

Madsen CT, Sylvestersen KB, Young C *et al.* Biotin starvation causes mitochondrial protein hyperacetylation and partial rescue by the SIRT3-like deacetylase Hst4p. *Nat Commun* 2015;**6**.

Magliano P, Flipphi M, Arpat BA *et al.* Contributions of the peroxisome and β-oxidation cycle to biotin synthesis in fungi. *J Biol Chem* 2011;**286**:42133-40.

Mans R, van Rossum HM, Wijsman M *et al.* CRISPR/Cas9: a molecular Swiss army knife for simultaneous introduction of multiple genetic modifications in *Saccharomyces cerevisiae*. *FEMS Yeast Res* 2015;**15**::fov004.

Mortimer RK, Johnston JR. Genealogy of principal strains of the yeast genetic stock center. *Genetics* 1986;**113**:35-43.

Mumberg D, Müller R, Funk M. Yeast vectors for the controlled expression of heterologous proteins in different genetic backgrounds. *Gene* 1995;**156**:119-22.

Nijkamp JF, van den Broek M, Datema E *et al. De novo* sequencing, assembly and analysis of the genome of the laboratory strain *Saccharomyces cerevisiae* CEN. PK113-7D, a model for modern industrial biotechnology. *Microb Cell Fact* 2012a;**11**:36.

Nijkamp JF, van den Broek MA, Geertman J-MA *et al. De novo* detection of copy number variation by co-assembly. *Bioinformatics* 2012b;**28**:3195-202.

Oud B, Guadalupe-Medina V, Nijkamp JF *et al.* Genome duplication and mutations in *ACE2* cause multicellular, fast-sedimenting phenotypes in evolved *Saccharomyces cerevisiae*. *Proc Natl Acad Sci U S A* 2013;**110**:E4223-E31.

Paalme T, Kahru A, Elken R *et al.* The computer-controlled continuous culture of *Escherichia coli* with smooth change of dilution rate (A-stat). *J Microbiol Methods* 1995;**24**:145-53.

Phalip V, Kuhn I, Lemoine Y *et al.* Characterization of the biotin biosynthesis pathway in *Saccharomyces cerevisiae* and evidence for a cluster containing *BIO5*, a novel gene involved in vitamer uptake. *Gene* 1999;**232**:43-51.

Piper MD, Daran-Lapujade P, Bro C *et al.* Reproducibility of oligonucleotide microarray transcriptome analyses - an interlaboratory comparison using chemostat cultures of *Saccharomyces cerevisiae*. *J Biol Chem* 2002;**277**:37001-8.

Piper P, Mahé Y, Thompson S *et al.* The Pdr12 ABC transporter is required for the development of weak organic acid resistance in yeast. *EMBO J* 1998;**17**:4257-65.

Pronk JT. Auxotrophic yeast strains in fundamental and applied research. *Appl Environ Microbiol* 2002;**68**:2095-100.
Schmitt ME, Brown TA, Trumpower BL. A rapid and simple method for preparation of RNA from *Saccharomyces cerevisiae*. *Nucleic Acids Res* 1990;**18**:3091.
Schmittgen TD, Livak KJ. Analyzing real-time PCR data by the comparative CT method. *Nat Protoc* 2008;**3**:1101-8.
Schüller C, Mamnun YM, Mollapour M *et al.* Global phenotypic analysis and transcriptional profiling defines the weak acid stress response regulon in *Saccharomyces cerevisiae*. *Mol Biol Cell* 2004;**15**:706-20.
Shevchuk NA, Bryksin AV, Nusinovich YA *et al.* Construction of long DNA molecules using long PCR-based fusion of several fragments simultaneously. *Nucleic Acids Res* 2004;**32**:e19-e.
Solis-Escalante D, Kuijpers NG, Bongaerts N *et al. amdSYM*, a new dominant recyclable marker cassette for *Saccharomyces cerevisiae*. *FEMS Yeast Res* 2013;**13**:126-39.
Stolz J, Hoja U, Meier S *et al.* Identification of the plasma membrane H+-biotin symporter of *Saccharomyces cerevisiae* by rescue of a fatty acid-auxotrophic mutant. *J Biol Chem* 1999;**274**:18741-6.
Streit W, Entcheva P. Biotin in microbes, the genes involved in its biosynthesis, its biochemical role and perspectives for biotechnological production. *Appl Microbiol Biotechnol* 2003;**61**:21-31.
Sumrada RA, Cooper TG. Urea carboxylase and allophanate hydrolase are components of a multifunctional protein in yeast. *J Biol Chem* 1982;**257**:9119-27.
Tanabe Y, Maruyama J-i, Yamaoka S *et al.* Peroxisomes are involved in biotin biosynthesis in *Aspergillus* and *Arabidopsis*. *J Biol Chem* 2011;**286**:30455-61.
Tomitori H, Kashiwagi K, Sakata K *et al.* Identification of a gene for a polyamine transport protein in yeast. *J Biol Chem* 1999;**274**:3265-7.
Torigata K, and Y. Akiyama. Tests of sake brewing by yeasts after cultured with ventilation. 1. Rising and falling of vitamins contained in sake moromi and preservative tests of yeasts. *J Brew Soc Japan* 1968;**63**:60-3.
Uemura T, Tachihara K, Tomitori H *et al.* Characteristics of the polyamine transporter *TPO1* and regulation of its activity and cellular localization by phosphorylation. *J Biol Chem* 2005;**280**:9646-52.
Van Rossum HM, Kozak BU, Niemeijer MS *et al.* Requirements for carnitine shuttle-mediated translocation of mitochondrial acetyl moieties to the yeast cytosol. *MBio* 2016;**7**:e00520-16.
Verduyn C, Postma E, Scheffers WA *et al.* Energetics of *Saccharomyces cerevisiae* in anaerobic glucose-limited chemostat cultures. *Microbiology* 1990;**136**:405-12.
Verhoeven MD, Lee M, Kamoen L *et al.* Mutations in *PMR1* stimulate xylose isomerase activity and anaerobic growth on xylose of engineered *Saccharomyces cerevisiae* by influencing manganese homeostasis. *Sci Rep, in press*.
Wakil SJ, Titchener EB, Gibson DM. Evidence for the participation of biotin in the enzymic synthesis of fatty acids. *Biochim Biophys Acta* 1958;**29**:225-6.
Walker BJ, Abeel T, Shea T *et al.* Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One* 2014;**9**:e112963.
Wodicka L, Dong H, Mittmann M *et al.* Genome-wide expression monitoring in *Saccharomyces cerevisiae*. *Nat Biotechnol* 1997;**15**:1359-67.
Wolfe KH, Shields DC. Molecular evidence for an ancient duplication of the entire yeast genome. *Nature* 1997;**387**:708-12.
Wu H, Ito K, Shimoi H. Identification and characterization of a novel biotin biosynthesis gene in *Saccharomyces cerevisiae*. *Appl Environ Microbiol* 2005;**71**:6845-55.
Yang Y, Lang N, Yang G *et al.* Improving the performance of solventogenic *clostridia* by reinforcing the biotin synthetic pathway. *Metab Eng* 2016;**35**:121-8.

Outlook

Compared to the well-developed processes for industrial first-generation ethanol production, second generation ethanol production is still a nascent technology. Just as high-efficiency petrochemical refineries did not appear overnight, optimizing the performance of the current plants does not only depend on advances in metabolic engineering of "yeast cell factories" – which were the subject of this thesis – but also on further process engineering breakthroughs. For second-generation biofuels to make a sizeable contribution to the reduction of greenhouse emissions and global warming, their production must be able to compete on cost with classical petrochemistry-based transport fuels. In addition to further progress in science and technology, policy decisions can have a significant impact on the near- and short-term profitability of second-generation biofuels. In particular, based on the lower carbon footprint of second-generation biofuels compared to petrochemical fuels, implementation of CO_2 credits for companies and products could rapidly and decisively increase their economic competitiveness.

Metabolic engineering strategies to further improve yeast performance in second generation bioethanol processes are already being explored. Non-*Saccharomyces* yeasts with industrially interesting inherent properties, such as low-pH or high-temperature tolerance are being investigated and an increase in their genetic accessibility via the implementation of CRISPR-based genome editing technologies may pave the way towards the usage of non-*Saccharomyces* yeast in the future. Alternative research efforts focus on transferring advantageous properties into the well-studied host *Saccharomyces cerevisiae* via rational metabolic engineering or adaptive laboratory evolution.

Multiple studies have shown that pentose sugars are not being recognized as fermentable substrates by *S. cerevisiae.* Consequently, once all yeast-native substrates are metabolized

during a fermentation, the elicited effect on global gene expression levels affecting transporter landscapes and pathway (de-)regulations are, in most cases, unfavorable for the fermentation of non-native carbon sources. Manipulation of glucose-sensing and signaling may greatly improve co-consumption of glucose and non-native substrates resulting in increased ethanol productivities. Furthermore, as production volumes increase, the economic relevance of the conversion of minor, potentially fermentable substrates that occur in lignocellulosic hydrolysates, such as galacturonic acid and deoxysugars, will increase. Co-feeding of additional, low-value carbon sources can be explored as a strategy to further increase ethanol yields. For example, glycerol, derived from fermentation stills or biodiesel manufacturing can be considered as a potential co-substrate. Fixation of CO_2 as an external electron acceptor or the reduction of the inhibitor acetic acid to ethanol by *S. cerevisiae* have already been achieved and may become integral parts of industrial strains.

A potentially transformative technological innovation is consolidated bioprocessing, i.e., the combination of microbial production of saccharolytic enzymes for hydrolysis of pretreated biomass with fermentation of the released sugars in a single bioreactor. When successful at industrial scale, this approach may greatly lower the capital and operational costs of bioethanol production plants.

Current scenarios for ‘zero carbon’ economies foresee that liquid fuels for combustion engines (including plane and ship engines) will continue to be used at the end of this century (2100), even when road transport will progressively be electrified. Until the latter technology has advanced sufficiently and the current infrastructure has been adapted worldwide, society will depend on a readily available fossil fuel alternative, such as biofuels, that is compatible with our existing infrastructure and vehicles.

In parallel with further developments in carbohydrate-based production of biofuels, technologies for harnessing alternative energy sources are advancing steadily, and the subsequent expected drop in electricity prices will create new opportunities for microbial biotechnology. For instance, hydrogen, obtained via electricity-powered electrolysis of water, can serve as a cheap electron donor. Furthermore, surplus electricity generated from sunlight or wind could be stored via electrochemical reduction of dissolved CO_2 into formic acid. The formic acid in turn, may serve as a potential carbon source for metabolically engineered microorganisms. It is likely that the insight gained and technologies developed for the production of 2nd generation bioethanol by *S. cerevisiae* will serve as a “case-study” and pave the way for the biotechnological production of other products via conversion of agricultural and household wastes, such as pharmaceuticals, bioplastics, and ingredients for the food and cosmetic industry.

Thanks to (i) the fast advances in strain construction tools via CRISPR-based genome editing tools, (ii) dropping prices for whole genome sequencing, (iii) dropping prices for *de novo* DNA synthesis, and (iv) the emerging opportunities in implementation of robotics and artificial intelligence to increase “man-power” and tremendously accelerate stain construction and testing, opportunities for the field of metabolic engineering have never been higher. To quote a former, famous Delft microbiology professor: “Happy are those who start now”!

Acknowledgements

In the past four years, when mentioning to others that I was doing a PhD, many have said that it must be hard to work on a project for four years all by myself. On these occasions, I answered that I was not working alone at all during those four years. Now that I am approaching the end of my PhD, I realise that, what *is* hard, is that I will miss all the people that I had the privilege to work and interact with during my time as a PhD student in Delft. I would like to herewith thank everyone, in a spontaneous order, and apologize in case I missed some names.

First and foremost, I would like to thank **Jack, Ton, Jean-Marc** and **Pascale** for having had the trust in me to offer me this exciting PhD project. Thank you, **Ton** for giving me very clear and rational guidance at the start of my project. Having been a little less rational than you, this really helped me with taking the right decisions and not "getting lost" during the slightly overwhelming start of a PhD. I also valued your "office walk-in" policy which helped me to, at least somewhat, become accustomed to the much flatter hierarchies in The Netherlands (which still turned out to be easier to get used to than the geographical flatness!). Thank you, **Jack**, for taking over full supervision with such enthusiasm after Ton moved to Sweden. You always managed to free up your busy schedule for regular subgroup meetings and *very fast* manuscript feedback. You were a tremendous help during my last year when I had to take difficult project-related decisions and you went the extra mile with me in order that I was able finish in time! I value that you encouraged me to address concerns to you, which everytime I did, made me leave your office stronger and optimistic.

Thank you, **Dani, Maarten, Ioannis, Anja, Frank, Harmen, Robert, Tim, Barbara, Stefan, Matthijs, Pilar, Nick, Susan, Niels** and **Angel** for a warm welcome at IMB when I started!

Having such a welcoming team that, despite setting high-standards, managed to radiate a homey atmosphere and willingness to help, tremendously helped with getting started at IMB. **Maarten**, thank you for being a very supportive and educative, close colleague who clearly made my start at IMB much smoother than it would have been without a project mate – I truly enjoyed our cosy office sessions back in the old Kluyver lab! It was a great pleasure to have a close project mate with whom I could discuss ideas and weird theories, share successes and failed experiments, prepare for meetings together the evening before in the office and celebrate successes together!! Not to forget everything you taught me about The Netherlands, Dutch history and Dutch culture. I'd like to thank the **Ethanolics team**, Ioannis, Maarten, Laura, Jack and Ton, for all the input in my project. I learned a lot from those meetings where we all shared our results and thoughts on our related projects! **Robert** thank you for the endless sparring of potential startup ideas which were our plan B, C, and D "in case we would not make it in science" and thank you for sharing a very peculiar type of daily humour with me! **Erik**, thank you for all the fermentation-troubleshooting help and your patience with me, especially when I started as a PhD student that had never run a fermentation ever before. Thank you also for all the climbing tips and tricks and organizing the bouldering sessions in Delft! Thank you, **Marcel** for your patient help with analysing my whole genome sequencing data - it was very exciting for me to get insight and hands-on experience in genome mapping. I also enjoyed our less professional discussions about hiking and meditation! Thank you **Alex** for answering lots of bioinformatics questions I had when I started! Thank you, **Pilar** for keeping the molecular lab so well organized. Even more, I valued you as a colleague and friend very much and hope we will stay in contact. Thank you, **Marijke**, for managing our analytical labs and help with setting up the dusty CHEF gel equipment back in the Kluyver lab! Thank you **Jolanda** for helping me with my very last two fermentations! Absolutely no science would happen in our labs without the MSD team that carefully provides most of the utilities – thank you **Apilena, Janie,** and **Astrid.** Many thanks also for the countless times you postponed the autoclave time for me when I still needed the classic "five more minutes" to finish preparing my bioreactors.

I would like to thank my collaboration partners in Groningen, **Arnold, Jeroen, Dick, Misun & Hanna,** for our joint work and the inputs in my project. Special thanks for hosting me and Maarten to carry out transport assays in your labs. Many thanks to my collaboration partners at DSM, **Paul K., Paul dW., Hans, Mickel, Sander,** and **Roel** – I valued the insight you provided me into the application of my project - knowing that my findings can potentially directly contribute to 2nd generation biofuel production via DSM greatly fuelled my enthusiasm and energy I had for my project.

Supervision of students goes far beyond the pleasure of having someone that runs your PCRs for you and cleans your bioreactors. Working together with my students allowed me

to realize that I like to work in a team and share my knowledge and enthusiasm with other people. In this sense, I would like to thank all my students that worked with me. Thank you, **Sophie**, **Stephanie**, and **Charlotte** for working with me on the biotin project and being my very first students (where I had to overplay that I myself still felt like a student). Thank you **Wijb** and **Oscar** for contributing a lot to my main research line!

Thank you, **Arthur** for being a colleague and friend to have fun with inside and outside of work. Paschmini valued your trust in her, all the lunch, tea, or sushi sessions (and beyond) and that you are a person she was able to count on! **Nicolo,** already when I "interviewed" you during an informal lunch at the TU, we had a nice and fun time which continued throughout the rest of our colleagueship! Arhur and Nicolo, I am really happy you accepted to be my Paranymphs – really appreciated! Thank you, **Xavier** for sharing offices during my last few months and entertaining me with the "which vegetable is he eating"-game ("or is it a fruit?"). Thank you **Aurin** for helping me with translating my English Thesis Summary to Dutch! Thank you **Jonna, Anna, Paola, Anja** and **Marlous** for our cosy gatherings. As my paragraph is getting longer and longer, I will herewith thank all remaining current IMB members – you are a lovely crowd! Let's not let those friendships end after my graduation.

Even though I was supported from afar, I would not have been able to complete my project in Delft without a nice environment outside of the lab! Many thanks to my housemates **Dawn, Zita, Hester, Marta, Sofia, Helena, and Ella** - I miss our cosy huisje**! Sabine, Norbert, Fardin and Nikos,** I am so glad we all shared our journies since we met at the very beginning at the "PhD startup course". I would like to thank the **Lindy hop community**, and especially **Hedwig**, **Lauris, and Eve** for all the dancing and for giving me a feeling of integration in The Netherlands. I hope we will keep dancing and being friends beyond my PhD time in Delft. **Martijn** thank you for turning me into a yoga-enthusiast. Thank you **Julian** for the countless gym classes we attended together and for (trying to) convince me that surfing in the cold and stormy waters of Scheveningen is fun.

A lot of people were *very* important for me long before I was embarking on my PhD adventure. I would like to thank not only the people that supported me educationally, but also the ones that enabled me to get to who and where I am. **Mami** and **Papi**, you were a great support during both, my studies and my PhD and you never got tired of listening to and supporting me when I was trying to "find myself". Thank you for raising me in such a snug environment where I was able to develop my awareness and curiosity for nature and for teaching me the priceless ability to enjoy the little things in life. **Opa, Oma, Grossmami** & **Grosspapi,** I really appreciate that I had the chance to have you as such supportive and caring grandparents that always believed in me and gave me strength. Thank you Grossmami for being an amazing example for being "small but powerful". **Bianca** and **Anna**, thank you for supporting me from afar, for all the endless phone calls,

your visits in Delft and for making me still feel home whenever I was visiting back home. Thank you **Simon, Michi, Petra, Kim, Selina K, Selina Z, Aaron, Jöni, and Ralph** for always welcoming me back home without giving me any feeling of losing connection by having moved abroad (and for promising me 1000 times that this won't happen). Thank you, **Yannick** for being a very important person for me in my early years. **Domi,** I am very grateful for everything we shared during our study time and for respecting and supporting my decision to perform my PhD research in The Netherlands.

Thank you **Heinz** and **Christian** for being very motivating Chemistry and Biology teachers, respectively, which clearly influenced my decision on my field of study. Thank you **Jake** for accepting my wish to come to your newly set up lab in Norwich as an intern and introducing the unexperienced me to "real life" science ("always triple check") – looking back, this must have been a challenge! Special thanks to my MSc thesis supervisor **Conrad**, for significantly contributing to raising me as a young scientist and patiently and professionally guiding me through my fist (clumsy) steps in the lab (those Western Blots!). Thanks for preparing me for my PhD ("The difference between being a MSc or PhD student is that as a MSc student you swim in a pond wearing swimmies, and as a PhD student you get thrown into a swimming pool") – but I now learned to swim! Similarly, thank you **Christian** and **Patrick** for making our lab atmosphere very "homey" – it felt like the living room of a shared house, but with running lots of experiments in the kitchen. Thank you, **Giovanni, Thomas, Fanny, Sumire**, **Harald, Patrick and Corina** for helping me with my first steps with yeast biotechnology and especially with answering all my industry-and career-related questions to help me decide to pursue a PhD which paved my way to Delft.

Ioannis, thank you for teaching me Greek-style "savoir vivre" and for putting a smile on my face every time I was struggling with something. You really helped me through the harder times of my thesis. You'll always have a special place in my heart and I am incredibly proud of you.

Curriculum vitae

Jasmine Melanie Bracher was born on October 10, 1989 in Basel, Switzerland. She followed her pre-university education at the "Gymnasium Kirschgarten" in Basel and graduated in 2008. Jasmine subsequently enrolled for the Bachelor studies in Molecular Biology at the University of Basel. During the latter, she completed a 1-month research internship at the John Innes Centre of the University of East Anglia, Norwich (GB), where she joined the group of Dr. Jacob Malone working on the production of deletion constructs for selected genes in *Steptomyces coelicolor.* In 2012, she continued at the University of Basel with the MSc programme in Molecular Biology, with a Major in Biochemistry and Cell Biology. For her MSc thesis, Jasmine joined the group of Prof. dr. Erich A. Nigg for 1.5 years at the Laboratory of Cell Cycle Control Signalling and Genome Stability at the Department of Biochemistry and Cell Biology of the Biozentrum in Basel. Directly supervised by Dr. Conrad von Schubert, Jasmine was working on a project focusing on cell cycle control signalling in human cells, in particular the role and regulation of the spindle assembly checkpoint kinases Mps1 and Plk1 which are involved in cancer development. After her graduation in 2013, she continued with a 6-month internship at the company Evolva, where she got trained in yeast biotechnology and developed her interest in industrial microbiology. Jasmine then continued with another 2-month internship at a governmental institution ("Kantonales Labor, BL") where she worked on a project in collaboration with the Swiss Federal Institute of Aquatic Science and Technology (ETH Zurich) on microbial water safety. Jasmine then moved to Delft, The Netherlands in 2014 to start a PhD project within the group of Industrial Microbiology supervised by Prof. dr. Jack T. Pronk and Prof. dr. ir. Ton van Maris. The research project, the results of which are described in this thesis, focused on the improvement of yeast strains for the production of 2nd generation biofuels.

List of publications

1. **Bracher JM**, Verhoeven MD, Wisselink HW, Crimi B, Nijland JG, Driessen AJ, Klaassen P, Van Maris AJ, Daran J-MG, Pronk JT. The *Penicillium chrysogenum* transporter *Pc*AraT enables high-affinity, glucose-insensitive L-arabinose transport in *Saccharomyces cerevisiae*. *Biotechnology for Biofuels*. 2018;11:63.
2. Jansen ML, **Bracher JM**, Papapetridis I, Verhoeven MD, de Bruijn H, de Waal PP, Van Maris AJA, Klaassen P, Pronk JT. *Saccharomyces cerevisiae* strains for second-generation ethanol production: from academic exploration to industrial implementation. *FEMS Yeast Research*. 2017; doi:10.1093/femsyr/fox044.
3. **Bracher JM**, de Hulster E, Koster CC, van den Broek M, Daran J-MG, van Maris AJ, Pronk JT. Laboratory evolution of a biotin-requiring *Saccharomyces cerevisiae* strain for full biotin prototrophy and identification of causal mutations. *Applied and Environmental Microbiology*. 2017; doi:10.1128/AEM.00892-17.
4. Verhoeven MD, **Bracher JM**, Nijland JG, Bouwknegt J, Daran J-MG, Driessen AJ, van Maris AJ, Pronk JT. Laboratory evolution of a glucose-phosphorylation-deficient, arabinose-fermenting *S. cerevisiae strain* reveals mutations in *GAL2* that enable glucose-insensitive L-arabinose uptake. *FEMS Yeast Research*. 2018; doi:10.1093/femsyr/foy062.
5. **Bracher JM**, Martinez-Rodriguez OA, Dekker WJC, Verhoeven MD, van Maris AJA, Pronk JT. Reassessement of requirements for anaerobic xylose fermentation by engineered, non-evolved *Saccharomyces cerevisiae* strains. *FEMS Yeast Research.* 2018; doi: https://doi.org/10.1093/femsyr/foy104.
6. von Schubert C, Cubizolles F, **Bracher JM**, Sliedrecht T, Kops GJ, Nigg EA. Plk1 and Mps1 cooperatively regulate the spindle assembly checkpoint in human cells. *Cell Reports.* 2015;12:66-78.

978-94-6380-178-2